

# THÈSE

présentée pour obtenir le grade de docteur

de l'École Nationale Supérieure  
des Télécommunications

Département : Eurécom ; Communications mobiles

LUC DENEIRE

## ESTIMATION AVEUGLE DE CANAL ET ACCÈS MULTIPLE PAR RÉPARTITION SPATIALE.

Soutenue le 21 décembre 1998 devant le jury composé de

Président	Claude Gueguen
Rapporteurs	Pierre Comon Philippe Loubaton
Examineurs	Pierre Duhamel Arthur Fawe Constantinos Papadias
Directeur de Thèse	Dirk Slock

**École Nationale Supérieure des Télécommunications**  
Enseignement supérieur des Télécommunications

Écrit avec la classe These-CRIN

*Quand vient le temps béni des remerciements, passé le soulagement d'en arriver à une tâche aussi ultime, vient la peur de l'oubli, de la froideur des mots et de la banalité. Au risque de succomber à cette dernière, je voudrais crier merci, de toutes mes forces, à mes parents, à ma famille, à mes amis qui m'ont permis de faire cette thèse, de quelque manière que ce soit.*

*A mon épouse, qui m'a su porter tout ce temps,*

*A Antonin, qui n'a pas eu trop de peine à me détourner de temps à autres de mon travail,*

*A celui qui vient en même temps que cette thèse,*

*A Dirk Slock, qui m'a supporté tout ce temps,*

*A Elisabeth, qui a partagé mes coups de blues ...,*

*A l'équipe « aveugle » qui m'a accompagné sur mon chemin,*

*A tous ceux d'Eurécom qui ont contribué à la bonne ambiance de travail,*

*Et enfin, à ceux qui me font le plaisir de juger cette thèse,*

**Merci**



*Attendre d'en savoir assez pour agir en toute lumière, c'est se condamner à l'inaction.*

*Anonyme.*



# Table des matières

<b>Chapitre 1 Introduction</b>	<b>1</b>
1.1 Le contexte : les communications mobiles . . . . .	1
1.1.1 Le canal multichemins à évanouissement . . . . .	2
1.1.2 Le cas particulier du GSM . . . . .	4
1.1.3 La modélisation du canal par un filtre FIR . . . . .	4
1.2 Accès Multiple par Répartition Spatiale . . . . .	5
1.3 Les choix de départ et les contributions . . . . .	5
<b>Publications associées à la thèse</b>	<b>7</b>
<b>Chapitre 2 Modèles</b>	<b>11</b>
2.1 Notations, modèle multicanal . . . . .	11
2.1.1 Modèle continu . . . . .	11
2.1.2 Modèle multicanal . . . . .	12
2.1.3 Modèle échantillonné vectoriel . . . . .	13
2.2 Modèle déterministe . . . . .	15
2.3 Modèle gaussien . . . . .	16
2.4 Liste des hypothèses . . . . .	16
2.4.1 Nomenclature des hypothèses . . . . .	16
<b>Chapitre 3 Identifiabilité pour un seul utilisateur</b>	<b>19</b>
3.1 Définition d'identifiabilité . . . . .	19
3.2 Notions préliminaires . . . . .	19
3.2.1 Notion de modes d'excitation . . . . .	19
3.2.2 Canal irréductible . . . . .	20
3.2.3 Paramétrisation minimale de l'espace bruit . . . . .	21
3.3 Modèle déterministe . . . . .	21
3.4 Modèle gaussien . . . . .	23
3.5 Bornes de Cramer-Rao . . . . .	25
3.5.1 Modèle déterministe . . . . .	25

3.5.2	Modèle gaussien . . . . .	26
<b>Chapitre 4</b>	<b>Identifiabilité en multi-utilisateurs</b>	<b>27</b>
4.1	Identifiabilité de la méthode par séquence d'apprentissage . . . . .	27
4.2	Longueur minimale de l'égaliseur FIR Zero-Forcing (ZF) . . . . .	28
4.2.1	Quelques propriétés des matrices polynomiales . . . . .	28
4.2.2	Quelques résultats additionnels. . . . .	29
4.2.3	Longueur minimale de l'égaliseur ZF . . . . .	30
4.2.4	Une borne plus lâche . . . . .	31
4.3	Quelques canaux dégénérés . . . . .	32
4.3.1	Nombre effectif de canaux . . . . .	32
4.3.2	Multitrajets distants. . . . .	32
4.3.3	Canaux à colonnes non réduites en CDMA . . . . .	33
4.3.4	Comparaison entre les bornes. . . . .	33
4.4	Identifiabilité pour les méthodes aveugles . . . . .	34
4.4.1	Modèle déterministe . . . . .	34
4.4.2	Modèle Gaussien . . . . .	36
4.5	Identifiabilité des méthodes semi-aveugles . . . . .	38
4.5.1	Modèle déterministe . . . . .	38
4.5.2	Modèle Gaussien . . . . .	40
4.6	Conclusions . . . . .	41
<b>Chapitre 5</b>	<b>Ajustement des sous-espaces</b>	<b>43</b>
5.1	Introduction . . . . .	43
5.2	Modèle utilisé: Déterministe . . . . .	43
5.3	La méthode . . . . .	43
5.3.1	Ajustement du sous-espace signal . . . . .	44
5.3.2	Paramétrisations linéaires du sous-espace bruit . . . . .	45
5.3.3	Ajustement du sous-espace bruit . . . . .	46
5.4	Méthodes sans décomposition en vecteurs propres . . . . .	46
5.4.1	Ajustement du sous-espace signal . . . . .	46
5.4.2	Ajustement du sous-espace bruit . . . . .	51
5.4.3	D'autres méthodes utilisant la paramétrisation du sous-espace bruit . . . . .	51
5.4.4	Équivalences . . . . .	54
5.4.5	Simulations . . . . .	54
5.5	Robustesse à la surestimation de la longueur du canal . . . . .	55
5.5.1	L'approche de base . . . . .	57
5.5.2	Robustesse . . . . .	57
5.5.3	Simulations . . . . .	59

---

5.6	Conclusions . . . . .	59
<b>Chapitre 6 Prédiction linéaire</b>		<b>61</b>
6.1	Introduction . . . . .	61
6.2	Modèle utilisé : gaussien . . . . .	61
6.2.1	Prédiction linéaire du signal sans bruit . . . . .	62
6.2.2	Le prédicteur linéaire en tant qu'égaliseur Zero-Forcing . . . . .	63
6.3	Prédiction linéaire et identification de canal . . . . .	63
6.3.1	La méthode de la prédiction linéaire pondérée. . . . .	64
6.4	L'inverse de $R_{YY}$ dans la prédiction linéaire. . . . .	64
6.4.1	La méthode «Levinson» . . . . .	64
6.4.2	La méthode «Pseudo-inverse» . . . . .	65
6.4.3	La méthode «Cadzow» . . . . .	65
6.5	Surestimation de la longueur du canal . . . . .	66
6.5.1	Méthode «Pseudo-inverse» . . . . .	66
6.5.2	Méthode «Levinson» . . . . .	67
6.6	Estimation d'ordre . . . . .	68
6.6.1	Estimation de la longueur du canal basée sur les $\lambda_i$ . . . . .	68
6.6.2	Estimation de l'ordre d'un processus VAR . . . . .	68
6.7	Algorithmes conjoints d'estimation d'ordre et de canal . . . . .	69
6.8	Simulations . . . . .	70
6.8.1	Estimation de l'ordre du processus AR . . . . .	70
6.8.2	Estimation du canal . . . . .	71
6.9	Une paramétrisation minimale du sous-espace bruit . . . . .	76
6.10	Conclusions . . . . .	76
<b>Chapitre 7 Identification aveugle par l'algorithme de Schur</b>		<b>79</b>
7.1	Décomposition LDU de la matrice de covariance débruitée . . . . .	79
7.2	Utilisation de l'algorithme de Schur Généralisé . . . . .	80
7.2.1	Notions de base . . . . .	80
7.2.2	Application à la factorisation LDU de $\hat{R}_{YY}$ . . . . .	81
7.3	Complexité . . . . .	82
7.4	Simulations . . . . .	83
7.5	Conclusions . . . . .	85
<b>Chapitre 8 Méthodes au maximum de vraisemblance</b>		<b>87</b>
8.1	Introduction . . . . .	87
8.2	DML : Méthode au Maximum de Vraisemblance dans le modèle Déterministe . . . . .	88
8.2.1	DIQML : un algorithme itératif débruité . . . . .	88

8.2.2	PQML : une méthode itérative Pseudo Quadratique . . . . .	90
8.2.3	DML semi-aveugle . . . . .	90
8.2.4	Simulations . . . . .	91
8.3	GML : Méthode au Maximum de Vraisemblance pour le modèle gaussien . . . . .	93
8.3.1	GML : gaussian Maximum Likelihood, méthodes approchées. . . . .	94
8.3.2	Une méthode optimale en développement . . . . .	96
8.4	Conclusions . . . . .	97
<b>Chapitre 9 Méthodes basées sur les statistiques cycliques</b>		<b>99</b>
9.1	Hypothèses . . . . .	99
9.2	Introduction aux statistiques cycliques . . . . .	99
9.3	Méthode basée sur le modèle cyclostationnaire . . . . .	100
9.3.1	Modèle cyclostationnaire . . . . .	100
9.4	Application de la méthode des sous-espaces . . . . .	101
9.4.1	La méthode classique . . . . .	101
9.4.2	Algorithme sans décomposition vecteurs/valeurs propres . . . . .	102
9.5	Prédiction Linéaire . . . . .	103
9.6	Modèle stationnaire multivarié . . . . .	103
9.6.1	Théorème de Gladyshev et Processus de Miamee . . . . .	103
9.7	Expression de la sortie du nouveau modèle . . . . .	104
9.8	Méthodes basées sur le modèle stationnaire multivarié . . . . .	105
9.9	Identifiabilité . . . . .	106
9.10	Charge de calcul et aspects algorithmiques . . . . .	106
9.11	Simulations . . . . .	107
9.11.1	Modèle cyclostationnaire et méthode sous-espace . . . . .	107
9.11.2	Modèle cyclostationnaire et méthode de la prédiction linéaire . . . . .	110
9.11.3	Modèle stationnaire multivarié et méthode sous-espace . . . . .	111
9.11.4	Modèle stationnaire multivarié et méthode de la prédiction linéaire . . . . .	113
9.12	Conclusions . . . . .	114
<b>Chapitre 10 Simulations</b>		<b>117</b>
10.1	Canaux . . . . .	117
10.2	Méthodes sous-espace . . . . .	118
10.3	Prédiction Linéaire . . . . .	120
10.3.1	Résultats sans surestimation de la longueur du canal . . . . .	122
10.3.2	Résultats avec surestimation de la longueur du canal . . . . .	124
10.4	Méthode de Schur . . . . .	127

---

<b>Chapitre 11 Conclusions</b>	<b>131</b>
11.1 Revue des résultats . . . . .	131
11.2 Directions de recherches futures . . . . .	132
<b>Annexe A Bornes de Cramer-Rao</b>	<b>133</b>
A.1 Quelques notions de base . . . . .	133
A.1.1 Distributions Gaussiennes . . . . .	133
A.1.2 Dérivées par rapport à une variable complexe . . . . .	133
A.2 CRBs pour des variables circulaires complexes . . . . .	133
<b>Bibliographie</b>	<b>135</b>



# Chapitre 1

## Introduction

### 1.1 Le contexte : les communications mobiles

L'objet de cette thèse consiste à explorer des pistes permettant d'améliorer la qualité des communications mobiles et/ou d'augmenter, par des techniques de traitement de signal, la capacité des réseaux. Pour ce faire, nous essayons d'identifier le mieux possible le canal de communication. Dans la définition de ce canal, nous incluons le filtre de mise en forme des signaux après modulation, ce qui nous mène au système schématique suivant :

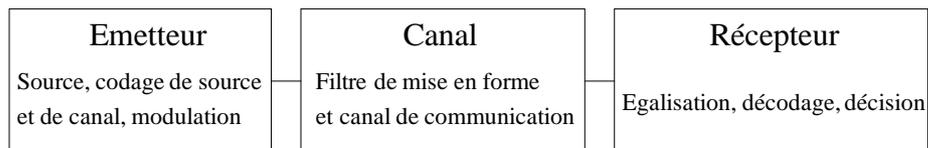


FIG. 1.1 – *Schéma du système de communication.*

D'autre part, dans le contexte des communications radio avec les mobiles, le canal de communication a un certain nombre de caractéristiques qui sont dues principalement, d'une part, aux réflexions sur les obstacles (montagnes, forêts, bâtiments, . . .), qui donnent lieu à la modélisation du signal reçu comme étant une superposition de signaux arrivant à des instants différents, d'autre part à la vitesse des mobiles, qui génère un effet Doppler. La section suivante explicite quelque peu la manière classique de caractériser les canaux mobiles.

### 1.1.1 Le canal multichemins à évanouissement

#### Bande et Temps de cohérence d'un canal, diversités.

Les notions de bande de cohérence et de temps de cohérence d'un canal de transmission permettent une mesure des espacements respectivement fréquentiel et temporel nécessaires entre différentes versions d'un même signal émis pour que celles-ci soient non corrélées ; elles constituent par conséquent des caractéristiques essentielles à connaître lors de l'utilisation des techniques de diversité spatio-temporelle ou fréquentielle.

**Bande de cohérence d'un canal de transmission.** Pour introduire le concept de bande de cohérence [62], on considère un canal de transmission caractérisé par un nombre  $L$  de chemins conduisant de l'émetteur au récepteur. Le canal peut donc être caractérisé par une réponse impulsionnelle équivalente en bande de base de la forme :

$$h(t; \tau) = \sum_{i=1}^L \beta_i(t) e^{-i2\pi(f_c \tau_i(t) + \phi_i(t))} \delta(\tau - \tau_i(t)) \quad (1.1)$$

où  $\beta_i(t)$ ,  $\tau_i(t)$  et  $\phi_i(t)$  sont respectivement les atténuation, variation de phase et délai introduits par le  $i^{\text{ème}}$  chemin. Toutes ces grandeurs sont susceptibles de varier au cours du temps.

Pour caractériser le canal, nous exprimerons les corrélations entre les divers chemins d'une part et les autocorrélations des paramètres à l'intérieur d'un même chemin d'autre part. En admettant que  $h(t, \tau)$  est une fonction aléatoire (avec les conditions d'ergodisme habituelles), on peut écrire sa fonction d'autocorrélation et la transformée de Fourier de celle-ci :

$$\phi_C(\tau; \Delta t) = 0.5E \{h(t; \tau)h^*(t + \Delta t; \tau)\} \quad (1.2)$$

$$\Phi_C(\Delta f; \Delta t) = \int_{-\infty}^{\infty} \phi_C(\tau; \Delta t) e^{-i2\pi f \tau} d\tau \quad (1.3)$$

où nous avons fait l'hypothèse vraisemblable que le canal était stationnaire<sup>1</sup>.

Dans la plupart des cas réels, l'ensemble des valeurs de  $\phi_C(\tau)$ , encore appelé spectre de puissance des délais et de  $\Phi_C(\Delta f; \Delta t = 0) = \Phi_C(\Delta f)$  ont l'allure donnée à la figure 1.2 [62]. On appelle  $T_m$  l'étalement multichemin et  $(\Delta f)_c \simeq \frac{1}{T_m}$  la bande de cohérence du canal.

---

1. Dans le cas contraire, on aurait  $\phi_C(\tau_1, \tau_2; \Delta t) = 0.5E \{h(t; \tau_1)h^*(t + \Delta t; \tau_2)\}$ , i.e. que les atténuations et les variations de phases associées à des chemins différents sont corrélées.

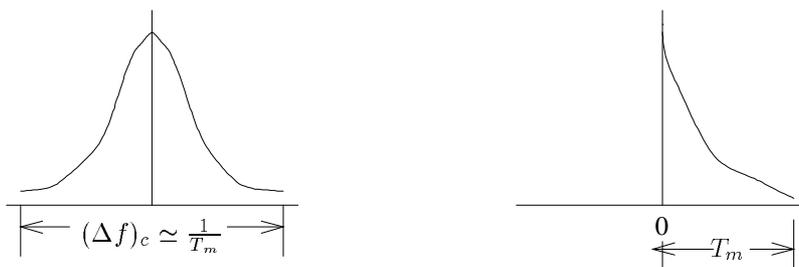


FIG. 1.2 – Étalement multichemin et bande de cohérence.

**Temps de cohérence d'un canal de transmission.** Le concept de temps de cohérence est introduit en suivant une démarche similaire. On définit dans un premier temps la fonction suivante, transformée de Fourier de  $\Phi_c(\Delta f; \Delta t)$ , dont l'allure est donnée en figure 1.3 [62] :

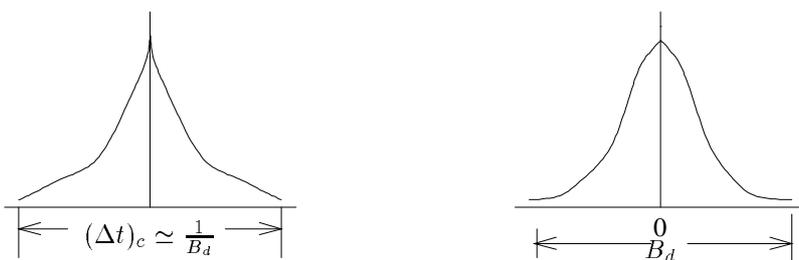


FIG. 1.3 – Temps de cohérence et bande d'étalement Doppler.

$$\psi_C(\Delta f; \lambda) = \int_{-\infty}^{\infty} \Phi_c(\Delta f; \Delta t) e^{-i2\pi\lambda\Delta t} d\Delta t \quad (1.4)$$

Cette fois-ci, nous nous attacherons à  $\psi_C(\Delta f = 0; \lambda) = \psi_C(\lambda)$ . La variable  $\lambda$  est appelée fréquence Doppler car elle est liée aux variations  $\Delta t$  qui tiennent compte des modifications de la réponse impulsionnelle du canal de transmission liée aux effets Doppler.  $B_d$  est la *bande d'étalement Doppler* du canal. Quant à  $(\Delta t)_C \simeq \frac{1}{B_d}$ , il est appelé *temps de cohérence* du canal de transmission. Sa signification est la suivante : si deux signaux sont émis à des instants  $t_1$  et  $t_2$  séparés d'une valeur supérieure à  $(\Delta t)_C$ , ils seront transmis dans un canal présentant deux configurations différentes puisque la réponse impulsionnelle au temps  $t_1$  sera différente de celle au temps  $t_2$ . Les effets du passage dans le canal de transmission de ces signaux seront alors non corrélés.

Si la durée d'émission d'un symbole  $T_b$  d'un signal est inférieure à  $(\Delta t)_C$ , tous les paramètres de gains, de variations de phase et de délais pourront être considérés comme constants. On parle alors de **canal à évanouissement progressif (slowly fading channel)**.

La bande de cohérence d'un canal de transmission est quant à elle une mesure de l'espace-temps fréquentiel nécessaire pour que deux signaux identiques soient non-corrélés. Dès lors, si un signal émis occupe une bande de fréquence  $B = \frac{1}{T_c} < B_C$ , il sera seulement caractérisé par trois paramètres : un gain, une variation de phase et un délai. Nous sommes donc en présence d'un seul chemin *résolvable*. Lorsque  $B < B_C$ , on parle de **canal à évanouissement non sélectif**, sinon, on parle de **canal à évanouissement sélectif**. Le nombre  $L$  de chemins nécessaires à la modélisation de l'ensemble des chemins contenus dans le canal est alors donné par :

$$L = \left\lfloor \frac{W}{B_C} \right\rfloor + 1 = \left\lfloor \frac{T_m}{T_c} \right\rfloor + 1 \quad (1.5)$$

Les  $L - 1$  chemins supplémentaires sont appelés chemins résolubles. Leur existence peut être mise à profit pour obtenir artificiellement plusieurs versions différentes d'un même signal ; ils permettent donc d'obtenir une forme de *diversité* fréquentielle.

### 1.1.2 Le cas particulier du GSM

Dans le cadre européen, c'est la norme GSM pour les radio-mobiles qui nous intéresse. Le spectre utilisé est divisé en 124 canaux d'une largeur de 200 kHz pour la liaison montante (du mobile vers la station de base), situés entre 890 et 915 MHz et de même pour la liaison descendante, entre 935 et 960 MHz. La technique d'accès multiple est une combinaison d'accès par répartition temporelle et fréquentielle avec, en outre, des sauts de fréquences périodiques permettant d'éviter les effets catastrophiques des évanouissements lents et profonds.

En GSM, un mobile reçoit 148 bits de durée  $3.69 \mu s$ , dans une trame de longueur 0.577 ms, ces bits étant modulés en GMSK, un bit correspond à un symbole. En suivant le modèle GSM [30], à une vitesse de 100 km/h, la bande Doppler est de 83 Hz, ce qui correspond à un temps de cohérence de l'ordre d'une dizaine de millisecondes. On en déduit que, pendant la durée d'une trame, le canal peut être considéré comme étant stationnaire (et donc, les techniques bloc que nous avons considérées sont valides). D'autre part, l'étalement multichemin, pour les divers modèles proposés (correspondants aux divers terrains : montagneux, rural, citadin) est de l'ordre de  $20 \mu s$ , soit au maximum six périodes symbole.

En fonction de ces données, nous pouvons dire que le canal GSM est un canal à évanouissement sélectif et progressif.

### 1.1.3 La modélisation du canal par un filtre FIR

Un canal de transmission peut être considéré, en l'absence de bruit, comme un simple filtre. Donc, la réponse impulsionnelle de ce filtre étant, dans notre modèle, finie, on peut simplement

considérer le canal, dans sa version échantillonnée, comme étant un filtre FIR (Finite Impulse Response). Le lien entre la version paramétrisée comme dans [30] et le filtre FIR se faisant simplement par rééchantillonnage de la réponse impulsionnelle continue du canal.

## 1.2 Accès Multiple par Répartition Spatiale

Si on considère les trois techniques actuelles d'accès multiple (par répartition fréquentielle, par répartition temporelle et par répartition de code), on constate que seule la dimension spatiale n'est pas prise en compte, du moins explicitement. En effet, l'agencement cellulaire, voire la sectorisation des antennes de certaines stations de base, est en fait une manière, primitive certes, de faire de l'accès multiple par répartition spatiale (SDMA : Space Division Multiple Access).

Le problème posé par le SDMA peut s'énoncer de la manière suivante: "*Nous voulons séparer plusieurs utilisateurs émettant à la même fréquence, en même temps et sans étalement de spectre, en fonction de l'endroit où ils se trouvent.*" L'objectif du SDMA étant clairement d'augmenter la capacité du réseau radio-mobile en investissant au mieux la dimension spatiale. La seconde partie de cette thèse s'attaquera à ce problème, dans la limite des choix de départ que nous nous fixons.

## 1.3 Les choix de départ et les contributions

Devant la littérature abondante dans ce domaine et les différentes approches, nous avons opté pour certaines voies que nous avons explorées et auxquelles nous avons apporté quelques contributions. Pour des raisons essentiellement "historiques" (liées aux activités du labo), nous avons tout d'abord opté pour les méthodes utilisant les statistiques d'ordre deux, que ce soit explicitement (méthodes sous-espace et assimilées, méthodes de prédiction linéaire et de Schur) soit implicitement (méthodes au maximum de vraisemblance basées) conjointement à l'utilisation explicite du signal reçu.

D'autre part, nous nous sommes attachés aux méthodes blocs, i.e. que les méthodes développées fonctionnent dans le cadre d'un canal à évanouissement sélectif et progressif (canal stationnaire pendant la durée de communication).

Dans ce champ d'investigation, nous avons principalement été préoccupé par certaines "idées fixes" :

1. La complexité de calcul, qui nous a mené à développer un algorithme d'ajustement des sous-espaces sans décomposition vecteurs/valeurs propres, aboutissant également sur un certain nombre d'équivalences entre différents algorithmes décrits dans la littérature. Cette

préoccupation nous a également mené à investiguer les différents algorithmes (rapides pour Levinson, moins pour les autres) liés à la méthode de prédiction linéaire et à la détermination du canal par l'algorithme de Schur généralisé.

2. La robustesse à la surestimation de la longueur du canal : certaines méthodes (du type sous-espace) sont réputées non robustes, dans le sens où elles cessent de donner un estimateur consistant du canal lorsque sa longueur est surestimée, d'autres (du type prédiction linéaire) sont réputées robustes, avec parfois quelques divergences sur le sens à apporter à ce terme. Dans ce cas, nous avons investigué plus avant le sens de cette robustesse, débouchant sur un algorithme rapide fournissant conjointement la longueur du canal et l'estimée du canal. En se basant sur cette approche, nous avons également proposé un algorithme de type sous-espace bénéficiant de certaines propriétés de robustesse similaires à celle des méthodes de prédiction linéaire.
3. L'indépendance par rapport à la couleur du bruit nous a incité à explorer les méthodes basées sur les statistiques cycliques et a donné naissance, d'une part à des algorithmes donnant des résultats indépendants de cette couleur, mais de performances médiocres ; d'autre part à un modèle de canal permettant d'utiliser directement les algorithmes dérivés pour les statistiques d'ordre deux classiques.

Chaque méthode proposée sera accompagnée d'un ensemble de simulations permettant d'illustrer les caractéristiques et les performances des algorithmes utilisés. Enfin, un ensemble de simulations plus complet (du point de vue Monte-Carlo) sera présenté en fin de thèse, pour permettre de comparer les différentes méthodes proposées.

# Publications associées à la thèse

- [1] Luc Deneire and Dirk T.M. Slock. «Blind channel identification based on cyclic statistics». *IEE Proceedings - Radar, Sonar and Navigation, Special issue on Antenna array processing techniques*, 145(1):58–62, February 1988.

## Article de Journal en préparation

- [2] Luc Deneire, Jahouar Ayadi, and Dirk T.M. Slock. «Weighted and Unweighted Subspace Fitting without Eigendecomposition». *To be submitted to Signal Processing*, 1998.

## Article de Journal en soumission

- [3] Luc Deneire, Irfan Ghauri, and Dirk T.M. Slock. «Minimum-Length Zero-Forcing Equalization for Multi-user communications». *Submitted to Transactions on Signal Processing*, 1998.

## Articles de conférence

- [4] Luc Deneire and Dirk T.M. Slock. «Second-order Cyclic Statistics based blind channel identification and equalization». In *Proceedings of Signal Processing Advances in Wireless Communications workshop, Paris, France*, pages 21–24. IEEE, April 1997.
- [5] Luc Deneire, Jaouhar Ayadi, and Dirk T.M. Slock. «Subspace Fitting Without Eigendecomposition». In *13<sup>th</sup> international Conference on Digital Signal Processing*, Santorini, Greece, July 1997.
- [6] Luc Deneire and Dirk T.M. Slock. «Linear Prediction And Subspace Fitting Blind Channel Identification Based On Cyclic Statistics.». In *13<sup>th</sup> international Conference on Digital Signal Processing*, Santorini, Greece, July 1997.
- [7] Jaouhar Ayadi, Luc Deneire, and Dirk T.M. Slock. «Channel-based Blind Multichannel Identification without Order Overestimation Problems». In *13<sup>th</sup> international Conference on Digital Signal Processing*, Santorini, Greece, July 1997.

- [8] Jaouhar Ayadi, Luc Deneire, and Dirk T.M. Slock. «Une Approche Robuste à la Surdétermination de l'Ordre du Canal pour l'Estimation Aveugle des Canaux Multiples RIF». In *Seizième colloque GRETSI sur le traitement du signal et des images*, Grenoble, France, September 1997.
- [9] Luc Deneire and Dirk T.M. Slock. «On robustness of Linear Prediction based blind identification». In *IEEE 5 th Symposium on Communications and Vehicular Technolgy in the Benelux*, Twente, Netherlands, October 1997.
- [10] Luc Deneire, Jahouar Ayadi, and Dirk T.M. Slock. «Weighted and Unweighted Subspace Fitting without Eigendecomposition». In *IEEE 5 th Symposium on Communications and Vehicular Technolgy in the Benelux*, Twente, Netherlands, October 1997.
- [11] Luc Deneire and Dirk Slock. «Gaussian Maximum Likelihood Blind Multichannel Multiuser Identification». In *48th annual Vehicular Technology Conference*, Ottawa, Canada, May, 18–21 1998.
- [12] Elisabeth de Carvalho, Luc Deneire, and Dirk Slock. «Blind and Semi-Blind Maximum Likelihood Techniques for Multiuser Multichannel identification». In *European Association for Signal Processing EUSIPCO-98*, Island of Rhodes, Greece, September, 8–11 1998.
- [13] Luc Deneire, Elisabeth de Carvalho and Dirk Slock. «Identifiability conditions for blind and semi-blind multiuser multichannel identification». In *9th IEEE Signal Processing Workshop On Statistical Signal And Array Processing*, Portland, Oregon, USA, September 1998.

**Article de conférence en soumission**

- [14] Luc Deneire and Dirk Slock. «A Schur method for multiuser multichannel blind identification,». In *Icassp 99*.

---

## Notations

De manière à donner un sens aux choix typographiques, et à distinguer au mieux la nature des grandeurs rencontrées, nous avons opté pour les conventions reprises dans le tableau ci-dessous.

Nature de la grandeur	Notation	Exemple
Domaine temporel	Italique	$h(k), \mathbf{A}(k)$
Domaine en z	Droit	$h(z), \mathbf{A}(z)$
Signaux scalaires	Minuscules	$y_i(k), y(z)$
Signaux vecteurs	Minuscules gras	$\mathbf{y}(k), \mathbf{y}(z)$
Vecteurs de signaux	Majuscules	$A_{j,N_j}(k), \mathbf{A}_N(k)$
Matrices	Majuscules simples	$R_{YY}$

### Notations diverses

$\text{ap}1$	:	Avec probabilité 1
$I, I_m$	:	Matrice identité, matrice identité de taille $m \times m$
$(.)^*$	:	Complexe conjugué
$(.)^T$	:	Transposée
$(.)^H$	:	Transposée hermitienne : i.e. $(.)^{*T}$
$(.)^t$	:	(.) étant formé de blocs, dénote la transposition des blocs de (.)
$(.)^\#$	:	Pseudoinverse (le plus souvent de Moore-Penrose)
$E\{.\}$	:	Espérance mathématique
$\mathcal{N}(m, c)$	:	Distribution normale de moyenne $m$ et de covariance $c$
$\lceil . \rceil$	:	Valeur entière supérieure
$\ .\ $	:	Norme

$\text{vec}(A)$	:	Vecteur comprenant les éléments de la matrice $A$ , colonne par colonne
$x^\dagger(z)$	:	Paraconjugué: $x^H(z^{-1})$ , i.e., si $x(z) = \sum_i x_i z^{-i}$ , $x^\dagger(z) = \sum_i x_i^H z^i$
$\mathcal{T}_L([\mathbf{h}(0) \cdots \mathbf{h}(N-1)])$	:	$= \begin{bmatrix} \mathbf{h}(0) & \mathbf{h}(1) & \cdots & 0 & 0 \\ 0 & \mathbf{h}(0) & \cdots & 0 & 0 \\ \vdots & \cdots & \cdots & \vdots & \vdots \\ 0 & 0 & \cdots & \mathbf{h}(N-2) & \mathbf{h}(N-1) \end{bmatrix}_{L \text{ lignes}}$ Matrice de convolution mono-utilisateur
$\mathcal{T}_L^p([\mathbf{H}_1 \cdots \mathbf{H}_p])$	:	$= [\mathcal{T}_L(\mathbf{H}_1) \cdots \mathcal{T}_L(\mathbf{H}_p)]$ Matrice de convolution pour $p$ utilisateurs
$\otimes$	:	Produit de Kronecker: $X \otimes Y = \begin{bmatrix} X_{11} Y & \cdots & X_{1n} Y \\ \vdots & \ddots & \vdots \\ X_{n1} Y & \cdots & X_{nn} Y \end{bmatrix}$
$\Re(\cdot), \Im(\cdot)$	:	Parties réelle et imaginaire

## Acronymes

De manière à ne pas rendre la lecture plus difficile par la francisation systématique des acronymes, nous avons gardé la version anglaise de la plupart de ceux-ci.

SNR	:	Signal to Noise Ratio
GSM	:	Global System for Mobile Communications
FIR	:	Finite Impulse Response
GMSK	:	Gaussian Minimum-Shift Keying
SISO – SIMO	:	Single Input Single Output – Single Input Multiple Output
MIMO	:	Multiple Input Multiple Output
CRB	:	Cramer-Rao Bound
SDMA	:	Space Division Multiple Acces Spatiale
CDMA	:	Code Division Multiple Acces de code
DS-CDMA	:	DS-CDMA Direct Sequence-CDMA
UMTS	:	Universal Mobile Telecommunication System

# Chapitre 2

## Modèles

|| *Ce chapitre introduit les différents modèles utilisés, les notations y afférentes et l'ensemble des hypothèses reprises dans le corps de la thèse.*

### 2.1 Notations, modèle multicanal

#### 2.1.1 Modèle continu

Nous nous plaçons dans un contexte de communications, dans un premier temps avec un seul émetteur (mono-utilisateur) et ensuite avec  $p > 1$  émetteurs (multi-utilisateurs) émettant en même temps, sur la même fréquence et en bande étroite, le seul paramètre permettant de les distinguer étant leur localisation spatiale. D'autre part, nous utiliserons un certain nombre d'hypothèses, regroupées en section 2.4.1 et précisées chaque fois que c'est nécessaire.

Nous considérons une modulation numérique linéaire (l'extension au cas important du GSM peut être faite en suivant [74]) et un canal linéaire à bruit additif. La réception se fait par un réseau de  $m_a$  antennes. Les signaux reçus par chaque élément d'antenne peuvent être écrits, en équivalent bande de base :

$$y_i(t) = \sum_{j=1}^p \sum_k a_j(k) h_{ij}(t - kT) + v_i(t), \quad i = 1, \dots, m_a \quad (2.1)$$

où les  $a_j(k)$  sont les symboles transmis par la source  $j$ ,  $T$  est la période des symboles, commune à tous les utilisateurs,  $h_{ij}(t)$  est la réponse impulsionnelle de l'émetteur  $j$  à l'antenne réceptrice  $i$ . Nous supposons que  $\{a_j(k)\}$  et  $\{v_i(t)\}$  sont faiblement (i.e. à l'ordre 2) stationnaires et mutuellement faiblement stationnaires, de sorte que les processus  $\{y_i(t)\}$  sont faiblement cyclostationnaires de période  $T$ . Si les  $\{y_i(t)\}$  sont échantillonnés à la période  $T$ , le processus (vectoriel) échantillonné est faiblement stationnaire. Si les  $\{y_i(t)\}$  sont (sur-)échantillonnés au

rythme  $m_e/T$ , le processus est faiblement cyclostationnaire de période cyclique  $m_e$ . Ce type de suréchantillonnage nous donne également un modèle multicanal explicité ci-dessous.

Si nous sommes en présence d'un seul utilisateur, l'indice  $j$  est omis.

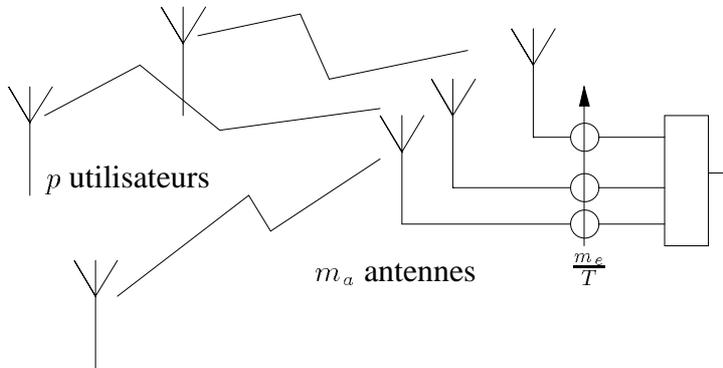


FIG. 2.1 – Schéma de principe de la réception

### 2.1.2 Modèle multicanal

La nature "multicanal" du modèle est claire pour le cas de plusieurs antennes (figure 2.2).

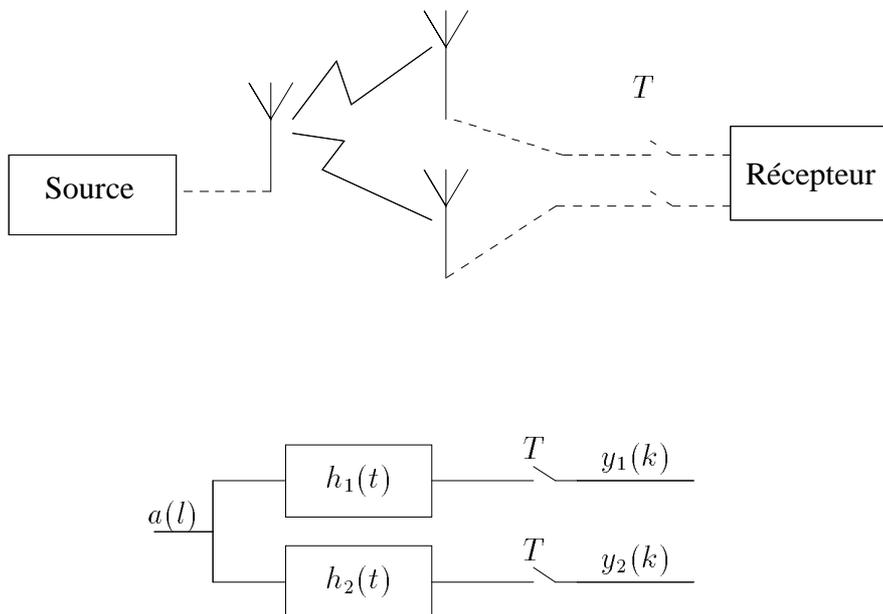
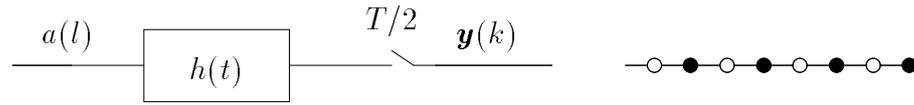
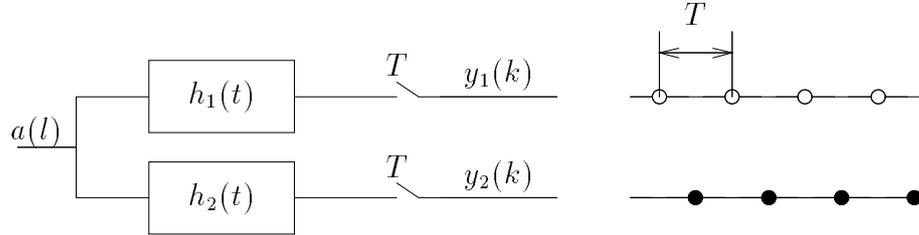


FIG. 2.2 – Exemple de deux antennes

Nous rappelons brièvement comment le suréchantillonnage mène également à un modèle multicanal (figure 2.3).



Modèle à un canal, au double du rythme des symboles.



Modèle à deux canaux, au rythme des symboles.

FIG. 2.3 – Exemple de suréchantillonnage par 2

Dans la figure 2.3, pour un système à une antenne et suréchantillonnage par deux, les séries  $\{y_i(k)\}$  et  $\{\mathbf{y}(k) = [y_1(k) y_2(k)]^T\}$  peuvent être écrites sous la forme :

$$y_i(k) = \sum_{l=-\infty}^{\infty} a(l)h_i(k-l) \quad i = 1, 2 \quad (2.2)$$

qui indique que le système original à une entrée et une sortie (SISO : Single Input Single Output) peut être remplacé par un système à une entrée et (ici) deux sorties (SIMO : Single Input Multiple Output). Dans le cas multi-utilisateurs, nous obtenons donc un système MIMO (Multiple Input Multiple Output).

Pour simplifier, nous pouvons donc considérer simplement que nous avons un système SIMO ou MIMO dont la nature des canaux multiples est spatiale et/ou temporelle due à la multiplicité des antennes et/ou au suréchantillonnage.

### 2.1.3 Modèle échantillonné vectoriel

Après échantillonnage, nous pouvons obtenir les équations suivantes, où les canaux multiples sont dus soit à la multiplicité des antennes, soit au suréchantillonnage, soit aux deux. Le nombre de canaux est dénoté  $m = m_e m_a$ . Nous supposons en outre que les canaux sont à réponse impulsionnelle finie (RIF ou FIR : Finite Impulse Response), cette hypothèse est bien remplie dans le contexte important des communications mobiles. Nous noterons  $N_j$  la longueur de la réponse impulsionnelle du canal vectoriel reliant la source  $j$  au récepteur multicanal et  $N = \sum_{j=1}^p N_j$ , la longueur totale des canaux. Nous pouvons supposer, sans perte de généralité, que

les canaux des différents utilisateurs ont leur première valeur non nulle au temps 0 et que  $N_1 \geq N_2 \geq \dots \geq N_p$ .

Le signal échantillonné au récepteur peut alors s'écrire :

$$\begin{aligned}
 \mathbf{y}(k) &= \sum_{j=1}^p \sum_{i=0}^{N_j-1} \mathbf{h}_j(i) a_j(k-i) + \mathbf{v}(k) \\
 &= \sum_{i=0}^{N_1-1} \mathbf{H}(i) \mathbf{a}(k-i) + \mathbf{v}(k) \\
 &= \sum_{j=1}^p \mathbf{H}_j \mathbf{A}_{j,N_j}(k) + \mathbf{v}(k) \\
 &= \mathbf{H} \mathbf{A}_N(k) + \mathbf{v}(k)
 \end{aligned} \tag{2.3}$$

$$\begin{aligned}
 \mathbf{y}(k) &= \begin{bmatrix} y_1(k) \\ \dots \\ y_m(k) \end{bmatrix}, \mathbf{v}(k) = \begin{bmatrix} v_1(k) \\ \dots \\ v_m(k) \end{bmatrix}, \mathbf{a}(k) = \begin{bmatrix} a_1(k) \\ \dots \\ a_p(k) \end{bmatrix}, \\
 \mathbf{A}_{jn}(k) &= \begin{bmatrix} a_j(k) \\ \dots \\ a_j(k-n+1) \end{bmatrix}, \mathbf{A}_{N+p+i}(k) = \begin{bmatrix} A_{1(N_1+i)}(k) \\ \dots \\ A_{p(N_p+i)}(k) \end{bmatrix}, \mathbf{h}_j(k) = \begin{bmatrix} h_{1j}(k) \\ \dots \\ h_{mj}(k) \end{bmatrix}, \\
 \mathbf{H}_j &= [\mathbf{h}_j(0) \dots \mathbf{h}_j(N_j-1)], \mathbf{H} = [\mathbf{H}_1 \dots \mathbf{H}_p], \\
 \mathbf{H}(k) &= [\mathbf{h}_1(k) \dots \mathbf{h}_p(k)], \mathbf{H}_{ij} = \text{ligne } i \text{ de } \mathbf{H}_j \\
 \mathbf{H}(z) &= \sum_i \mathbf{H}(i) z^{-i} \quad \mathbf{h}(z) = \sum_i \mathbf{H}_1(i) z^{-i}
 \end{aligned} \tag{2.4}$$

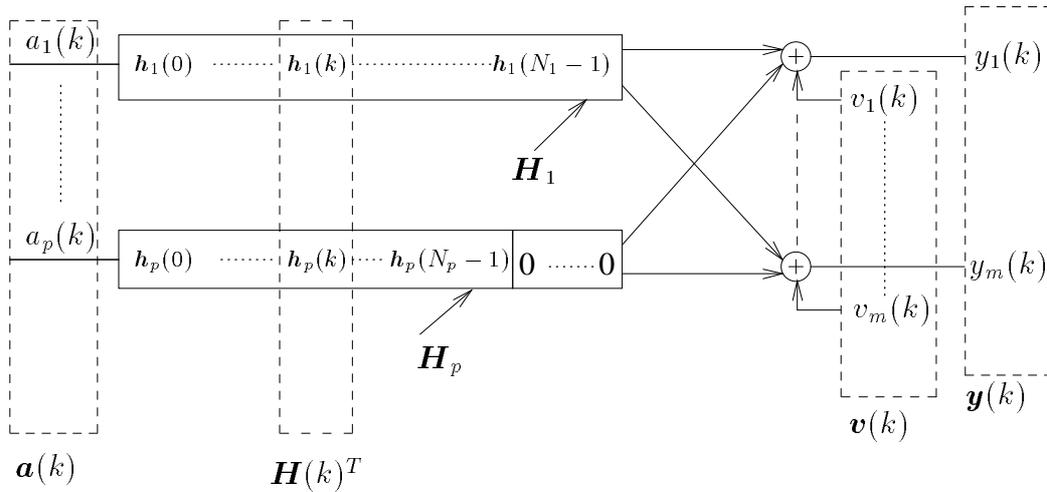


FIG. 2.4 – Schéma des notations.

Nous considérons un bruit additif stationnaire de matrice de corrélation  $R_{vv}(k-i) = E\{\mathbf{v}(k)\mathbf{v}^H(i)\}$  et de moyenne nulle. D'autre part, nous supposons que nous recevons  $M$  échantillons, soit :

$$\mathbf{Y}_M(k) = \mathcal{T}_M^p(\mathbf{H}) \mathbf{A}_{N+p(M-1)}(k) + \mathbf{V}_M(k) \quad (2.5)$$

où  $\mathbf{Y}_M(k) = [\mathbf{y}^H(k) \cdots \mathbf{y}^H(k-M+1)]^H$  et  $\mathbf{V}_M(k) = [\mathbf{v}^H(k) \cdots \mathbf{v}^H(k-M+1)]^H \cdot \mathcal{T}_M^p(\mathbf{H})$  est la matrice de convolution de  $\mathbf{H}$ , de taille  $(mM) \times (N+p(M-1))$ .

De façon plus détaillée, et pour expliciter la forme de la matrice de convolution multi-utilisateurs et multicanal, on peut réécrire l'expression l'équation  $\mathcal{T}_M^p(\mathbf{H})$  dans l'équation (2.5) sous la forme :

$$\mathcal{T}_M^p(\mathbf{H}) = [\mathcal{T}_M(\mathbf{H}_1) | \cdots | \mathcal{T}_M(\mathbf{H}_p)]$$

$$\mathcal{T}_M(\mathbf{H}_1) = \begin{bmatrix} \mathbf{h}_1(0) & \cdots & \mathbf{h}_1(N_1-1) & 0 & \cdots & 0 \\ 0 & \mathbf{h}_1(0) & \cdots & \mathbf{h}_1(N_1-1) & \cdots & 0 \\ \vdots & \ddots & \ddots & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & \ddots & \ddots & \vdots \\ 0 & \cdots & 0 & \mathbf{h}_1(0) & \cdots & \mathbf{h}_1(N_1-1) \end{bmatrix} \quad (2.6)$$

La structure de la matrice de covariance du signal reçu  $\mathbf{Y}(k)$  est

$$C_{YY} = \mathcal{T}_M^p(\mathbf{H}) C_{AA} \mathcal{T}_M^{pH}(\mathbf{H}) + C_{VV} \quad (2.7)$$

où  $C_{AA} = E\{\mathbf{A}_{N+p(M-1)}(k)\mathbf{A}_{N+p(M-1)}^H(k)\}$ ,  $C_{VV}$  et  $C_{YY}$  sont définis de façon similaire.

De manière à simplifier les notations au maximum, nous réécrivons l'équation (2.5) sous la forme :

$$\mathbf{Y} = \mathcal{H}\mathbf{A} + \mathbf{V} \quad (2.8)$$

## 2.2 Modèle déterministe

Dans l'établissement des équations du modèle multicanal, à part l'hypothèse de stationnarité du signal vectoriel  $\mathbf{Y}$  et du bruit  $\mathbf{V}$  (et, implicitement, l'hypothèse de stationnarité du canal pendant la durée de la transmission), nous n'avons pas émis d'hypothèses menant à un modèle à proprement parler. Le premier modèle que nous allons (brièvement) décrire est le modèle déterministe. Dans ce modèle, les symboles d'entrée sont considérés déterministes : i.e. ce sont des symboles donnés de valeur inconnue. La partie aléatoire du modèle provient uniquement du

bruit considéré simplement stationnaire. Dans la plupart des cas, ce bruit sera considéré comme spatialement et temporellement blanc.

D'autre part, dans le cas multi-utilisateurs, nous traiterons l'identification semi-aveugle. Pour ce faire, nous pouvons détailler l'équation (2.8) :

$$\mathbf{Y} = \mathcal{H} \mathbf{A} + \mathbf{V} = \mathcal{H}_u \mathbf{A}_u + \mathcal{H}_k \mathbf{A}_k + \mathbf{V} \quad (2.9)$$

où  $\mathbf{A}_u$  et  $\mathbf{A}_k$  sont respectivement les symboles inconnus et connus.

## 2.3 Modèle gaussien

Le modèle stochastique général considère que les séries temporelles formées par les symboles d'entrée sont des processus aléatoires. Stricto sensu, ce sont des processus, le plus souvent i.i.d. (indépendants et identiquement distribués), de densité de probabilité discrète.

Dans cette thèse, nous nous concentrons sur les méthodes du second ordre, ce qui nous amène à adopter un modèle gaussien, où les symboles sont considérés comme étant i.i.d. et de densité de probabilité gaussienne.

$$\mathbf{A} \sim \mathcal{N}(\mathbf{A}_k, C_{AA}) \quad (2.10)$$

Le bruit est également considéré comme gaussien:  $\mathbf{V} \sim \mathcal{N}(0, C_{VV})$ , ce qui complète le modèle et les signaux reçus ont donc la distribution :

$$\mathbf{Y} \sim \mathcal{N}(\mathcal{H}_k \mathbf{A}_k, C_{YY}) \quad ; \quad C_{YY} = \mathcal{H}_u C_{AA} \mathcal{H}_u + C_{VV} \quad (2.11)$$

D'autre part, nous supposons la plupart du temps que les symboles sont mutuellement i.i.d. et ont la même puissance.

$$C_{AA} = \sigma_a^2 I. \quad (2.12)$$

Cette approche à été considérée (pour le cas mono-utilisateur) dans [19].

## 2.4 Liste des hypothèses

### 2.4.1 Nomenclature des hypothèses

«– **Hstat** :  $\{a_j(k)\}$  et  $\{v_i(t)\}$  sont faiblement stationnaires et mutuellement faiblement stationnaires.

- «– **H $\mathcal{H}$ full** : La matrice  $\mathcal{H} = \mathcal{T}_M^p(\mathbf{H})$  est de rang colonne plein, pour  $M \geq \bar{L}$ . Les chapitres concernant l'identifiabilité précisent cette notion.
- «– **Hbblanc** :  $\{\mathbf{V}(k)\}$  est un bruit vectoriel spatialement et temporellement blanc de matrice de covariance  $R_{VV} = \sigma_v^2 I$  et de moyenne nulle.
- «– **Hpex** : Les sources  $\{a_j(k)\}$  présentent un nombre de modes d'excitation suffisants. Cette notion est précisée dans les chapitres concernant les conditions d'identifiabilité.
- «– **Hgauss** : Les sources et le bruit sont considérés gaussiens.



## Chapitre 3

# Identifiabilité pour un seul utilisateur

|| Ce chapitre dérive les conditions d'identifiabilité sur les canaux et les symboles d'entrée, dans le cas aveugle et pour un seul utilisateur. Les extensions au cas semi-aveugle sont données par de Carvalho dans [17]. D'autre part, nous donnons également les bornes de Cramer-Rao.

### 3.1 Définition d'identifiabilité

Nous adoptons la définition d'identifiabilité présentée dans [45].

**Définition 3.1** Soit  $\theta$ , le paramètre à estimer, et  $Y$ , le vecteur d'observations :

$$\theta \text{ est identifiable ssi : } \forall Y, f(Y|\theta) = f(Y|\theta') \Rightarrow \theta = \theta' \quad (3.1)$$

Dans les modèles que nous avons adoptés,  $f(Y|\theta)$  décrit une distribution de probabilité gaussienne et l'identifiabilité au sens (3.1) équivaut à l'identifiabilité à partir de la moyenne et de la covariance de  $Y$ .

### 3.2 Notions préliminaires

#### 3.2.1 Notion de modes d'excitation

La notion de modes d'excitation d'une séquence finie correspond à la notion de fréquence dans un signal continu, i.e., les modes d'excitation forment des fonctions de base qui décrivent toute séquence de longueur finie. Nous reprenons ici essentiellement les définitions de [46].

**Définition 3.2** Un mode est une séquence finie de la forme  $m(k) = k^n z^k$ , pour un nombre complexe  $z$  et  $n$  entier positif ou nul. On appelle  $z$  la racine du mode et  $n$  son ordre.

Une séquence  $\{s(k), k = -L, \dots, N - 1\}$  de longueur  $N + L$  a  $o$  modes (distincts) si on peut écrire  $s(k)$  sous la forme d'une combinaison linéaire de  $o$  modes de longueur  $N + L$  :

$$s(k) = \sum_{i=1}^o c_i m_i(k) \quad (3.2)$$

où, si plusieurs modes ont la même racine (non nulle), seul le mode d'ordre le plus élevé doit avoir un coefficient non nul. Par exemple, la séquence  $c_1 k^2 z_1^k + c_2 k z_1^k + c_3 z_1^k + c_4 z_2^k$  a quatre modes tant que  $c_1$  et  $c_4$  sont non nuls.

Les modes jouissent des propriétés suivantes :

- «– Tout ensemble de  $o$  modes distincts de longueur  $r$  sont indépendants tant que  $o \leq r$ .
- «– Toute séquence de longueur  $r$  peut s'exprimer comme étant la combinaison linéaire de  $r$  modes de longueur  $r$ .
- «– Une séquence de longueur  $r$  a, avec une probabilité unitaire,  $r$  modes de longueur  $r$ .
- «– Soit un mode de racine  $z$  et d'ordre  $n$  au temps  $k + 1$ , ce mode est une combinaison linéaire de modes de la même racine  $z$  et d'ordres  $0, 1, \dots, n$  à l'instant  $k$  :

$$(k + 1)^n z^{k+1} = \sum_{i=0}^n \frac{n!}{(n - i)! i!} k^i z^k \quad (3.3)$$

### 3.2.2 Canal irréductible

**Définition 3.3** *Un canal  $\mathbf{h}(z)$  est dit irréductible s'il est de rang plein  $\forall z$ .*

On se convainc sans peine que  $\mathbf{h}(z)$  est irréductible si les sous-canaux  $\mathbf{h}_i(z)$ ,  $i = 1 \dots m$  n'ont pas de zéros en commun. Si on considère la matrice  $\mathcal{T}_M(\mathbf{h})$ , de taille  $mM \times M + N - 1$  avec  $mM \geq M + N - 1$ , la théorie des matrices de Sylvester (voir e.g. [49]) nous apprend que  $\mathcal{T}_M(\mathbf{h})$  est de rang plein si et seulement si  $M \geq N - 1$  et  $\mathbf{h}_i(z)$ ,  $i = 1 \dots m$  n'ont pas de zéros en commun, i.e. si  $\mathbf{h}(z)$  est irréductible. D'autre part, on peut énoncer le théorème suivant :

**Théorème 3.1** *Soit  $\mathbf{H}(z)$ , et  $\mathcal{T}_M(\mathbf{H})$  où  $M$  est tel que  $\mathcal{T}_M(\mathbf{H})$  a plus de lignes que de colonnes, alors  $\mathcal{T}_M(\mathbf{H})$  est de rang colonne plein avec probabilité 1.*

Intuitivement, un canal choisi aléatoirement remplit un espace de dimension  $m \times N$ . Les conditions générées par le fait de forcer  $\mathcal{T}_M(\mathbf{H})$  à ne pas être de rang plein vont confiner le canal dans un sous-espace de l'espace global, et donc, la probabilité qu'un canal choisi aléatoirement soit dans ce sous-espace sera nulle. Pour une démonstration plus formelle, voir [43].

◁

Donc,  $\mathcal{T}_M(\mathbf{H})$  est de rang colonne plein ssi

«-  $\mathbf{h}(z)$  est irréductible ;

«-  $M \geq N - 1$ .

et  $\mathcal{T}_M(\mathbf{H})$  est de rang colonne plein avec probabilité 1 (ap1) ssi

«-  $\mathbf{h}(z)$  est irréductible ;

«-  $M \geq \underline{L} \triangleq \left\lceil \frac{N-1}{m-1} \right\rceil$ .

Dans la suite du document, si  $\mathcal{T}_{\underline{L}}(\mathbf{H})$  est de rang plein, nous dirons que le canal est *fortement irréductible*.

### 3.2.3 Paramétrisation minimale de l'espace bruit

Sous les conditions **[H $\mathcal{H}$ full]**, considérant l'espace signal de  $\mathbf{Y}$ , identique à l'espace de  $\mathcal{H}$  et son orthogonal que nous appellerons l'espace bruit, nous montrons plus loin, en suivant [67], qu'une paramétrisation minimale (en terme de dimensions du filtre) de l'espace bruit est donnée par  $\overline{\mathbf{P}}_{\underline{L}}$  de taille  $(m-1) \times m(\underline{L}+1)$ .

## 3.3 Modèle déterministe

Pour ce modèle, l'identifiabilité a été explorée par Hua dans [47] et l'équivalence entre notre approche et celles de [56] pour la méthode d'ajustement des sous-espaces et de [79] est détaillée dans [17].

Dans le modèle déterministe, la partie aléatoire provient uniquement du bruit, la fonction de densité de probabilité s'écrit :

$$f(\mathbf{Y}|\mathbf{H}) = \frac{1}{\pi\sigma_v^2} \exp \left[ -\frac{1}{\sigma_v^2} \|\mathbf{Y} - \mathcal{H}\mathbf{A}\|^2 \right] \quad (3.4)$$

Dans ce cas, l'identifiabilité est basée uniquement sur la moyenne :  $\overline{\mathbf{Y}} = \mathbb{E}\{\mathbf{Y}\} = \mathcal{H}\mathbf{A}$ , la matrice de covariance étant constante (ce qui revient à étudier l'identifiabilité d'un système sans bruit). De ce fait, la propriété d'identifiabilité s'écrit :

$$\overline{\mathbf{Y}} = \mathcal{H}\mathbf{A} = \mathcal{H}'\mathbf{A}' \Rightarrow \mathbf{H} = \alpha\mathbf{H}' \quad (3.5)$$

$\alpha$  est un scalaire complexe et constitue une ambiguïté fondamentale qu'on ne peut pas lever de manière aveugle.

D'autre part, l'identifiabilité de  $\mathbf{H}$  est ici équivalente à l'identifiabilité de  $(\mathbf{H}, \mathbf{A})$ .

**Théorème 3.2 (Condition Nécessaire et Suffisante)** *Dans le modèle déterministe,  $\mathbf{H}$  et  $\mathbf{A}$  sont identifiables de manière aveugle à un facteur scalaire près ssi les 3 conditions suivantes sont remplies :*

(i) *Le canal est fortement irréductible ;*

(ii)  $M \geq 2\underline{L} + N$  ;

(iii)  $o \geq N + \underline{L}$ ,

où  $\underline{L} = \left\lceil \frac{N-1}{m-1} \right\rceil$  et  $o$  est le nombre de modes d'excitation des symboles d'entrée.

**Condition Suffisante :** Pour la condition suffisante, il suffit de démontrer qu'une méthode permet de déterminer le canal, à l'ambiguïté fondamentale près, sous les conditions indiquées. Nous nous pencherons sur la méthode du sous-espace bruit avec sa paramétrisation minimale.

(i) Le canal est fortement irréductible ;

En utilisant la paramétrisation minimale du sous-espace bruit, nous pouvons écrire :

$$\overline{\mathbf{P}}_{\underline{L}} \mathcal{Y}_M = \overline{\mathbf{P}}_{\underline{L}} \mathcal{T}_{\underline{L}+1}(\mathbf{H}) \mathcal{A} = 0 \quad (3.6)$$

où  $\mathcal{Y}_M$  est de taille  $m(\underline{L} + 1) \times M - \underline{L}$  et

$$\mathcal{Y}_M = \begin{bmatrix} \mathbf{y}(\underline{L}) & \cdots & \mathbf{y}(M-1) \\ \vdots & \ddots & \vdots \\ \mathbf{y}(0) & \cdots & \mathbf{y}(M-\underline{L}-1) \end{bmatrix},$$

$$\mathcal{A} = \begin{bmatrix} a(\underline{L}) & \cdots & a(M-1) \\ \vdots & \ddots & \vdots \\ a(1-N) & \cdots & a(M-N-\underline{L}) \end{bmatrix}.$$

Si  $\mathcal{A}$  est de rang colonne plein, (3.6) implique que  $\overline{\mathbf{P}}_{\underline{L}} \mathcal{T}_{\underline{L}+1}(\mathbf{H}) = 0$ . Sous la condition (i),  $\mathcal{T}_{\underline{L}+1}(\mathbf{H})$  est de rang plein et le canal peut être identifié.

(ii)  $M \geq 2\underline{L} + N$

Cette condition est directement impliquée par le fait que  $\mathcal{A}$ , de dimension  $(N + \underline{L}) \times (M - \underline{L})$ , doit être de rang ligne plein.

(iii)  $o \geq N + \underline{L}$

En utilisant les propriétés des modes d'excitation, le Lemme 1 de [47] nous apprend que

$$\text{rang}(\mathcal{A}) = \min(o, M - \underline{L}, N + \underline{L}) \quad (3.7)$$

soit, sous la condition (ii)  $\text{rang}(\mathcal{A}) = \min(o, N + \underline{L})$  et donc, pour que  $\mathcal{A}^T$  soit de rang plein,  $o \geq N + \underline{L}$ .

**Condition Nécessaire :** Nous introduisons les notations :

$$\mathbf{Y}_M = \mathcal{T}_M(\mathbf{H}) \mathbf{A} = \mathbf{A}_{N+(M-1)} \mathbf{h} \quad (3.8)$$

où

$$\mathbf{h} = [\mathbf{h}_1(0)^H \cdots \mathbf{h}_1(N_1 - 1)^H]^H, \quad (3.9)$$

$$\text{et } \mathbf{A}_K = \begin{bmatrix} a(K - N) & \cdots & a(K - 2N + 1) \\ \vdots & \ddots & \vdots \\ a(0) & \cdots & a(1 - N) \end{bmatrix} \otimes I_m.$$

- (i) Si le canal est réductible, on peut trouver  $\mathbf{A}'$  tel que  $\mathbf{A}' = \mathbf{A} + \mathbf{A}''$  avec  $\mathbf{A}'' \mathbf{h} = 0$  et  $\mathbf{A}''$  indépendant de  $\mathbf{A}$ , ce qui montre la nécessité de (i).
- (ii) -(iii) Si  $\mathcal{A}$  n'est pas de rang plein, alors, comme  $\bar{\mathbf{P}}$  est calculé par  $\bar{\mathbf{P}}' \mathcal{T}(\mathbf{H}) \mathcal{A} \mathcal{A}^H \mathcal{T}^H(\mathbf{H}) \bar{\mathbf{P}}^H = 0$ , on peut trouver  $\bar{\mathbf{P}}' \neq \bar{\mathbf{P}}$  tel que  $\bar{\mathbf{P}}' \mathcal{Y} = 0$  et  $\bar{\mathbf{P}}' \mathcal{T}(\mathbf{H}) \neq 0$ . On obtient donc  $\mathbf{H}' \neq \mathbf{H}$  tel que  $\bar{\mathbf{P}}' \mathcal{T}(\mathbf{H}') = 0$  et  $\mathbf{Y} = \mathcal{T}(\mathbf{H}') \mathbf{A}$ , ce qui montre la nécessité des conditions (ii-iii).

### 3.4 Modèle gaussien

Il peut paraître choquant d'utiliser un modèle dit gaussien, où les sources sont considérées gaussiennes, pour un problème de communications numériques, où les sources sont loin d'être gaussiennes. Cependant, si on adopte un modèle stochastique général et qu'on se limite à utiliser les statistiques d'ordre deux, cela revient à utiliser un modèle gaussien. L'avantage de cette manière de voir est qu'on peut en déduire les conditions générales d'identifiabilité pour un modèle stochastique et l'ensemble des méthodes, connues ou pas, utilisant uniquement les statistiques d'ordre un et deux.

Pour l'identification aveugle de canal par le modèle gaussien, la moyenne est nulle et la covariance est donnée par :

$$C_{YY}(\theta) = \sigma_a^2 \mathcal{H} \mathcal{H}^H + \sigma_v^2 I \quad (3.10)$$

avec  $\theta = [\mathbf{h} \ \sigma_v^2]$ .

Ici, il existe une ambiguïté fondamentale qui consiste en un facteur scalaire de norme un, en effet, si  $\mathbf{h}' = e^{j\alpha} \mathbf{h}$ ,  $\mathcal{H}'\mathcal{H}'^H = \mathcal{H}\mathcal{H}^H$  et  $\mathbf{h}$  ne peut être déterminé qu'à une phase près.

**Théorème 3.3 (Condition nécessaire)** *Dans le modèle gaussien,  $\mathbf{h}$  est identifiable de manière aveugle à un facteur scalaire de norme un près seulement si :*

$$M \geq \frac{N}{m} + \frac{1}{2} \quad (3.11)$$

Le nombre de paramètres scalaires réels à identifier est de  $2Nm + 1 - 1 = 2Nm$ , où  $2Nm$  est le nombre de paramètres correspondant à  $\mathbf{h}$ ,  $+1$  pour  $\sigma_v^2$  et  $-1$  pour l'ambiguïté qui ne peut être levée de façon aveugle. En considérant l'équation (3.10), étant donné la nature bloc Toeplitz de  $C_{YY}(\theta)$ , il suffit de considérer la première ligne bloc, qui contient  $2.Mm^2$  valeurs réelles, auxquels il convient de retrancher  $m^2$ , en tenant compte de la nature hermitienne du premier bloc ( $C_{YY}(0)$ ) et en supposant que chaque élément bloc de  $C_{YY}$  est de rang plein.

Nous avons donc  $2.Nm$  paramètres pour  $(2M-1)m^2$  équations, ce qui conduit à la condition  $M \geq \frac{N}{m} + \frac{1}{2}$ .

Ensuite, pour vérifier que  $C_{YY}(i)$ ,  $i = 0 \dots M_{min}$  est de rang plein, il faut que  $M_{min} \leq N - m$ , soit  $N \geq m$ , ce qui est toujours vrai si on considère  $m$  comme étant le nombre de canaux effectifs.

**Théorème 3.4 (Condition suffisante)** *Dans le modèle gaussien,  $\mathbf{h}$  est identifiable de manière aveugle à un facteur scalaire de norme un près si les deux conditions suivantes sont remplies :*

(i) *Le canal est fortement irréductible ;*

(ii)  $M \geq \underline{L} + 1$ .

Si le canal est fortement irréductible, et si  $M \geq \underline{L} + 1$ ,  $\mathcal{H}$  est de rang colonne plein (=  $\underline{L} + N$ ) et la matrice  $C_{YY}$ , de taille  $m(\underline{L} + 1) \times m(\underline{L} + 1)$ , avec  $m(\underline{L} + 1) > \underline{L} + N$  a ses valeurs propres les plus petites égales à  $\sigma_v^2$ , ce qui permet de construire la matrice de covariance débruitée  $\sigma_a^2 \mathcal{H}\mathcal{H}^H$  et de déterminer  $\mathbf{h}$  par prédiction linéaire par exemple.

**Théorème 3.5 (Condition suffisante)** *Dans le modèle gaussien, la partie irréductible de  $\mathbf{h}$  est identifiable de manière aveugle à un facteur scalaire de norme un près si :*

(i)  $M \geq \underline{L} + 1$ .

*De plus, la partie réductible est identifiable de manière aveugle pour  $M$  infini.*

Soit  $\mathbf{H}(z) = \mathbf{H}_o(z)\mathbf{h}_c(z)$  où  $\mathbf{H}_o(z)$  est appelé la partie irréductible de  $\mathbf{H}(z)$  et  $\mathbf{h}_c(z)$  est un filtre scalaire représentant les zéros en communs des sous-canaux de  $\mathbf{H}(z)$  et peut être, sans perte de généralité, considéré comme étant monique. Les longueurs des canaux  $\mathbf{H}(z)$ ,  $\mathbf{H}_o(z)$  et  $\mathbf{h}_c(z)$  sont respectivement  $N$ ,  $N_o$  et  $N_c$ . Alors, si  $M \geq \underline{L}+1$ ,  $\mathcal{H}$  est de rang colonne  $\underline{L}+N_o$  et la matrice  $C_{YY}$ , de taille  $m(\underline{L}+1) \times m(\underline{L}+1)$ , avec  $m(\underline{L}+1) > \underline{L}+N$  a ses valeurs propres les plus petites égales à  $\sigma_v^2$ , ce qui permet de construire la matrice de covariance débruitée  $\sigma_v^2 \mathcal{H}\mathcal{H}^H$  et de déterminer  $\mathbf{h}_o$  par une technique de sous-espaces par exemple [56] ou par prédiction linéaire, à condition d'utiliser la méthode pondérée. (voir chapitre 6).

Ensuite, comme l'inverse de  $\mathbf{H}_o(z)$  est un filtre FIR,  $\mathbf{h}_c(z)$  peut être déduit par factorisation spectrale. En temporel, soit  $\mathcal{T}_M(\mathbf{h}) = \mathcal{T}_M(\mathbf{h}_o)\mathcal{T}_{M+N_o-1}(\mathbf{h}_c)$ , on obtient  $\mathbf{h}_c(z)$  par décomposition LU de  $\mathcal{T}_M(\mathbf{h}_o)^\#(C_{YY} - \sigma_v^2 I)\mathcal{T}_M^H(\mathbf{h}_o)^\#$ .

### 3.5 Bornes de Cramer-Rao

On notera  $\hat{\theta}$  l'estimée du paramètre,  $\tilde{\theta} = \theta - \hat{\theta}$  l'erreur d'estimation et la matrice de covariance d'erreur  $C_{\tilde{\theta}\tilde{\theta}} = \mathbb{E}\{\tilde{\theta}\tilde{\theta}^T\}$ .

La borne de Cramer-Rao est une borne inférieure de la matrice de covariance d'erreur de tout estimateur non-biaisé d'un paramètre  $\theta$  [66]. Cette borne est calculée par l'inverse de la matrice d'information de Fisher (FIM : Fisher Information Matrix).

La FIM est définie par :

$$J_{\theta\theta} = \mathbb{E}_{Y|\theta} \left( \frac{\partial \ln f(\mathbf{Y}|\theta)}{\partial \theta} \right) \left( \frac{\partial \ln f(\mathbf{Y}|\theta)}{\partial \theta} \right)^T \quad (3.12)$$

où  $\theta$  est le vecteur de paramètres réels. La borne de Cramer-Rao (CRB : Cramer-Rao Bound) est donc :

$$C_{\tilde{\theta}\tilde{\theta}} \geq J_{\theta\theta}^{-1} \triangleq \text{CRB} \quad (3.13)$$

Si  $\theta$  sont des paramètres complexes, on considérera le vecteur  $\theta_R = [\Re(\theta) \Im(\theta)]$ .

Pour une discussion plus approfondie sur les bornes de Cramer-Rao complexes, ainsi que sur les différentes bornes de Cramer-Rao pour les cas non-aveugles, les liaisons avec les questions d'identifiabilité et les régularisations de FIMs singulières, voir l'excellent travail d'Elisabeth de Carvalho [21]. En annexe, nous reproduisons les formules utiles pour le cas des bornes concernant les variables complexes.

#### 3.5.1 Modèle déterministe

Pour ce modèle, si  $\theta = [\mathbf{h} \ \mathbf{a}]$ , on a

$$J_{\theta\theta} = \frac{1}{\sigma_v^2} [\mathcal{T}(\mathbf{h}) | \mathcal{A}]^H [\mathcal{T}(\mathbf{h}) | \mathcal{A}] \quad (3.14)$$

En ce qui concerne la borne sur l'erreur d'estimation de  $\mathbf{h}$  [59] :

$$C_{\hat{\mathbf{h}}\hat{\mathbf{h}}} \geq \sigma_v^2 \left( \mathcal{A}^H P_{\mathcal{T}(\mathbf{h})}^\perp \mathcal{A} \right)^{-1} \quad (3.15)$$

où  $P_{\mathcal{T}(\mathbf{h})}^\perp = I - P_{\mathcal{T}(\mathbf{h})} = I - \mathcal{T}(\mathbf{h})(\mathcal{T}(\mathbf{h})^H \mathcal{T}(\mathbf{h}))^{-1} \mathcal{T}(\mathbf{h})^H$ .

### 3.5.2 Modèle gaussien

#### Cas de variables complexes circulaires

La borne de Cramer-Rao est déterminée à l'aide de (A.4) et

$$J_{\theta\theta}(i, j) = \text{trace} \left\{ C_{YY}^{-1} \left( \frac{\partial C_{YY}}{\partial \theta_i^*} \right) C_{YY}^{-1} \left( \frac{\partial C_{YY}}{\partial \theta_j^*} \right)^H \right\} \quad (3.16)$$

$$J_{\theta\theta^*}(i, j) = \text{trace} \left\{ C_{YY}^{-1} \left( \frac{\partial C_{YY}}{\partial \theta_i^*} \right) C_{YY}^{-1} \left( \frac{\partial C_{YY}}{\partial \theta_j^*} \right) \right\} \quad (3.17)$$

$$\text{où} \begin{cases} \frac{\partial C_{YY}}{\partial h_i^*} = \mathcal{T}(h) R_{AA} \mathcal{T}^H \left( \frac{\partial h}{\partial h_i^*} \right) \\ \frac{\partial C_{YY}}{\partial \sigma_v^2} = \frac{I}{2} \end{cases} \quad (3.18)$$

#### Cas de variables réelles

La borne de Cramer-Rao est déterminée par

$$J_{\theta\theta}(i, j) = \frac{1}{2} \text{trace} \left\{ C_{YY}^{-1} \left( \frac{\partial C_{YY}}{\partial \theta_i} \right) C_{YY}^{-1} \left( \frac{\partial C_{YY}}{\partial \theta_j} \right)^T \right\} \quad (3.19)$$

$$\text{où} \begin{cases} \frac{\partial C_{YY}}{\partial h_i} = \mathcal{T}(h) R_{AA} \mathcal{T}^T \left( \frac{\partial h}{\partial h_i} \right) + \mathcal{T} \left( \frac{\partial h}{\partial h_i} \right) R_{AA} \mathcal{T}^T(h) \\ \frac{\partial C_{YY}}{\partial \sigma_v^2} = I \end{cases} \quad (3.20)$$

## Chapitre 4

# Identifiabilité en multi-utilisateurs

|| Ce chapitre dérive les conditions d'identifiabilité sur les canaux et les symboles d'entrée, dans les cas aveugles, semi-aveugles et à séquence d'entrée connue, pour plusieurs utilisateurs.

### 4.1 Identifiabilité de la méthode par séquence d'apprentissage

De manière à pouvoir comparer les avantages et inconvénients de toutes les méthodes, du point de vue de l'identifiabilité, nous allons dériver les conditions pour la méthode par séquence d'apprentissage (TS : Training Sequence). Dans ce cas, le signal reçu  $\mathbf{y}^{(k)}$  contient uniquement des symboles connus ( $\mathbf{A} = \mathbf{A}_K$ ) et du bruit :

$$\mathbf{Y}_M = \mathcal{T}_M^p(\mathbf{H}) \mathbf{A} = \mathbf{A}_{N+p(M-1)} \mathbf{h} + \mathbf{V}_M \quad (4.1)$$

où

$$\begin{aligned} \mathbf{A}_{N+p(M-1)} &= [\mathbf{A}_{N_1+M-1,1} \cdots \mathbf{A}_{N_p+M-1,p}] \\ \mathbf{h} &= [\mathbf{h}_1(0)^H \cdots \mathbf{h}_1(N_1-1)^H \cdots \mathbf{h}_p(0)^H \cdots \mathbf{h}_p(N_p-1)^H]^H, \\ \text{et } \mathbf{A}_{K,i} &= \begin{bmatrix} a_i(K-N_i) & \cdots & a_i(K-2N_i+1) \\ \vdots & \ddots & \vdots \\ a_i(0) & \cdots & a_i(1-N_i) \end{bmatrix} \otimes I_m. \end{aligned} \quad (4.2)$$

**Théorème 4.1 (Condition nécessaire et suffisante)** *Le canal  $\mathbf{H}$  est identifiable par une séquence d'apprentissage ssi*

- (i)  $M \geq N$ ,  $(K_i \geq N + N_i - 1, i = 1 \cdots p)$ ;
- (ii)  $o_i \geq N_i, i = 1 \cdots p$ .

D'une manière similaire à ce qui est fait dans le chapitre 3, ces conditions découlent de ce que  $\mathbf{A}$  doit être de rang colonne plein. De toute évidence, il n'y a pas de conditions sur le canal, si ce n'est dans le cas dégénéré:  $\mathbf{Y}_M = \mathcal{T}_M^p(\mathbf{H}) \mathbf{A} = 0$ .

## 4.2 Longueur minimale de l'égaliseur FIR Zero-Forcing (ZF)

La propriété principale à l'origine du succès des méthodes basées sur la formulation multicanal du problème d'identification aveugle est l'inversibilité d'un canal par un égaliseur de longueur finie, en l'absence de bruit. Cet inverse est également appelé égaliseur Zero-Forcing (ZF), dans la mesure où il force l'interférence entre symboles et l'interférence entre utilisateurs à être nulle, en l'absence de bruit. Dans la mesure où les conditions d'identifiabilité dépendent de la longueur minimale de cet égaliseur ZF, nous plaçons ce résultat dans le chapitre traitant de l'identifiabilité. De plus, pour établir ces résultats, nous faisons appel à quelques notions concernant les matrices polynomiales, notions qui seront également utiles par la suite.

### 4.2.1 Quelques propriétés des matrices polynomiales

Les développements qui suivent se basent sur les propriétés des matrices polynomiales suivantes, tirées de [49, chap. 6] :

#### Définition 4.1

1. Un canal  $\mathbf{H}(z)$  est appelé **irréductible** si  $\mathbf{H}(z)$  est de rang plein  $\forall z$ .
2. Un canal  $\mathbf{H}(z)$  est appelé **à colonnes réduites** si  $[\mathbf{h}_1(N_1 - 1) \cdots \mathbf{h}_p(N_p - 1)]$  est de rang plein.

**Définition 4.2** Une matrice polynomiale  $\mathbf{H}_{un}(z)$  est **unimodulaire** si elle est régulière et a un déterminant constant.

**Propriété 4.1** Une matrice  $\mathbf{H}_{un}(z)$  est **unimodulaire** si et seulement si son inverse  $\mathbf{H}_{un}^{-1}(z)$  est également unimodulaire.

**Propriété 4.2** Toute matrice à colonnes non réduites peut être rendue à colonnes réduites par multiplication par une matrice unimodulaire.

#### Propriété 4.3 Prédicibilité du degré des matrices à colonnes réduites

Soit  $\mathbf{H}(z)$ , une matrice polynomiale de rang colonne plein, et pour tout vecteur polynomial  $u(z)$ , soit

$$t(z) = \mathbf{H}(z)u(z) \quad (4.3)$$

$\mathbf{H}(z)$  est à colonnes réduites si et seulement si

$$\deg t(z) = \max_{i: u_i(z) \neq 0} [\deg u_i(z) + k_i] \quad (4.4)$$

où  $u_i(z)$  est la  $i^{\text{ème}}$  ligne de  $u(z)$  et  $k_i$  est le degré de la  $i^{\text{ème}}$  colonne de  $\mathbf{H}(z)$ .

**Propriété 4.4 Invariance du degré des matrices à colonnes réduites .** Soit  $\mathbf{H}_{cr}(z)$  et  $\bar{\mathbf{H}}_{cr}(z)$  deux matrices irréductibles et à colonnes réduites, si

$$\mathbf{H}_{cr}(z) = \bar{\mathbf{H}}_{cr}(z)\mathbf{T}(z) \quad ; \quad \mathbf{T}(z) \text{ unimodulaire} \quad (4.5)$$

alors  $\mathbf{H}_{cr}(z)$  et  $\bar{\mathbf{H}}_{cr}(z)$  ont leurs colonnes de même degré.

#### 4.2.2 Quelques résultats additionnels.

**Théorème 4.2** Soit  $\bar{\mathbf{H}}(z)$  et  $\mathbf{H}(z)$ , reliés par une matrice unimodulaire.  $\bar{\mathbf{H}}(z)$  et  $\mathbf{H}(z)$  sont de même degrés colonnes, classés par ordre décroissant.

$$\bar{\mathbf{H}}(z) = \mathbf{H}(z)\mathbf{T}(z) \quad (4.6)$$

si  $\mathbf{H}(z)$  est à colonnes réduites, alors  $\bar{\mathbf{H}}(z)$  est à colonnes réduites et la matrice unimodulaire  $\mathbf{T}(z)$  a la structure triangulaire dynamique particulière<sup>2</sup>

$$\mathbf{T}(z) = \begin{bmatrix} T_{11} & 0 & \cdots & 0 \\ T_{12}(z) & T_{22} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ T_{1l}(z) & T_{2l}(z) & \cdots & T_{ll} \end{bmatrix} \quad (4.7)$$

$T_{rs}(z)$  sont des matrices polynomiales de dimension  $\nu_r \times \nu_s$  de degré  $L_r - L_s$  si  $r < s$ .  $T_{rr}$  sont des matrices constantes régulières.  $L_r, r = 1, \dots, l$  sont les différentes valeurs prises dans la séquence  $N_1, \dots, N_p$  et  $\nu_r$  est leur multiplicité.

En effet, soit  $\mathbf{H}_{l1}(z) = [\mathbf{H}_1(z) \cdots \mathbf{H}_{\nu_1}(z)]$  et  $\mathbf{H}_{li}(z) = [\mathbf{H}_{\nu_{i-1}+1}(z) \cdots \mathbf{H}_{\nu_l}(z)]$  pour  $i = 2 \dots, l$ , on déduit aisément de la propriété de prédictibilité du degré :

$$\deg[\mathbf{H}_{kj}(z)] = \max_{i: \mathbf{T}_{kji}(z)} [\deg[\mathbf{T}_{kji}(z)] + L_i] = L_i \quad ; \quad k = 1 \dots l \quad (4.8)$$

ce qui implique la structure de l'équation (4.7).

De plus, la matrice contenant les éléments  $[\bar{\mathbf{h}}_1(N_1) \cdots \bar{\mathbf{h}}_p(N_p)]$  peut s'écrire sous la forme

$$[T_{11}\mathbf{h}_{l1}(L_1) | T_{12}\mathbf{h}_{l1}(L_1) + T_{22}\mathbf{h}_{l2}(L_2) | \cdots | T_{1p}\mathbf{h}_{l1}(L_1) + \cdots + T_{pp}\mathbf{h}_{lp}(L_p)] \quad (4.9)$$

et on en déduit immédiatement que  $\text{rang}[\bar{\mathbf{h}}_1(N_1) \cdots \bar{\mathbf{h}}_p(N_p)] = \text{rang}[\mathbf{h}_1(N_1) \cdots \mathbf{h}_p(N_p)]$ .

◁

2. Cette structure se retrouve, dans un contexte légèrement différent, dans [39].

D'une manière un peu plus intuitive, pour deux utilisateurs, on peut écrire :

$$\begin{aligned} & [\bar{\mathbf{h}}_1(0) + \bar{\mathbf{h}}_1(1)z^{-1} + \bar{\mathbf{h}}_1(2)z^{-2} \quad \bar{\mathbf{h}}_2(0) + \bar{\mathbf{h}}_2(1)z^{-1}] \\ = & \begin{bmatrix} a & 0 \\ b_0 + b_1z^{-1} & c \end{bmatrix} [\mathbf{h}_1(0) + \mathbf{h}_1(1)z^{-1} + \mathbf{h}_1(2)z^{-2} \quad \mathbf{h}_2(0) + \mathbf{h}_2(1)z^{-1}] \end{aligned} \quad (4.10)$$

Le corollaire suivant découle immédiatement de la discussion précédente.

**Corollaire 4.1** Soit  $\bar{\mathbf{H}}(z)$  et  $\mathbf{H}(z)$ , reliés par une matrice unimodulaire et de degrés colonnes différents,

$$\bar{\mathbf{H}}(z) = \mathbf{H}(z)\mathbf{T}(z) \quad (4.11)$$

si  $\mathbf{H}(z)$  est irréductible et à colonnes réduites, alors  $\bar{\mathbf{H}}(z)$  n'est pas à colonnes réduites et de degré supérieur ou égal à celui de  $\mathbf{H}(z)$ .

**Théorème 4.3** Soit  $\mathbf{H}(z)$ , et  $\mathcal{T}_M^p(\mathbf{H})$  où  $M$  est tel que  $\mathcal{T}_M^p(\mathbf{H})$  a plus de lignes que de colonnes,  $\mathcal{T}_M^p(\mathbf{H})$  est de rang colonne plein avec probabilité 1.

De manière intuitive, des canaux générés de manière aléatoire remplissent un espace de dimensions  $m \times N$ . Les conditions nécessaires à imposer pour que le rang de  $\mathcal{T}_M^p(\mathbf{H})$  soit inférieur à son nombre de colonnes vont réduire la liberté du choix du canal dans un sous-espace de l'espace de dimensions  $m \times N$ , et donc, la probabilité qu'un canal généré aléatoirement soit dans ce sous-espace est nulle. On trouvera une démonstration plus formelle dans [43].

◁

### 4.2.3 Longueur minimale de l'égaliseur ZF

De manière à inclure le cas des canaux irréductibles mais pas nécessairement à colonnes réduites, nous écrivons le canal sous la forme :  $\mathbf{H}_{ncr}(z) = \mathbf{H}_{cr}(z)\mathbf{H}_{un}(z)$ , où les indices  $ncr$ ,  $cr$ ,  $un$  signifient respectivement : à colonnes non réduites, à colonnes réduites et unimodulaire. Soit  $N_{un} - 1$  le degré de  $\mathbf{H}_{un}(z)$  le degré maximum de l'inverse FIR de  $\mathbf{H}_{un}(z)$  vaut  $N_{inv} = (N_{un} - 1) \cdot (p - 1)$ , qui découle de  $\mathbf{H}_{un}^{-1}(z) = [\text{Adj } \mathbf{H}_{un}(z)] / \det \mathbf{H}_{un}(z)$ .

En introduisant la notation  $\mathcal{Z} = \text{diag}\{z^{-n_1} \dots z^{-n_p}\}$ , la condition pour obtenir un égaliseur  $\mathbf{F}(z)$  zero-forcing est :

$$\mathbf{F}(z)\mathbf{H}_{cr}(z)\mathbf{H}_{un}(z) = \mathcal{Z}$$

$$\mathbf{F}(z)\mathbf{H}_{cr}(z) = \mathbf{H}_{un}^{-1}(z)\mathcal{Z}$$

$$F\mathcal{T}_M^p(\mathbf{H}_{cr}) = \mathbf{H}_{un}^{-1} \quad (4.12)$$

où, sans perte de généralité, nous considérons un égaliseur à délai nul. On peut alors transcrire les conditions sous la forme :

1. De  $F\mathcal{T}_M^p(\mathbf{H}_{cr}) = \mathbf{H}_{un}^{-1}$  où  $F$  est de taille  $p \times mM$  et  $\mathcal{T}_M^p(\mathbf{H}_{cr})$  est de taille  $mM \times N_{cr,tot} + p(M-1)$ , pour obtenir un système exactement ou sous-déterminé, il faut

$$M \geq \underline{L} = \left\lceil \frac{N_{cr,tot} - p}{m - p} \right\rceil \quad (4.13)$$

$N_{cr,tot} - p$  étant la somme des degrés des colonnes de  $\mathbf{H}_{cr}(z)$ .

2. Le nombre de colonnes du terme de gauche de l'équation (4.12) doit être supérieur ou égal au nombre de colonnes du terme de droite, ce qui conduit à la condition :

$$\begin{aligned} N_{cr,tot} + p(M-1) &\geq p((N_{un}-1)(p-1) + 1) \\ M &\geq \lceil ((N_{un}-1) \cdot (p-1) + 1) - N_{cr,tot}/p + 1 \rceil \triangleq \underline{\underline{L}} \end{aligned} \quad (4.14)$$

La longueur minimale de l'égaliseur zero-forcing est donc donnée (ap1) par :

$$\boxed{M \geq \bar{L} \triangleq \max\{\underline{L}, \underline{\underline{L}}\} \quad \begin{aligned} \underline{L} &= \left\lceil \frac{N_{cr,tot} - p}{m - p} \right\rceil \\ \underline{\underline{L}} &= \lceil ((N_{un}-1) \cdot (p-1) + 1) - N_{cr,tot}/p + 1 \rceil \end{aligned}} \quad (4.15)$$

Il est clair que pour un grand nombre de cas, la condition  $M \geq \underline{L}$  sera valable.

#### 4.2.4 Une borne plus lâche

Les résultats classiques concernant les matrices résultantes généralisées de Sylvester donnent une borne plus lâche pour le cas général des matrices irréductibles et à colonnes réduites [8].

$$\boxed{M \geq N_{cr,tot} - p} \quad (4.16)$$

Pour des canaux à colonnes non réduites, la borne  $\underline{\underline{L}}$  est valide, et la condition générale est  $M \geq \bar{L} \triangleq \max\{N_{cr,tot} - p, \underline{\underline{L}}\}$ .

Dans la suite de ce document,  $\bar{L}$  représentera toujours la longueur minimale de l'égaliseur ZF, quelles que soient les propriétés du canal.

### 4.3 Quelques canaux dégénérés

Il peut paraître curieux d'arguer que la matrice de canal  $\mathcal{H}$  est de rang colonne plein avec probabilité 1, et en même temps, de traiter les cas particuliers où ce n'est pas le cas. Outre les canaux réductibles, qui ont fait l'objet d'interprétations bien connues dans [76, 29, 63], on peut identifier un certain nombre de canaux « dégénérés » qui ont une origine physique précise et qui peuvent réellement exister, avec probabilité non nulle.

#### 4.3.1 Nombre effectif de canaux

De toute évidence, si  $\mathbf{H}$  n'est pas de rang colonne plein,  $\mathcal{T}_{\underline{L}}(\mathbf{H})$  ne le sera pas non plus. En effet, le nombre de colonnes de  $\mathcal{T}_{\underline{L}}(\mathbf{H})$  est le même que son nombre de lignes, et la matrice carrée est singulière. C'est le cas lorsque l'un des sous-canaux est une combinaison linéaire des autres. Dans ces circonstances, le canal peut être irréductible et à colonnes réduites et  $\mathcal{T}_{\underline{L}}(\mathbf{H})$  ne sera pas de rang plein. La longueur minimale de l'égaliseur ZF sera plus grande que  $\underline{L}$ , elle sera en effet celle d'un égaliseur ZF pour un multicanal équivalent comprenant un nombre réduit de sous-canaux. Cette interprétation milite en faveur de l'utilisation d'un "nombre de canaux effectifs", où chaque sous-canal apporte une information additionnelle sur le système global. Au contraire, les canaux "non-effectifs" apportent une "information redondante".

#### 4.3.2 Multitrajets distants.

Quand le canal consiste en des multitrajets distants, comme dans le cas de canaux en milieu montagneux, ou dans le cas d'échos rapprochés et distants pour le canal téléphonique, le canal peut ressembler à :

$$\mathbf{H} = \left[ \begin{array}{ccc|ccc|ccc} * & \dots & * & 0 & \dots & 0 & * & \dots & * \\ \vdots & \ddots & \vdots & \vdots & \dots & \vdots & \vdots & \ddots & \vdots \\ * & \dots & * & 0 & \dots & 0 & * & \dots & * \end{array} \right] \quad (4.17)$$

Ainsi, lorsque les deux parties non nulles de la réponse impulsionnelle sont éloignées,  $\mathcal{T}_{\underline{L}}(\mathbf{H})$  aura des colonnes nulles et donc ne sera pas de rang plein. On pourrait alors trouver l'égaliseur ZF de longueur minimale en paramétrant les longueurs des différentes parties en les considérant comme des canaux distincts sur un système présentant  $2p$  utilisateurs pour chaque canal. Cette approche ne sera cependant utile que dans des cas très particuliers où certaines caractéristiques des canaux sont bien connues a priori.

### 4.3.3 Canaux à colonnes non réduites en CDMA

Dans des situations particulières, où le nombre de sous-canaux est grand, on peut rencontrer des cas spéciaux. Par exemple en DS-CDMA synchrone, supposons que nous avons  $p$  utilisateurs actifs et supposons que les canaux de propagation sont de même longueur pour tous les utilisateurs. Notons le canal composite  $G = CH$ , où  $C$  est la matrice comprenant les codes d'étalement de chaque utilisateur et  $H$  est le canal de propagation. Ces canaux composites auront également la même longueur pour chaque utilisateur:  $N_1 = N_2 \dots = N_p$ . De plus, supposons que les canaux sont tels que le dernier coefficient  $h_j(N_i - 1), \forall j$  ont un petit nombre de coefficients non nuls, soit  $J$  d'entre eux, avec  $J < p$ . Alors, il existe au plus  $J$  coefficients  $h_j(N_i - 1)$  linéairement indépendants. Dans ce cas, il est clair que le canal n'est pas à colonnes réduites et  $\mathcal{T}_{\underline{L}}(H)$  n'est pas de rang plein. Pour plus de détails, voir [35].

### 4.3.4 Comparaison entre les bornes.

De manière à illustrer l'intérêt de cette discussion, nous fournissons des comparaisons entre la longueur minimale et la borne plus lâche. Dans un premier temps, nous nous plaçons dans un contexte SDMA, avec 2 utilisateurs, de 3 à 10 sous-canaux, la longueur totale du canal variant de  $N = 7$  à  $N = 19$ . Le premier point de la courbe représente la borne lâche ( $= N - p$ ). Les courbes de la figure 4.3.4 montrent clairement qu'il y a avantage à utiliser des canaux multiples ainsi que l'utilité de considérer, dans la conception des égaliseurs, la longueur  $\underline{L}$ .

Dans un contexte CDMA, nous considérons une situation où l'étalement varie de 10 à 90 chips (ce qui correspond à un étalement de  $20 \mu\text{secs}$  pour une longueur de chip de 250 ns, comme en UMTS) et une période de symbole de 4 à 256 chips. Nous ne considérons pas le cas dégénéré à colonnes non réduites.

Ces courbes montrent clairement qu'on a avantage à considérer la longueur la plus courte, en effet, en général, la borne lâche est beaucoup plus grande que  $\underline{L}$  et l'utilisation d'un égaliseur de longueur beaucoup trop grande affectera les performances de l'égalisation de manière défavorable, quel que soit l'algorithme utilisé.

Dans cette figure, l'axe des abscisses représente la longueur du code et l'axe des ordonnées représente la longueur de l'égaliseur ZF en nombre de symboles. Les courbes différentes représentent des étalements de longueurs différentes. La situation de réseau chargé correspond à 1 utilisateur sur 4 possibles, soit 1 utilisateur pour des symboles de durée équivalente à 4 chips, 2 pour des symboles de durée de 8 chips, etc. La situation de réseau non chargé correspond à 1 utilisateur sur 16 possibles (ou au minimum 1 pour les symboles de longueur courte).

Les deux courbes inférieures donnent le rapport entre la longueur minimale et la borne lâche.

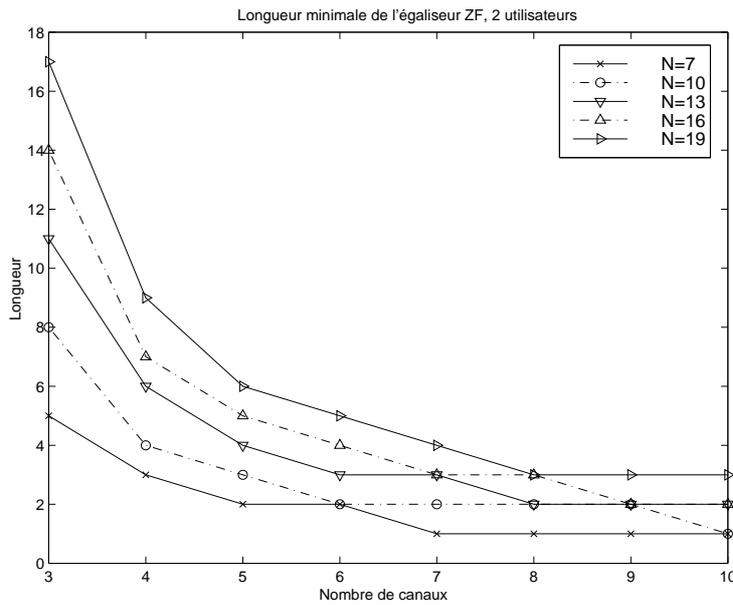


FIG. 4.1 – Égaliseur ZF de longueur minimale en SDMA

## 4.4 Identifiabilité pour les méthodes aveugles

### 4.4.1 Modèle déterministe

Dans ce cas,  $\mathbf{A} = \mathbf{A}_U$ , et, d'une manière similaire à ce qui a été fait en section 3.3, l'identifiabilité signifie que  $(\mathbf{H}, \mathbf{A})$  sont identifiables de façon univoque, à une indétermination fondamentale près, à partir de  $\overline{\mathbf{Y}}$ .

#### Indétermination fondamentale

Comme dans le cas d'un seul utilisateur, où le canal ne peut être déterminé qu'à un facteur scalaire, il y a une indétermination fondamentale (on peut la trouver dans le cas de la méthode sous-espace dans [39]). En effet, sous les conditions d'identifiabilité détaillées ci-dessous, les méthodes déterministes ne peuvent identifier le canal qu'à un facteur triangulaire dynamique près. Ainsi, dans le cas idéal sans bruit, toute méthode est susceptible d'identifier un canal  $\tilde{\mathbf{H}}(z)$

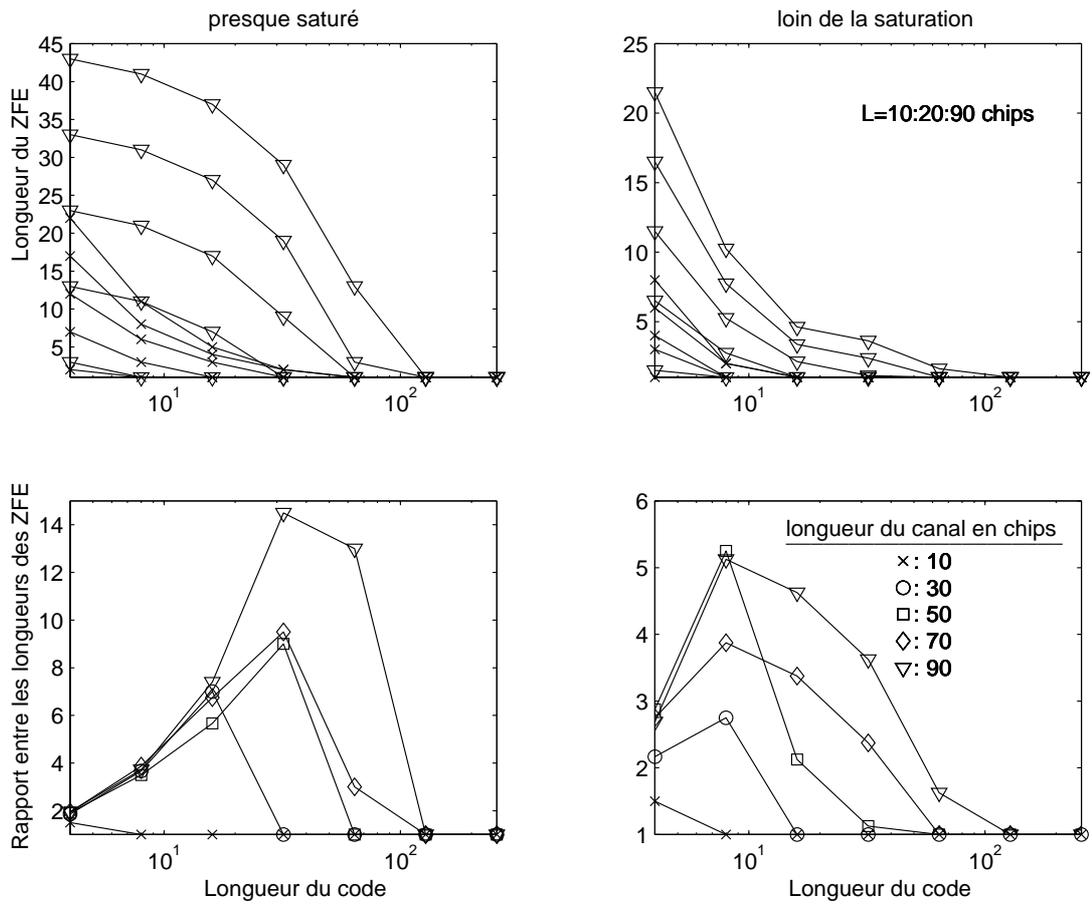


FIG. 4.2 – Égaliseur ZF de longueur minimale en CDMA

tel que  $\tilde{\mathbf{H}}(z) = \mathbf{H}(z)\mathbf{T}(z)$ , où

$$\mathbf{T}(z) = \begin{bmatrix} \mathbf{T}_{11} & 0 & \cdots & 0 \\ \mathbf{T}_{12}(z) & \mathbf{T}_{22} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{T}_{1l}(z) & \mathbf{T}_{2l}(z) & \cdots & \mathbf{T}_{ll} \end{bmatrix} \quad (4.18)$$

$\mathbf{T}_{rs}(z)$  sont des matrices polynomiales de dimension  $\nu_r \times \nu_s$  de degré  $L_r - L_s$  si  $r < s$ .  $\mathbf{T}_{rr}$  sont des matrices constantes régulières.  $L_r, r = 1, \dots, l$  sont différentes valeurs prises dans la séquence  $N_1, \dots, N_p$  et  $\nu_r$  est leur multiplicité.

Cette indétermination fondamentale se déduit immédiatement du théorème 4.2.

### Conditions d'identifiabilité

**Théorème 4.4 (Condition nécessaire et suffisante)** *Dans le modèle déterministe,  $\mathbf{H}$  et  $\mathbf{A}$  sont identifiables de manière aveugle à un facteur triangulaire dynamique  $\mathbf{T}(z)$  près ssi les 3 conditions suivantes sont remplies :*

- (i) *Le canal est irréductible,*
- (ii)  $M \geq \bar{L}(p+1) + N$
- (iii)  $o_i \geq N_i + \bar{L}, \quad i = 1, \dots, p,$

La démonstration est une extension immédiate du cas mono-utilisateur (voir section 3.3).

### 4.4.2 Modèle Gaussien

Dans ce cas, les symboles sont considérés Gaussiens, les paramètres à estimer sont donc le canal et la variance du bruit. L'identifiabilité pour ce modèle équivaut à l'identifiabilité pour tout modèle stochastique en n'utilisant que la moyenne (nulle dans le cas aveugle) et la covariance :

$$C_{YY}(\theta) = \mathcal{T}^p(\mathbf{H})\mathcal{T}^{pH}(\mathbf{H})\sigma_a^2 + \sigma_v^2 I \quad (4.19)$$

### Indétermination fondamentale

Pour les besoins de cette section, nous introduisons les notations :

$$\begin{aligned} \check{\mathbf{h}}(i) &= [\mathbf{h}_1(i), \mathbf{h}_2(i), \dots, \mathbf{h}_p(i)], \quad i = 1 \cdots N_1; \\ \check{\mathbf{H}} &= [\check{\mathbf{h}}(0), \check{\mathbf{h}}(1) \cdots \check{\mathbf{h}}(N_1)] \end{aligned} \quad (4.20)$$

où  $\check{\mathbf{h}}(i)$  est de dimension  $m \times p$ . On en déduit aisément que  $C_{YY}(\theta) = \mathcal{T}(\check{\mathbf{H}})\mathcal{T}^H(\check{\mathbf{H}})\sigma_a^2 + \sigma_v^2\mathbf{I}$ ,

$$\text{où } \mathcal{T}(\check{\mathbf{H}}) = \begin{bmatrix} \check{\mathbf{h}}(0) & \check{\mathbf{h}}(1) & \check{\mathbf{h}}(2) & \cdots & 0 \\ 0 & \check{\mathbf{h}}(0) & \check{\mathbf{h}}(1) & \cdots & 0 \\ 0 & 0 & 0 & \cdots & \check{\mathbf{h}}(N_1) \end{bmatrix}.$$

Les éléments de la séquence de covariance du signal (qui forment la matrice de covariance) s'écrivent :

$$\begin{aligned} C_{YY}(\tau) &= \sigma_a^2 \sum_{i=0}^{i=N_1} \check{\mathbf{h}}(i)\check{\mathbf{h}}(\tau - i) + \sigma_v^2\mathbf{I}_m \\ &= \sigma_a^2 \sum_{i=0}^{i=N_1} \check{\mathbf{h}}(i)U\check{\mathbf{h}}(\tau - i) + \sigma_v^2\mathbf{I}_m \end{aligned} \quad (4.21)$$

où  $U$  est une matrice unitaire de dimension  $p \times p$ . Nous sommes donc en présence d'une ambiguïté fondamentale telle que tout canal  $U\check{\mathbf{H}}(z)$  est une solution du problème (4.19). Nous dirons que le canal est identifiable à une matrice unitaire près.

Notons que si les longueurs de canaux sont différentes, on peut réduire la matrice unitaire à une matrice bloc diagonale de même structure que  $\mathbf{T}(z)$  dans le cas déterministe, à part les éléments dynamiques qui sont nuls.

En effet, considérons le cas de deux canaux de longueur connue et différente  $N_1 > N_2$ , connaissant  $\mathbf{h}_2(N_1 - 1) = 0$ , et en tenant compte de l'orthonormalité de  $U$  :

$$\tilde{\mathbf{H}}(z) = U\mathbf{H}(z) \Rightarrow \tilde{\mathbf{H}}(N_1 - 1) = U\check{\mathbf{H}}(N_1 - 1) \Rightarrow U = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}. \quad (4.22)$$

L'extension au cas général est immédiate.

### Conditions d'identifiabilité

**Théorème 4.5 (Condition nécessaire)** *Dans le modèle Gaussien,  $\mathbf{h}$  est identifiable de manière aveugle à une matrice unitaire près seulement si :*

$$M \geq \frac{N}{m} + \frac{1}{2} + \frac{2 - p(p+1)}{4m^2} \quad (4.23)$$

Le nombre de paramètres scalaires réels à identifier est de  $2.Nm + 1 - p(p+1)/2$ , où  $2Nm$  est le nombre de paramètres correspondant à  $\mathbf{h}$ ,  $+1$  pour  $\sigma_v^2$  et  $-p(p+1)/2$  pour l'ambiguïté qui ne peut être levée de façon aveugle. On arrive à la condition donnée de la même manière que dans le cas mono-utilisateur.

**Théorème 4.6 (Condition suffisante)** *Dans le modèle Gaussien,  $\mathbf{h}$  est identifiable de manière aveugle à une matrice unitaire près si :*

- (i) *Le canal est irréductible;*

(ii)  $M \geq \bar{L} + 1$ .

**Théorème 4.7 (Condition suffisante)** *Dans le modèle Gaussien, la partie irréductible de  $\mathbf{h}$  est identifiable de manière aveugle à une matrice unitaire près si :*

(i)  $M \geq \bar{L} + 1$ .

De plus, la partie réductible est identifiable de manière aveugle pour  $M$  infini.

Les démonstrations de ces deux théorèmes sont des extensions immédiates de ce qui a été fait dans le cas d'un seul utilisateur.

## 4.5 Identifiabilité des méthodes semi-aveugles

### 4.5.1 Modèle déterministe

Dans un premier temps, nous nous attachons à identifier  $\mathbf{T}(z)$ .

**Théorème 4.8 (Condition suffisante)** *Sous les conditions du théorème 4.4  $\mathbf{T}(z)$  est identifiable par une séquence d'apprentissage  $\mathbf{A}_K$  de longueur  $M_k$  et ayant des modes d'excitation en nombre  $o_{ki}$  pour chaque utilisateur  $i$  si :*

(i)  $M_k \geq \max_i \nu_i (1 + \sum_{j=2}^i (L_j - L_{j-1} + 1))$

(ii)  $o_{ki} \geq \begin{cases} \nu_{i'} & i = 1 \\ \nu_{i'} (L_i - L_{i-1} + 1) & i > 1 \end{cases}$

où  $i'$  est l'indice qui maximise la première expression et les  $\nu_i$  sont définis en 4.7.

Considérons le signal  $\mathbf{Y} = \mathcal{T}(\mathbf{H})\mathbf{A}_K$ , où  $\mathcal{T}(\mathbf{H}) = \mathcal{T}(\widetilde{\mathbf{H}})\mathcal{T}(\mathbf{T})$ . Nous pouvons toujours construire :

$$\widetilde{\mathbf{Y}} = \mathcal{T}(\mathbf{T})\mathbf{A}_K \quad (4.24)$$

et appliquer l'identification par séquence d'apprentissage à ce signal.

La matrice  $\mathcal{T}(\mathbf{T})$  est une matrice de convolution multicanal multivariable de taille  $M + N - p \times M + N - p$  et a la structure particulière :

$$\mathcal{T}(\mathbf{T}) = \begin{bmatrix} \mathcal{T}(T_{11}) & 0 & \cdots & 0 \\ \mathcal{T}(T_{12}) & \mathcal{T}(T_{22}) & \cdots & 0 \\ \vdots & \ddots & \ddots & 0 \\ \mathcal{T}(T_{1l}) & \mathcal{T}(T_{2l}) & \cdots & \mathcal{T}_{ll} \end{bmatrix} \quad (4.25)$$

Notons également que si  $N_1 = N_2$ , alors  $\mathcal{T}(T_{11}) = T_{11} \otimes I_{M+N_1-1}$ . D'autre part, si  $T_{12}(z)$  est de taille  $2 \times 2$ , (soit  $T_{12}(z) = \begin{bmatrix} a(z) & b(z) \\ c(z) & d(z) \end{bmatrix}$ ) alors  $\mathcal{T}(T_{12}) = \begin{bmatrix} \mathcal{T}(a) & \mathcal{T}(b) \\ \mathcal{T}(c) & \mathcal{T}(d) \end{bmatrix}$ . La généralisation pour des multiplicités  $\nu_i$  quelconques est immédiate.

Par exemple, pour  $p = 2$  et  $N_1 = 3$  et  $N_2 = 2$

$$\mathcal{T}(\mathbf{T}) = \left[ \begin{array}{cccc|ccc} T_{11} & 0 & \cdots & 0 & 0 & \cdots & 0 \\ 0 & T_{11} & \ddots & 0 & 0 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & T_{11} & 0 & 0 & 0 \\ \hline T_{12}(0) & T_{12}(1) & \cdots & 0 & T_{22} & \cdots & 0 \\ 0 & T_{12}(0) & \ddots & 0 & 0 & \ddots & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & T_{12}(1) & 0 & \cdots & T_{22} \end{array} \right] \quad (4.26)$$

$\tilde{\mathbf{Y}}$  peut se diviser en  $l$  blocs correspondants aux canaux  $\tilde{\mathbf{H}}_{li}$ . Il suffit d'appliquer les conditions d'identifiabilité TS à chaque bloc et d'adopter les plus contraignantes. Pour chaque bloc  $i = 1, \dots, l$ :

$$N'_i = \nu_i \left( 1 + \sum_{j=2}^i (L_j - L_{j-1} + 1) \right) \quad (4.27)$$

Il suffit donc d'adopter  $M_k \geq N' = \max_i N'_i$  et, si l'indice  $i$  qui maximise cette expression est appelé  $i'$ , il faut que  $o_{ki} \geq \nu_{i'}(L_i - L_{i-1} + 1)$ .

**Théorème 4.9 (Condition suffisante pour tout canal)** *Un canal quelconque  $\mathbf{H}(z) = \underline{\mathbf{H}}(z)T(z)$  où  $\underline{\mathbf{H}}(z)$  est irréductible et à colonnes réduites, et  $T(z) = [t_1(z) \dots t_p(z)]$ , est identifiable si*

- (i) *Les conditions du théorème 4.4 sont remplies pour  $\underline{\mathbf{H}}(z)$ ;*
- (ii) *les conditions du théorème 4.1 sont remplies pour  $\mathbf{T}(z)T(z)$ .*

où  $t_i(z)$  est un polynôme de dimension  $p \times 1$  et d'ordre  $N_i - \underline{N}_i$ ,  $\underline{N}_i$  étant l'ordre de  $\underline{\mathbf{H}}_i(z)$ .

La démonstration découle directement de la décomposition de  $\mathbf{H}(z)$  en partie irréductible et réductible. L'admissibilité de cette décomposition découle de la propriété de prédictibilité des degrés des polynômes irréductibles et à colonnes réduites.

## 4.5.2 Modèle Gaussien

### Identification aveugle suivie d'une identification par séquence d'apprentissage

Dans ce paragraphe, nous supposons que le canal a été identifié de manière aveugle à une matrice unitaire près, cette matrice étant ensuite identifiée par une séquence d'apprentissage.

#### Partie aveugle

Sous les conditions du théorème 4.6, tant que  $\mathbf{Y}$  contient un bloc d'au moins  $\bar{L} + 1$  échantillons  $\mathbf{y}(k)$  généré par des symboles inconnus, le canal peut être identifié à une matrice unitaire près.

#### Séquence d'apprentissage : identification de la matrice unitaire

**Théorème 4.10 (Condition suffisante)** *Dans le modèle Gaussien, le canal  $\mathbf{H}$  est identifiable si*

- (i) *Le canal est irréductible ;*
- (ii)  $M_u \geq \bar{L} + 1 ;$
- (iii) *chaque utilisateur émet un symbole connu, apparaissant à des moments différents pour chaque utilisateur.*

En effet, soit  $k_1$  l'instant auquel le premier utilisateur émet un symbole connu, de  $\bar{\mathbf{Y}}_{\bar{L}+1}(k_1) = \mathcal{T}_{\bar{L}+1}(\mathbf{H})\mathbf{A}_K$ , nous pouvons déterminer  $(\hat{\mathbf{h}}^H(0)\hat{\mathbf{h}}(0))^{-1}\hat{\mathbf{h}}^H(0)\mathbf{P}_{\bar{L}}\bar{\mathbf{Y}}_{\bar{L}+1}(k_1) = Q\mathbf{a}(k_1)$  où  $\mathbf{P}_{\bar{L}}$  est le prédicteur linéaire et  $Q$  est la matrice de mélange unitaire ( $\mathbf{h}(0) = \hat{\mathbf{h}}(0)Q$ ). Supposons qu'il en est de même pour tout utilisateur, à des instants  $k_i$  tous distincts. Nous pouvons déterminer  $Q$  de

$$Q = \left(\hat{\mathbf{h}}^H(0)\hat{\mathbf{h}}(0)\right)^{-1}\hat{\mathbf{h}}^H(0)\mathbf{P}_{\bar{L}}\left[\bar{\mathbf{Y}}_{\bar{L}+1}(k_1) \dots \bar{\mathbf{Y}}_{\bar{L}+1}(k_p)\right]\left[\mathbf{a}(k_1) \dots \mathbf{a}(k_p)\right]^{-1}. \quad (4.28)$$

#### Identification semi-aveugle pour un canal quelconque

**Théorème 4.11 (Condition suffisante)** *Dans le modèle Gaussien, le canal  $\mathbf{H}$  est identifiable si*

- (i)  $M_k \geq N ;$
- (ii) *chaque utilisateur  $i$  a un symbole connu et ce symbole est distant d'au moins  $N_i$  des autres utilisateurs.*

En effet, dans ce cas,  $\bar{\mathbf{Y}}$  contient tous les échantillons de la réponse impulsionnelle du canal, multipliée par le(s) symbole(s) connu(s).

---

**Théorème 4.12 (Condition suffisante)** *Dans le modèle Gaussien, le canal  $H$  est identifiable si*

(i) *Les symboles connus sont contigus et les trames des symboles connus pour les différents utilisateurs commencent toutes au même instant ;*

(ii)  $M_k \geq N \quad (K_i \geq N - N_i + 1)$ .

Sous la première condition, nous sommes dans une situation similaire au cas de l'identification par séquence d'apprentissage, pour le signal  $\bar{Y}$ . En plus, les  $N_i - 1$  zéros (qui sont en fait la moyenne des  $N_i - 1$  symboles inconnus) apparaissant avant et après le bloc des symboles connus font également office de séquence d'apprentissage. La seconde condition du théorème 4.12 en découle immédiatement.

## 4.6 Conclusions

Nous avons dérivé des conditions nécessaires et suffisantes pour l'identification de canal aveugle d'un multicanal, ceci pour les modèles déterministe et Gaussien et en ne recourant qu'aux statistiques d'ordre deux. Nous avons également donné des conditions suffisantes pour le cas semi-aveugle, soit en levant les ambiguïtés fondamentales des méthodes aveugles (mais sans s'affranchir de la non-identifiabilité de certains canaux), soit en identifiant n'importe quel canal.

En explorant les conditions d'identifiabilité, nous avons naturellement exploré la longueur minimale de l'égaliseur Zero-Forcing, et montré que l'utilisation de la borne générale, applicable aux canaux irréductibles et à colonnes réduites, conduisait à une surestimation importante de la longueur de l'égaliseur ZF, et donc, à une perte possible de performance.

Les résultats de cette analyse appellent deux commentaires importants :

- «– Les approches semi-aveugles sont intrinsèquement capables de lever les ambiguïtés aveugles ou d'identifier n'importe quel canal à l'aide de peu de symboles connus.
- «– Les approches stochastiques sont capables d'identifier des classes de canaux nettement plus larges que leurs homologues déterministes.

Notons que toutes les conditions dérivées pour le modèle Gaussien sont des conditions suffisantes pour le modèle stochastique général, puisque le modèle Gaussien peut être considéré comme un modèle stochastique général, mais avec des méthodes ne recourant qu'aux statistiques de premier et second ordre.

Les bornes de Cramer-Rao pour le cas multi-utilisateur sont des extensions directes du cas mono-utilisateur.



# Chapitre 5

## Ajustement des sous-espaces

*Ce chapitre récapitule les principaux résultats obtenus jusqu'ici pour l'identification de canal recourant aux sous-espaces signal et bruit de la matrice de corrélation. Une paramétrisation minimale du sous-espace bruit obtenue par le biais de la prédiction linéaire, utile pour l'identifiabilité, y est présentée. D'autre part, un algorithme bloc rapide y est présenté ainsi qu'une ébauche d'algorithme robuste à la surestimation de la longueur du canal.*

### 5.1 Introduction

La méthode d'ajustement des sous-espaces, initiée dans le domaine de l'identification de canal par Moulines et al. dans [56], est basée sur la séparation en sous-espaces "signal" et "bruit" de la matrice de corrélation du signal reçu. Une analyse en a été proposée dans [1]. On y déduit les conditions d'identifiabilité, qui coïncident avec les conditions générales d'identifiabilité pour le modèle déterministe. D'autre part, une méthode pondérée y est dérivée.

### 5.2 Modèle utilisé : Déterministe

Les hypothèses utilisées dans le cadre de la méthode par ajustement des sous-espaces sont celles correspondant au modèle déterministe. En effet, cette méthode consiste à ajuster le sous-espace signal ou bruit de  $Y = HA + V$  au sous-espace correspondant de  $\mathcal{H}$ .

### 5.3 La méthode

La méthode d'ajustement des sous-espaces consiste à ajuster le sous-espace signal ou bruit des données reçues au sous-espace correspondant de la matrice de canal. De manière à se ramener à un problème quadratique et à s'affranchir de l'influence du bruit additif supposé blanc, cette méthode utilise la matrice de corrélation des données reçues et ses sous-espaces signal et bruit.

### 5.3.1 Ajustement du sous-espace signal

La matrice de corrélation du signal reçu s'écrit :

$$R_{YY} = E \{ \mathbf{Y} \mathbf{Y}^H \} = \mathcal{H} \mathbf{A} \mathbf{A}^H \mathcal{H}^H + \sigma_v^2 \mathbf{I} \quad (5.1)$$

en considérant un bruit additif spatialement et temporellement blanc.

Nous pouvons écrire la décomposition en vecteurs/valeurs propres de cette matrice :

$$R_{YY} = V_S \Lambda_S V_S^H + V_N \Lambda_N V_N^H \quad (5.2)$$

où  $V_S$  a les mêmes dimensions que  $\mathcal{H}$  et  $\Lambda_N = \sigma_v^2 \mathbf{I}$ . Ces valeurs propres correspondant à la variance du bruit,  $V_N$  sera appelé "*espace bruit*" tandis que  $V_S$  représente "*l'espace signal*". Des équations (5.1) et (5.2), on déduit immédiatement la propriété suivante du sous-espace signal :

$$\text{range}\{V_S\} = \text{range}\{\mathcal{H}\} \quad (5.3)$$

Le problème d'ajustement des sous-espaces signal consiste à minimiser la distance entre l'espace engendré par  $\mathcal{H}$  et l'estimation du sous-espace signal  $\hat{V}_S$ , la matrice  $Q$  dans l'équation (5.4) jouant le rôle de changement de base.

$$\min_{\mathbf{H}, Q} \|\mathcal{H} - \hat{V}_S Q\|_F^2 \quad (5.4)$$

Ce problème est séparable en  $\mathbf{H}$  et  $Q$ . La minimisation en  $Q$  mène à  $Q = \hat{V}_S^H \mathcal{H}$  (en se souvenant que  $\hat{V}_S^H \hat{V}_S = \mathbf{I}$ )

De  $\hat{V}_S^H \hat{V}_N = 0$ , et grâce à la commutativité de l'opération de convolution on débouche sur

$$\min_{\mathbf{H}} h^t \left[ \sum_{i=D^\perp}^{Mm} \mathcal{T}_N(\hat{V}_i^{Ht}) \mathcal{T}_N^H(\hat{V}_i^{Ht}) \right] h^{tH} \quad (5.5)$$

où  $V_i$  est la colonne  $i$  de  $\hat{V} = [\hat{V}_S \hat{V}_N]$ ,  $h = \text{vec}(\mathbf{H})$  et  $D^\perp = N + M$ . Sous la contrainte  $\|h\| = 1$ ,  $\hat{h}^{tH}$  est le vecteur propre correspondant à la valeur propre minimale (on l'appellera vecteur propre minimal) de la matrice entre crochets. On peut diminuer la charge de calcul en utilisant  $D^\perp > N + M$  avec cependant une perte de performances (voir par exemple [56, 70]).

De manière équivalente, en utilisant les vecteurs propres de l'espace signal, de par :

$$\mathbf{P}_{V_N} = \mathbf{P}_{V_S}^\perp = \mathbf{I} - \mathbf{P}_{V_S} = \mathbf{I} - V_S (V_S^H V_S)^{-1} V_S^H \quad (5.6)$$

nous obtenons la maximisation :

$$\max_{\mathbf{H}} h^t \left[ \sum_{i=1}^{D^\perp-1} \mathcal{T}_N(V_i^{Ht}) \mathcal{T}_N^H(V_i^{Ht}) \right] h^{tH} \quad (5.7)$$

Une version pondérée optimalement peut être obtenue en effectuant une minimisation pondérée dans (5.4) [1].

### 5.3.2 Paramétrisations linéaires du sous-espace bruit

Nous présentons ici différentes paramétrisations linéaires du sous-espace bruit, nécessaires au développement des méthodes d'ajustement du sous-espace bruit, en terme de canal orthogonal.

Considérons d'abord le cas simple de deux canaux :  $m = 2$ , où, comme nous sommes dans le cas d'un seul utilisateur,  $\mathbf{h}_i$  signifie le sous-canal  $i$  de  $\mathbf{H}$ . Pour un signal reçu sans bruit :  $\mathbf{h}_2(z)y_1(k) - \mathbf{h}_1(z)y_2(k) = 0$ , soit, sous forme matricielle :  $[\mathbf{h}_2(z) \ -\mathbf{h}_1(z)] \mathbf{y}(k) = \mathbf{H}^{\perp\dagger}(z) \mathbf{y}(k) = 0$ . La matrice  $\mathbf{H}^{\perp\dagger}(z)$  est exprimée en fonction de la réponse impulsionnelle du canal et satisfait la propriété  $\mathbf{H}^{\perp\dagger}(z) \mathbf{H}(z) = 0$ . Dans le domaine temporel,  $\mathcal{T}^H(\mathbf{H}^\perp)$  génère le complément orthogonal de l'espace de  $\mathcal{T}(\mathbf{H})$  et satisfait la propriété  $\mathcal{T}(\mathbf{H}^\perp) \mathcal{T}(\mathbf{H}) = 0$ . Pour  $m > 2$ , on peut construire plusieurs canaux orthogonaux  $\mathbf{H}^{\perp\dagger}(z)$  en considérant les paires de sous-canaux.

De toute évidence, le choix des différentes paires de sous-canaux conduit à une pléthore de  $\mathbf{H}^{\perp\dagger}(z)$  différents. D'une part, le nombre de paires de sous-canaux à considérer peut varier de  $m - 1$  à  $\frac{m(m-1)}{2}$ . D'autre part, le choix d'un canal orthogonal tel que

$$\text{trace} \{ \mathbf{H}^{\perp\dagger}(z) \mathbf{H}^\perp(z) \} = \alpha \mathbf{H}^\dagger(z) \mathbf{H}(z) \quad (5.8)$$

où  $\alpha$  est un scalaire réel et  $\mathbf{H}^\dagger(z) = \mathbf{H}^H(1/z)$  apporte, dans le cas de la méthode SRM (Sub-channel Response Matching : voir section 5.4.3) par exemple, des avantages importants (i.e. un estimateur non-biaisé du canal). Nous appellerons les canaux orthogonaux de ce type des "*canaux orthogonaux équilibrés*" et nous les flanquerons de l'indice *bal* (pour "balanced").

Le nombre minimal<sup>3</sup> de lignes pour un canal orthogonal équilibré  $\mathbf{H}^{\perp\dagger}(z)$  est  $m$ . Nous avons par exemple :

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

$$\mathbf{H}_{bal,min}^{\perp\dagger}(z) = \begin{bmatrix} -\mathbf{h}_2(z) & \mathbf{h}_1(z) & 0 & \cdots & 0 \\ 0 & -\mathbf{h}_3(z) & \mathbf{h}_2(z) & \cdots & \vdots \\ \vdots & & \ddots & \ddots & 0 \\ \mathbf{h}_m(z) & 0 & \cdots & 0 & -\mathbf{h}_1(z) \end{bmatrix} \quad (5.9)$$

3. Notons que la notation  $\mathbf{H}_{min}^{\perp\dagger}(z)$  signifie un canal orthogonal paramétrisé linéairement par  $\mathbf{H}(z)$  et ayant un nombre minimal de lignes, le degré de  $\mathbf{H}_{min}^{\perp\dagger}(z)$  n'est pas minimal (le degré minimal étant donné par  $\bar{P}$  introduit au chapitre 6 est toujours égal à  $\bar{L}$ )

### 5.3.3 Ajustement du sous-espace bruit

En se basant sur la paramétrisation linéaire du sous-espace bruit et la décomposition en vecteurs propres de la matrice de corrélation (5.2), nous pouvons écrire le problème d'ajustement du sous-espace bruit :

$$\min_{\mathbf{H}, T} \left\| \mathcal{T}^H(\mathbf{H}^\perp) - V_N Q \right\|_F \quad (5.10)$$

Cette minimisation est également séparable, et après optimisation par rapport à  $Q$ , nous obtenons :

$$\min_{\|\mathbf{H}\|_2=1} \text{trace} \{ \mathcal{T}(\mathbf{H}^\perp) \mathbf{P}_{V_N}^\perp \mathcal{T}^H(\mathbf{H}^\perp) \} = \min_{\|\mathbf{H}\|_2=1} \mathbf{H}^H A \mathbf{H} \quad (5.11)$$

ou

$$\max_{\|\mathbf{H}\|_2=1} \text{trace} \{ \mathcal{T}(\mathbf{H}^\perp) \mathbf{P}_{V_N} \mathcal{T}^H(\mathbf{H}^\perp) \} = \max_{\|\mathbf{H}\|_2=1} \mathbf{H}^H A \mathbf{H} \quad (5.12)$$

où la matrice  $A$  peut être aisément construite à partir des vecteurs propres de la matrice de corrélation et  $\widehat{\mathbf{H}}$  est obtenu par le vecteur propre minimal (ou maximal) de  $A$ .

## 5.4 Méthodes sans décomposition en vecteurs propres

Nous utiliserons l'acronyme SFWE : Subspace Fitting Without Eigendecomposition.

### 5.4.1 Ajustement du sous-espace signal

#### La méthode

Nous pouvons écrire la matrice de covariance "débruitée" :

$$R_{YY} - \sigma_v^2 I = R = \mathcal{H} A A^H \mathcal{H}^H = V_S \Lambda_S' V_S^H \quad (5.13)$$

où  $\Lambda_S' = \Lambda_S - \sigma_v^2 I$ . Cette expression nous indique que les espaces colonne de  $\mathcal{H}$  et  $R$  sont identiques. Nous pouvons donc écrire le problème d'ajustement de sous-espaces suivant :

$$\min_{\mathbf{H}, Q} \left\| \mathcal{H} - \hat{R} B Q \right\|_F^2 \quad (5.14)$$

où  $\hat{R}$  est une estimée consistante de  $R$  et  $F = \hat{R} B$  est une estimée de l'espace signal. La matrice  $B$  a les mêmes dimensions que  $\mathcal{H}$  et est fixée, nous verrons plus loin comment le choix de  $B$  influence les performances de l'estimation de canal. Nous pouvons prendre  $\hat{R} = \hat{R}_{YY} - \hat{\sigma}_v^2 I =$

$\hat{R}_{YY} - \lambda_{\min}(\hat{R}_{YY})I$  où  $\lambda_{\min}(\cdot)$  est la valeur propre minimale. Une autre approche (qui permet de s'affranchir de tout calcul de valeurs propres) consiste à utiliser l'estimée de la variance du bruit données par la méthode SRM (5.4.3). De plus, les simulations montrent que l'utilisation de  $\hat{R} = \hat{R}_{YY}$  peuvent donner des résultats très honorables.

Le problème (5.14) est séparable en  $\mathbf{H}$  et  $Q$ . La minimisation par rapport à  $Q$  conduit à :

$$Q = (F^H F)^{-1} F^H \mathcal{H}. \quad (5.15)$$

En substituant dans (5.14), nous avons le problème équivalent :

$$\min_{\mathbf{H}} \|\mathbf{P}_F^\perp \mathcal{H}\|_F^2 = \min_{\mathbf{H}} \text{trace} \left\{ \mathcal{H}^H \mathbf{P}_F^\perp \mathcal{H} \right\}. \quad (5.16)$$

Soit, sous la contrainte  $\|\mathbf{H}\| = 1$  :

$$\begin{aligned} \hat{\mathbf{H}} &= \arg \min_{\|\mathbf{H}\|=1} \text{trace} \left\{ \mathcal{H}^H \mathbf{P}_F^\perp \mathcal{H} \right\} \\ &= \arg \min_{\|\mathbf{H}\|=1} \mathbf{H}^t \underbrace{\mathcal{B}_N^H(\mathcal{P}_F^\perp) \mathcal{B}_N(\mathcal{P}_F^\perp)}_{=\mathcal{F}} \mathbf{H}^{Ht} \end{aligned} \quad (5.17)$$

où  $\mathbf{P}_F^\perp = [\mathbf{P}_F^\perp(0), \dots, \mathbf{P}_F^\perp(L-1)]$ ,  $\mathbf{P}_F^\perp(\tau)$  sont des matrices de taille  $Mm \times m$  et  $\mathcal{B}_N(\mathbf{P})$  est une matrice bloc Toeplitz dont la première colonne bloc vaut  $\begin{bmatrix} \mathbf{P} \\ 0 \end{bmatrix}$ .

La solution est donc  $V_{\max}(\mathcal{F})$ , le vecteur propre maximal de  $\mathcal{F}$ . Comme  $R = \mathcal{H} \mathbf{A} \mathbf{A}^H \mathcal{H}^H$ ,  $B$  ne peut pas être quelconque, par exemple, si les colonnes de  $B$  sont dans le sous-espace bruit de  $R$ , alors  $F = 0$  pour  $\hat{R} = R$ . Intuitivement, le choix optimal pour  $B$  devrait être  $B = \mathcal{H}$ , ce qui correspond au filtrage adapté de  $\mathcal{H}^H$  par  $\mathcal{H}$  (la post-multiplication de  $B$  par une matrice carrée non singulière n'a pas d'influence, puisqu'elle peut être absorbée dans  $Q$ ). Ces considérations mènent à l'algorithme suivant, en deux étapes.

Étape 1 Dans un premier temps,  $B$  est choisi de manière arbitraire. Cette première étape fournit une estimée consistante du canal (si  $\mathcal{H}^H B$  est non singulière). Il est à noter que cette première étape pourrait être remplacée par une procédure d'estimation de canal SRM, qui, outre qu'elle nous fournit une estimée consistante du canal, nous donne également une estimée consistante de la variance du bruit.

Étape 2 Dans un second temps, la première estimée consistante du canal est utilisée pour former  $\hat{\mathcal{H}}$  et nous résolvons (5.14) avec  $B = \hat{\mathcal{H}}$ .

Pour la première étape, le choix de  $B^H = [I \ 0]$  fournit un algorithme assez proche de la méthode de «Pisarenko Rectangulaire» de Fuchs [31]. D'un point de vue pratique, l'utilisation d'une matrice  $B$  ayant la même structure bloc Toeplitz que  $\mathcal{H}$ , mais formée à partir d'un canal généré aléatoirement, donne de meilleurs résultats (c'est ce choix qui a été fait dans les simulations).

**Analyse asymptotique : estimation exacte des statistiques**

Asymptotiquement,  $\hat{R} = R$ . Nous obtenons donc  $F = RB = \mathcal{H}\mathbf{A}\mathbf{A}^H\mathcal{H}^HB$ . Dans les conditions d'identifiabilité,  $\mathbf{A}\mathbf{A}^H > 0$  et si  $\mathcal{H}^HB$  est non singulière :

$$\mathbf{P}_F = \mathbf{P}_{\mathcal{H}} = \mathbf{P}_{V_S}. \quad (5.18)$$

De plus, si nous disposons d'une estimée consistante du canal, nous pouvons asymptotiquement considérer  $B = \mathcal{H}$ . Dans ce cas, l'utilisation de  $R = R_{YY}$  et donc de  $F = R_{YY}\mathcal{H}$  conduit également à l'équation (5.18).

**Analyse au premier ordre**

Pour la première étape de l'algorithme, nous avons une estimée consistante du canal si  $\mathcal{H}^HB$  est non singulière. Nous pouvons poursuivre l'analyse asymptotique en appliquant une perturbation à la matrice de projection. Si  $\hat{F} = F + \Delta F$ , alors l'approximation au premier ordre en  $\Delta F$  de la matrice de projection donne  $\mathbf{P}_{\hat{F}} = \mathbf{P}_F + \Delta\mathbf{P}_F$  où

$$\Delta\mathbf{P}_F = 2Sym\left(\mathbf{P}_F^\perp \Delta F (F^H F)^{-1} F^H\right) \quad (5.19)$$

avec  $2Sym(\mathbf{X}) = \mathbf{X} + \mathbf{X}^H$ .

**Optimalité de  $B = \mathcal{H}$**  Soit  $\hat{R} = R + \Delta R$ , en utilisant la décomposition en vecteurs propres de  $R$ , nous obtenons l'approximation au premier ordre suivante :

$$\begin{aligned} \Delta R &= \Delta V_S \Lambda'_S V_S^H + V_S \Delta \Lambda'_S V_S^H + V_S \Lambda'_S \Delta V_S^H \\ &+ \Delta V_N \Lambda'_N V_N^H + V_N \Delta \Lambda'_N V_N^H + V_N \Lambda'_N \Delta V_N^H \end{aligned} \quad (5.20)$$

De  $\hat{F} = \hat{R}B$  on déduit :

$$\begin{aligned} \Delta\mathbf{P}_F &= 2Sym\left(\mathbf{P}_F^\perp \Delta F (F^H F)^{-1} F^H\right) \\ &= 2Sym\left(\mathbf{P}_{V_N} \Delta F (V_S^H \mathcal{H})^{-1} (\Lambda'_S)^{-1} V_S^H\right) \end{aligned} \quad (5.21)$$

soit, en posant  $B = V_S B_S + V_N B_N$  où  $B_S$  est supposé non-singulier.

$$\Delta\mathbf{P}_F = 2Sym\left(\mathbf{P}_{V_N} \left[\Delta V_S \Lambda'_S + V_N \Delta \Lambda'_N B_N B_S^{-1}\right] (\Lambda'_S)^{-1} V_S^H\right) \quad (5.22)$$

ce qui montre l'optimalité de  $B_N = 0$ , en effet, dans ce cas,

$$\Delta\mathbf{P}_F|_{\hat{F}=\hat{R}B} = \Delta\mathbf{P}_F|_{\hat{F}=\hat{V}_S} = 2Sym\left(\mathbf{P}_{V_N} \Delta V_S V_S^H\right) \quad (5.23)$$

**Équivalence asymptotique de l'approche SFWE et de l'approche d'ajustement du sous-espace signal** En calculant les matrices de projection avec  $\hat{F} = (\hat{R}_{YY} - \lambda_{\min}(\hat{R}_{YY})I)\hat{\mathcal{H}}$ , où  $\hat{R}_{YY}$  et  $\hat{\mathcal{H}}$  sont des estimées consistantes, on montre aisément que  $\Delta\mathbf{P}_F$  est identique si on utilise  $\hat{F} = \hat{V}_S$  ou  $\hat{F} = \hat{R}B$ .

En effet, pour  $F = R\mathcal{H}$ , de manière similaire à (5.21),

$$\begin{aligned}\Delta\mathbf{P}_F &= 2Sym(\mathbf{P}_F^\perp \Delta F (F^H F)^{-1} F^H) \\ &= 2Sym(\mathbf{P}_{V_N} \Delta F (V_S^H \mathcal{H})^{-1} (\Lambda_S')^{-1} V_S^H)\end{aligned}\quad (5.24)$$

La perturbation au premier ordre de  $\hat{F} = (\hat{R}_{YY} - \lambda_{\min}(\hat{R}_{YY})I)\hat{\mathcal{H}}$  vaut :

$$\Delta F = (\Delta R - \Delta\sigma_v^2)\mathcal{H} + (R)\Delta\mathcal{H} = (\Delta R - \Delta\sigma_v^2)\mathcal{H} + (V_S \Lambda_S' V_S^H)\Delta\mathcal{H}\quad (5.25)$$

Le second terme de  $\Delta F$  n'a pas d'influence (de par :  $\mathbf{P}_{V_N} V_S = 0$ ). Le terme en  $\Delta\sigma_v^2\mathcal{H}$  est également annulé par  $\mathbf{P}_{V_N} \mathcal{H} = 0$ , donc :

$$\begin{aligned}\Delta\mathbf{P}_F &= 2Sym(\mathbf{P}_{V_N} \Delta V_S \Lambda_S' V_S^H \mathcal{H} (V_S^H \mathcal{H})^{-1} (\Lambda_S')^{-1} V_S^H) \\ &= 2Sym(\mathbf{P}_{V_N} \Delta V_S V_S^H)\end{aligned}\quad (5.26)$$

ce qui est bien l'expression de la matrice de projection pour  $\hat{F} = \hat{V}_S$ .

Donc, au premier ordre :

$$\mathbf{P}_{(\hat{R}_{YY} - \lambda_{\min}(\hat{R}_{YY})I)\hat{\mathcal{H}}} = \mathbf{P}_{(\hat{R}_{YY} - \sigma_v^2 I)\mathcal{H}} = \mathbf{P}_{\hat{V}_S}.\quad (5.27)$$

Cette équation montre que notre méthode sous-espace sans décomposition en vecteurs propres donne asymptotiquement les mêmes performances que la « vraie » méthode d'ajustement du sous-espace signal. De plus, tant que des estimées consistantes de  $\sigma_v^2$  et  $\mathcal{H}$  sont utilisées, les erreurs d'estimation sur ces grandeurs n'ont pas d'influence au premier ordre.

**Méthode simplifiée** En utilisant  $\hat{F} = \hat{R}_{YY}\hat{\mathcal{H}}$  :

$$\Delta\mathbf{P}_F = 2Sym(\mathbf{P}_{V_N} (\Delta V_S + \Delta\mathcal{H}(V_S^H \mathcal{H})^{-1} \Lambda_N \Lambda_S^{-1}) V_S^H)\quad (5.28)$$

L'utilisation de  $\hat{R}_{YY}$  en lieu et place de  $\hat{R}$  fait apparaître un second terme dont l'importance est proportionnelle à  $\Lambda_N \Lambda_S^{-1}$ . Ce terme est donc négligeable à SNR élevé.

### Méthode pondérée

Gorokhov a montré [1] que la matrice de pondération optimale pour l'ajustement du sous-espace signal, correspondant à :

$$\widehat{\mathbf{H}} = \arg \min_{\|\mathbf{H}\|=1} \mathbf{H}^t \mathcal{B}_N^H(\mathbf{P}_F^\perp) W \mathcal{B}_N(\mathbf{P}_F^\perp) \mathbf{H}^{Ht} \quad (5.29)$$

est  $W_o = \Sigma^\#$  où  $\Sigma$  est la matrice de covariance de  $\mathcal{B}_N(\mathbf{P}_{V_N}) \mathbf{H}^{Ht}$ , et peut être exprimée en fonction de  $\mathbf{P}_{V_N}$ ,  $\mathbf{A}$  et  $\mathbf{H}$ . De plus,  $\Sigma$  peut être remplacé par une estimée consistante  $\widehat{\Sigma}$  sans influence sur les performances asymptotiques. Il s'ensuit qu'on peut aisément implémenter une méthode sous-espace pondérée sans décompositions en vecteurs propres. Il suffit en effet de calculer la matrice de pondération à partir des valeurs déterminées dans la première étape et de l'intégrer dans la seconde étape.

### Algorithme rapide

Si on utilise des matrices de grande taille, la charge de calcul de la nouvelle méthode sera fortement influencée par la manière dont l'inversion de la matrice de projection  $I - F(F^H F)^{-1} F^H$  sera effectuée, les autres optimisations de l'algorithme (e.g. les convolutions rapides) étant identiques que pour l'algorithme classique. On peut recourir aux algorithmes rapides d'inversion des matrices presque-Toeplitz (voir [48] et les articles associés).

**Définition 5.1** *Le déplacement d'une matrice Hermitienne  $n \times n$  est défini par*

$$\Delta R_{\mu,\nu} \triangleq R - Z^\mu R Z^{\nu T}, \quad (5.30)$$

où  $Z$  est une matrice de décalage de dimension  $n \times n$  avec des 1 sur la 1<sup>ère</sup> sous-diagonale et des zéros ailleurs<sup>4</sup>.

Le rang  $r$  de  $\Delta R$  est appelé rang de déplacement et on peut montrer aisément que le rang de déplacement de  $R_{YY}(-\sigma_v^2 I)$  vaut  $2m$ .

**Lemme 5.1** *Le rang de déplacement de  $F = RB$  est au plus égal à  $3m$ , i.e.,*

$$\text{rang}(\Delta_{\{Z^m, Z^1\}} F) = 3m \text{ où}$$

$$\Delta_{\{Z^m, Z^1\}} F = F - Z^m F Z^T \quad (5.31)$$

et  $Z^i$  est une matrice de taille adéquate, dont les éléments sont nuls partout, sauf la  $i^{\text{ème}}$  sous-diagonale dont les éléments valent 1.

---

4. On trouvera des définitions plus générales dans [48, 12].

En utilisant les propriétés des matrices Toeplitz et presque-Toeplitz, nous savons que le rang de déplacement de  $F = RB$  est le rang de déplacement de  $\mathcal{M} = \begin{bmatrix} I & B \\ R & 0 \end{bmatrix}$ .

De  $\text{rang}(\Delta_{\{Z^m, Z^m\}}R) = 2m$  et  $\text{rang}(\Delta_{\{Z^m, Z^1\}}B) = m$ , on déduit aisément

$\text{rang}(\Delta_{\{Z^m \oplus Z^m, Z^m \oplus Z^1\}}\mathcal{M}) \leq 3m$  où  $A \oplus B = \begin{bmatrix} A & 0 \\ 0 & B \end{bmatrix}$ , ce qui termine la démonstration.

◇

**Lemme 5.2** *Le rang de déplacement de  $\mathbf{P}_F^\perp$  est au plus égal à  $6m$ .*

Soit la matrice  $\mathcal{M} = \begin{bmatrix} F^H F & F^H \\ F & I \end{bmatrix}$ , on peut aisément déterminer son rang de déplacement :  $\text{rang}(\Delta_{\{Z^1 \oplus Z^m, Z^1 \oplus Z^m\}}\mathcal{M}) \leq 6m$ .  $\mathbf{P}_F^\perp$  est le complément de Schur de  $F^H F$  dans  $\mathcal{M}$  et a donc le même rang de déplacement par rapport à  $\{Z^m, Z^m\}$ .

◇

Ayant déterminé le rang de déplacement de la matrice de projection, [48] nous apprend que nous pouvons construire un algorithme rapide de complexité  $O((6m)^3 K \log(K))$  où  $K$  est la taille de  $\hat{R}$  en nombre de blocs (le recours à ce type d'algorithme n'est donc intéressant que pour des tailles de matrice importantes, en comparant avec  $O(m^3 K^3)$ ).

### 5.4.2 Ajustement du sous-espace bruit

Une formulation de l'ajustement du sous-espace bruit sans décomposition en vecteurs propres de la matrice de corrélation est la suivante (sous une contrainte de non-trivialité):

$$\min_{\mathbf{H}} \|\mathcal{T}(\mathbf{H}^\perp)R^{\frac{1}{2}}\|_F^2 \quad (5.32)$$

où la racine carrée de la matrice de corrélation est de la forme  $R^{\frac{1}{2}} = V_S \Lambda_S'^{\frac{1}{2}} Q$  pour une matrice unitaire  $Q$  ( $\mathcal{T}(\mathbf{H})\mathbf{A}\mathbf{A}^H \mathcal{T}^H(\mathbf{H}) = V_S \Lambda_S' V_S^H + V_N \Lambda_N' V_N^H$  où  $\Lambda_S' = \Lambda_S - \sigma_v^2 I$  et  $\Lambda_N' = 0$ ).

De toute évidence, l'équation (5.31) est une version pondérée par  $\Lambda_S'$  de l'équation par (5.11) et les performances asymptotiques ne seront donc pas identiques. Par contre, la version pondérée optimalement donnera les mêmes résultats, ce qui nous suffit pour tirer des conclusions sur les équivalences entre les diverses méthodes sous-espaces et autres.

### 5.4.3 D'autres méthodes utilisant la paramétrisation du sous-espace bruit

#### Ajustement croisé des sous-canaux (SRM : Subchannel Response Matching)

En se basant sur les développements conduisant aux canaux orthogonaux, nous avons  $\mathcal{T}(\mathbf{H}^\perp)\mathbf{Y} = 0$  et il est naturel de vouloir résoudre le problème

$$\min_{\mathbf{H}} \left\| \mathcal{T}(\mathbf{H}^\perp) \mathbf{Y} \right\|_2^2 \quad (5.33)$$

C'est la méthode SRM, proposée initialement par Gurelli et Nikias [40] et réinventée sous des formes légèrement différentes par Xu et al. [79] et Baccala et Roy [7].

Pour  $\mathbf{H}_{bal,min}^{\perp\dagger}(z)$ , le critère de la méthode SRM peut s'écrire sous la forme (sous une contrainte de non-trivialité):

$$\begin{aligned} \text{trace} \{ \mathcal{T}(\mathbf{H}^\perp) \mathbf{Y} \mathbf{Y}^H \mathcal{T}^H(\mathbf{H}^\perp) \} &= \text{trace} \left\{ \mathbf{H}^\perp \left( \sum_{k=N-1}^{M-1} \mathbf{Y}_N(k) \mathbf{Y}_N^H(k) \right) \mathbf{H}^{\perp H} \right\} \\ &= (M-N+1) \text{trace} \{ \mathbf{H}^\perp \hat{R}_{YY} \mathbf{H}^{\perp H} \} \end{aligned} \quad (5.34)$$

Par une transformation relativement simple, nous pouvons réécrire le critère SRM, pour une certaine matrice  $A$ , sous la forme :

$$\min_h h^H A h \quad (5.35)$$

Si la matrice de covariance  $R_{YY}$  exacte est utilisée, la contribution du bruit dans (5.33) est donnée par

$$\text{trace} \{ \sigma_v^2 \mathbf{H}^\perp \mathbf{H}^{\perp H} \} = 2\sigma_v^2 \|\mathbf{H}\|^2 \quad (5.36)$$

grâce à la forme particulière du canal orthogonal équilibré de taille minimale. Il s'ensuit que, en considérant la longueur du canal comme étant connue, la minimisation de (5.34) sous la contrainte de non trivialité  $\|\mathbf{H}\| = 1$  fournit une estimée  $\widehat{\mathbf{H}} = V_{min}(B)$  consistante.

### Méthode dérivée de la prédiction linéaire

Cette méthode utilise une paramétrisation du sous-espace bruit qui n'est pas linéaire en fonction de la réponse du canal ( $\bar{\mathbf{P}}(z)$  proposé dans [70]). On considère que  $\mathbf{P}(z) = \sum_{i=0}^L \mathbf{P}(i) z^{-i}$ , où  $\mathbf{P}(0) = I_m$ , est le prédicteur d'ordre  $L$  pour le signal reçu  $\mathbf{y}(k)$  supposé sans bruit. Si  $L \geq \bar{L}$ , alors (voir chap. 6)

$$\mathbf{P}(z) \mathbf{H}(z) = \mathbf{H}(0) . \quad (5.37)$$

Nous pouvons donc considérer que  $\mathbf{H}(z)$  et  $[\mathbf{P}(z), \mathbf{H}(0)]$  sont des paramétrisations équivalentes. Nous construisons alors la matrice  $\mathbf{H}^\perp(0)$  de taille  $m \times (m-1)$  et de rang plein telle que  $\mathbf{H}^{\perp H}(0) \mathbf{H}(0) = 0$ . Alors,  $\bar{\mathbf{P}}(z) = \mathbf{H}^{\perp H}(0) \mathbf{P}(z)$  est un polynôme matriciel qui satisfait :

$$\bar{\mathbf{P}}(z)\mathbf{H}(z) = 0. \quad (5.38)$$

A partir de cette équation, nous pouvons poser un problème de minimisation quadratique similaire à ceux rencontrés précédemment. On notera que c'est une méthode pour identifier le canal par prédiction linéaire dans le cadre du modèle déterministe.

### Équivalences entre méthodes

La comparaison des critères de la méthode sous-espace bruit sans décomposition en vecteurs propres  $\min_{\mathbf{H}} \|\mathcal{T}(\mathbf{H}^\perp) R^{\frac{1}{2}}\|_F^2$  et de la méthode SRM  $\text{trace} \{ \mathbf{H}^\perp R_{YY} \mathbf{H}^{\perp H} \}$  montre immédiatement que les deux méthodes sont équivalentes.

L'équivalence avec la méthode basée sur la paramétrisation par le prédicteur linéaire fait l'objet du théorème suivant :

**Théorème 5.1** *Soit  $\mathbf{H}^{\perp \dagger}(z) = \bar{\mathbf{P}}(z)$ . Le critère sous-espace bruit sans décomposition en vecteurs propres est équivalent à celui basé sur la paramétrisation par le prédicteur linéaire.*

En effet,

$$\begin{aligned} & \min_{\mathbf{H}} \|\mathcal{T}(\mathbf{H}^\perp) R^{\frac{1}{2}}\|^2 = \min_{\bar{\mathbf{P}}} \|\mathcal{T}(\bar{\mathbf{P}}) R^{\frac{1}{2}}\|^2 \\ &= \min_{\bar{\mathbf{P}}} \text{trace} \left\{ \mathcal{T}(\bar{\mathbf{P}}) R^{\frac{1}{2}} R^{\frac{H}{2}} \mathcal{T}^H(\bar{\mathbf{P}}) \right\} \\ &= \min_{\bar{\mathbf{P}}} \text{trace} \left\{ \mathcal{T}(\bar{\mathbf{P}}) R \mathcal{T}^H(\bar{\mathbf{P}}) \right\} \\ &= \min_{\bar{\mathbf{P}}} \text{trace} \left\{ \bar{\mathbf{P}} \left( \sum_{k=N-1}^{M-1} R(k) \right) \bar{\mathbf{P}}^H \right\} \\ &= \min_{\bar{\mathbf{P}}} (M - N + 1) \text{trace} \left\{ \bar{\mathbf{P}} \hat{R}_{YY} \bar{\mathbf{P}}^H \right\} \\ &= \min_P (M - N + 1) \\ & \quad \text{trace} \left\{ \mathbf{H}^{\perp H}(0) P \hat{R}_{YY} P^H \mathbf{H}^\perp(0) \right\} \end{aligned} \quad (5.39)$$

Cette équation coïncide avec les équations normales :

$$\min_{P(1:L)} [P(0:L)] \hat{R}_{YY} [P(0:L)]^H \Rightarrow \quad (5.40)$$

$$\left\{ \begin{array}{l} \left[ \begin{array}{cc} I & P(1:L) \end{array} \right] \hat{R}_{YY} = \left[ \begin{array}{ccc} \sigma_y^2 & 0 & \dots & 0 \end{array} \right] \\ \min_{\mathbf{H}(0)} \mathbf{H}^{\perp H}(0) \sigma_y^2 \mathbf{H}^\perp(0) \Rightarrow \mathbf{H}(0) = V_{max}(\sigma_y^2) \end{array} \right. \quad (5.41)$$

où  $V_{max}(A)$  est le vecteur propre maximal de  $A$

### 5.4.4 Équivalences

En fonction des sections précédentes, nous pouvons établir le tableau d'équivalences suivant, où = signifie une équivalence stricte et  $\sim$  signifie une équivalence asymptotique (i.e. nombre de données infini). Remarquons que les équivalences pour la prédiction linéaire pondérée optimalement dépend de la paramétrisation utilisée.

		WSSF	WEFSSF	WNSF	WEFNSF	WSRM	WLP
WSSF	=	$\sim$	$\sim$	$\sim$	$\sim$		
WEFSSF		=	$\sim$	$\sim$	$\sim$		
WNSF			=	$\sim$	$\sim$	$\sim$	
WEFNSF				=	=	$\sim$	
WSRM					=	$\sim$	
WLP							=

- WSSF: Weighted Signal Subspace Fitting
- WEFSSF: Weighted Eigendecomposition Free SSF
- WNSF: Weighted Noise Subspace Fitting
- WEFNSF: Weighted Eigendecomposition Free NSF
- WSRM: Weighted Subchannel Response Matching
- WLP: Weighted Linear Prediction

### 5.4.5 Simulations

#### Sous-espace signal

Nous avons utilisé un canal de longueur  $6T$  généré aléatoirement comprenant  $m = 3$  canaux. Nous traçons le NRMSE (Normalized Root Mean Squared Error) défini par

$$\text{NRMSE} = \sqrt{\frac{1}{N_{mc}} \sum_{l=1}^{N_{mc}} \|\hat{\mathbf{H}}^{(l)} - \mathbf{H}\|_F^2 / \|\mathbf{H}\|_F^2} \quad (5.42)$$

où  $\hat{\mathbf{H}}^{(l)}$  est le canal estimé lors de la  $l^e$  réalisation et  $N_{mc}$  est le nombre de réalisations effectuées.

Le rapport signal/bruit est défini par  $\text{SNR} = (\|\mathbf{H}\|_F^2 \sigma_a^2) / (mM\sigma_v^2)$ . La longueur des trames est de  $M = 100$  symboles QAM-4 et nous avons effectué  $N_{mc} = 100$  tirs Monte-Carlo.

Nous traçons les courbes de performances pour la première étape de l'algorithme, pour la seconde étape et pour l'algorithme classique (avec décompositions en vecteurs propres). Ces courbes montrent que l'algorithme proposé a les mêmes performances que l'algorithme original.

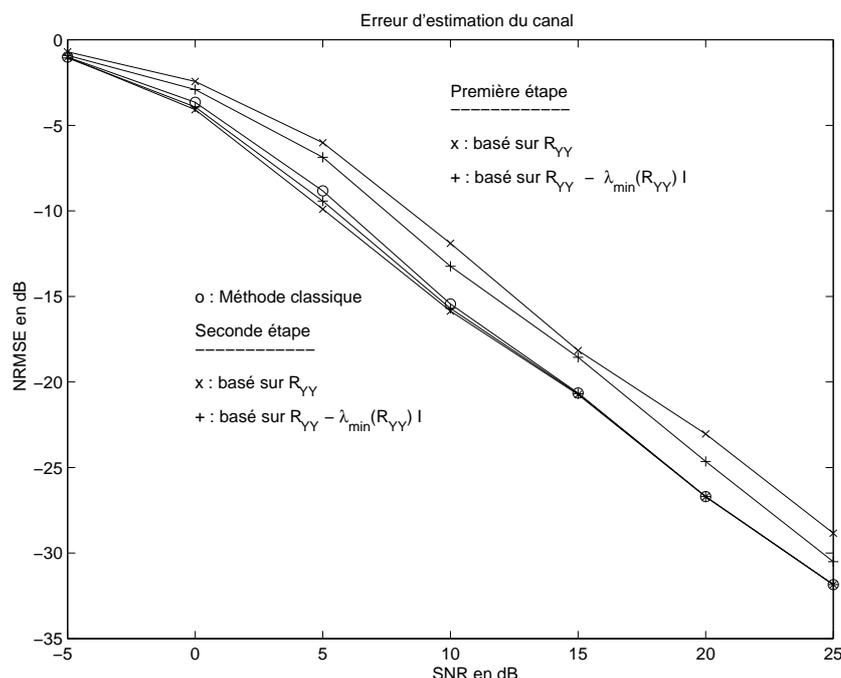


FIG. 5.1 – Performance des divers algorithmes d’ajustement du sous-espace signal

De plus, nous avons effectué les simulations avec  $\hat{R}_{YY}$  and  $\hat{R}_{YY} - \lambda_{\min}(\hat{R}_{YY})I$ . Les courbes sont également confondues à l’issue de la seconde étape.

Pour illustrer la faible influence du débruitage de la matrice de corrélation, nous avons également tracé le NRMSE de l’estimation de canal quand on utilise les statistiques exactes (i.e.  $R_{YY}$ ). Ces courbes montrent qu’à la deuxième étape, l’influence de  $\sigma_v^2$  est devenue négligeable.

### Sous-espace bruit

Nous comparons la performance des méthodes sous-espace bruit classique et sans décompositions en vecteurs propres. La comparaison des performances montre que les deux algorithmes donnent des résultats très voisins.

## 5.5 Robustesse à la surestimation de la longueur du canal

Nous présentons ici une tentative de "robustification" de la méthode sous-espace à la surestimation de la longueur du canal. En effet, les méthodes sous-espace exigent une connaissance a priori de la longueur du canal, sous peine de ne pas obtenir un estimateur consistant du canal (voir par exemple [56]). Partant de la constatation que la méthode de la prédiction linéaire est robuste à la surestimation de l’ordre, l’idée est d’essayer de s’inspirer de ce qui est fait dans ce cadre et de

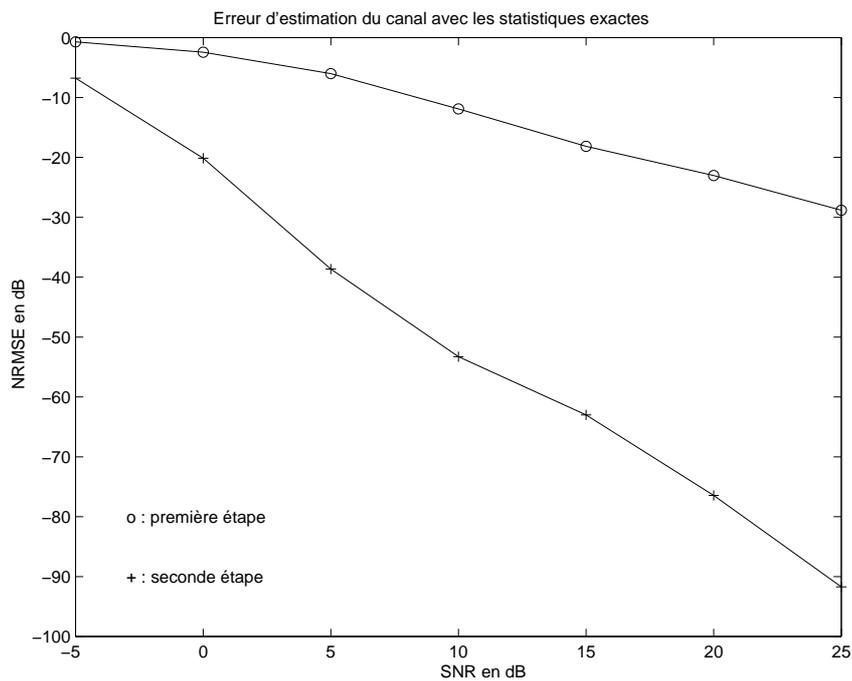


FIG. 5.2 – Influence du non-débruitage de  $R_{YY}$

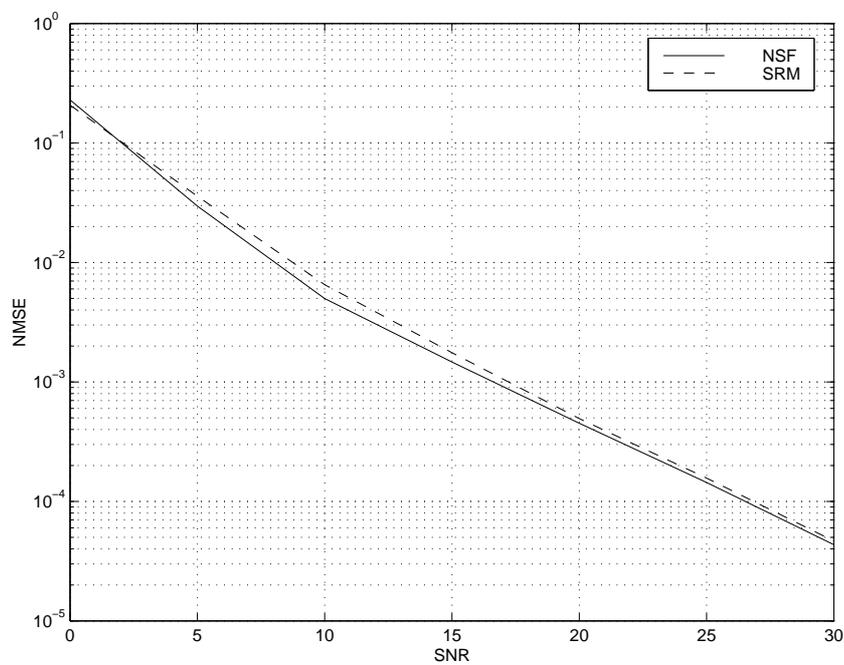


FIG. 5.3 – Performances des méthodes sous-espace bruit.

l'appliquer, d'un point de vue algorithmique, aux méthodes sous-espace. En particulier, la robustesse de la prédiction linéaire en termes de robustesse à la surestimation de l'ordre du prédicteur est intimement liée au fait que le premier terme du prédicteur est l'identité. La transposition, dans le cadre de la méthode sous-espace, consistera à décomposer le canal en un produit comprenant un canal statique et un canal monique (ce type de décomposition a été introduit également dans [14, 13]).

### 5.5.1 L'approche de base

L'approche que nous considérons consiste à écrire  $\mathbf{H}(z)$  comme  $\mathbf{Q}(z)\mathbf{H}(0)$  où  $\mathbf{H} = [\mathbf{H}^H(N-1) \dots \mathbf{H}^H(0)]^H$  comme  $\mathbf{Q}\mathbf{H}(0) = [\mathbf{q}^H(N-1) \dots \mathbf{q}^H(1) \mathbf{I}_m]^H \mathbf{H}(0)$  où la matrice carrée  $\mathbf{Q}(z)$  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 toutes les techniques d'estimation aveugle décrites précédemment. 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  $\alpha$  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

$$\min_{\|\mathbf{H}(0)\|_2=1} \mathbf{H}^H(0) \left( \min_{\mathbf{q}(0)=\mathbf{I}_m} \mathbf{Q}^H A \mathbf{Q} \right) \mathbf{H}(0). \quad (5.43)$$

### 5.5.2 Robustesse

#### Première approche

Dans [6], une première approche utilisant cette formulation du problème a été introduite. L'obtention de la robustesse à la surestimation d'ordre se fait en deux étapes. Dans une première

étape, nous réécrivons  $A$  et  $\mathbf{H}$  sous la forme partitionnée suivante :  $A = \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix}$ ,  $\mathbf{H} =$

$\begin{bmatrix} \widetilde{\mathbf{H}} \\ \mathbf{H}(0) \end{bmatrix}$  où  $A_{11}$  est  $m(N-1) \times m(N-1)$  et  $A_{22}$  est  $m \times m$ . La solution de (5.42) est :

$$\begin{aligned} \mathbf{H}(0) &= V_{min} (A_{22} - A_{21} A_{11}^{-1} A_{12}), \\ \widetilde{\mathbf{H}} &= \left[ \mathbf{H}^H(N-1) \dots \mathbf{H}^H(1) \right]^H = -A_{11}^{-1} A_{12} \mathbf{H}(0). \end{aligned} \quad (5.44)$$

L'ancrage de la réponse impulsionnelle en soi ne permet pas d'obtenir la robustesse à la surestimation d'ordre. L'utilisation de la pseudo-inverse de Moore-Penrose  $A_{11}^\#$  pour  $A_{11}$  aboutira à des problèmes de surestimation de l'ordre du canal. La deuxième étape consiste donc à considérer la factorisation triangulaire UDL pour  $A_{11} = UDU^T$  et à forcer les diagonales bloc singulières qui se répètent dans  $D$  à zéro et mettre les éléments de la colonne correspondante dans  $U$ , qui se trouvent en haut de l'élément diagonal, à zéro. Dans [6], nous démontrons que cette procédure en deux étapes permet de garantir la robustesse à la surestimation de l'ordre du canal. Malheureusement, une implémentation rigoureuse de cet algorithme nécessite le recours à des tests statistiques (pour la mise à zéro d'éléments petits) dont le développement nous a paru peu intéressant au regard de la seconde approche.

### Seconde approche

Dans les deux cas traités ci-dessous (SRM et sous-espace), on peut aisément conclure que le noyau de la matrice  $A$  est, dans le cas exact, de dimension  $N' - N + 1$ , où  $N'$  est l'ordre (sur)estimé (la généralisation aux autres méthodes ne pose pas de problème particulier). De plus, cette matrice est Toeplitz par blocs, ce qui permet d'en effectuer une factorisation UDL par l'algorithme rapide de Levinson multicanal adapté au cas singulier. Pour  $N' = N$ , le dernier 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 [69], les éléments  $Q(n)$ ,  $n > N$  deviennent nuls. D'autre part, l'élément  $\mathbf{H}(0)$  est déterminé comme étant le vecteur propre minimal du dernier élément de la matrice bloc diagonale et  $\mathbf{H} = \mathbf{QH}(0)$ . La robustesse à la surestimation 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<sup>5</sup> (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 de l'ordre nous donnant "gratuitement" l'estimée du canal.

---

5. en fait, il conviendrait de poursuivre en utilisant des critères de type AIC, MDL et consorts, pour obtenir un algorithme complet. Ce travail ayant été fait en collaboration avec Jahouar Ayadi, nous lui laissons le soin de finaliser ce projet.

### 5.5.3 Simulations

Dans les simulations que nous présentons, nous mesurons la performance des méthodes à travers l'Erreur Quadratique Moyenne Normalisée (NRMSE), moyennée sur 300 réalisations Monte Carlo, et donnée par  $\text{NRMSE} = \frac{1}{300} \sqrt{\sum_{i=1}^{300} \|\widehat{\mathbf{H}}^{(i)} - \mathbf{H}\|^2 / \|\mathbf{H}\|^2}$  où l'ambiguïté fondamentale a été levée dans  $\widehat{\mathbf{H}}$  de telle manière que le NRMSE soit minimal. Nous utilisons un canal complexe généré de façon aléatoire  $\mathbf{H}$  avec  $N = 3$  et  $m = 3$ . Les symboles sont i.i.d. BPSK, et la longueur de la trame des données est  $M = 100$ .

Dans la figure 5.5.3, nous illustrons la robustesse de notre approche dans le cas de l'ajustement du sous-espace signal. Nous avons utilisé un ensemble non complet de vecteurs de l'espace bruit pour estimer  $\mathbf{H}$  en supposant que son ordre est  $N' = 5 > 3 = N$ . Les coefficients  $q(N'-1) \cdots q(1)$  sont calculés par l'intermédiaire de l'algorithme de Levinson multicanal. Dans cet algorithme, les inverses des matrices représentant l'équivalent des puissances d'erreurs dans le cas de la prédiction linéaire sont remplacés par des pseudo-inverses. En-deçà de ce seuil, les valeurs propres sont considérées comme étant nulles et les vecteurs propres associés comme faisant partie du noyau de la matrice. En fait, si on continue l'algorithme, les coefficients  $Q(N)$  suivants sont, en théorie, nuls. Nous pouvons donc interpréter l'apparition d'une singularité dans les matrices de puissance d'erreur de prédiction comme étant une détection (implicite) de l'ordre du canal. Dans la simulation ci-dessous, nous avons implémenté cette stratégie, qui donne le bon ordre 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 NRMSE par rapport à la méthode classique, alors que l'échec de l'utilisation de la contrainte  $\|\mathbf{H}\| = 1$  est clair.

## 5.6 Conclusions

Nous avons proposé un nouvel algorithme pour résoudre le problème d'identification aveugle de canal par la méthode du sous-espace signal, ne requérant pas de décompositions en vecteurs propres. Une analyse asymptotique montre que cet algorithme offre des performances similaires à celles de la méthode classique, ce que confirment les simulations. De plus, la possibilité d'implémenter des algorithmes rapides a été introduite.

Ensuite, en développant les expressions des critères de différentes méthodes proposées dans la littérature, nous avons démontré les équivalences asymptotiques entre ces différents algorithmes.

Dans la dernière partie de ce chapitre, nous avons proposé une approche qui permet de rendre bon nombre de méthodes robuste à la surdétermination de la longueur du canal. Il est cependant à noter que la méthode proposée ressemble à une détermination implicite de la longueur du canal.

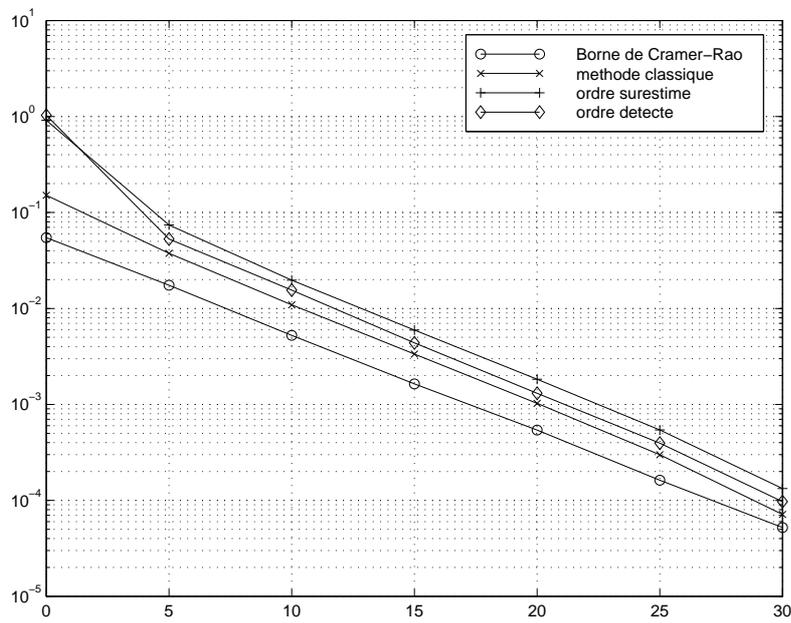


FIG. 5.4 – SSF : robustesse à la surestimation de l'ordre du canal avec les contraintes  $\|\mathbf{H}\|_2 = 1$  et  $\|\mathbf{H}(0)\|_2 = 1$

# Chapitre 6

## Prédiction linéaire

*Ce chapitre récapitule les principaux résultats obtenus jusqu'ici pour l'identification de canal par prédiction linéaire. Nous y présentons une discussion sur la robustesse des méthodes de prédiction linéaire aboutissant à un algorithme de détermination de la longueur du canal rapide et performant. Les algorithmes présentés s'appliquent aussi bien au cas de plusieurs utilisateurs que d'un seul utilisateur.*

### 6.1 Introduction

Les méthodes d'identification de canal par prédiction linéaire, initiées par D. Slock dans [68], malgré leurs performances réputées médiocres (par rapport aux méthodes sous-espaces) dans un premier temps [2], connaissent un succès grandissant grâce aux caractéristiques suivantes :

- «– elles sont réputées robustes à la surdétermination de l'ordre du canal ;
- «– on peut dériver des algorithmes rapides ;
- «– les approches pondérées sont asymptotiquement statistiquement équivalentes aux méthodes sous-espace pondérées [4], en levant les ambiguïtés supplémentaires liées aux méthodes déterministes ;
- «– elles sont robustes à une perte de disparité des canaux [73], ce qui est consistant avec le fait qu'on peut considérer cette approche comme utilisant le modèle gaussien.

### 6.2 Modèle utilisé : gaussien

Les hypothèses utilisées dans le cadre de la prédiction linéaire, outre [**Hstat**], sont la blancheur du bruit et des sources et [**Hgauss**], ce qui nous place dans le modèle gaussien. D'autre part, le paragraphe suivant montre que la technique d'identification fonctionne si on peut trouver

un égaliseur Zero-Forcing, ce qui mène à imposer l'utilisation d'un canal irréductible (ce qui est une condition légèrement moins contraignante que  $[\mathbf{H}\mathcal{H}\text{full}]$ , puisqu'on accepte les canaux qui ne sont pas à colonnes réduites).

### 6.2.1 Prédiction linéaire du signal sans bruit

Nous considérons le problème consistant à prédire  $\mathbf{y}(k)$  à partir des  $L$  échantillons précédents  $\mathbf{Y}_L(k-1)$  où les  $Y_M$  sont considérés sans bruit (nous discuterons du cas bruité plus loin). L'erreur de prédiction s'écrit :

$$\tilde{\mathbf{y}}(k)|_{\mathbf{Y}_L(k-1)} = \mathbf{y}(k) - \hat{\mathbf{y}}(k)|_{\mathbf{Y}_L(k-1)} = \mathbf{P}_L \mathbf{Y}_{L+1}(k) \quad (6.1)$$

avec  $\mathbf{P}_L = [\mathbf{P}_{L,L} \cdots \mathbf{P}_{L,1} \mathbf{P}_{L,0}]$ ,  $\mathbf{P}_{L,0} = \mathbf{I}_m$ .

Le prédicteur optimal est celui qui minimise la variance de l'erreur de prédiction, ce qui mène au problème suivant :

$$\mathbf{P}_L : \min_{\mathbf{P}_{L,0} = \mathbf{I}_m} \mathbf{P}_L R_{YY} \mathbf{P}_L^H = \sigma_{\tilde{\mathbf{y}},L}^2 \quad (6.2)$$

soit :

$$\mathbf{P}_L R_{YY} = [0 \cdots 0 \ \sigma_{\tilde{\mathbf{y}},L}^2]. \quad (6.3)$$

où la première partie de l'équation (impliquant les zéros) exprime la minimisation par rapport à  $\mathbf{P}_L$  tandis que la seconde exprime simplement l'expression de la variance de l'erreur de prédiction.

Ces expressions ne sont valables que pour  $L \geq \underline{L}$ , en effet, pour que l'équation (6.3) puisse être vérifiée, il faut que  $R_{YY}$  ait le noyau (null-space) de son espace colonne non vide (i.e. son sous-espace bruit), i.e. que la matrice  $R_{YY}$  de taille  $mL \times N + p(L-1)$  ait plus de colonnes que de lignes, ce qui conduit à (4.13). D'autre part, en raisonnant toujours sur la taille de l'espace bruit de  $R_{YY}$ , on obtient l'expression du rang de  $\sigma_{\tilde{\mathbf{y}},L}^2$  en fonction de  $L$  :

$$\text{rang}(\sigma_{\tilde{\mathbf{y}},L}^2) \begin{cases} = p & , L \geq \underline{L} \\ = m - \underline{m} \in \{p+1, \dots, m\} & , L = \underline{L} - 1 \\ = m & , L < \underline{L} - 1 \end{cases} \quad (6.4)$$

où  $\underline{m} = (m-p)\underline{L} - N - p \in \{0, 1, \dots, m-1-p\}$  est le degré de singularité de  $R_{YY,\underline{L}}$ .

D'autre part, pour  $\underline{L} \leq L \leq \bar{L}$ , la puissance associée au processus de dimension  $p$  décroît, et la longueur du prédicteur doit être d'au moins  $\bar{L}$  avant que la puissance ne devienne constante.

### 6.2.2 Le prédicteur linéaire en tant qu'égaliseur Zero-Forcing

Nous considérons toujours le signal reçu sans bruit, qui peut être interprété comme étant un processus MA (à moyenne mobile) multivarié singulier, pour  $L = \bar{L}$  :

$$\mathbf{y}(k) + \sum_{i=1}^{\bar{L}} P_{\bar{L},i} \mathbf{y}(k-i) = \tilde{\mathbf{y}}_{\bar{L}}(k) = \mathbf{h}(0) \mathbf{a}(k) \quad (6.5)$$

de telle sorte que l'erreur de prédiction est un bruit blanc singulier. L'extension pour les sources colorées peut être trouvée dans [59]. Cela signifie que le signal reçu (sans bruit)  $\mathbf{y}(k)$  est également un processus AR multivarié singulier (cette équivalence entre MA et AR est connue, dans le domaine des systèmes multivariés, comme étant l'équation de Bezout [49]). On en déduit :

$$\mathbf{P}_L = [\dots 0 \ \mathbf{P}_{\bar{L}}] , \quad \sigma_{\tilde{\mathbf{y}},L}^2 = \sigma_{\tilde{\mathbf{y}},\bar{L}}^2 , \quad L > \bar{L} . \quad (6.6)$$

Pour  $L \geq \bar{L}$ ,  $\sigma_{\tilde{\mathbf{y}}}^2 = \sigma_a^2 \mathbf{h}^H(0) \mathbf{h}(0)$ , ce qui permet de déterminer  $\mathbf{h}(0)$  à une matrice unitaire près. De (6.1) et de  $\tilde{\mathbf{y}}(k) | \mathbf{Y}_{L \geq \bar{L}}(k-1) = \mathbf{h}(0) \mathbf{a}(k)$  on déduit que  $\mathbf{h}^H(0) (\mathbf{h}^H(0) \mathbf{h}(0))^\# \mathbf{P}_{\bar{L}}$  est un égaliseur ZF à délai nul, ce qui nous donne l'égaliseur ZF de longueur minimale.

## 6.3 Prédiction linéaire et identification de canal

L'équation

$$\mathbf{P}_{\bar{L}} \mathbf{E} \{ \mathbf{Y}_{\bar{L}+1}(k) \mathbf{Y}_N^H(k+N-1) \} = \sigma_a^2 \mathbf{h}(0) [\mathbf{h}^H(0) \dots \mathbf{h}^H(N-1)] \quad (6.7)$$

permet de déterminer le canal [2]. L'équation  $\mathbf{P}_{\bar{L}}(z) \mathbf{H}(z) = \mathbf{h}(0) \Rightarrow \mathbf{H}(z) = \mathbf{P}_{\bar{L}}^{-1}(z) \mathbf{h}(0)$ , en utilisant la paramétrisation en treillis du filtre  $\mathbf{P}_{\bar{L}}(z)$  obtenue par l'algorithme rapide de Levinson permet également d'obtenir une estimée du canal, c'est un des algorithmes sous-tendus dans l'article original [68].

Imaginons que l'ordre du canal soit correctement estimé, alors, on peut avoir un estimateur consistant du prédicteur et l'équation fondamentale s'écrit :

$$\mathbf{P}(z) \mathbf{H}(z) = \mathbf{h}(0) \quad (6.8)$$

où  $\mathbf{h}(0)$  est déterminé, à une matrice unitaire près (appelons la  $U$ ) à partir de  $\sigma_{\tilde{\mathbf{y}}}^2 = \sigma_a^2 \mathbf{h}(0) \mathbf{h}^H(0)$ . En effet,  $\mathbf{H}'(z) = U \mathbf{H}(z)$  vérifie (6.8), ce qui est cohérent avec les conditions d'identifiabilité dans le modèle gaussien. L'identification de la matrice unitaire doit être faite soit en faisant appel aux statistiques d'ordre supérieur, en appliquant une méthode de séparation de sources non convolutives aux innovations du processus AR, soit en utilisant des symboles connus (méthode semi-aveugle). Dans le cas où les ordres des canaux sont différents, le recours aux statistiques supérieures n'est (en théorie) pas nécessaire (voir les conditions d'identifiabilité en multi-utilisateurs).

### 6.3.1 La méthode de la prédiction linéaire pondérée.

Une méthode de détermination du canal à partir du prédicteur linéaire consiste,  $\mathbf{h}(0)$  étant déterminé (e.g. à partir de la variance de l'erreur de prédiction), à résoudre  $\mathbf{P}(z)\mathbf{H}(z) = \mathbf{h}(0)$ , en utilisant une procédure de minimisation par les moindres carrés pondérés. [4] :

$$\widehat{\mathbf{H}} = \arg \min_{\mathbf{H}} \|\mathbf{W}^{-5}(\mathcal{T}^{tH}(\mathbf{P})\mathbf{H} - [\hat{\mathbf{h}}(0)^H, 0 \dots 0]^H)\|^2 \quad (6.9)$$

Nous retiendrons ici la matrice de pondération « pratique » proposée par Gorokhov, dont l'expression est :

$$\mathbf{W} = \mathbf{I} \otimes (\hat{\sigma}_y^2 + \hat{\sigma}_v^2 \mathbf{I}_m)^{-1} \quad (6.10)$$

Elle représente une pondération entre le sous-espace signal des innovations, qui serait suffisant si l'ordre du canal était connu, et le sous-espace bruit, ce qui confère une certaine robustesse à la surestimation de la longueur du canal en régularisant le système d'équations.

## 6.4 L'inverse de $R_{YY}$ dans la prédiction linéaire.

### 6.4.1 La méthode « Levinson »

L'approche la plus naturelle pour obtenir le prédicteur linéaire est de résoudre les équations normales (6.3) par l'algorithme Levinson multicanal (aussi connu sous le nom "Levinson-Wiggins-Robinson"), adapté au cas singulier. Cette approche correspond à une factorisation triangulaire bloc d'une inverse généralisée de  $R_{YY}$  :

$$L_L R_{YY} L_L^H = D_L, \quad (L_L)_{i,j} = \mathbf{P}_{i-1,i-j}, \quad (D_L)_{i,i} = \sigma_{y,i-1}^2 \quad (6.11)$$

où  $L_L$  est bloc-triangulaire inférieure et  $D_L$  est bloc diagonale. Si  $R_{YY,\bar{L}}$  est singulier, il faut prendre des précautions dans la détermination des derniers coefficients blocs  $P_{\bar{L} \dots \bar{L}, \bar{L} \dots \bar{L}}$ , qui ne sont pas uniques [61]. En outre, cet algorithme nous donnant les "vrais" coefficients de prédiction linéaire (par opposition au prédicteur trouvé dans la méthode que nous appellerons, par abus de langage, de la pseudo-inverse), de l'équation (6.6), on déduit que  $L_L$  est une matrice asymptotiquement Toeplitz (Toeplitz à partir de la ligne  $\bar{L}m + 1$ ).

### 6.4.2 La méthode «Pseudo-inverse»

En supposant la longueur du canal connue, l'équation (6.3) peut se réécrire :

$$[-\mathbf{Q}_{\bar{L}}|I_m] = \left[ \begin{array}{c|c} R_{YY,\bar{L}} & r^H \\ \hline r & r_0 \end{array} \right] = [0 \cdots 0 | \sigma_y^2] \quad (6.12)$$

ce qui fournit la solution :

$$\begin{cases} \sigma_y^2 = r_0 - r(R_{YY,\bar{L}})^{-1}r^H \\ \mathbf{Q}_{\bar{L}} = r(R_{YY,\bar{L}})^{-1} \end{cases} \quad (6.13)$$

Dans cette méthode, l'inverse est remplacée par une inverse généralisée quand l'ordre du canal est surestimé, en adoptant la pseudo-inverse de Moore-Penrose, on aboutit à la méthode de la «Pseudo-inverse». Cette méthode donne des résultats différents de la méthode Levinson, en ceci qu'elle correspond à une solution à norme minimale, alors que l'algorithme de Levinson correspond à une solution à longueur minimale. En effet, dans ce dernier cas,  $P_{L,L_o} = 0$  si  $L_o \geq \bar{L}$ .

Le principal inconvénient de cette méthode est son coût de calcul. Son principal avantage, outre sa robustesse à la surestimation de la longueur du canal, est qu'elle permet l'utilisation d'un facteur de fenêtrage  $K$  (i.e. une matrice de corrélation  $mK \times mK$  de taille supérieure à  $m\bar{L} \times m\bar{L}$ ) qui, joint à l'utilisation de la pseudo-inverse, effectue un nettoyage de  $R_{YY}$  par la séparation entre son espace signal et son espace bruit. L'algorithme global est voisin de [75], développé dans le cadre des sinusoides dans le bruit.

### 6.4.3 La méthode «Cadzow»

Pour obtenir cet effet de nettoyage, sans recourir à une SVD, nous proposons la méthode simplifiée suivante, inspirée de [9] qui, quand l'ordre de canal est correctement estimé, se base sur :

$$[0 \cdots 0 | -\mathbf{Q}_{\bar{L}} | I_m] \begin{array}{c} \overbrace{\left[ \begin{array}{ccc|c} * & \cdots & * & * \\ & \ddots & & \vdots \\ * & \cdots & * & * \\ \hline & & R_{\text{rect}} & * \\ \hline & & & \vdots \\ & & r & * \end{array} \right]}^{=R_{YY}-\sigma_y^2 I} \\ \\ \end{array} = [0 \cdots 0 | \sigma_y^2] \quad (6.14)$$

soit, en minimisant par les moindres carrés :

$$\mathbf{Q}_{\bar{L}} = (R_{\text{rect}} R_{\text{rect}}^H)^{-1} R_{\text{rect}} r^H \quad (6.15)$$

De par l'utilisation de plus de données de la matrice de corrélation, nous obtenons une meilleure estimation du prédicteur que pour la méthode de Levinson, à un coût relativement faible. Cependant, on peut s'attendre à une faible robustesse à la surestimation de l'ordre, en effet, dans ce cas, on utilise une matrice  $R_{\text{rect}}$  trop grande, et il n'y a pas de séparation réelle entre espace bruit et signal des innovations comme dans la méthode recourant à la SVD.

L'étape ultérieure est le recours à une méthode des moindres carrés pondérée.

## 6.5 Surestimation de la longueur du canal

La robustesse à la longueur du canal provient, soit de la robustesse de la détermination du prédicteur linéaire à la surestimation de l'ordre (propriété de l'algorithme de Levinson, sur laquelle nous nous appuyerons pour la détermination d'ordre), soit de l'utilisation d'un algorithme séparant espace signal et bruit pour la méthode de la pseudo-inverse.

### 6.5.1 Méthode « Pseudo-inverse »

Dans cette méthode, l'équation (6.13) devient :

$$\begin{cases} \sigma_y^2 = r_0 - r(R_{YY,\bar{L}})^{\#} r^H \\ \mathbf{Q}_{\bar{L}} = r(R_{YY,\bar{L}})^{\#} \end{cases} \quad (6.16)$$

où  $A^{\#}$  est l'inverse généralisée de Moore-Penrose de  $A$ .

Cette pseudo-inverse se base sur une décomposition entre l'espace signal et bruit de la matrice de corrélation (usuellement par la SVD). Si on utilise la vraie valeur de la puissance du bruit (ou

une estimée calculée selon le principe du maximum de vraisemblance, qui est donnée par la moyenne des valeurs singulières associées à l'espace bruit), les simulations et des considérations théoriques (e.g. appendice F de [4]) montrent qu'une erreur dans l'estimation de la taille des espaces signal et bruit n'a (asymptotiquement) pas d'influence sur l'estimation du canal. De façon plus précise, [4] argumente la proposition :

$$\hat{P}_* = \hat{P} + \hat{P}^\perp + O\left(\frac{1}{\sqrt{L}}\right) \quad ; \quad \hat{P}^\perp R_{YY} = 0 \quad (6.17)$$

où  $\hat{P}_*$  est l'estimée obtenue avec une éventuelle erreur sur l'ordre et  $\hat{P}$  l'estimée obtenue avec la longueur de canal exacte.

Malheureusement, les simulations s'éloignent quelque peu de cette conclusion. En effet, si la longueur du canal est surestimée ( $N' > N$ ), la puissance du bruit est sous-estimée :

$$\hat{\sigma}_v^2 = \frac{1}{mL - N' - L + 2} \sum_{i=N'+L-1}^{mL} \lambda_i < \frac{1}{mL - N - L + 2} \sum_{i=N+L-1}^{mL} \lambda_i = \hat{\sigma}_{v,ML}^2 \quad (6.18)$$

où les  $\lambda_i$  sont les valeurs propres de  $R_{YY}$ , classées par ordre décroissant, ce qui donne :

$$\hat{P}_* = \hat{P} + \hat{P}^\perp + \Delta \hat{\sigma}_v^2 I + O\left(\frac{1}{\sqrt{L}}\right) \quad ; \quad P^\perp R_{YY} = 0 \quad (6.19)$$

D'où

$$\hat{\sigma}_y^2 = \sigma_a^2 \mathbf{h}(0) \mathbf{h}^H(0) + \Delta \hat{\sigma}_v^2 r r^H \quad (6.20)$$

ce qui introduit une erreur d'estimation sur  $\mathbf{h}(0)$ . Cette erreur est une fonction du canal lui-même via  $r$  et peut être difficilement caractérisée de manière statistique. De plus, l'erreur sur le filtre prédicteur affectera également l'estimée du canal via l'estimateur WLS.

### 6.5.2 Méthode « Levinson »

Dans la méthode de Levinson, dans un contexte strictement sans bruit, les coefficients de prédiction au-delà de  $\bar{L}$  deviennent nuls. Sans entrer dans les détails de l'algorithme, les calculs se basent sur les variances d'erreur de prédiction avant et arrière (cette dernière correspondant à  $\sigma_y^2$ ), et sur une inverse généralisée et le rang de ces variances. Pour obtenir le bon prédicteur linéaire, la détection de ce rang est nécessaire, même si le type d'inverse généralisée utilisé n'a pas d'influence sur les variances d'erreur. Si ce rang n'est pas correctement estimé, et donc si le prédicteur n'est pas correctement estimé, cette méthode ne jouit pas des mêmes propriétés de

robustesse que celles de la «Pseudo-inverse». On peut considérer que l'utilisation de la méthode de «Levinson» équivaut à une inverse généralisée de  $R_{YY}$  à longueur minimale : en effet, le prédicteur est de longueur minimale, ce qui mène à une décomposition triangulaire où le nombre de paramètres est minimal. La méthode «pseudo-inverse» par contre correspond à une inverse généralisée de  $R_{YY}$  à norme minimale (voir e.g. [66]), qui possède plus de vertus de robustesse. Cependant, les valeurs des variances d'erreur de prédiction nous donneront une bonne méthode pour la détermination de l'ordre du processus AR, et leur rang pour la détermination finale de l'ordre du canal.

## 6.6 Estimation d'ordre

### 6.6.1 Estimation de la longueur du canal basée sur les $\lambda_i$

Une des méthodes "naturelles" d'estimation de la longueur du canal est l'estimation du rang de la matrice de corrélation par l'inspection des valeurs propres. Dans le contexte du traitement d'antenne, Wax a développé une méthode [77] basée sur des critères tirés de la théorie de l'information. Malheureusement, cette méthode donne de piètres résultats pour les mélanges convolutifs. Wax a également développé une méthode générale [78] basée sur le principe du maximum de vraisemblance et qui fait appel à une optimisation non linéaire. De part sa complexité numérique, elle perd son intérêt dans le contexte de la prédiction linéaire.

### 6.6.2 Estimation de l'ordre d'un processus VAR

L'estimation de l'ordre d'un processus VAR( $L < \infty$ ) (processus auto-régressif vectoriel) se fait soit par l'utilisation de tests statistiques sur les coefficients de prédiction idéalement nuls ( $L \geq \bar{L}$ ), ou sur les variances d'erreur. Un autre groupe d'algorithmes se base sur des principes tirés de la théorie de l'information. Ces procédures ont été initialement développées par Akaike (AIC : An Information theoretic Criterion) et Rissanen (MDL : Minimum Description Length) pour les processus scalaires et Hannan et Quinn (HQ) pour les processus vectoriels. [64] fournit une synthèse intéressante de ces diverses méthodes. Nous allons utiliser ces résultats en les adaptant au cas de processus singuliers.

Les critères classiques de la théorie de l'information sont :

$$\begin{array}{ll}
 \text{AIC} & \log |\sigma_{y,k}^2| + \frac{2km^2}{L} \\
 \text{HQ} & \log |\sigma_{y,k}^2| + \frac{2km^2 \log \log L}{L} \\
 \text{MDL} & \log |\sigma_{y,k}^2| + \frac{km^2 \log L}{2L}
 \end{array} \tag{6.21}$$

où  $k$  sont les ordres estimés et  $|A|$  représente le déterminant de  $A$ .

Le premier terme de ces expressions maximise le logarithme de la vraisemblance de l'erreur de prédiction, tandis que le second terme est un terme de correction de biais tenant compte du nombre de paramètres ( $km^2$ ) et du nombre de données utilisées ( $L$ ). Le critère MDL et le critère HQ donnent des estimées consistantes de l'ordre pour un vrai processus AR( $L$ ), tandis que le critère AIC a tendance à surestimer l'ordre. En fonction de la procédure d'estimation du canal utilisée, nous choisirons l'une ou l'autre de ces estimées (MDL ou HQ pour la méthode Levinson, AIC pour la méthode de la pseudo-inverse).

Dans le cas de processus singuliers, en utilisant les résultats classiques concernant les distributions normales singulières multivariées, ces critères peuvent être étendus tels quels, mais donnent des résultats médiocres, ce qui nous a motivé à modifier légèrement les critères, en nous basant sur les interprétations "physiques" des grandeurs en présence.

Nous avons  $\tilde{y}(n) = \mathbf{h}(0)\mathbf{a}(n)$  et  $R_{aa} = \sigma_a^2 I_m$ . Nous pouvons donc raisonner sur le processus source équivalent  $\mathbf{h}^H(0)\tilde{\mathbf{y}}(n) = \mathbf{h}^H(0)\mathbf{h}(0)\mathbf{a}(n)$ , de dimension plus faible, et sur son erreur de prédiction. Dans un premier temps, il paraît donc naturel de remplacer simplement  $|\sigma_{\tilde{y},k}^2|$  par  $|\sigma_{\tilde{y}_p,k}^2|$ , où l'indice  $p$  signifie simplement qu'on se limite au sous-espace de dimension  $p$  de  $\tilde{\mathbf{y}}$ . Cependant, l'application directe de ce critère donne encore des résultats médiocres, parce que nous n'avons pas pris en compte que le prédicteur linéaire nous fait passer progressivement d'un espace de dimension  $m$  à un espace de dimension  $p$ . Nous proposons simplement d'ajouter un terme  $\text{trace}(\sigma_{\tilde{y}_{m-p,k}}^2)$ , calculé à partir des  $m-p$  valeurs propres les plus petites de  $\sigma_{\tilde{y},k}^2$ . En effet, ce terme aura pour effet de pénaliser le critère quand on a pas atteint la taille de sous-espace  $p$ . Une fois que cette dimension est atteinte, on permet encore de tenir compte d'une coloration du signal de dimension  $p$  (par exemple dans le cas d'un canal à colonnes non réduites).

Des méthodes plus fines, mais également plus lourdes en terme de calculs, sont présentées (pour le cas non-singulier) dans [10].

## 6.7 Algorithmes conjoints d'estimation d'ordre et de canal

Les algorithmes conjoints proposés utilisent la détermination de l'ordre du processus AR suivi d'une inspection des valeurs propres de  $\sigma_{\tilde{y},L-1}^2$ , pour obtenir une estimation de la longueur du canal. Ensuite, nous continuons avec une procédure d'estimation du canal. L'algorithme sera réellement conjoint dans le cas où on utilise Levinson. Les autres procédures donnent de meilleures performances, au prix d'une complexité de calcul accrue.

## 6.8 Simulations

### 6.8.1 Estimation de l'ordre du processus AR

Nous avons simulé un système à  $m = 4$  canaux,  $p = 1$  utilisateur et une longueur de canal de  $N = 6$ , ce qui mène à un processus AR d'ordre 2, le canal étant généré aléatoirement. Les résultats sont très bons, même pour des SNR relativement bas. L'hypothèse de départ est que l'ordre du processus AR est au maximum de 6.

AIC	SNR						
ord.	25	20	15	10	5	0	-5
1	0	0	0	0	0	1	69
2	99	99	82	11	2	23	15
3	0	0	18	87	86	61	3
4	0	0	0	0	1	2	0
5	0	0	0	0	1	1	2
6	1	1	0	2	10	12	11

MDL	SNR						
ord.	25	20	15	10	5	0	-5
1	0	0	0	0	0	2	99
2	100	100	100	56	33	79	1
3	0	0	0	44	67	19	0
6	0	0	0	0	0	0	0

HQ	SNR						
ord.	25	20	15	10	5	0	-5
1	0	0	0	0	0	17	100
2	100	100	100	86	78	81	0
3	0	0	0	14	22	2	0
6	0	0	0	0	0	0	0

Nous nous sommes ensuite placés dans le cas de  $p = 2$  utilisateurs avec  $m = 4$  canaux et une longueur totale de  $N = 12$ , soit un processus AR d'ordre 5. Les résultats sont légèrement moins

convaincants à SNR bas. L'hypothèse de départ est celle d'un AR(10).

AIC	SNR						
ord.	25	20	15	10	5	0	-5
1	0	0	0	0	0	70	100
2	0	0	0	0	0	1	0
3	0	0	0	0	5	9	0
4	0	0	0	0	0	20	0
<b>5</b>	83	90	100	99	94	0	0
6	17	10	0	1	0	0	0
10	0	0	0	0	1	0	0

MDL	SNR						
ord.	25	20	15	10	5	0	-5
1	0	0	0	0	9	100	100
2	0	0	0	0	40	0	0
3	0	0	0	0	37	0	0
4	0	0	0	0	0	0	0
<b>5</b>	100	100	100	100	14	0	0
10	0	0	0	0	0	0	0

HQ	SNR						
ord.	25	20	15	10	5	0	-5
1	0	0	0	0	72	100	100
2	0	0	0	0	24	0	0
3	0	0	0	0	4	0	0
4	0	0	0	0	0	0	0
<b>5</b>	100	100	100	100	0	0	0
10	0	0	0	0	0	0	0

### 6.8.2 Estimation du canal

Pour illustrer la robustesse de l'estimation de canal par l'approche LP, nous avons effectué des simulations illustrant les effets des variations des différents paramètres sur l'estimation de la longueur du canal et sur l'estimation du canal lui-même.

La performance est caractérisée par l'erreur quadratique moyenne normalisée (NRMSE :

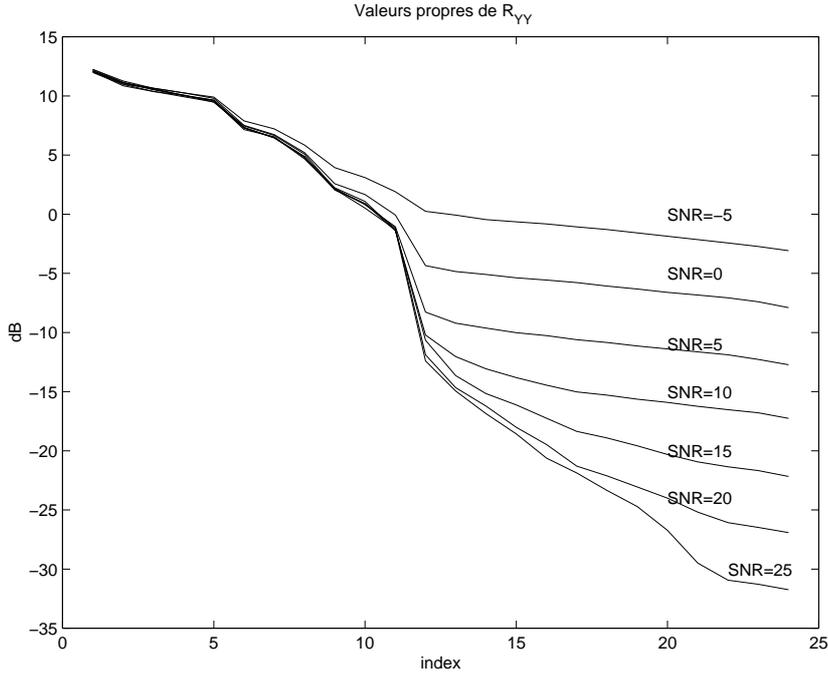


FIG. 6.1 – Valeurs propres de  $R_{YY}$ .

Normalized Root Mean Square Error), calculée sur 100 réalisations indépendantes :

$$\text{NRMSE} = \sqrt{\frac{1}{100} \sum_{i=1}^{100} \mathbf{h}^H P_{\hat{\mathbf{h}}^{(i)}}^\perp \mathbf{h} / \|\mathbf{h}\|^2} \quad (6.22)$$

où  $\mathbf{h}^H P_{\hat{\mathbf{h}}}^\perp \mathbf{h} = \min_{\alpha} \|\alpha \hat{\mathbf{h}} - \mathbf{h}\|^2$ . Nous utilisons le canal réel  $\mathbf{h}$  où  $N = 6$ ,  $m = 4$  et  $p = 1$  utilisé dans [1].

Les symboles sont i.i.d., QPSK et on utilise  $M = 250$  symboles en entrée. Le SNR est défini par  $\text{SNR} = (\|\mathbf{h}\|^2 \sigma_a^2) / (mM\sigma_v^2)$ .

Le spectre des valeurs propres de la matrice de corrélation est reproduit ci-dessous, de manière à donner une idée intuitive de la difficulté à déterminer la longueur du canal aux différents SNR.

**De l'utilisation de  $\hat{\sigma}_{v,ML}^2$**  Les simulations confirment la proposition selon laquelle, en utilisant l'estimée ML de la puissance de bruit, il n'y a pas de perte de performance due à la surestimation de la longueur du canal. La taille de la fenêtre est  $K = 6$  (i.e.  $R_{YY}$  de taille  $mK \times mK$ ).

**De l'utilisation de  $\hat{\sigma}_v^2$**  Ici, les simulations illustrent clairement l'influence de  $\Delta \hat{\sigma}_v^2$ , principalement à un SNR élevé, où l'estimation de l'ordre est la plus facile à faire. Nous montrons d'abord

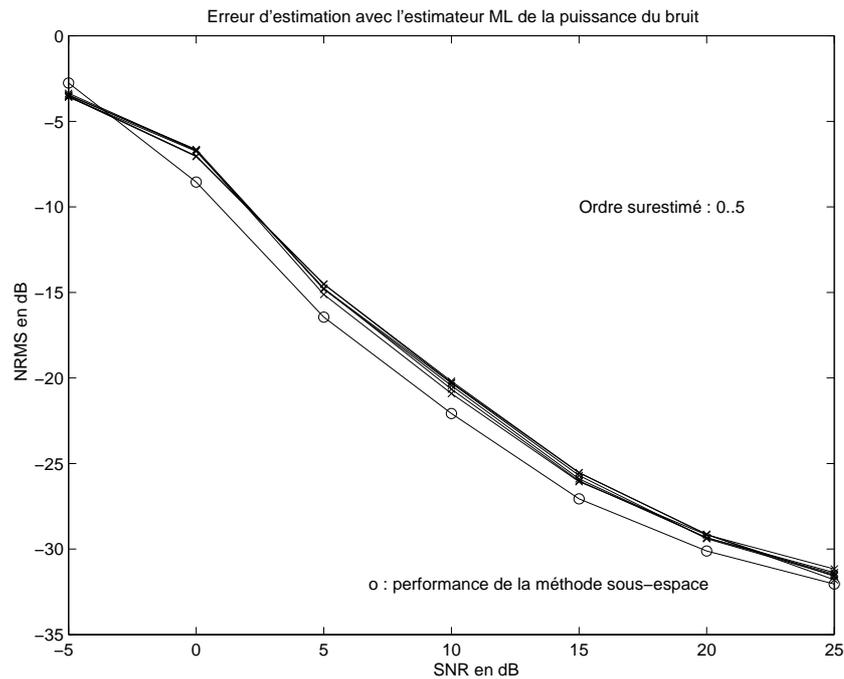


FIG. 6.2 – Performance de WLP en utilisant  $\hat{\sigma}_{v,ML}^2$ .

les résultats avec une fenêtre de taille minimale et ensuite avec une fenêtre de taille  $K = \hat{N}$ . Cette comparaison nous incite sans équivoque à utiliser une fenêtre de taille non minimale.

**Utilisation de l'estimation de l'ordre** Dans cette dernière figure, nous comparons les différents algorithmes d'estimation de canal par prédiction linéaire :

- «– WLP sans estimation de l'ordre du canal ;
- «– WLP avec estimation de l'ordre du processus AR ;
- «– WLP avec estimation de la longueur du canal ;
- «– estimation de la longueur du canal et méthode de Cadzow ;
- «– estimation de la longueur du canal et méthode de Levinson ;
- «– pour comparaison, la méthode sous-espace.

Les résultats confirment les analyses effectuées plus haut, en ceci que l'estimation préalable de la longueur du canal améliore nettement les performances et que la méthode WLP utilisant la pseudo-inverse est nettement meilleure que les autres et a des performances comparables à la méthode sous-espace.

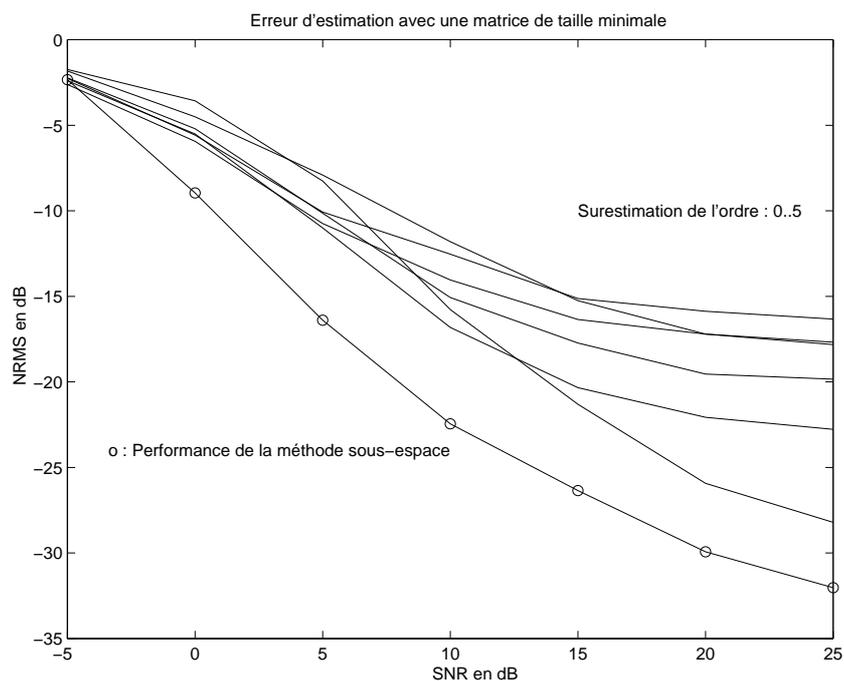


FIG. 6.3 – Performance de WLP sans connaissance ni estimation de  $N$ ,  $K$  minimal

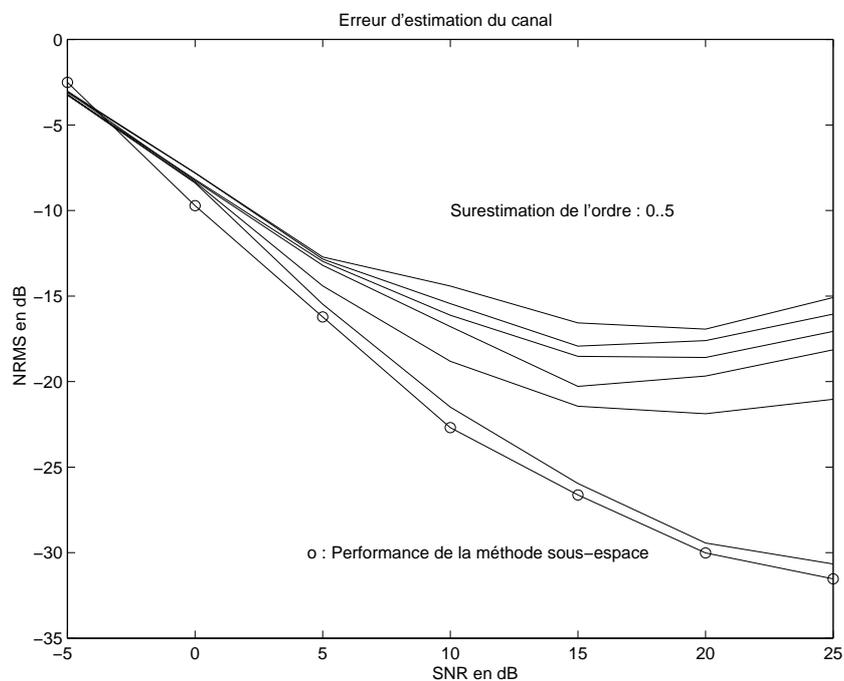


FIG. 6.4 – Performance de WLP sans connaissance ni estimation de  $N$ ,  $K = \hat{N}$

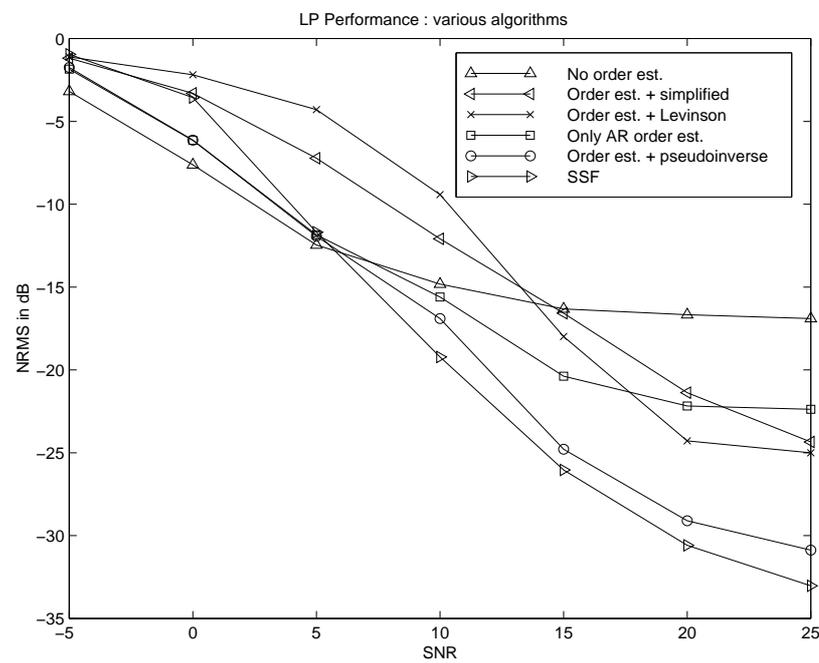


FIG. 6.5 – Identification par prédiction linéaire : influence de l'estimation de la longueur du canal

## 6.9 Une paramétrisation minimale du sous-espace bruit

Pour  $L \geq \bar{L}$ ,  $\tilde{\mathbf{y}}(k+L)|_{\mathbf{Y}_{L(k)}} = \mathbf{H}(0)\mathbf{a}(k+L)$ . Si on considère la prédiction des quantités scalaires  $y_i(k)$ , de l'équation précédente, on déduit que  $\tilde{y}_i(k+L) = 0, i = p+1, \dots, m$  et les  $m-p$  lignes correspondantes de  $\mathbf{P}_L$  sont des filtres de prédiction singulier (i.e. tels que  $\mathbf{P}_{L,i}\mathcal{T}(\mathbf{H}) = 0$ ). En utilisant ces  $m-p$  filtres singuliers et en éliminant toutes les dépendances entre éléments, on obtient un filtre  $\bar{\mathbf{P}}$  de la forme :

$$\begin{array}{c}
 \overbrace{\hspace{1.5cm}}^{\bar{L}m} \qquad \qquad \qquad \underbrace{\hspace{1.5cm}}_m \\
 \left. \begin{array}{ccc|ccc}
 * * * & *.. & *.. & 10 & 0 & 0 & 0 & 0..0 \\
 * * * & *.. & *.. & 01 & 0 & 0 & 0 & 0..0 \\
 \hline
 * * * & *.. & *.. & 00 & * & 1 & 0 & 0..0 \\
 * * * & *.. & *.. & 00 & * & * & 1 & 0..0
 \end{array} \right\} \begin{array}{l} m \\ m-p-m \end{array} \qquad (6.23) \\
 \underbrace{\hspace{1.5cm}}_m
 \end{array}$$

avec  $Nm - p^2$  paramètres indépendants et  $m-p$  valeurs 1.

D'autre part, on peut montrer que la matrice formée avec cette paramétrisation minimale du sous-espace bruit génère effectivement ce dernier [67]. Cette paramétrisation particulière est utilisée dans les chapitres traitant des conditions d'identifiabilité et le chapitre traitant des méthodes au maximum de vraisemblance.

## 6.10 Conclusions

Nous avons investigué les caractéristiques de robustesse à la surestimation de la longueur du canal des méthodes d'identification de canal basées sur la prédiction linéaire. Pour ce faire, nous avons confronté deux approches de la robustesse :

- «– La première [68] constate que la méthode de Levinson fournit le prédicteur du bon ordre, même si celui-ci est surestimé a priori.
- «– La seconde [1, 4] constate que, si l'ordre du prédicteur est surestimé, la méthode de la pseudo-inverse fournit une estimée du prédicteur non consistante mais une estimée du canal consistante.

La seconde méthode, pour jouir pleinement de sa robustesse, nécessite une bonne estimée de la puissance du bruit. La première méthode débouche naturellement sur un algorithme de détermination de l'ordre performant, fournissant l'estimée du canal sans coût additionnel, mais avec une

performance légèrement dégradée. Pour pallier cet inconvénient, nous avons exploré la méthode de "Cadzow", qui, quoique plus lourde que celle de Levinson, permet d'effectuer un nettoyage de la matrice de corrélation à un coût inférieur au coût de la pseudo-inverse d'une matrice singulière. Enfin, nous avons introduit une paramétrisation minimale du sous-espace bruit, qui est particulièrement utile pour définir les conditions d'identifiabilité et développer une méthode d'initialisation pour certaines méthodes au maximum de vraisemblance.



## Chapitre 7

# Identification aveugle par l'algorithme de Schur

|| Ce chapitre montre comment le canal peut être identifié par décomposition LU de la matrice de covariance du signal, en nous limitant au modèle gaussien.

### 7.1 Décomposition LDU de la matrice de covariance débruitée

Nous considérons que, dans un premier temps, nous avons obtenu une estimée de la variance du bruit. Nous pouvons donc raisonner sur le signal reçu  $Y$  supposé sans bruit. Soit  $\tilde{Y}$ , l'erreur de prédiction, sous les conditions du théorème 4.6,  $Y$  peut être prédit de façon parfaite. En reprenant l'équation 6.11 :

$$L_L R_{YY} L_L^H = D_L, (L_L)_{i,j} = P_{i-1,i-j}, (D_L)_{i,i} = \sigma_{\tilde{y},i-1}^2 \quad (7.1)$$

où, de (6.1), on a que  $L_L = R_{\tilde{Y}Y} R_{YY}^{-1}$ , (7.1) se réécrit :  $R_{\tilde{Y}Y} R_{YY}^{-1} R_{YY} = R_{\tilde{Y}Y}$ , sans tenir compte des singularités éventuelles de  $R_{YY}$ . On en déduit immédiatement :

$$R_{YY} = R_{\tilde{Y}Y} R_{\tilde{Y}Y}^{\#} R_{\tilde{Y}Y} = U^H D U \quad (7.2)$$

où  $\#$  est une inverse généralisée quelconque.

Dans un premier temps, considérons que nous opérons une triangularisation par blocs. En se basant simplement sur le rang de  $R_{YY}$ , nous déduisons que les  $\underline{L} - 1$  premiers blocs de  $D$  sont de rang plein, que le  $\underline{L}^e$  bloc est de rang  $(1 - \underline{L})(m - p) + N$  et que les blocs suivants sont de rang  $p$ .

La prédiction de  $Y = Y_M(k) = [\mathbf{y}^H(k) \cdots \mathbf{y}^H(k - M + 1)]^H$  étant parfaite à partir du  $(\bar{L} + 1)^e$  instant (correspondant au  $(\bar{L} + 1)^e$  bloc),  $\tilde{\mathbf{y}}(k - t)$  contient, en  $t > \bar{L}$ , les symboles

émis, à un mélange instantané près. D'autre part, de  $U^H = R_{Y\tilde{Y}}$ , nous avons, en représentant le bloc  $(i, j)$  de  $U$  par  $U(i, j)$ ,

$$\begin{aligned} U^H(i, j) &= \mathbb{E} \{ \mathbf{y}(k-i) \tilde{\mathbf{y}}^H(k-j) \} \\ &= \sum_{l=0}^{N_1-1} \mathbf{H}(l) \mathbb{E} \{ \mathbf{a}(k-i-l) u \mathbf{a}^H(k-j) \} \quad \text{pour } i, j > \bar{L} \quad (7.3) \\ &= u \mathbf{H}(j-i) \sigma_a^2 \end{aligned}$$

où  $u$  est une matrice unitaire. Il s'ensuit que nous pouvons identifier le canal, à une matrice unitaire près, par triangularisation de la matrice de covariance du signal reçu.

## 7.2 Utilisation de l'algorithme de Schur Généralisé

### 7.2.1 Notions de base

Nous rappelons ici la définition de rang de déplacement :

**Définition 7.1** Le déplacement d'une matrice Hermitienne  $n \times n$  est défini par

$$\Delta R_{\mu, \nu} \triangleq R - Z^\mu R Z^{\nu T}, \quad (7.4)$$

où  $Z$  est une matrice de décalage de dimension  $n \times n$  avec des 1 sur la 1<sup>ère</sup> sous-diagonale et des zéros ailleurs.

Le rang  $r$  de  $\Delta R$  est appelé rang de déplacement et on peut montrer aisément que le rang de déplacement de  $R_{YY}(-\sigma_v^2 I)$  vaut  $2m$ .

De plus, on peut factoriser  $\nabla R_{YY}$  sous la forme  $\nabla R_{YY} = G \Sigma G^H$ , où  $\Sigma = (I_m \oplus -I_m)$  est une matrice dite de signature et  $G$  est appelé le générateur de  $R_{YY}$ . On vérifie aisément que, en notant  $r_i \triangleq R_{YY}^{-1/2}(0) R_{YY}(i)$ ,  $R_{YY}$  de dimension  $Km \times Km$  et  $\mathcal{R}_{YY} = [r_{i-j}]_{i,j}$ .

$$\nabla \mathcal{R}_{YY} = \begin{bmatrix} r_0 & 0 \\ r_1 & r_1 \\ \vdots & \vdots \\ r_{K-1} & r_{K-1} \end{bmatrix} \begin{bmatrix} I_m & 0 \\ 0 & -I_m \end{bmatrix} \begin{bmatrix} r_0 & 0 \\ r_1 & r_1 \\ \vdots & \vdots \\ r_{K-1} & r_{K-1} \end{bmatrix}^H$$

Dans sa version bloc, l'algorithme de Schur généralisé est initialisé par  $G^{(0)} = G$ , et

$$\begin{aligned} G^{(1)H} &= S^{(1)} \begin{bmatrix} 0 & r_0^{(0)} & r_1^{(0)} & \cdots & r_{K-1}^{(0)} \\ 0 & r_1^{(0)} & r_2^{(0)} & \cdots & r_K^{(0)} \end{bmatrix} \\ &= \begin{bmatrix} 0 & r_0^{(1)} & r_1^{(1)} & \cdots & r_{K-1}^{(1)} \\ 0 & 0 & \tilde{r}_2^{(1)} & \cdots & \tilde{r}_K^{(1)} \end{bmatrix} \quad (7.5) \end{aligned}$$

où  $S^{(1)}$  est une transformation hyperbolique de Householder (telle que  $S$  est  $\Sigma$ -unitaire : i.e.  $S^{(1)}\Sigma S^{(1)H} = \Sigma$ ). On peut alors montrer que  $G^{(1)}$  est le générateur du complément de Schur de  $\mathcal{R}_{YY}$  par rapport à  $r_0$  (le complément de Schur de  $A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$  par rapport à  $a$  étant  $d - ca^{-1}b$ ).

En continuant ce processus, on obtient  $\mathcal{R}_{YY} = U^H D U$  où  $U = \begin{bmatrix} r_0^{(0)} & r_1^{(0)} & \cdots & r_{K-1}^{(0)} \\ 0 & r_0^{(1)} & \ddots & r_{K-2}^{(1)} \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & r_0^{(K-1)} \end{bmatrix}$

### 7.2.2 Application à la factorisation LDU de $\hat{R}_{YY}$

Après la  $\bar{L}^{\text{ème}}$  itération, si on utilise les statistiques exactes (i.e.  $R_{YY} - \sigma_v^2 I$ ), les  $p$  premières colonnes du générateur contiennent le canal, les  $m - p$  colonnes suivantes étant nulles.

L'algorithme de Schur généralisé est uniquement applicable aux matrices fortement non singulières, (i.e. dont tous les mineurs principaux sont définis positifs). Des solutions ont été proposées pour des matrices singulières, [32, 65] mais elles introduisent une complexité additionnelle que nous voulons éviter. De ce fait, le choix de l'estimateur de la matrice de corrélation du signal est crucial. Nous avons opté pour l'estimateur "biaisé" de la séquence de corrélation

$$\hat{R}_{YY}(i) = \frac{1}{M} \sum_{t=1}^{M-i} \mathbf{y}(t) \mathbf{y}^H(t+i),$$

parce que la matrice formée par cet estimateur est, par construction, bloc Toeplitz et définie positive. L'estimateur non biaisé (où le facteur  $1/M$  est remplacé par  $1/(M - i)$ ) donne une matrice indéfinie tandis que la matrice de covariance empirique ( $\hat{R}_{YY} = \frac{1}{M} \sum_{t=0}^{M-1} \mathbf{Y}(t) \mathbf{Y}^H(t)$ , où  $\mathbf{y}(k) = 0$  pour  $k \geq M$ ) a un rang de déplacement  $2(m + 1)$ , et elle a donc une structure légèrement différente de la structure de la vraie matrice de covariance.

De plus, à cause du biais introduit par l'estimateur, le spectre des valeurs singulières ne présente pas une cassure abrupte au rang théorique de  $R_{YY}$ , mais ces valeurs singulières décroissent de façon relativement douce. Cette propriété, qui peut paraître indésirable (l'application d'une méthode sous-espace basée sur cet estimateur, par exemple, donnera des résultats médiocres), nous sera extrêmement utile, en ce qu'elle nous assurera un bon conditionnement des compléments de Schur jusqu'à un stade avancé dans l'algorithme et nous ne serons donc pas obligés d'utiliser les techniques applicables aux matrices singulières. En outre, nous pourrons également, en utilisant une matrice  $\hat{R}_{YY}$  suffisamment grande (2 à 3 ( $\bar{L}m \times \bar{L}m$ )), surestimer la longueur du canal sans problème numérique. En effet, nous pourrons faire des itérations plus loin que le pas  $\bar{L}$ .

L'inconvénient de ce comportement de  $\widehat{R}_{YY}$  est que, après l'itération  $\bar{L}$ , le  $(\bar{L} + 1)^{\text{ème}}$  bloc de la diagonale bloc sera loin d'être singulière, comme elle devrait l'être théoriquement et donc, les  $p$  premières colonnes du générateur ne peuvent prétendre contenir le canal exactement. Une méthode simple pour circonvier ce problème consiste à faire une SVD sur les  $m$  premières colonnes du générateur et de prendre les  $p$  premières colonnes du résultat, qui correspondent à la partie non singulière de  $D_{\bar{L}+1}$ , au prix d'un calcul de complexité  $\mathcal{O}(pm^2mK)$ .

#### IDENTIFICATION PAR L'ALGORITHME DE SCHUR

1. Calculer l'estimée de  $R_{YY}$

$$\widehat{R}_{YY}(i) = \frac{1}{M} \sum_{t=1}^M \mathbf{y}(t) \mathbf{y}^H(t+i).$$

et  $\widehat{R}_{YY} - \lambda_{\min}(\widehat{R}_{YY})I$

2. Calculer  $\widehat{R}_{YY}^{1/2}(0)$ .
3. Appliquer l'algorithme de Schur généralisé (7.5) jusqu'à la  $\bar{L}^{\text{ème}}$  itération.
4. Soit
  - «- collecter les  $p$  premières colonnes du générateur.
  - «- calculer les  $p$  premiers vecteurs singuliers à gauche du générateur.

### 7.3 Complexité

Dans ce paragraphe, nous évaluons l'ordre de grandeur de la complexité de calcul de l'algorithme de Schur, sans prendre en compte le calcul de  $\widehat{R}_{YY}$ , qui est commune à tous les algorithmes.

La complexité se décompose en

1.  $\lambda_{\min}(\widehat{R}_{YY}) : \mathcal{O}(Km)$  ;
2.  $\widehat{R}_{YY}^{1/2}(0) : \mathcal{O}(m^2)$  ;
3. multiplication de la première colonne bloc :  $\mathcal{O}(Km^3)$  ;
4.  $\bar{L}$  itérations de Schur<sup>6</sup> :  $\mathcal{O}(\bar{L}m(2m)Km)$  ;
5. SVD éventuelle :  $\mathcal{O}(Km^3)$ .

---

<sup>6</sup> pour une itération la complexité est  $\mathcal{O}(rmKm)$  où  $r$  est le rang de déplacement

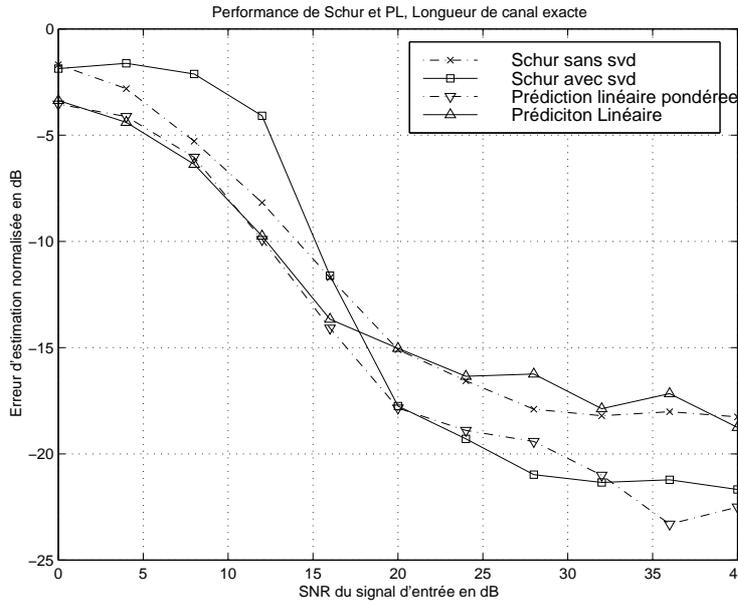
Au total, on obtient une complexité qui est linéaire dans la taille de la matrice de covariance ( $\mathcal{O}(2(\bar{L} + 1)Kmm^2)$ ).

Si on compare avec les algorithmes de prédiction linéaire qui ne font pas appel à l'algorithme de Levinson, la complexité de ceux-ci étant dominée par le calcul d'une Pseudo-inverse ( $\mathcal{O}((Km)^3)$ ) et par le calcul de la solution d'un problème de moindres carrés impliquant des matrices structurées ( $\mathcal{O}(4p(Km)^2)$ ), nous obtenons un algorithme de complexité nettement réduite avec des performances comparables à des SNR modérés.

## 7.4 Simulations

Pour évaluer les performances des algorithmes proposés, nous fournissons le NMSE des canaux estimés, moyennés sur 200 tirs Monte-Carlo. Nous avons utilisé un canal généré aléatoirement avec  $p = 2$  utilisateurs, des longueurs de canal différentes ( $N_1 = 3$  et  $N_2 = 4$ ) et  $m = 4$  sous-canaux. Les symboles sont i.i.d. QPSK et la longueur de trame est de  $M = 250$  symboles. La matrice de mélange a été estimée a posteriori par  $T = \frac{\mathbf{H}_t \mathbf{H}_t^H}{\widehat{\mathbf{H}}_t \widehat{\mathbf{H}}_t^H}$  pour la méthode de Schur (où  $\mathbf{H}_t \triangleq [\mathbf{H}_{1t}^H \cdots \mathbf{H}_{pt}^H]^H$  and  $\mathbf{H}_{it} \triangleq [\mathbf{h}_i^T(j)]_{j=1, N_p}$ ) et de manière similaire pour les méthodes de prédiction linéaire (mais basé sur  $\mathbf{h}(0)$ ).

Nous évaluons la performance de l'algorithme de Schur, sans et avec le calcul de la SVD du générateur et les comparons avec les algorithmes de prédiction linéaire sans et avec pondération (PLP-WLP : Plain Linear Prediction et Weighted Linear Prediction) [4].

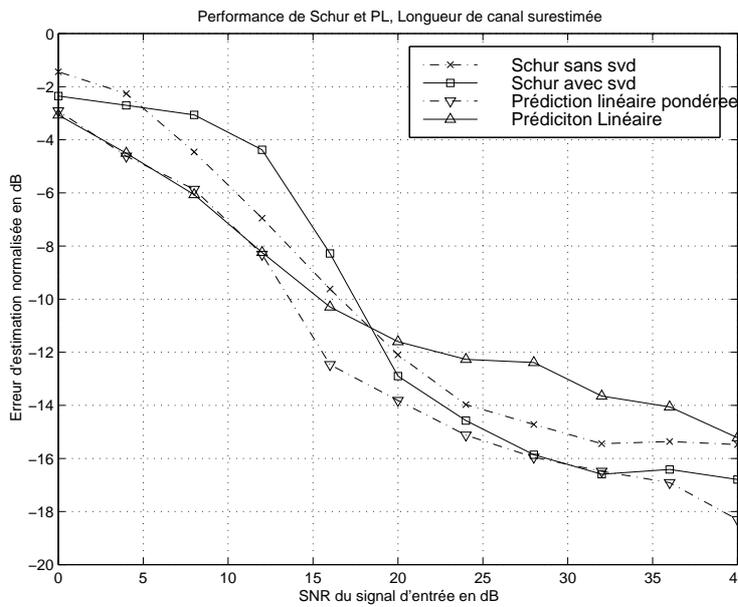


**Comparaison entre les algorithmes de Schur et WLP**

**Longueur du canal exacte.**

Ces courbes montrent que l'algorithme de Schur, quoique très simple, présente des performances comparables aux méthodes de prédiction linéaire.

Les résultats de simulation suivants explorent la robustesse des algorithmes à une surestimation de la longueur du canal ( $\hat{N}_1 = \hat{N}_2 = 6$ ),

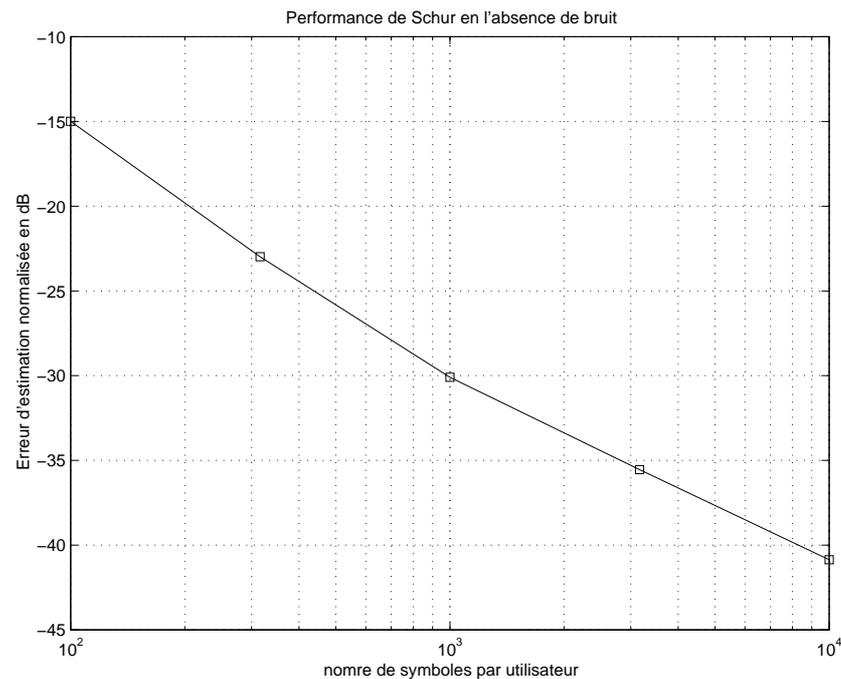


**Comparaison entre les algorithmes de Schur et WLP**

**Longueur du canal surestimée.**

Encore une fois, les performances de l'algorithme de Schur et de prédiction linéaire sont assez proches, ce qui indique que la robustesse de notre algorithme à la surestimation de l'ordre est bonne.

Dans ces deux figures, on peut voir que le calcul de la SVD du générateur donne lieu à des performances meilleures à SNR élevé, auquel cas nous atteignons les performances de la prédiction linéaire pondérée<sup>7</sup>. A SNR faible, la performance est moins bonne.



### Performance de l'algorithme de Schur en fonction du nombre de symboles

Cette courbe montre qu'à un SNR élevé, la performance de notre méthode dépend essentiellement du nombre de symboles de la trame, ce qui est consistant avec l'hypothèse stochastique et le fait qu'on utilise l'estimateur biaisé où on fait implicitement l'hypothèse  $R_{AA} = \sigma_a^2 I$ .

## 7.5 Conclusions

Nous avons introduit une méthode de Schur pour identifier des canaux multiples de façon aveugle dans un cadre SDMA, et ceci avec une faible complexité de calcul. Cet algorithme est récursif dans l'ordre et peut être couplé à un estimateur de longueur de canal et un détecteur de nombre de sources en examinant la diagonale de la décomposition LDU de  $R_{YY} - \sigma_v^2 I$ . Même

7. Nous avons utilisé l'estimateur biaisé de la matrice de corrélation dans les deux cas. Les méthodes PLP-WLP utilisés avec la matrice de covariance empirique ne souffriraient pas de l'effet de seuil constaté à SNR élevé, et la courbe de performance continuerait à descendre.

si la longueur du canal est surestimée, l'algorithme fournit une estimée consistante du canal. De plus, une version adaptative pourrait être développée en utilisant des techniques standard.

Les performances de cet algorithme simple semblent être proches de celle de la méthode de prédiction linéaire. Une étude théorique des performances est en cours de réalisation.

La complexité de calcul de l'algorithme de Schur est linéaire dans la taille de la matrice de covariance, ce qui, en comparant avec la complexité cubique des méthodes de prédiction linéaire, rend notre algorithme très attractif.

## Chapitre 8

# Méthodes au maximum de vraisemblance

*Dans ce chapitre, nous présentons des méthodes basées sur le principe du maximum de vraisemblance en SDMA, que ce soit pour le modèle déterministe ou le modèle gaussien. Nous présentons des algorithmes DML (Deterministic Maximum Likelihood) et des simulations pour le meilleur algorithme tandis que les méthodes gaussiennes, encore en développement à cette heure, sont ébauchées.*

### 8.1 Introduction

Nous investiguons les techniques aveugles et semi-aveugles d'estimation de canal basées sur le principe du maximum de vraisemblance. Deux méthodes DML, basées sur les filtres de prédiction cycliques, ont été présentées dans [67]. L'algorithme IQML (Iterative Quadratic ML) est utilisé dans [67] pour résoudre le problème DML : cette stratégie ne fonctionne bien qu'à des SNR élevés et donne des estimées biaisées du canal. Nous proposons une modification de l'algorithme IQML qui tienne compte explicitement du bruit : le DIQML (Denoised IQML). D'autre part, nous adaptons l'idée de l'algorithme PQML (Pseudo-Quadratic ML), utilisée par [42]. Cette approche consiste à écrire le gradient du critère DML sous une forme linéaire (en fonction du canal) et de se baser sur cette forme pour implémenter un algorithme itératif. La prise en compte explicite du bruit permet également de développer un algorithme PQML consistant et globalement convergent. Des simulations illustrent la performance des différents algorithmes, qui approche les bornes de Cramer-Rao.

D'autre part, nous introduisons le critère GML, ainsi que des suggestions de méthodes à employer pour résoudre le problème GML, entre autres une méthode optimale qui est en cours de développement.

## 8.2 DML : Méthode au Maximum de Vraisemblance dans le modèle Déterministe

La méthode DML a été introduite pour la première fois dans le contexte de l'estimation de canal aveugle dans [68, 69] et [67].

Dans le modèle déterministe, les coefficients du canal et les sources sont considérées comme étant des grandeurs déterministes et estimés conjointement par le critère :

$$\max_{\mathbf{A}, \mathbf{H}} f(\mathbf{Y}|\mathbf{H}) \Leftrightarrow \min_{\mathbf{A}, \mathbf{H}} \|\mathbf{Y} - \mathcal{T}(\mathbf{H})\mathbf{A}\|^2 \quad (8.1)$$

où  $f(\mathbf{Y}|\mathbf{H})$  est la densité de probabilité du signal reçu, le canal étant considéré connu. En résolvant (8.1) par rapport à  $\mathbf{A}$  et en réinjectant la solution dans (8.1), on obtient :

$$\min_{\mathbf{H}} \mathbf{Y}^H P_{\mathcal{T}(\mathbf{H})}^\perp \mathbf{Y} \quad (8.2)$$

où  $P_{\mathcal{T}(\mathbf{H})}^\perp$  est la projection orthogonale sur le sous-espace bruit.

En utilisant la paramétrisation minimale du sous-espace bruit introduite en section 6.9, et de  $P_{\mathcal{T}(\mathbf{H})}^\perp = P_{\mathcal{T}^H(\overline{\mathbf{P}})}$ , (8.2) s'écrit :

$$\min_{\overline{\mathbf{P}}} \mathbf{Y}^H \mathcal{T}^H(\overline{\mathbf{P}}) \mathcal{R}^{-1} \mathcal{T}(\overline{\mathbf{P}}) \mathbf{Y} \quad (8.3)$$

où  $\mathcal{R} = \mathcal{T}(\overline{\mathbf{P}}) \mathcal{T}^H(\overline{\mathbf{P}})$ .

On peut aisément construire  $\mathcal{Y}$  à partir du vecteur d'observations  $\mathbf{Y}$  tel que  $\mathcal{T}(\overline{\mathbf{P}}) \mathbf{Y} = \mathcal{Y} \overline{\mathbf{p}}$ , où  $\overline{\mathbf{p}}$  est un vecteur groupant les éléments de  $\overline{\mathbf{P}}$ . Alors (8.3) s'écrit :

$$\min_{\overline{\mathbf{p}}} \overline{\mathbf{p}}^H \mathcal{Y}^H \mathcal{R}^{-1} \mathcal{Y} \overline{\mathbf{p}} \quad (8.4)$$

D'une manière similaire à [67], on peut introduire un vecteur  $G_N$  contenant les paramètres libres de  $\overline{\mathbf{P}}$  et, en utilisant  $\overline{\mathbf{g}} = [1 \ G_N^T]^T$ , nous obtenons l'équation équivalente

$$\min_{\|\overline{\mathbf{g}}\|=1} \overline{\mathbf{g}}^H \mathcal{Y}_g^H \mathcal{R}^{-1} \mathcal{Y}_g \overline{\mathbf{g}} \quad (8.5)$$

où  $\mathcal{T}(\overline{\mathbf{P}}) \mathbf{Y} = \mathcal{Y}_g \overline{\mathbf{g}}$ . La contrainte  $\overline{\mathbf{g}}(0) = 1$ , est équivalente à  $\|\overline{\mathbf{g}}\| = 1$  pour la minimisation de (8.5), donc  $\hat{\overline{\mathbf{g}}}$  est le vecteur propre minimal de  $\mathcal{Y}_g^H \mathcal{R}^{-1} \mathcal{Y}_g$ .

### 8.2.1 DIQML : un algorithme itératif débruité

Dans [67], l'algorithme IQML est utilisé pour résoudre (8.5) : le dénominateur  $\mathcal{R}$ , calculé à partir de l'itération précédente, est considéré comme étant constant et (8.5) devient quadratique.

On peut montrer que l'estimée obtenue de cette manière est consistante à SNR élevé. Par contre, à SNR plus faible, le vrai canal n'est pas un point stationnaire de l'algorithme et l'estimée du canal est biaisée.

Découlant naturellement de cette constatation, la méthode DIQML va tenter de débruiter le mieux possible les grandeurs en présence, ce faisant, le nouvel algorithme fournira une estimée consistante du canal.

Par la loi des grands nombres, si le nombre de données  $M$  tend vers l'infini, (8.4) est équivalent à son espérance mathématique dont l'expression est :

$$\text{trace} \{P_{\mathcal{T}^H(\bar{\mathbf{P}})} E(\mathbf{Y}\mathbf{Y}^H)\}.$$

La stratégie de débruitage consiste à retrancher le terme de bruit présent dans  $E(\mathbf{Y}\mathbf{Y}^H)$ , i.e.  $\sigma_v^2 I$  ; le critère DIQML devient alors :

$$\min_{\|\bar{\mathbf{p}}\|=1} \text{trace} \{P_{\mathcal{T}^H(\bar{\mathbf{P}})} (\mathbf{Y}\mathbf{Y}^H - \sigma_v^2 I)\} \quad (8.6)$$

On notera que cette opération (i.e.  $-\sigma_v^2 I$ ) ne change pas le critère DML, puisque  $\sigma_v^2 \text{trace} \{P_{\mathcal{T}^H(\bar{\mathbf{P}})}\}$  est une constante sous la condition  $\|\bar{\mathbf{p}}\| = 1$ .

On résout alors (8.6) en considérant que  $\mathcal{R}$  est constant à chaque itération :

$$\min_{\bar{\mathbf{p}}} \bar{\mathbf{p}}^H \left\{ \mathcal{Y}^H \mathcal{R} \# \mathcal{Y} - \sigma_v^2 D \right\} \bar{\mathbf{p}} \quad (8.7)$$

où  $\bar{\mathbf{p}}^H D \bar{\mathbf{p}} = \text{trace} \{\mathcal{T}^H(\bar{\mathbf{P}}) \mathcal{R} \# \mathcal{T}(\bar{\mathbf{P}})\}$ . DIQML est asymptotiquement (en  $M$ ) globalement convergent [20] : le critère DIQML est équivalent au critère "débruité" :

$$\min_{\bar{\mathbf{p}}} \bar{\mathbf{p}}^H \mathcal{X}^H \mathcal{R}^{-1} \mathcal{X} \bar{\mathbf{p}} \quad (8.8)$$

où  $\mathcal{X} \bar{\mathbf{p}} = \mathcal{T}(\bar{\mathbf{P}}) \mathbf{X}$ . Quand on utilise  $\bar{\mathbf{g}}$ , la matrice  $\mathcal{X}_g^H \mathcal{R}^{-1} \mathcal{X}_g$  a exactement une singularité et la solution est alors le vecteur propre minimal. L'utilisation de  $\bar{\mathbf{p}}$  produit une matrice  $\mathcal{X}^H \mathcal{R} \# \mathcal{X}$  présentant  $(m - p)^2$  singularités, correspondant aux  $(m - p)^2$  ambiguïtés de  $\bar{\mathbf{P}}$  si ses valeurs fixes (1 et 0) ne sont pas prises en compte. Une minimisation classique permet de lever ces indéterminations. En pratique,  $M$  est fini, et le Hessien de (8.7) est indéfini : pour résoudre ce problème, on retranche une quantité  $\lambda D$  (et non  $\sigma_v^2 D$ ) pour rendre ce Hessien positif (semi-)défini, ce qui conduit à résoudre une minimisation contrainte sur  $\bar{\mathbf{p}}$  et  $\lambda$ . Si on utilise  $\bar{\mathbf{g}}$ ,  $\lambda$  est choisi de telle manière que  $\mathcal{Y}_g^H \mathcal{R}^{-1} \mathcal{Y}_g - \lambda D_g$  soit semi-défini.  $\lambda$  est donc la valeur propre généralisée de  $\mathcal{Y}_g^H \mathcal{R}^{-1} \mathcal{Y}_g$  et  $D$  ;  $\bar{\mathbf{g}}$  est le vecteur propre généralisé correspondant.

La convergence asymptotique globale a été démontrée dans [5, 20] pour un seul utilisateur et peut être étendue pour le cas multi-utilisateurs.

### 8.2.2 PQML : une méthode itérative Pseudo Quadratique

Le principe de PQML a été appliqué à la méthode DML par [42], où le critère DML était paramétrisé par les coefficients du canal. Dans notre cas, ce critère doit être paramétrisé par le prédicteur linéaire, dans la mesure où il n'existe pas de paramétrisation linéaire du sous-espace bruit par les coefficients du canal.

On peut écrire le gradient de la fonction de coût DML (8.4) sous la forme  $\mathcal{P}(\bar{\mathbf{p}})\bar{\mathbf{p}}$ , où  $\mathcal{P}(\bar{\mathbf{p}})$  est (idéalement) positive semi-définie. La solution ML vérifie la relation  $\mathcal{P}(\bar{\mathbf{p}})\bar{\mathbf{p}} = 0$ , que l'on résout itérativement comme suit : dans un premier temps,  $\mathcal{P}(\bar{\mathbf{p}})$  est supposé constant, ce qui donne un problème quadratique,  $\mathcal{P}(\bar{\mathbf{p}})$  étant positive semi-définie, le Hessien de (8.4) l'est également et on peut trouver une solution aisément. Cette solution est utilisée pour réévaluer  $\mathcal{P}(\bar{\mathbf{p}})$  et d'autres itérations peuvent être faites.

La principale difficulté consiste à trouver la bonne matrice  $\mathcal{P}(\bar{\mathbf{p}})$ , en particulier, le Hessien de la fonction de coût de PQML doit rester défini positif. Dans notre cas :

$$\mathcal{P}(\bar{\mathbf{p}}) = \mathcal{Y}^H \mathcal{R}^{-1} \mathcal{Y} - \mathcal{B}^H \mathcal{B} \quad (8.9)$$

$\mathcal{T}^H(\bar{\mathbf{P}})\mathcal{B} = \mathcal{B}^* \bar{\mathbf{p}}^*$  avec  $\mathcal{B} = [\mathcal{T}(\bar{\mathbf{P}})\mathcal{T}^H(\bar{\mathbf{P}})]^{-1} \mathcal{T}(\bar{\mathbf{P}})\mathcal{Y}$ . Le Hessien de (8.4) est indéfini pour un nombre  $M$  de symboles fini : la solution proposée par [42] est de prendre le vecteur propre minimal, mais cette stratégie ne fonctionne que pour des SNR très élevés.

On peut mettre en relation PQML et IQML, en notant que le premier terme de la matrice du milieu dans (8.7) et (8.9) sont les mêmes  $E(\mathcal{B}^H \mathcal{B}) = \sigma_v^2 D$ . Par analogie avec IQML, nous introduisons un facteur  $\lambda$  arbitraire, le critère PQML devient :

$$\min_{\|\bar{\mathbf{p}}\|=1, \lambda} \bar{\mathbf{p}}^H \{ \mathcal{Y}^H \mathcal{R}^{-1} \mathcal{Y} - \lambda \mathcal{B}^H \mathcal{B} \} \bar{\mathbf{p}} \quad (8.10)$$

avec la contrainte que le Hessien soit positif (semi-)défini.

Si on travaille avec  $\bar{\mathbf{g}}$ ,  $\lambda$  est choisi de telle manière que  $\mathcal{Y}_g^H \mathcal{R}^{-1} \mathcal{Y}_g - \lambda \mathcal{B}_g^H \mathcal{B}_g$  soit positif semi-défini.  $\lambda$  est alors la valeur propre généralisée minimale de  $\mathcal{Y}_g^H \mathcal{R}^{-1} \mathcal{Y}_g$  et  $\mathcal{B}_g^H \mathcal{B}_g$  ;  $\bar{\mathbf{g}}$  est son vecteur propre généralisé. La convergence asymptotique est démontrée de la même manière que pour DIQML [20]. Les points stationnaires de PQML sont ceux de DML, et la performance de PQML est donc la même que celle de DML.

### 8.2.3 DML semi-aveugle

Nous divisons la trame reçue suivant  $\mathbf{Y} = [\mathbf{Y}_{ts}^H \mathbf{Y}_b^H]^H$  où  $\mathbf{Y}_{ts}^H = \mathcal{T}(\mathbf{H})\mathbf{A}_k + \mathbf{V}_{ts}$  contient les sorties générées par les symboles connus uniquement.  $\mathbf{Y}_b$  contient les observations générées par les symboles inconnus et un mélange de symboles connus et inconnus. La solution du critère

DML basée sur  $\mathbf{Y}$  implique une estimation conjointe du filtre de prédiction singulier et du canal, estimation qui devrait donc être faite sous la contrainte d'orthogonalité  $\overline{\mathbf{P}}\mathcal{T}(\mathbf{H}) = 0$ .

Nous proposons un algorithme plus simple où  $\overline{\mathbf{P}}$  est estimé par un algorithme PQML, en se servant des observations  $\mathbf{Y}_b^H$ . Le canal est ensuite estimé en utilisant le filtre de prédiction singulier par la relation  $\overline{\mathbf{P}}\mathcal{T}(\mathbf{H}) = 0$  et la séquence d'apprentissage. Cette stratégie donne lieu à un critère combiné (notons que  $\mathbf{Y}_{ts}$  et  $\mathbf{Y}_b$  sont décorrélés, ce qui justifie l'utilisation du critère combiné).

$$\min_{\mathbf{H}} (M - K)h^H \mathcal{T}(\overline{\mathbf{P}}^t) \mathcal{T}^H(\overline{\mathbf{P}}^t)h + K \|\mathbf{Y}_{ts} - \mathcal{T}(\mathbf{H})\mathbf{A}_k\|^2 \quad (8.11)$$

où  $^t$  signifie qu'on transpose les blocs d'une matrice ou d'un vecteur bloc.

$$\hat{h} = K \left( (M - K) \cdot \mathcal{T}(\overline{\mathbf{P}}^t) \mathcal{T}^H(\overline{\mathbf{P}}^t) + K \cdot \mathcal{A}_k^H \mathcal{A}_k \right)^{-1} \cdot \mathcal{A}_k^H \mathbf{Y}_{ts} \quad (8.12)$$

où  $h = \text{vec}(\mathbf{H})$ ,  $K$  est le nombre de symboles connus et  $\mathcal{A}_k h = \mathcal{T}(\mathbf{H})\mathbf{A}_k$ . Les facteurs  $K$  et  $(M - K)$  sont utilisés pour pondérer les contributions "aveugles" et "d'apprentissage" dans le critère.

### 8.2.4 Simulations

Nous avons considéré des symboles BPSK, des trames de longueur  $M = 200$ ,  $p=2$  canaux à coefficients réels de longueur 5 et  $m = 4$  sous-canaux, générés aléatoirement. La performance est mesurée en Normalise Root MSE et est estimée sur 100 réalisations Monte-Carlo. Pour le cas aveugle, la matrice de mélange  $U$  est déterminée par, en notant  $\mathbf{H}(i) = [\mathbf{H}_1(i) \cdots \mathbf{H}_p(i)]$  et  $\check{\mathbf{H}} = [\mathbf{H}(0) \cdots \mathbf{H}(N_1)]$ ,  $U = \check{\mathbf{H}}^H \check{\mathbf{H}} / (\check{\mathbf{H}}^H \widehat{\mathbf{H}})$ .

L'application des stratégies PQDML et IQDML basées sur  $\overline{\mathbf{g}}$  ont donné des résultats de simulation relativement décevants dans un premier temps. Nous avons donc implémenté ces algorithmes en se basant sur  $\overline{\mathbf{p}}$ , mais en utilisant  $\lambda$  comme étant la valeur propre généralisée de  $(\mathcal{Y}^H \mathcal{R}^{-1} \mathcal{Y}, \mathcal{B}^H \mathcal{B})$ , ce qui donne les excellents résultats que nous avons reproduit ci-dessous. Cette valeur de  $\lambda$  assure la positivité du Hessien, mais n'est pas optimale. la valeur de  $\lambda$  optimale pour ce problème étant de toute évidence la même que pour le problème paramétrisé par  $\overline{\mathbf{g}}$ , les simulations nous menaient à la conclusion paradoxale qu'une valeur légèrement sous-optimale de  $\lambda$  donnait de meilleurs résultats. En fait, il semble que le  $\lambda$  optimal provoque certaines instabilités dans l'algorithme, similaires aux instabilités qu'on peut trouver dans les algorithmes du gradient.

L'algorithme semi-aveugle a été implémenté pour 20 symboles connus (par utilisateur). La matrice de mélange ne doit pas être estimée séparément. Ces simulations (voir figure 8.3) montrent

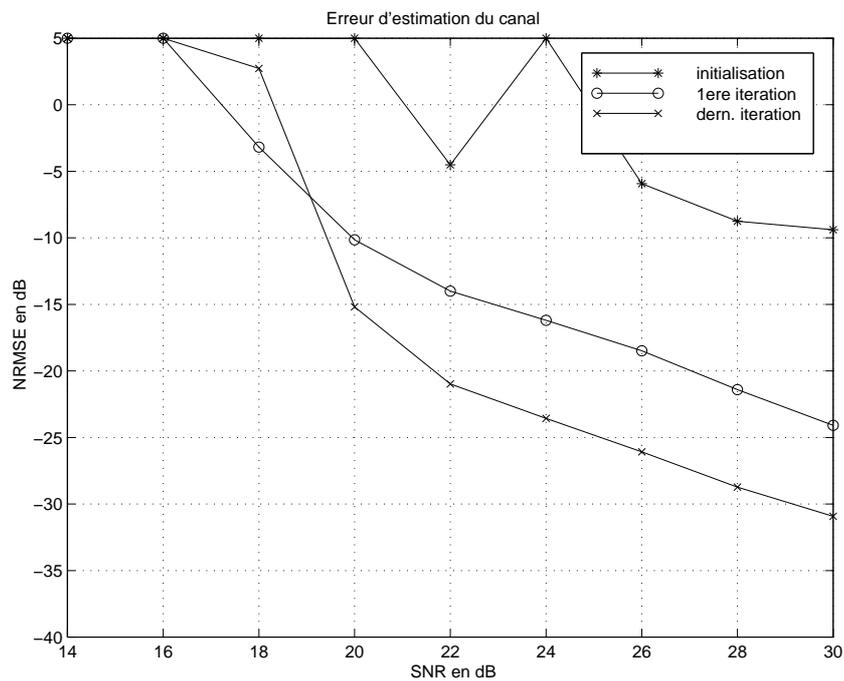


FIG. 8.1 – Performance de PQML aveugle

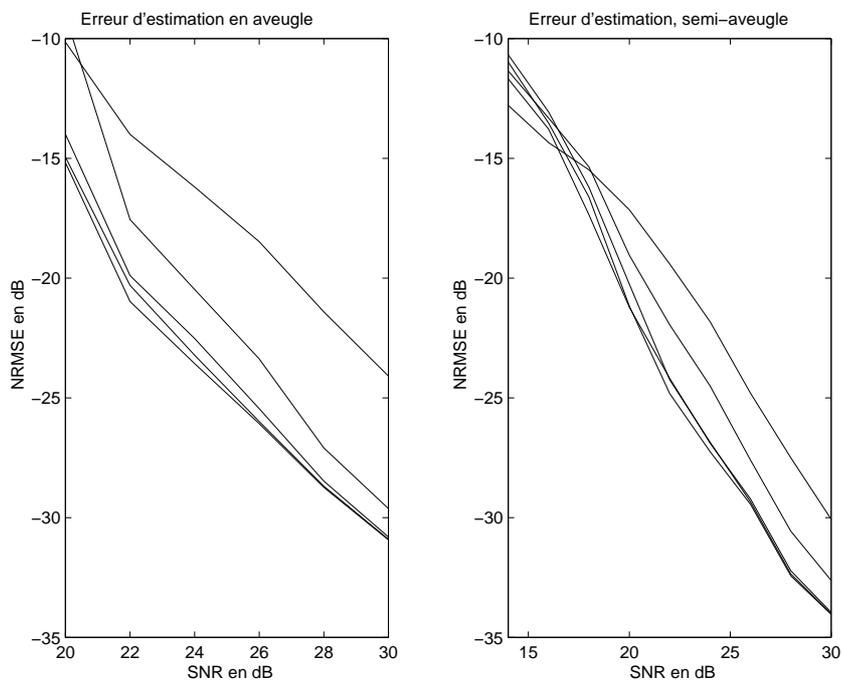


FIG. 8.2 – Itérations successives en PQML

clairement l'avantage de l'utilisation de l'algorithme semi-aveugle par rapport aux algorithmes aveugles et avec séquence d'apprentissage.

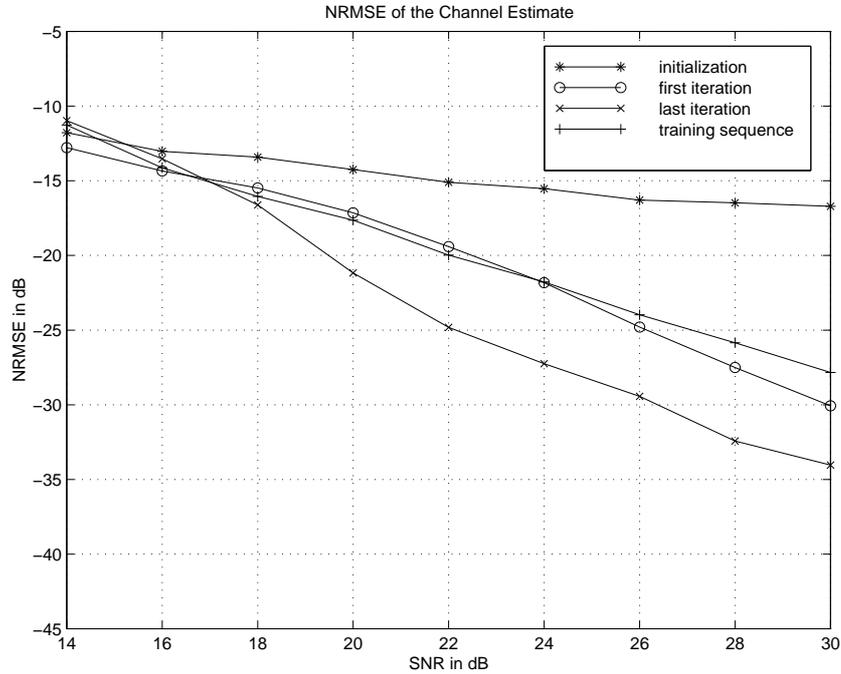


FIG. 8.3 – Performance de PQML semi-aveugle

La comparaison des CRB's pour l'estimation aveugle (où la matrice de mélange est supposée connue), semi-aveugle avec peu de symboles connus (ici 10, ce qui est insuffisant pour faire de l'estimation avec séquence d'apprentissage) et semi-aveugle avec 20 symboles connus.

Les deux premières courbes de la figure (8.4) sont relativement proches, ce qui indique que les premiers symboles connus utilisés par l'algorithme semi-aveugle sont utilisés principalement pour estimer la matrice de mélange. Les symboles suivants nous fournissent une meilleure estimation du canal.

### 8.3 GML : Méthode au Maximum de Vraisemblance pour le modèle gaussien

Le développement de méthodes GML est motivé par les constatations suivantes :

1. Les méthodes GML, développées dans le cas mono-utilisateur [19], donnent de meilleures performances que les méthodes DML. On s'attend à ce que ce résultat soit toujours valable dans le cas multi-utilisateurs. L'hypothèse de sources gaussiennes a également été utilisée

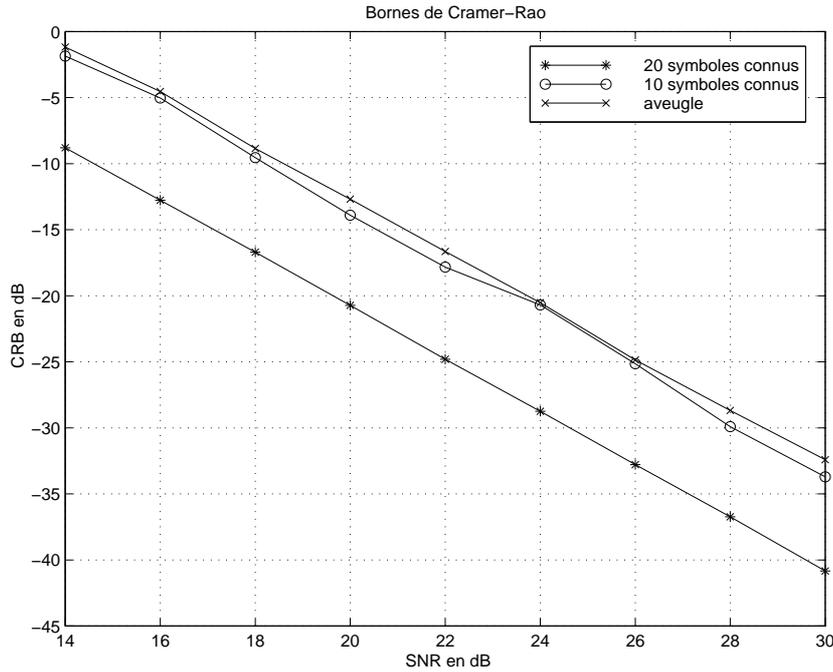


FIG. 8.4 – CRB's de DML aveugle et semi-aveugle

avec succès pour le problème de la détermination de direction d'arrivée de source [57].

2. Les conditions d'identifiabilité développées précédemment montrent que l'utilisation du modèle gaussien génère moins d'ambiguïtés que l'utilisation du modèle déterministe.
3. La robustesse des méthodes de prédiction linéaire et de Schur à la surestimation de la longueur du canal indiquent que le modèle gaussien donne lieu à des méthodes robustes.

L'ensemble de ces considérations indique que l'utilisation du modèle gaussien pour l'identification aveugle de canal en multi-utilisateurs est nettement préférable à l'utilisation du modèle déterministe.

### 8.3.1 GML : gaussian Maximum Likelihood, méthodes approchées.

Réécrivons l'équation  $\mathbf{Y} = \mathcal{T}(\mathbf{H})\mathbf{A} + \mathbf{V}$ , avec les hypothèses gaussiennes  $\mathbf{V} \sim \mathcal{N}(0, R_{VV})$  et  $\mathbf{A} \sim \mathcal{N}(0, R_{AA})$ . Nous désirons maximiser  $f(\mathbf{Y}|\mathbf{H})$ ,  $\mathbf{Y} \sim \mathcal{N}(0, R_{YY})$  et le critère devient :

$$\min_{\mathbf{H}} \{ \ln(\det R_{YY}) + \mathbf{Y}^H R_{YY}^{-1} \mathbf{Y} \}. \quad (8.13)$$

### GML basé sur le canal orthogonal

Considérons que l'on filtre le signal reçu par le canal orthogonal (voir chapitre 5) du premier utilisateur :

$$\begin{aligned} \mathcal{T}(\mathbf{H}^{1\perp\uparrow})\mathbf{Y} &\sim \mathcal{N}(0, R) \\ R &= \mathcal{T}(\mathbf{H}^{1\perp\uparrow}) \left[ \sigma_a^2 \sum_{i=2}^p \mathcal{T}(\mathbf{H}^i)\mathcal{T}^H(\mathbf{H}^i) + \sigma_v^2 I \right] \mathcal{T}^H(\mathbf{H}^{1\perp\uparrow}) \end{aligned} \quad (8.14)$$

On peut alors effectuer la minimisation ML :

$$\min_{\mathbf{H}^1} (\ln(\det R) + \mathbf{Y}^H \mathcal{T}^H(\mathbf{H}^{1\perp\uparrow}) R^{-1} \mathcal{T}(\mathbf{H}^{1\perp\uparrow}) \mathbf{Y}) \quad (8.15)$$

où on considère que les canaux des autres utilisateurs sont connus (estimée précédente). On peut alors poursuivre par une procédure itérative, en utilisant les estimées de l'étape précédente pour initialiser  $R$ .

### Approche DML/GML combinée

En supposant que l'utilisateur 1 est l'utilisateur d'intérêt, on applique une stratégie DML sur le premier utilisateur et GML sur les autres. Donc, on considère que la première source est déterministe et que les autres sont du bruit blanc de puissance  $\sigma_a^2$  :

$$\mathbf{Y} \sim \mathcal{N}(\mathcal{T}(\mathbf{H}^1)A^1, \underbrace{\sigma_a^2 \sum_{i=2}^p \mathcal{T}(\mathbf{H}^i)\mathcal{T}^H(\mathbf{H}^i) + \sigma_v^2 I}_{=R}) \quad (8.16)$$

ce qui donne la minimisation :

$$\min_{\mathbf{H}^1, A^1} (\mathbf{Y} - \mathcal{T}(\mathbf{H}^1)A^1)^H R^{-1} (\mathbf{Y} - \mathcal{T}(\mathbf{H}^1)A^1) \quad (8.17)$$

qui peut être résolue séquentiellement en  $A^1$ ,  $\mathbf{H}^1$ . Dans un premier temps,

$\hat{A}^1 = (\mathcal{T}^H(\mathbf{H}^1)R^{-1}\mathcal{T}(\mathbf{H}^1))^{-1}\mathcal{T}^H(\mathbf{H}^1)R^{-1}\mathbf{Y}$ , soit, en combinant avec (8.17), et en utilisant la propriété  $\mathbf{P}_{R^{-\frac{1}{2}}\mathcal{T}(\mathbf{H}^1)}^\perp = \mathbf{P}_{R^{\frac{1}{2}}\mathcal{T}^H(\mathbf{H}^{1\perp\uparrow})}$ , comme  $\mathcal{T}(\mathbf{H}^{1\perp\uparrow})R^{\frac{1}{2}}R^{-\frac{1}{2}}\mathcal{T}(\mathbf{H}^1) = 0$ , nous obtenons le même critère que (8.15) au terme  $\ln(\det R)$  près.

Cette interprétation est intéressante, dans la mesure où elle indique que l'algorithme approché pour résoudre (8.15) consiste en l'utilisation d'une approche IQML (ou PQML) avec la présence d'un bruit important. Il s'ensuit que cette stratégie ne fonctionne correctement que pour des SINR élevés, i.e. quand l'utilisateur d'intérêt est nettement plus puissant que l'interfère. Dans le cadre du SDMA, cette procédure est donc inapplicable.

### GML basé sur le canal orthogonal global

On ne peut pas trouver de canal orthogonal à tous les canaux des utilisateurs et qui soit paramétrisé par  $\mathbf{H}$ . Cependant, on peut introduire  $\mathbf{H}^{\perp\dagger} = \begin{bmatrix} \mathbf{H}^{1\perp\dagger} \\ \vdots \\ \mathbf{H}^{p\perp\dagger} \end{bmatrix}$ , et considérer le signal reçu filtré par ce *canal orthogonal global*:

$$\begin{aligned} \mathcal{T}(\mathbf{H}^{\perp\dagger})\mathbf{Y} &\sim \mathcal{N}(0, R) \\ R &= \mathcal{T}(\mathbf{H}^{\perp\dagger}) \left[ \sigma_a^2 \sum_{i=1}^p \mathcal{T}(\mathbf{H}^i)\mathcal{T}^H(\mathbf{H}^i) + \sigma_v^2 I \right] \mathcal{T}^H(\mathbf{H}^{\perp\dagger}) \end{aligned} \quad (8.18)$$

On peut alors introduire la minimisation ML suivante :

$$\min_{\mathbf{H}} (\ln(\det R) + \mathbf{Y}^H \mathcal{T}^H(\mathbf{H}^{\perp\dagger}) R^{-1} \mathcal{T}(\mathbf{H}^{\perp\dagger}) \mathbf{Y}) \quad (8.19)$$

qui permet de faire la minimisation en une seule étape. On peut à nouveau appliquer une méthode itérative, auquel cas le terme  $\ln(\det R)$  disparaît.

### Simulations

Nous ne reproduisons pas ici les résultats de simulations partielles que nous avons effectué. Signalons que ces méthodes approchées, de par le problème "near-far" évoqué dans le paragraphe traitant de l'approche combinée DML/GML, ne donnent des résultats acceptables que pour des SNR très élevés, et sont, si on ne développe pas des algorithmes rapides, très lourds, dans la mesure où les matrices entrant en jeu sont de la taille de la trame de symboles multipliée par le nombre de canaux.

### 8.3.2 Une méthode optimale en développement

De par les avantages intrinsèques des méthodes basées sur le modèle gaussien, nous envisageons cependant de continuer le développement d'une méthode optimale pour résoudre GML, i.e. par l'algorithme de Scoring. Cette méthode est en cours de développement avec Elisabeth de Carvalho.

#### La méthode de Scoring : implémentation rapide

La méthode de Scoring est une méthode bien connue qui consiste à implémenter un algorithme itératif de Newton-Raphson, i.e., en notant  $\theta^{(i)}$ , l'estimée de  $\theta$  à l'itération  $i$  et  $f(\theta)$  la fonction à minimiser :

$$\theta^{(i)} = \theta^{(i-1)} - \frac{\frac{\partial f(\theta)}{\partial \theta^*}}{\frac{\partial}{\partial \theta^*} \left( \frac{\partial f(\theta)}{\partial \theta^*} \right)} \Bigg|_{\theta^{(i-1)}} \quad (8.20)$$

en approximant le Hessien par son espérance mathématique, i.e. la FIM [18]. Cette approximation est valable dans la mesure où le nombre de symboles est grand.

L'expression de cette FIM est :

$$FIM(i, j) = \text{trace} \left\{ C_{YY}^{-1} \left( \frac{\partial C_{YY}}{\partial \theta_i} \right) C_{YY}^{-1} \left( \frac{\partial C_{YY}}{\partial \theta_j} \right)^T \right\} \quad (8.21)$$

L'existence d'un algorithme "rapide" vient de la structure asymptotique de la FIM et également de la structure bloc Toeplitz de  $C_{YY}$  et de ce que  $\frac{\partial C_{YY}}{\partial \theta_i}$  sont des matrices creuses.

## 8.4 Conclusions

Nous avons proposé plusieurs méthodes pour résoudre les critères DML aveugles et semi-aveugles qui apparaissent dans l'estimation de canal multi-utilisateur.

On montre que la méthode PQML fournit asymptotiquement le minimum global ML et les simulations confirment que les performances sont assez proches des bornes de Cramer-Rao, que ce soit en aveugle ou en semi-aveugle, quoique DML ne puisse atteindre la CRB. Dans le cas aveugle, il semble qu'on puisse s'attendre à obtenir des performances acceptables à un SNR de 20 dB pour une trame de 200 symboles. Dans le cas semi-aveugle, l'algorithme donne de meilleurs résultats, mais la contribution de la partie aveugle devient nulle à partir d'un SNR de 15 dB.

D'autre part, les simulations confirment l'intuition naturelle selon laquelle les premiers symboles connus dans le cas semi-aveugle servent à démêler les utilisateurs, avant d'améliorer la performance de l'estimation.

Une piste explorable consisterait à exploiter l'orthogonalité du filtre prédicteur singulier et du canal en les estimant conjointement sous la contrainte  $\overline{\mathbf{P}}^T(\mathbf{H}) = 0$ .

Enfin, pour les méthodes GML, les algorithmes approchés basés sur le canal orthogonal donnent de piètres résultats, et justifient le développement d'une méthode optimale, qui peut, de plus, être implémentée grâce à des algorithmes rapides.



# Chapitre 9

## Méthodes basées sur les statistiques cycliques

*Dans ce chapitre, nous nous intéressons aux méthodes d'identification de canal aveugles basées sur les statistiques cycliques du signal reçu. Dans un premier temps, nous exploitons l'absence de la contribution du bruit dans les statistiques de fréquence cyclique non nulle pour développer des méthodes sous-espace et basées sur la prédiction linéaire [28]. Ensuite, nous développons un modèle multicanal qui permet d'exploiter directement les statistiques cycliques dans leur ensemble [25].*

### 9.1 Hypothèses

Dans un premier temps, les seules hypothèses utilisées seront la stationnarité des symboles, la cyclostationnarité du signal reçu et le fait que le bruit additif ne doit pas être cyclostationnaire de même fréquence cyclique que le signal reçu.

### 9.2 Introduction aux statistiques cycliques

Conformément aux hypothèses exprimées ci-dessus, la séquence de corrélation

$$R_{yy}(n, \tau) = \mathbb{E} \{ \mathbf{y}(n) \mathbf{y}^H(n - \tau) \} \quad (9.1)$$

est cyclique en  $n$  de période  $m$ . Nous pouvons aisément l'exprimer en fonction du canal sous la forme :

$$R_{yy}(n, \tau) = \sum_{\alpha=-\infty}^{\infty} \sum_{\beta=-\infty}^{\infty} \mathbf{H}(n - \alpha m) R_{aa}(\beta) \mathbf{H}^H(n - \alpha m + \beta m - \tau) + R_{vv}(\tau) \quad (9.2)$$

où  $\mathbf{H}(k) = [\mathbf{H}_{*1}(k) \cdots \mathbf{H}_{*m_a}(k)]$  et les  $\mathbf{H}_{*i}(k)$  sont les canaux reliant les sources au capteur  $i$  (voir figure 9.1) Nous pouvons alors définir la  $k^{\text{ème}}$  séquence corrélation cyclique (i.e. de fréquence cyclique  $k/T$ ):

$$R_{yy}^{\{k\}}(\tau) \triangleq \frac{1}{m} \sum_{l=0}^{m-1} R_{yy}(l, \tau) e^{-j\frac{2\pi lk}{m}} \triangleq \mathbf{E}^k \{ \mathbf{y}(l) \mathbf{y}^H(l - \tau) \} \quad (9.3)$$

soit, en fonction du canal :

$$R_{yy}^{\{k\}}(\tau) = \frac{1}{m} \sum_{\alpha=-\infty}^{\infty} \sum_{\beta=-\infty}^{\infty} \mathbf{H}(\alpha) R_{aa}(\beta) \mathbf{H}^H(\alpha + \beta m - \tau) e^{-j\frac{2\pi \alpha k}{m}} + R_{vv}(\tau) \delta(k) \quad (9.4)$$

avec  $k$  entier et  $R_{yy}^{\{k\}}(\tau) = R_{yy}^{\{k+m\}}(\tau)$  (habituellement, on utilise  $k \in [-\lfloor \frac{m}{2} \rfloor, \dots, \lfloor \frac{m}{2} \rfloor]$ ).

Nous introduisons également

$$D_{DFT}^{\{k,p\}} = \text{blockdiag}[I_p | e^{-j\frac{2\pi k}{m}} I_p | \cdots | e^{-j\frac{2\pi(N-1)k}{m}} I_p] \quad (9.5)$$

## 9.3 Méthode basée sur le modèle cyclostationnaire

### 9.3.1 Modèle cyclostationnaire

Dans ce modèle, les signaux suréchantillonnés ne sont pas vectorisés, ce qui nous donne simplement, dans le cas multi-utilisateurs et multi-antennes :

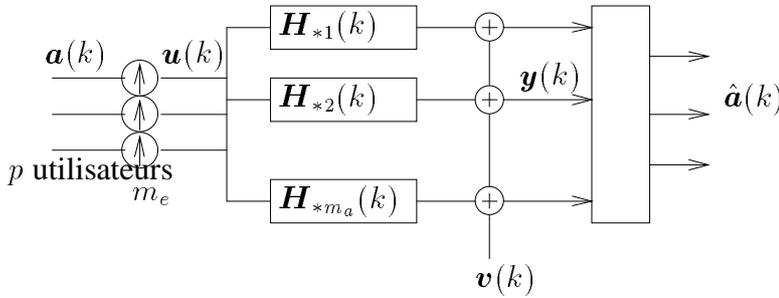


FIG. 9.1 – Modèle cyclostationnaire

Nous pouvons alors introduire la matrice de corrélation cyclique de fréquence cyclique  $k$  :

$$\begin{aligned} R_{yy}^{\{k\}} &\triangleq \begin{bmatrix} R_{yy}^{\{k\}}(0) & R_{yy}^{\{k\}}(1) & \cdots & R_{yy}^{\{k\}}(K-1) \\ R_{yy}^{\{k\}}(-1) & R_{yy}^{\{k\}}(0) & \cdots & R_{yy}^{\{k\}}(K-2) \\ \vdots & \vdots & \ddots & \vdots \\ R_{yy}^{\{k\}}(1-K) & R_{yy}^{\{k\}}(2-K) & \cdots & R_{yy}^{\{k\}}(0) \end{bmatrix} \\ &= \mathcal{T}_K(\mathbf{H}_N \mathbf{D}_{DFT}^{\{k\}}) \mathbf{R}_{uu}^{\{k\}} \mathcal{T}_K^H(\mathbf{H}_N) + \delta(k) \mathbf{R}_{vv} \end{aligned}$$

où  $\mathbf{R}_{uu}^{\{k\}} = \mathbf{R}_{aa} \otimes I_m$  En effet :

$$\begin{aligned} r_{uu}(n, \tau) &= \mathbf{E} \{ \mathbf{u}(n) \mathbf{u}(n - \tau) \} \\ &= \sum_{i=-\infty}^{\infty} \sum_{j=-\infty}^{\infty} \delta(n - im) \delta(n - jm - \tau) \mathbf{a}(i) \mathbf{a}^H(j) \end{aligned}$$

$$r_{uu}^{\{k\}}(\tau) = \sum_{l=0}^{m-1} \sum_i \sum_j \underbrace{\delta(l - im)}_{\Rightarrow l=im=0} \delta(n - jm - \tau) \mathbf{a}(i) \mathbf{a}^H(j) r_{aa}(j - i) w^{kl}$$

où  $w = e^{-i \frac{2\pi}{m}}$

$$r_{uu}^{\{k\}}(\tau) = \sum_{j=-\infty}^{\infty} \delta(jm + \tau) r_{aa}(j)$$

Donc,

$$\forall k, \mathbf{R}_{uu, Lm}^{\{k\}} = \mathbf{R}_{aa, L} \otimes I_m \triangleq \begin{bmatrix} r_{aa}(0)I_m & r_{aa}(1)I_m & \cdots & r_{aa}(K-1)I_m \\ r_{aa}(-1)I_m & r_{aa}(0)I_m & \ddots & r_{aa}(K-2)I_m \\ \vdots & \vdots & \ddots & \vdots \\ r_{aa}(1-K)I_m & r_{aa}(2-K)I_m & \cdots & r_{aa}(0)I_m \end{bmatrix}$$

où  $\otimes$  est un produit de Kronecker bloc.

Cette relation montre que les corrélations cycliques de fréquence cycliques  $k \neq 0$  ne sont pas affectées par un bruit additif, quelle que soit sa couleur, à condition qu'il ne soit pas  $k/T$  cyclostationnaire.

## 9.4 Application de la méthode des sous-espaces

### 9.4.1 La méthode classique

Nous revenons au cas mono-utilisateur et introduisons la notation  $\mathbf{H}'_N = \mathbf{H}_N D_{DFT}^{\{k\}}$ , l'exposant  $k$  étant déterminé par le contexte.

Nous pouvons écrire la forme compacte de la SVD de la matrice de corrélation cyclique  $R_{yy}^{\{k\}} = U D V^H$  où

$$\text{range}\{U\} = \text{range}\{\mathcal{T}_K(\mathbf{H}'_N)\}$$

and

$$\text{range}\{V\} = \text{range}\{\mathcal{T}_K(\mathbf{H}_N)\}$$

Sous les hypothèses  $[\mathbf{H}\mathcal{H}\text{full}]$ ,  $\mathcal{T}_K(\mathbf{H}'_N)$  est de rang plein (on peut vérifier facilement que  $\text{rank}\{\mathcal{T}_K(\mathbf{H}_N)\} = \text{rank}\{\mathcal{T}_K(\mathbf{H}'_N)\}$ ). Nous nous retrouvons alors avec le problème sous-espace :

$$\min_{\mathbf{H}'_N, T} \|\mathcal{T}_K(\mathbf{H}'_N) - \mathbf{U}T\|_F^2$$

En introduisant  $\mathbf{U}^\perp$  tel que  $[\mathbf{U}\mathbf{U}^\perp]$  est une matrice unitaire :

$$\min_{\mathbf{H}'_N} \mathbf{H}'_N{}^t \left[ \sum_{i=1}^{D^\perp} \mathcal{T}_N(\mathbf{U}_i^{\perp Ht}) \mathcal{T}_N^H(\mathbf{U}_i^{\perp Ht}) \right] \mathbf{H}'_N{}^{Ht}$$

où  $\mathbf{U}_i^\perp$  est un vecteur bloc de dimensions  $K \times 1$  avec des blocs  $M \times 1$ ,  $D^\perp = N + K - 1$ . L'estimée  $\widehat{\mathbf{H}}_N'^{Ht}$  est donnée par le vecteur propre minimal de la matrice entre crochets.

#### 9.4.2 Algorithme sans décomposition vecteurs/valeurs propres

L'algorithme sans décomposition vecteurs/valeurs propres présenté précédemment s'applique de manière très naturelle, en effet, pour  $k \neq 0$  :

$$R_{yy}^{\{k\}} = \mathcal{T}_K(\mathbf{H}'_N) R_{uu}^{\{k\}} \mathcal{T}_K^H(\mathbf{H}_N) \quad ,$$

la contribution du bruit disparaît. Il s'en suit que l'espace colonne à gauche de  $\mathcal{T}_K(\mathbf{H}'_N)$  et de  $R_{yy}^{\{k\}}$  sont identiques, ce qui mène au critère.

$$\min_{\mathbf{H}'_N, \mathbf{Q}} \|\mathcal{T}_K(\mathbf{H}'_N) - \widehat{R}_{yy}^{\{k\}} \mathbf{B}\mathbf{Q}\|_F^2 \quad (9.6)$$

La matrice  $\mathbf{B}$  joue le même rôle que plus haut, et, de la même manière, on obtient :

$$\min_{\mathbf{H}'_N} \|\mathbf{P}_F^\perp \mathcal{T}_K(\mathbf{H}'_N)\|_F^2$$

où  $\mathbf{P}_F^\perp = \mathbf{I} - \mathbf{P}_F$  et  $\mathbf{P}_F = \mathbf{F}(\mathbf{F}^H \mathbf{F})^{-1} \mathbf{F}^H$ . Sous contrainte  $\|\mathbf{H}_N\| = \|\mathbf{H}'_N\| = 1$  :

$$\begin{aligned} \widehat{\mathbf{H}}'_N &= \arg \max_{\|\mathbf{H}'_N\|=1} \mathcal{T}_K(\mathbf{H}'_N)^H \mathbf{P}_F \mathcal{T}_K(\mathbf{H}'_N) \\ &= \arg \max_{\|\mathbf{H}'_N\|=1} \mathbf{H}'_N{}^t \mathcal{F} \mathbf{H}'_N{}^{Ht} \end{aligned}$$

où  $\mathcal{F}$  se déduit de  $\mathbf{P}_F$ , L'algorithme est alors identique à celui décrit dans la section 5.4.

Les avantages principaux de cette procédure sont décrits en section 5.4, auxquels on ajoutera qu'il ne faut pas calculer la puissance du bruit, ce qui simplifie encore l'algorithme.

## 9.5 Prédiction Linéaire

Une manière relativement simple d'exploiter les corrélations cycliques est de réexprimer le problème de la prédiction linéaire du chapitre 6 en utilisant, dans l'expression du critère d'orthogonalité (6.3), les corrélations cycliques, ce qui mène à :

$$\mathbf{P}_{L_m} R_{\tilde{\mathbf{y}}\tilde{\mathbf{y}}}^{\{k\}} = \begin{bmatrix} 0 \cdots 0 & \sigma_{\tilde{\mathbf{y}}, L_m}^{2\{k\}} \end{bmatrix}. \quad (9.7)$$

où

$$\begin{aligned} \sigma_{\tilde{\mathbf{y}}, L_m}^{2\{k\}} &\triangleq \mathbf{E}^k \{ \tilde{\mathbf{y}}(n) \tilde{\mathbf{y}}^H(n) \} \\ &= \mathbf{P}_{L_m} \mathbf{E}^k \{ \mathbf{Y}_{L_m+1}(n) \mathbf{Y}_{L_m+1}^H(n) \} \mathbf{P}_{L_m}^H \\ &= \bar{\mathbf{P}}_{L_m} R_{\tilde{\mathbf{y}}\tilde{\mathbf{y}}}^{\{k\}} \bar{\mathbf{P}}_{L_m}^H \end{aligned}$$

L'expression de l'erreur de prédiction devient :

$$\tilde{\mathbf{y}}(n) = \mathbf{H}(0) \mathbf{u}(n) + (\bar{\mathbf{H}}_N - \mathbf{P}_{L_m} \mathcal{T}(\mathbf{H}_N)) \mathbf{U}_{L_m+N-1}(n-1) \quad (9.8)$$

où  $\bar{\mathbf{H}}_N = [\mathbf{H}(1) \cdots \mathbf{H}(N) 0 \cdots 0]$  et  $\mathbf{U}_{L_m+N-1}(n) = [\mathbf{u}(n)^H \cdots \mathbf{u}(n-L_m-N+1)^H]^H$ .

La cyclovariance de l'erreur s'écrit alors :

$$\mathbf{E}^k \{ \tilde{\mathbf{y}}(n) \tilde{\mathbf{y}}^H(n) \} = \frac{\sigma_a^2}{m} \mathbf{H}(0) \mathbf{H}^H(0) + \frac{\sigma_a^2}{m} \bar{\bar{\mathbf{H}}}_N \bar{\bar{\mathbf{H}}}_N^H$$

où  $\bar{\bar{\mathbf{H}}}_N = \bar{\mathbf{H}}_N + \mathbf{P}_{L_m} \mathcal{T}(\mathbf{H}_N)$ . En minimisant cette cyclovariance par rapport à  $\mathbf{P}_{L_m}$ , on débouche sur  $\bar{\bar{\mathbf{H}}}_N = 0$ .

Donc, en utilisant le prédicteur linéaire basé sur une cyclocorrélation cyclique, nous pouvons déterminer un égaliseur ZF et identifier le canal d'une manière similaire à ce qui a été fait dans le chapitre 6, avec un avantage important : le bruit ne doit pas être estimé et sa couleur n'a a priori aucune importance. L'inconvénient majeur est que ces méthodes ne peuvent fonctionner correctement que si la largeur de bande est substantiellement plus grande que la largeur strictement nécessaire.

## 9.6 Modèle stationnaire multivarié

### 9.6.1 Théorème de Gladyshev et Processus de Miamee

Le théorème de Gladyshev [37] affirme que :

**Théorème 9.1**  $R_{yy}(n, \tau)$  est la fonction de corrélation d'une séquence périodiquement corrélée si et seulement si la fonction matricielle :

$$\underline{\mathbf{R}}(\tau) = \left[ R_{yy}^{\{kk'\}}(\tau) \right]_{k,k'=0}^{m-1} \quad (9.9)$$

$$\text{où } R_{yy}^{\{kk'\}}(\tau) = R_{yy}^{\{k-k'\}}(\tau) e^{2\pi j k \tau / m} \quad (9.10)$$

est la fonction matricielle de corrélation d'une séquence stationnaire  $m$ -variée.

En tenant compte de la propriété classique des cyclocorrélations :  $R_{yy}^{\{k\}}(\tau) = R_{yy}^{\{m-k\}H}(-\tau)$ , on peut écrire la matrice suivante

$$\underline{\mathbf{R}} \triangleq \begin{bmatrix} \underline{\mathbf{R}}(0) & \underline{\mathbf{R}}(1) & \cdots & \underline{\mathbf{R}}(K-1) \\ \underline{\mathbf{R}}(-1) & \underline{\mathbf{R}}(0) & \cdots & \underline{\mathbf{R}}(K-2) \\ \vdots & \vdots & \ddots & \vdots \\ \underline{\mathbf{R}}(1-K) & \underline{\mathbf{R}}(2-K) & \cdots & \underline{\mathbf{R}}(0) \end{bmatrix} \quad (9.11)$$

qui est une matrice bloc Toeplitz de  $K \times K$  blocs de taille  $Mm \times Mm$ .

De Míamee [54] on déduit une expression du processus stationnaire multivarié associé :

$$\underline{\mathbf{Z}}_n = \left[ \mathbf{Z}_n^k \right]_{k=0}^{m-1} \quad \text{où } \mathbf{Z}_n^k = \bigoplus_{j=0}^{m-1} \mathbf{y}(n+j) e^{2\pi j k(n+j)/m} \quad (9.12)$$

où  $\oplus$  est la somme directe, i.e., avec  $w = e^{2\pi j/m}$

$$\mathbf{Z}_n^k = w^{kn} [\mathbf{y}(n), \mathbf{y}(n+1)w^k, \dots, \mathbf{y}(n+m-1)w^{k(m-1)}] \quad (9.13)$$

est défini dans un espace de Hilbert, où la corrélation est définie par le produit euclidien suivant :

$$\langle \mathbf{Z}_n^k, \mathbf{Z}_{n+l}^{k'} \rangle = \sum_{j=0}^{m-1} \mathbf{E} \left\{ \mathbf{Z}_n^k(j) \mathbf{Z}_{n+l}^{k'}{}^H(j) \right\} \quad (9.14)$$

et  $\underline{\mathbf{Z}}_n = [\mathbf{Z}_n^0{}^T \mathbf{Z}_n^1{}^T \cdots \mathbf{Z}_n^{m-1}{}^T]^T$  avec la définition classique de la corrélation pour les processus stationnaires multivariés.

## 9.7 Expression de la sortie du nouveau modèle

De  $\mathbf{Z}_n^k = \bigoplus_{j=0}^{m-1} \mathbf{y}(n+j) e^{2\pi j k(n+j)/m}$  et

$$\begin{aligned} \mathbf{y}(n+j) &= \sum_{k=0}^{L-1} \mathbf{H}(k) \mathbf{u}(n+j-k) + \mathbf{v}(n+j) \\ &= \mathbf{H}_N \begin{bmatrix} \mathbf{u}(n+j) \\ \mathbf{u}(n+j+1) \\ \vdots \\ \mathbf{u}(n+j-N+1) \end{bmatrix} + \mathbf{v}(n+j) \end{aligned} \quad (9.15)$$

Définissons  $\mathbf{U}_{n+j} = [\mathbf{u}(n+j)^H \cdots \mathbf{u}(n+j-N+1)^H]^H$  et  $\mathbf{H}_N^{\{k\}} = [w^{-kj} \mathbf{H}(j)]_{j=0}^{N-1}$ . Le processus de Miamee s'exprime alors :

$$\begin{aligned} \mathbf{Z}_n^k &= \bigoplus_{j=0}^{m-1} (\mathbf{H}_N^{\{-k\}} w^{kn} \mathbf{u}_{n+j} + \mathbf{v}(n+j) e^{2\pi j k \frac{n+j}{m}}) \\ &= \mathbf{H}_N^{\{-k\}} w^{kn} [\mathbf{u}_n \mathbf{u}_{n+1} \cdots \mathbf{u}_{n+m-1}] + \bigoplus_{j=0}^{m-1} \mathbf{v}(n+j) e^{2\pi j k \frac{n+j}{m}} \end{aligned} \quad (9.16)$$

$$\Rightarrow \underline{\mathbf{Z}}_n = \mathbf{H}_{tot} \mathcal{U}(n) + \mathcal{V}(n) \quad (9.17)$$

où  $\mathbf{H}_{tot} = [\mathbf{H}_N^{\{0\}H} \mathbf{H}_N^{\{-1\}H} \cdots \mathbf{H}_N^{\{1-m\}H}]^H$ ,

$$\mathcal{U}(n) = D_{DFT}^{\{n, pN\}} [\mathbf{u}_n \mathbf{u}_{n+1} \cdots \mathbf{u}_{n+m-1}] \quad (9.18)$$

$$\mathcal{V}(n) = \begin{bmatrix} \mathbf{v}(n) & \cdots & \mathbf{v}(n+m-1) \\ \mathbf{v}(n)w^n & \cdots & \mathbf{v}(n+m-1)w^{n+m-1} \\ \vdots & \ddots & \vdots \\ \mathbf{v}(n)w^{n(m-1)} & \cdots & \mathbf{v}(n+m-1)w^{(m-1)(n+m-1)} \end{bmatrix} \quad (9.19)$$

$$\Rightarrow \underline{\mathbf{Z}} = \mathcal{T}_{L+N-1}(\mathbf{H}_{tot}) \mathcal{U}_L + \mathcal{V}_L \quad (9.20)$$

où  $\mathcal{U}_L = [\mathcal{U}(n)]_{n=L-1}^0$  est un processus stationnaire dont la matrice de corrélation peut être aisément calculée à partir de  $\mathbf{R}_{aa}$ .

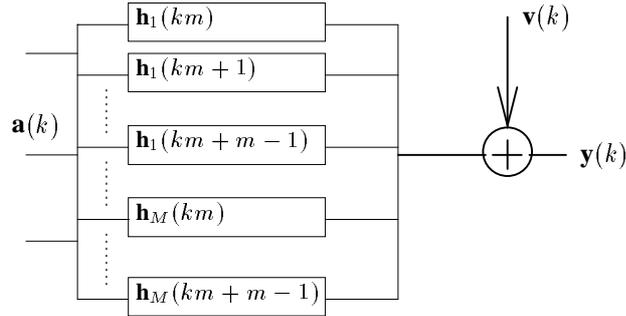


FIG. 9.2 – Modèle de canal classique

## 9.8 Méthodes basées sur le modèle stationnaire multivarié

En se basant sur la relation (9.20), nous pouvons appliquer directement les méthodes sous-espace et prédiction linéaire vues aux chapitres 5 et 6.

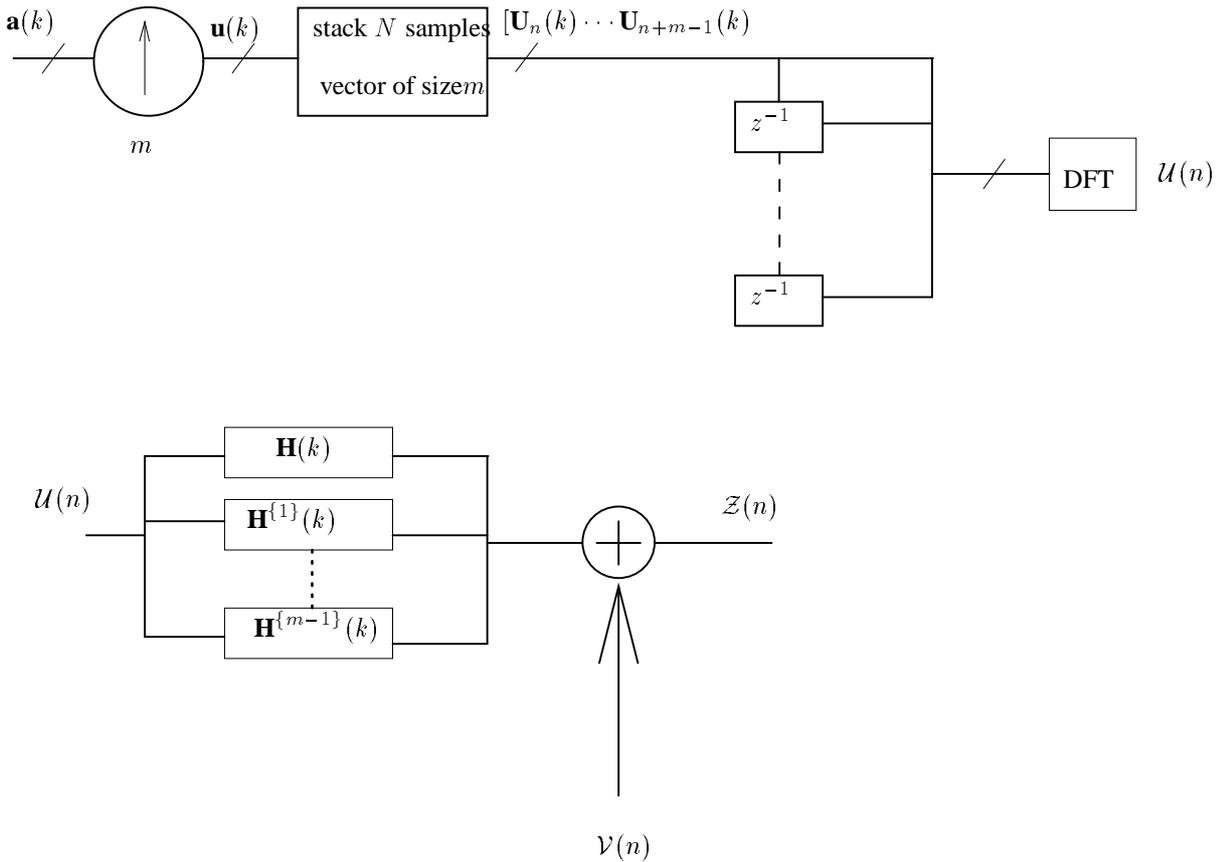


FIG. 9.3 – Nouveau modèle de canal

## 9.9 Identifiabilité

Les conditions d'identifiabilité sont les mêmes que pour les méthodes utilisant les statistiques classiques, en effet, le nouveau modèle se base simplement sur des canaux déphasés chacune d'un multiple de  $\frac{2\pi}{m}$ , et donc, la condition sur les zéros des sous-canaux n'est pas modifiée. Pour les conditions sur le nombre de données et les modes, le nouveau modèle de canal ne les modifie pas.

## 9.10 Charge de calcul et aspects algorithmiques

Il est clair que la matrice de corrélation  $\underline{R}$  construite à partir des corrélations cycliques est plus volumineuse (chaque scalaire dans  $R$  est remplacé par un bloc de taille  $m_e \times m_e$  dans  $\underline{R}$ ) que la matrice de corrélation construite dans le modèle classique obtenu par vectorisation des signaux sur-échantillonnés. Par contre, la matrice  $\underline{R}$  est plus structurée. En fait, on peut

montrer que les estimées  $\hat{\mathbf{H}}_N^{\{-k\}}$  sont strictement reliées (i.e.  $\hat{\mathbf{H}}_N^{\{-k\}} = [w^{-kj}\hat{\mathbf{h}}(j)]_{j=0}^{N-1}$  pour tout  $k$ ), ce qui indique que l'utilisation de cette structure devrait mener à des algorithmes de complexité réduite par rapport à l'application brute des algorithmes originaux. L'examen détaillé des expressions du prédicteur linéaire (en se basant sur [55]) le montre de façon évidente. D'autre part, comme indiqué dans [69], les problèmes de prédiction linéaire multicanal correspondent à une factorisation bloc triangulaire de  $\underline{R}$ , équivalente à une orthogonalisation des composants blocs du vecteur  $\underline{Z}$ . Si cette orthogonalisation se fait, dans le cas du modèle classique, sur les composants scalaires de  $\mathbf{y} = [\mathbf{y}(n) \cdots \mathbf{y}(n+K)]$ , on débouche sur une factorisation triangulaire vraie (non bloc) dont les éléments, d'une manière similaire à ce qui été fait dans le chapitre 6, sont les éléments du prédicteur cyclique. L'exploitation de ce résultat permet effectivement d'utiliser un algorithme rapide pour déterminer le prédicteur linéaire cyclique.

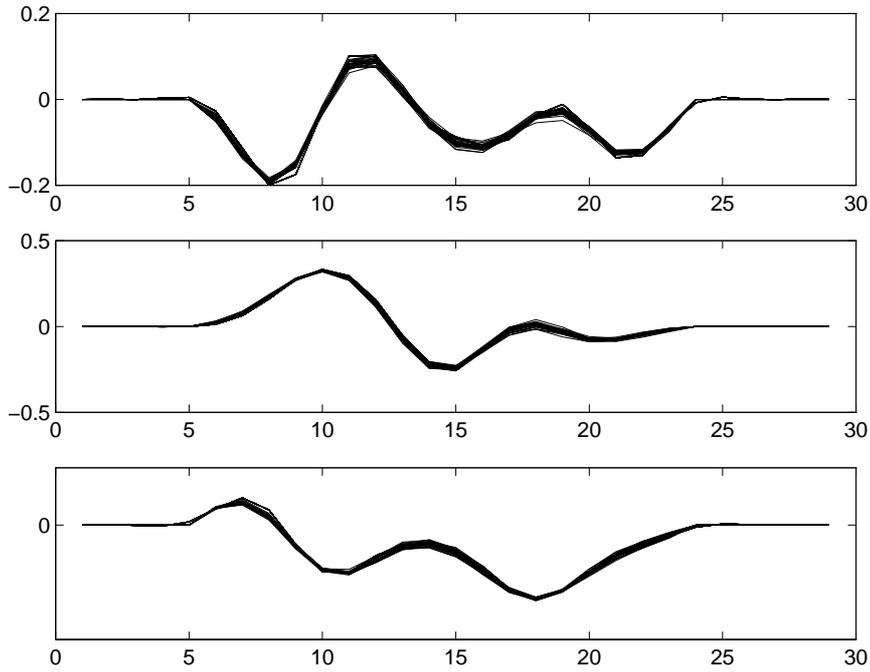
## 9.11 Simulations

Pour ces simulations, nous utilisons (dans un cas mono-utilisateur), un canal à coefficients réels, générés aléatoirement et de longueur  $5T$ , précédé d'un filtre en cosinus surélevé avec un excès de largeur de bande de 90 %.

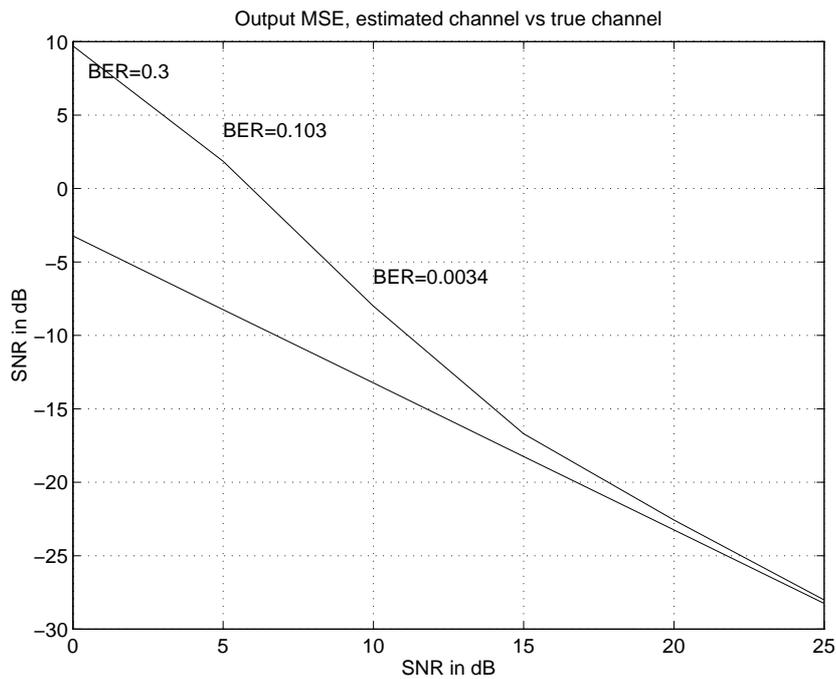
Le récepteur est composé de  $m_a = 3$  antennes et nous sur-échantillons par un facteur  $m_e = 3$ . Nous avons calculé le NRMSE du canal. Nous utilisons une trame de 100 symboles QAM-4 (il est à noter que si nous utilisions des sources réelles, nous pourrions utiliser la corrélation cyclique conjuguée, ce qui, en présence d'un bruit circulaire, permet également de s'affranchir de l'influence du bruit). Nous avons effectué 50 tirs Monte-Carlo.

### 9.11.1 Modèle cyclostationnaire et méthode sous-espace

L'estimation est basée sur la méthode sous-espace (non pondérée) à partir de la matrice de corrélation de fréquence cyclique  $k = 1$ . L'estimation de la longueur du canal sur-échantillonné de 18 (les 5 premiers et 6 derniers coefficients du canal étant considérés nuls). L'estimation du canal à 20 dB est donnée ci-dessous.



Nous avons également effectué une égalisation MMSE-ZF basée sur cette estimée du canal, ce qui nous donne le MSE à la sortie et une estimée du taux d'erreur (pour des SNR de 0, 5 et 10 dB), en utilisant 100000 symboles QAM-4 à chaque tir Monte-Carlo (i.e.  $10^7$  bits pour chaque SNR).

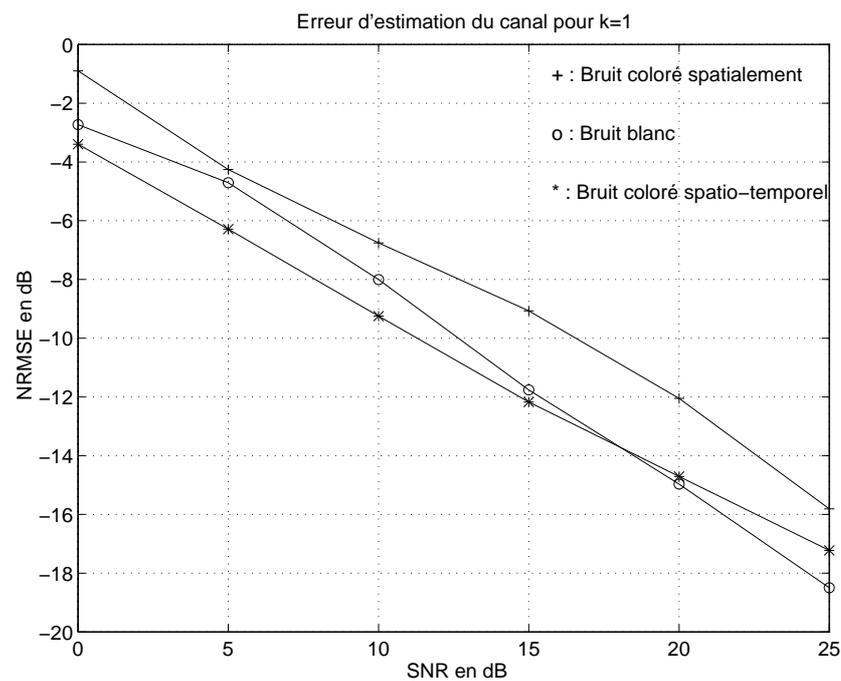


D'autre part, nous avons simulé deux type des bruit coloré. Dans le premier scénario, nous avons un bruit spatialement coloré de matrice de corrélation spatiale :

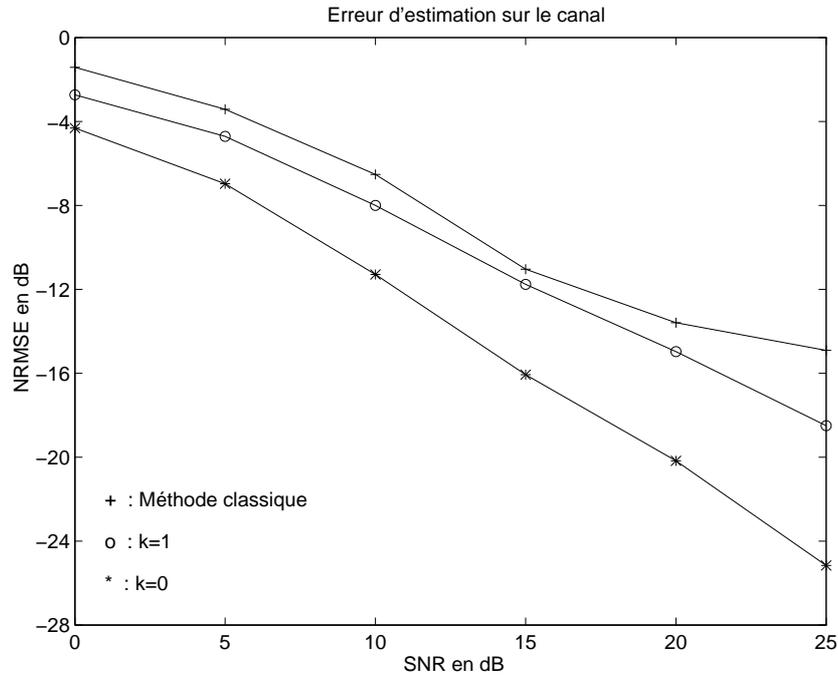
$$\begin{pmatrix} 1 & .7 & .49 \\ .7 & 1 & .7 \\ .49 & .7 & 1 \end{pmatrix}$$

Dans le second scénario, nous avons un bruit coloré spatialement et temporellement par

$$h_n = \begin{pmatrix} .346 & -.180 & .057 & -.057 & -.365 & -.358 \\ .298 & .068 & -.231 & .053 & .090 & -.229 \\ .387 & -.100 & .413 & .172 & .032 & .079 \end{pmatrix}$$

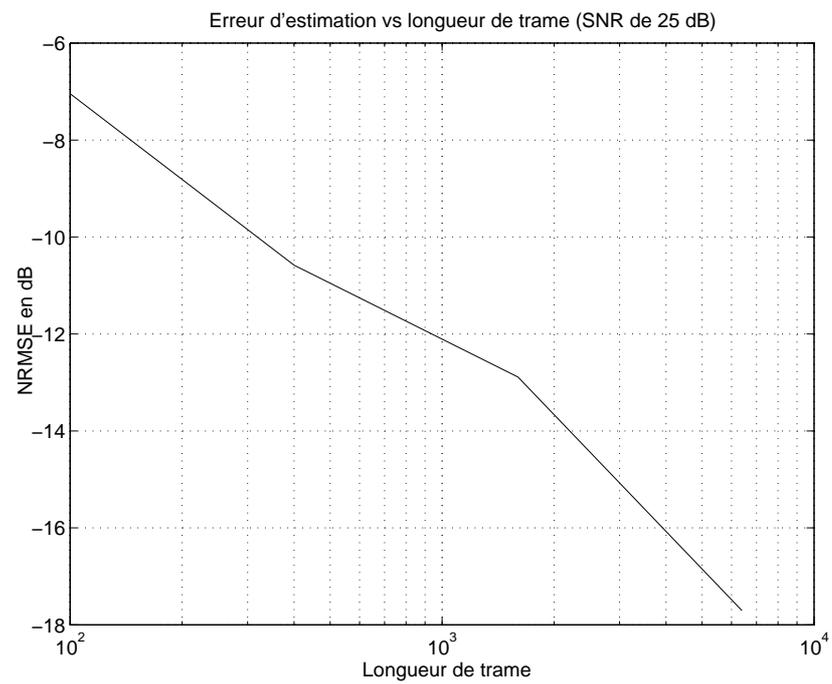


De manière à pouvoir comparer les performances, nous avons également effectué les simulations d'une part pour une fréquence cyclique  $k = 0$  et d'autre part pour l'algorithme basé sur les statistiques classiques.



### 9.11.2 Modèle cyclostationnaire et méthode de la prédiction linéaire

Les simulations mènent à des résultats médiocres (pour  $k = 1$ ), en ce sens qu'il faut des trames de longueur importante pour avoir une estimée correcte du canal. Pour illustration, nous montrons l'évolution de l'erreur d'estimation pour une trame de longueur grandissante sur un canal généré aléatoirement, avec  $m_a = 3$ ,  $m_e = 3$  et un SNR de 25 dB.



### 9.11.3 Modèle stationnaire multivarié et méthode sous-espace

Les estimées basées sur 25 réalisations, pour un SNR de 20 dB, sont reproduites ci-dessous.

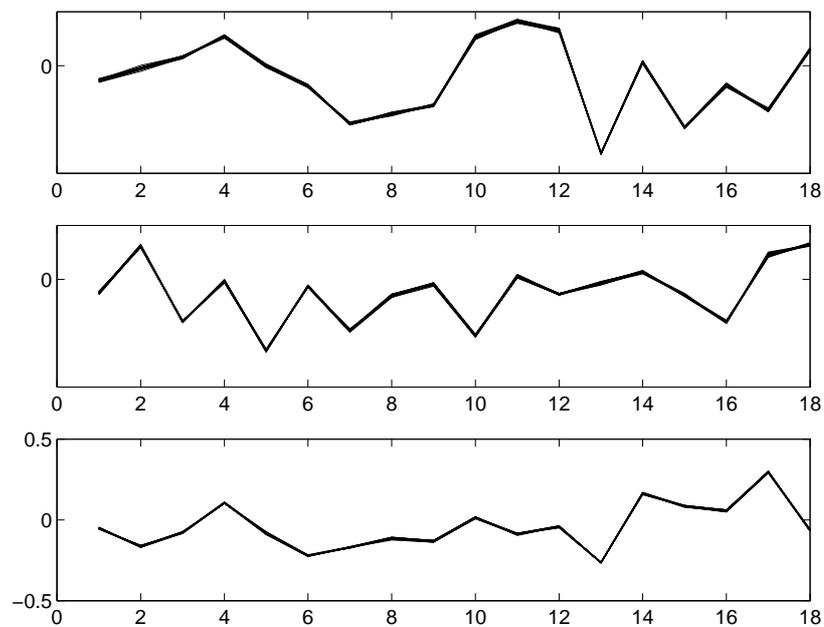


FIG. 9.4 – Estimée du canal à 20 dB

Nous comparons les résultats avec l’algorithme basé sur le modèle classique. Les résultats semblent favoriser l’approche classique, parce que nous avons pris les mêmes tailles de matrice dans les deux cas, ce qui génère un sous-espace bruit de dimension réduite pour l’approche cyclique. En théorie (en raison du lien biunivoque entre les statistiques cycliques et classiques), les performances, en utilisant des tailles de matrices adéquates, doivent être identiques, ce qui est illustré par la troisième courbe.

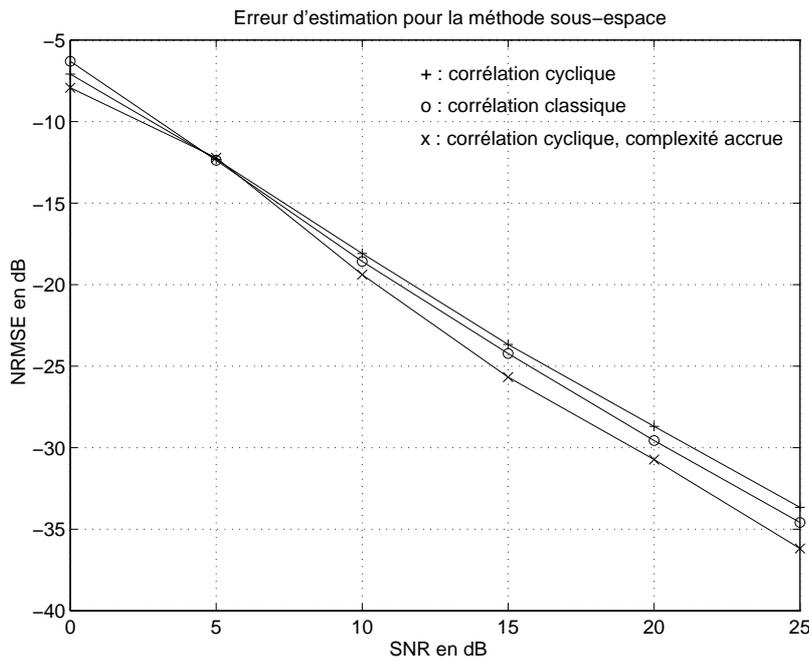


FIG. 9.5 – Erreur d’estimation du canal pour la méthode sous-espace

Dans le cas où le canal comprend le filtre de mise en forme à l’émetteur et le filtre de réception, l’approche cyclique peut donner de meilleurs résultats. Ce gain est dû principalement à l’affinement de la longueur du canal (à condition d’avoir une bonne estimée de la longueur du canal ... !). En effet, si on a un canal de la forme ( $m_e = m_a = 2$ )

$$\begin{pmatrix} \epsilon & * & * & * & * & * & \epsilon \\ \epsilon & * & * & * & * & * & \epsilon \end{pmatrix} \quad (9.21)$$

où  $\epsilon$  est une valeur quasi-nulle, l’approche cyclique va estimer la partie centrale du canal, tandis

que l'approche classique essaiera d'estimer le multicanal formé de  $m_e * m_a$  canaux élémentaires :

$$\begin{pmatrix} \epsilon & * & * \\ * & * & * \\ \epsilon & * & * \\ * & * & * \end{pmatrix} \quad \text{ou} \quad \begin{pmatrix} * & * & * \\ * & * & \epsilon \\ * & * & * \\ * & * & \epsilon \end{pmatrix} \quad (9.22)$$

avec deux paramètres supplémentaires à estimer, ce qui donnera une estimée globalement moins bonne. Les figures suivantes illustrent les performances pour un canal qui incorpore les filtres d'émission/réception.

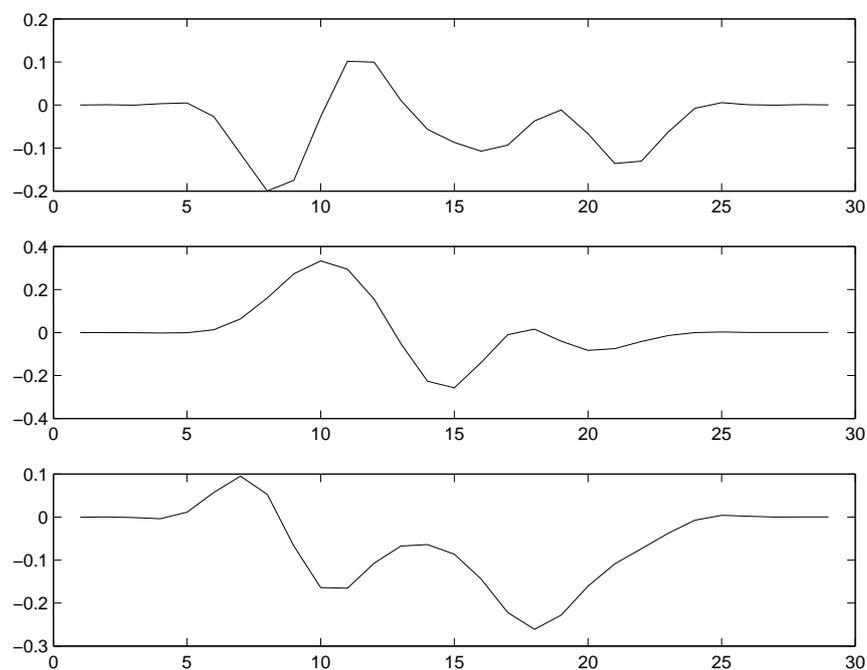


FIG. 9.6 – Canal + filtre d'émission/réception

#### 9.11.4 Modèle stationnaire multivarié et méthode de la prédiction linéaire

Dans le cas de la prédiction linéaire, l'influence du bruit est plus importante que pour la méthode sous-espace, en ceci que, comme on l'a vu dans le chapitre 6, l'estimation de ce bruit et le travail sur le débruitage de la matrice de corrélation joue un rôle prépondérant. Dans ce cadre, dans la mesure où la matrice  $\underline{R}$  comprend un nombre beaucoup plus important de paramètres insensibles (asymptotiquement) à l'influence du bruit, nous nous attendons à une performance qui devrait être légèrement meilleure que pour les algorithmes classiques.

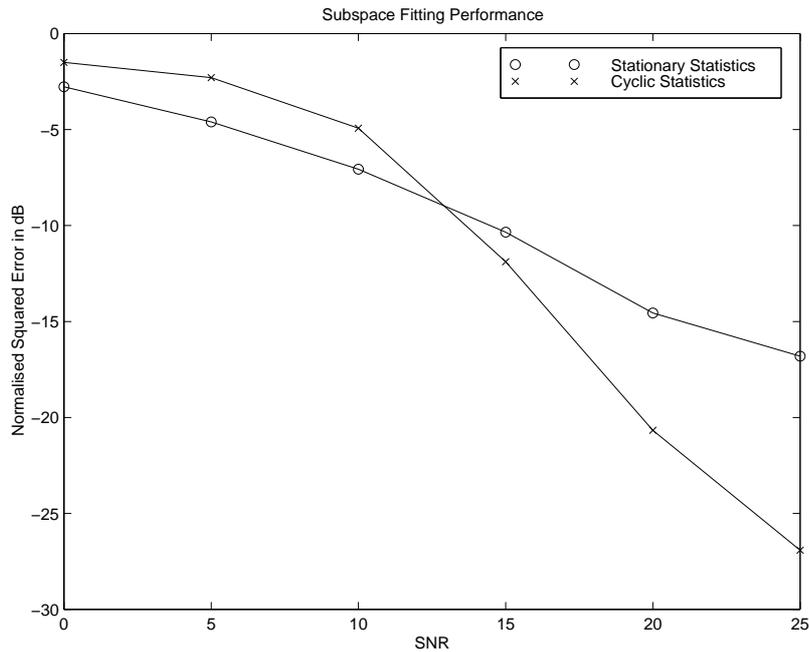


FIG. 9.7 – Erreur d'estimation du canal global par la méthode sous-espace

Ce raisonnement est confirmé par les simulations où, en ne débruitant pas  $\underline{R}$ , on obtient cependant des résultats corrects.

## 9.12 Conclusions

Nous avons, dans un premier temps, proposé deux méthodes d'identification de canal basées sur les statistiques cycliques à une fréquence cyclique non nulle, qui les rendent indépendantes de la couleur du bruit additif, pour peu que ce bruit ne soit pas lui-même cyclostationnaire de même fréquence cyclique que le signal utile.

L'approche sous-espace donne des performances correctes, mais le bénéfice dû à l'indépendance par rapport à la couleur du bruit est largement contrebalancé par la perte d'information due à la non-utilisation de la corrélation de fréquence cyclique nulle (pour un canal à bande limitée).

Par contre, à la fréquence cyclique nulle, l'affinement de la longueur du canal due à l'utilisation du signal sur-échantillonné non vectorisé permet un gain en performance, dans la mesure où l'estimation de la longueur du canal est correcte.

L'approche issue de la prédiction linéaire ne donne par contre pas de résultats probants.

Dans un second temps, nous avons utilisé la représentation par un processus stationnaire multivarié introduite par [37] et [54, 55]. Nous avons dérivé les expressions explicites de ce

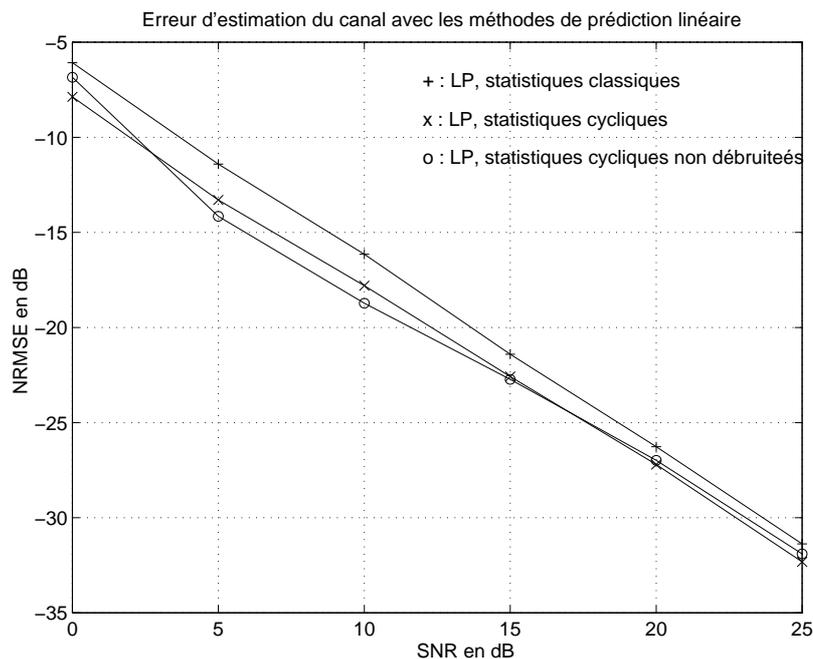


FIG. 9.8 – Erreur d'estimation du canal pour la méthode de la prédiction linéaire

processus, qui peut être vu comme étant la sortie d'un système comprenant le canal  $\mathbf{H}_{tot} = [\mathbf{H}_N^{\{0\}T} \mathbf{H}_N^{\{-1\}T} \dots \mathbf{H}_N^{\{1-m\}T}]^T$  et dont l'expression des entrées et sorties est aisément déduite de l'expression des grandeurs correspondantes dans le modèle classique. Nous avons alors appliqué à ce modèle les algorithmes d'identification aveugle de canal basés sur les approches sous-espace et prédiction linéaire.

Pour l'approche sous-espace, nous obtenons essentiellement les mêmes performances que pour le cas classique, si ce n'est un petit gain dû à un affinement de la durée du canal. Le désavantage principal de cette approche est un coût de calcul plus élevé.

Pour la prédiction linéaire, la prise en compte du passé le plus proche fournit une amélioration des performances. De plus, l'utilisation d'algorithmes de prédiction linéaire multicanal modulaire (tels que l'extension au cas singulier des algorithmes décrits dans [38]) permet d'utiliser des algorithmes rapides (et, au besoin, des implémentations adaptatives).



# Chapitre 10

## Simulations

### 10.1 Canaux

De manière à comparer les différentes méthodes que nous proposons, non seulement pour illustrer des points précis, comme dans le cours du texte, mais pour comparer les mérites respectifs des algorithmes proposés, nous avons fait un ensemble de simulations Monte-Carlo avec les caractéristiques suivantes :

#### «– Pour les méthodes mono-utilisateur

1. moyennage sur 10 canaux à coefficients réels, générés aléatoirement ;
2. 3 types de canaux :
  - «– 2 sous-canaux obtenus par suréchantillonnage,
  - «– 4 sous-canaux obtenus par 4 éléments d’antenne supposés indépendants,
  - «– 4 sous-canaux obtenus par 2 éléments d’antenne supposés indépendants et suréchantillonnage par 2 ;
3. chaque canal a une version à largeur de bande
  - «– infinie, de longueur de 6 temps symbole,
  - «– filtrée par un filtre à cosinus surélevé avec un excès de bande de 80 % ,
  - «– filtrée par un filtre à cosinus surélevé avec un excès de bande de 40 % ,
  - «– filtrée par un filtre à cosinus surélevé avec un excès de bande de 20 % ;
4. moyennage sur 100 réalisations de la trame de données et du bruit pour chaque réalisation du canal ;
5. une longueur de trame de 100 symboles QPSK ;

6. pour le calcul du taux d'erreur en méthode sous-espace,  $10^5$  symboles par réalisation (soit un taux d'erreurs basé sur  $10^8$  symboles).

«← **Pour les méthodes multi-utilisateurs**

1. 2 utilisateurs ;
2. moyennage sur 100 canaux à coefficients réels, générés aléatoirement ;
3. 2 types de canaux : l'un ayant les mêmes longueur de canal pour les deux utilisateurs (4 temps symbole) et l'autre ayant une longueur de 3 temps symbole pour un utilisateur et de 6 temps symbole pour le second ;
4. une largeur de bande infinie ;
5. 4 sous-canaux obtenus par 4 éléments d'antenne supposés indépendants ;
6. moyennage sur 100 réalisations de la trame de données et du bruit pour chaque réalisation du canal ;
7. une longueur de trame de 250 symboles QPSK.

D'autre part, nous avons toujours utilisé l'implémentation la plus naïve et honnête possible des algorithmes, i.e. sans ajouter de petites astuces qui rendent les algorithmes plus performants dans certaines circonstances.

## 10.2 Méthodes sous-espace

Nous avons estimé les performances des algorithmes sous-espace classique et sans décompositions vecteurs/valeurs propres. Dans les figures ci-dessous, les courbes indexées par un cercle correspondent à la méthode sans décompositions en vecteurs propres, les courbes indexées par un losange à la méthode classique et les courbes en pointillé donnent le taux d'erreurs de symbole pour un égaliseur MMSE-ZF basé sur l'estimation de canal obtenue. Ce taux d'erreur donne une idée de la qualité d'estimation nécessaire pour obtenir une bonne qualité de transmission.

Au niveau des simulations, ces algorithmes n'étant pas robustes à la surdétermination de l'ordre, nous avons, pour les largeurs de bande non infinies, déterminé la longueur du canal qui donnait les meilleures performances à un SNR de 30 dB.

Les enseignements que l'on peut tirer de ces courbes sont que :

- «← Notre méthode donne les mêmes performances que la méthode classique, voire légèrement mieux vers les bas SNRs.

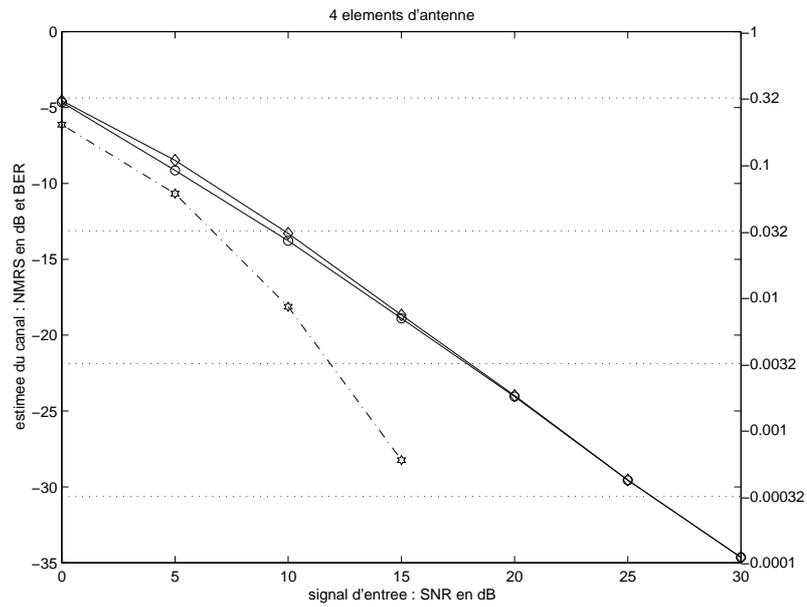


FIG. 10.1 – Performances des méthodes sous-espace pour 4 éléments d'antenne

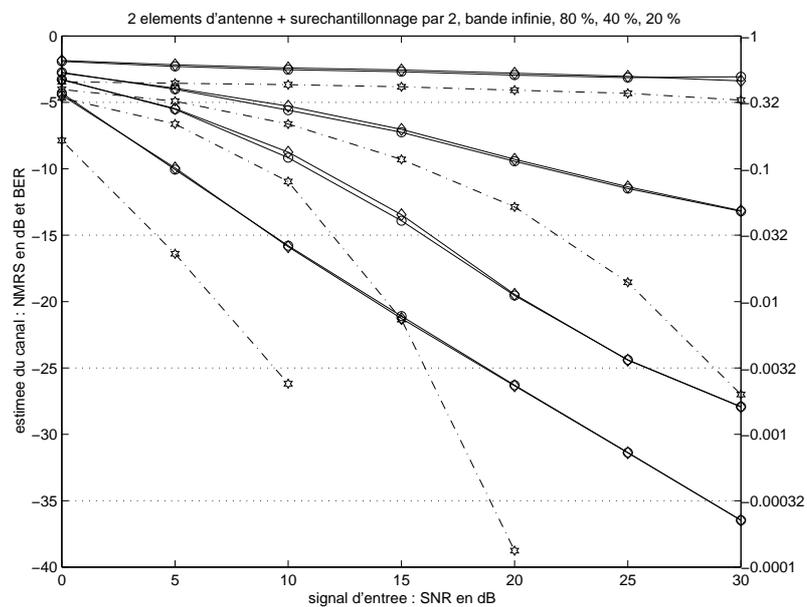


FIG. 10.2 – Performances des méthodes sous-espace pour 2 éléments d'antenne

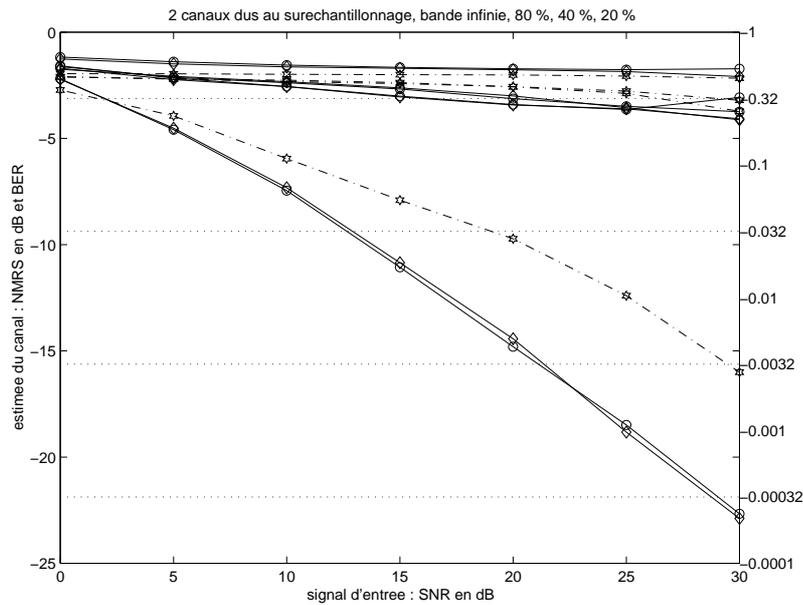


FIG. 10.3 – Performances des méthodes sous-espace pour 1 élément d’antenne

- «– Pour que l’algorithme fonctionne correctement dans les cas de suréchantillonnage, il faut un excès de bande confortable. Si les canaux multiples sont uniquement dus au suréchantillonnage, les performances sont mauvaises, dans le cas de 2 antennes, un excès de bande de 50 % est appréciable. Une des raisons pour lesquelles on obtient des performances médiocres, même avec 2 éléments d’antenne, est la difficulté d’obtenir une estimation de la longueur des canaux, et la présence de beaucoup d’éléments de faible grandeur en début et fin de canal.
- «– Les taux d’erreurs de symbole obtenus par un égaliseur simple augurent de performances acceptables si on utilise cet algorithme dans des cas favorables : multi-canal du à des antennes multiples et/ou excès de bande appréciable.

### 10.3 Prédiction Linéaire

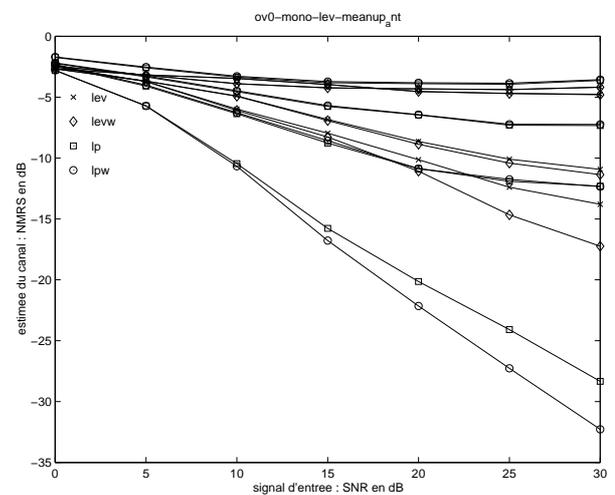
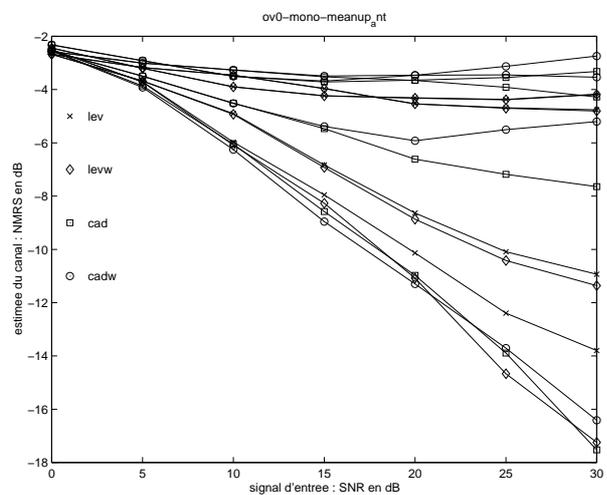
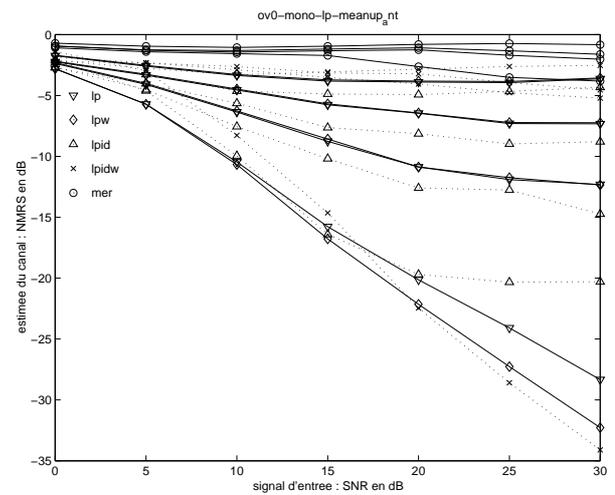
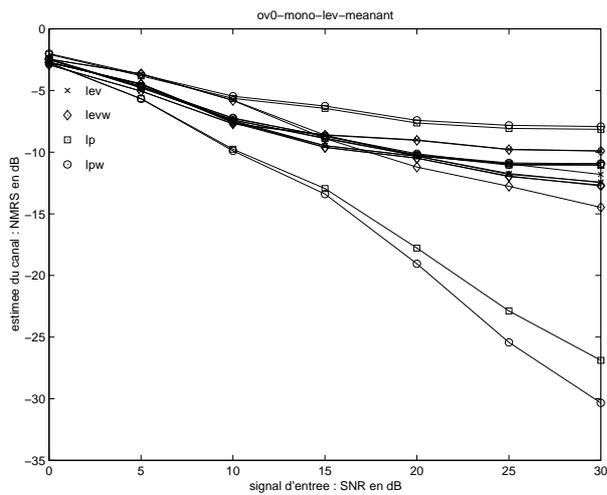
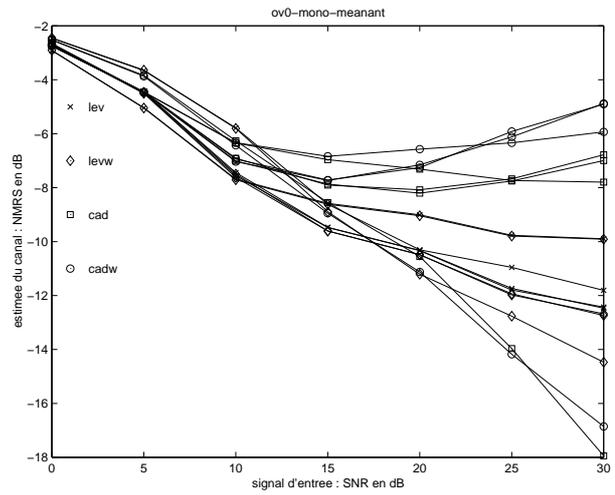
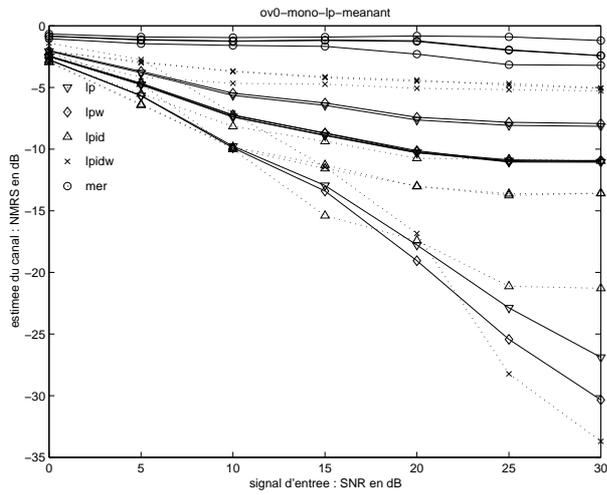
Pour la prédiction linéaire, nous explorons les différents algorithmes abordés et leur robustesse à la surdétermination de la longueur du canal. Les algorithmes simulés sont :

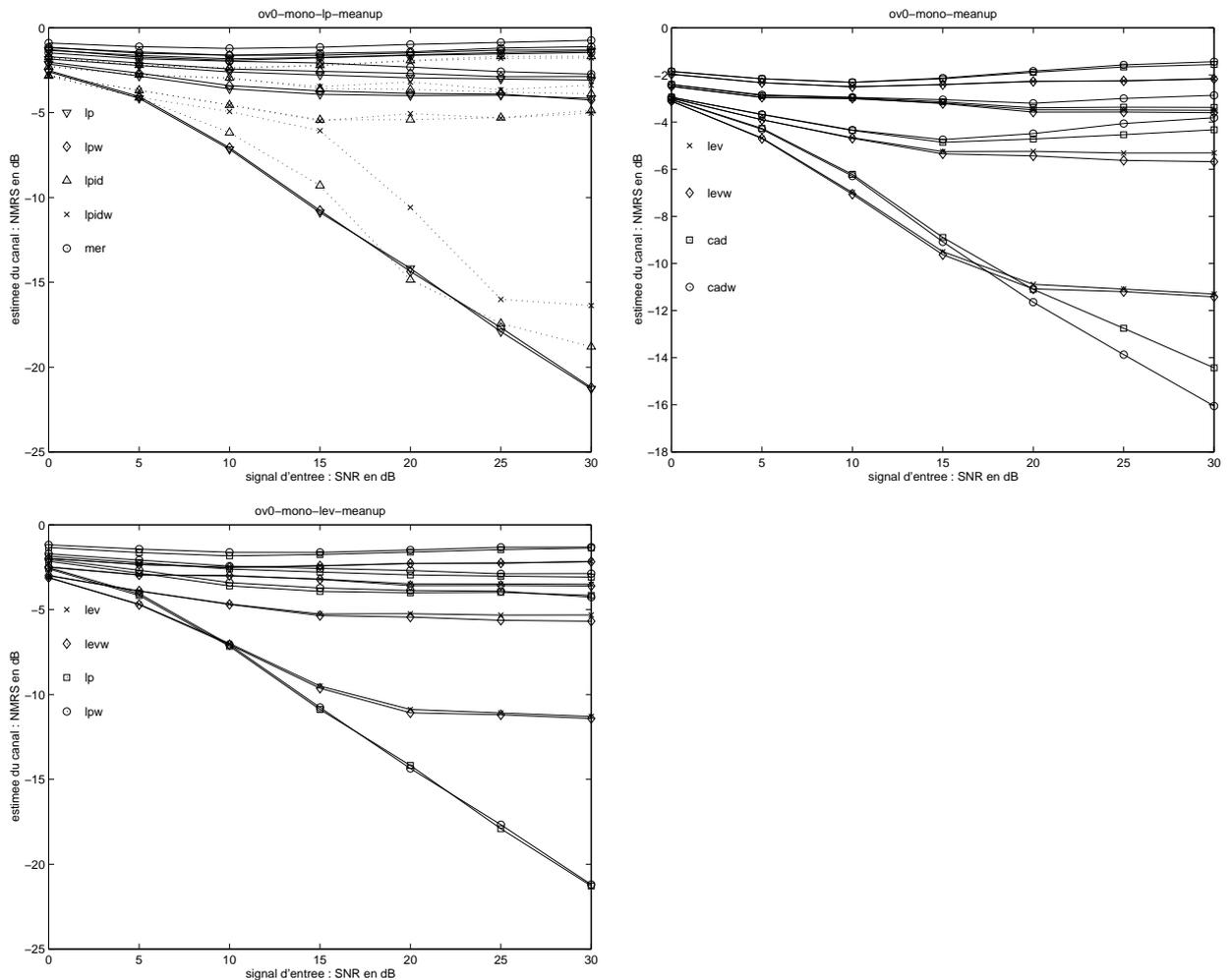
1. L’algorithme original de Abed-Meraim. Ses piètres performances sont dues à une longueur de trame courte pour cet algorithme (notons, pour rassurer le lecteur, que nous avons pu reproduire les courbes de performance de [1]). La légende pour cette méthode est *mer*.

2. La méthode basée sur la détermination du prédicteur par l'algorithme de Levinson multicanal, suivie d'une détermination du canal par une solution d'un problème de moindres carrés et de moindres carrés pondérés pour  $\mathbf{P}(z)\mathbf{H}(z) = \mathbf{h}(0)$ . Notons que l'estimateur de la matrice de corrélation est l'estimateur biaisé, qui a une structure bloc-Toeplitz qui permet l'utilisation de l'algorithme de Levinson. La légende pour cette méthode est `lev[w]`.
  
3. La méthode basée sur la détermination du prédicteur par l'algorithme de Cadzow. La légende pour cette méthode est `cad[w]`.
  
4. La méthode de Gorokhov [4], dans une version "normale", où l'estimateur de la variance du bruit est la valeur propre minimale de la matrice de covariance. La légende pour cette méthode est `lp[w]`.
  
5. La méthode de Gorokhov [4], dans une version "lourde", où l'estimateur de la variance du bruit est l'estimateur au maximum de vraisemblance (moyenne des valeurs propres correspondant à l'espace bruit). Si l'estimateur initial est  $\hat{R}_{YY} = S.D.V$ , où  $[SDV]$  est la décomposition en vecteurs/valeurs propres de  $\hat{R}_{YY}$ , alors on utilise  $\hat{R}_{YY} = S.D_o.V - \hat{\sigma}_v^2 I$ , où  $D_o$  contient les mêmes éléments diagonaux que  $D$ , sauf les éléments correspondant à l'espace bruit qui sont mis à zéro. La légende pour cette méthode est `lpid[w]`.

Les titres des figures contiennent la surestimation de la longueur du canal `ov[0,2,4]`, le type de comparaisons effectuées, `mean` pour rappeler que c'est un moyennage sur 10 canaux, `[ant, up_ant, up]` pour les types de canaux, respectivement à 4 antennes, 2 antennes et suréchantillonnage par 2 et 2 canaux dûs au suréchantillonnage par 2.

10.3.1 Résultats sans surestimation de la longueur du canal

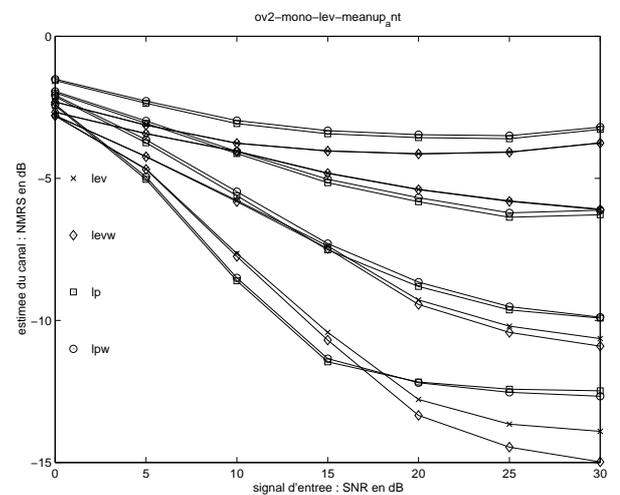
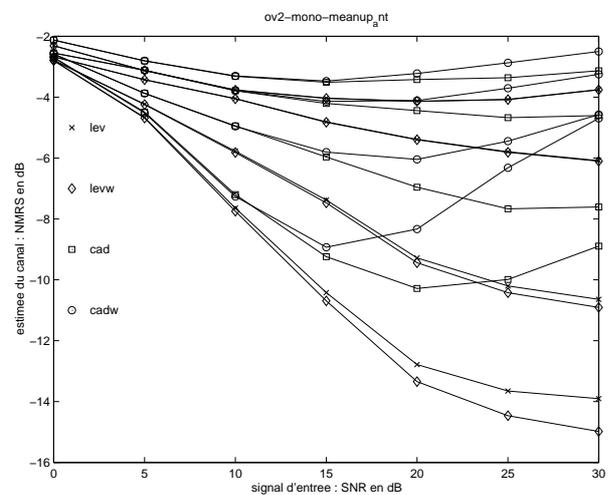
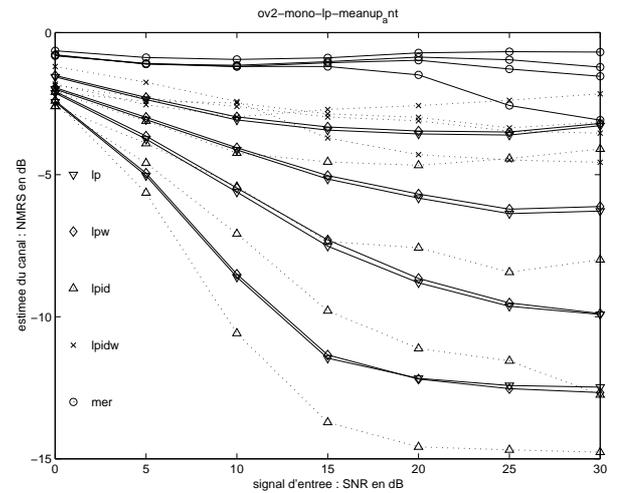
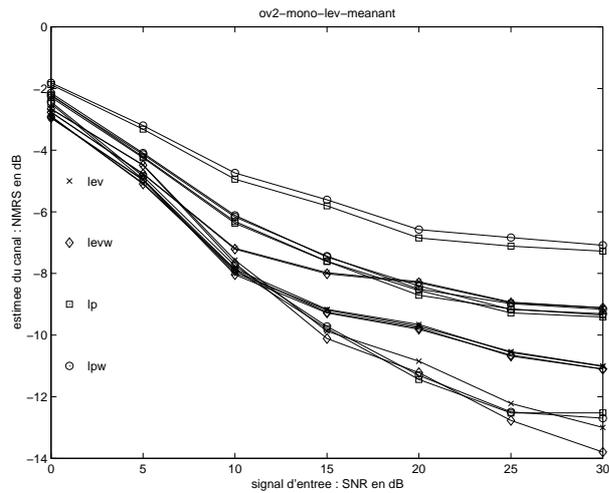
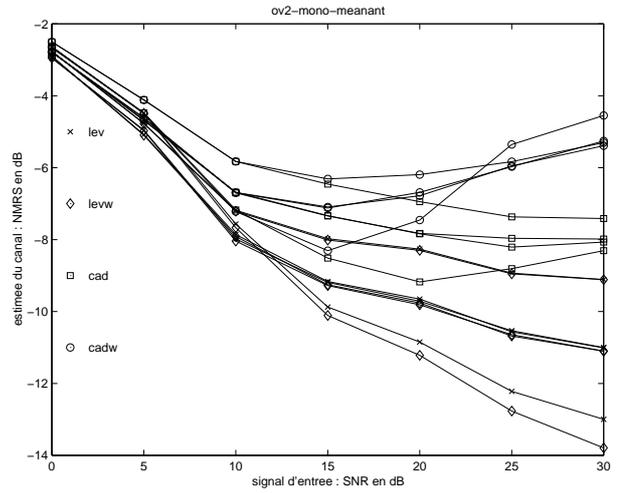
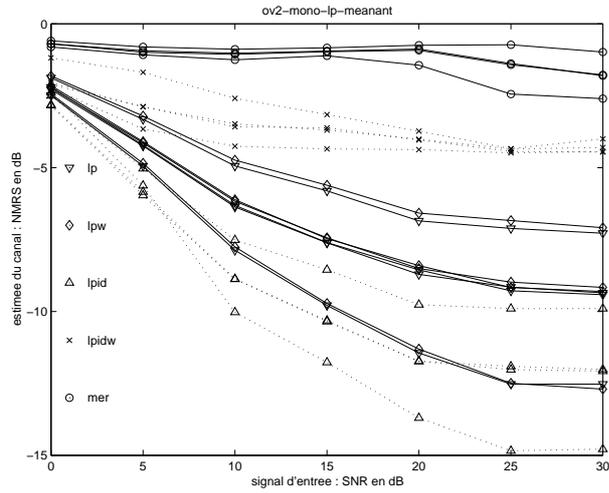


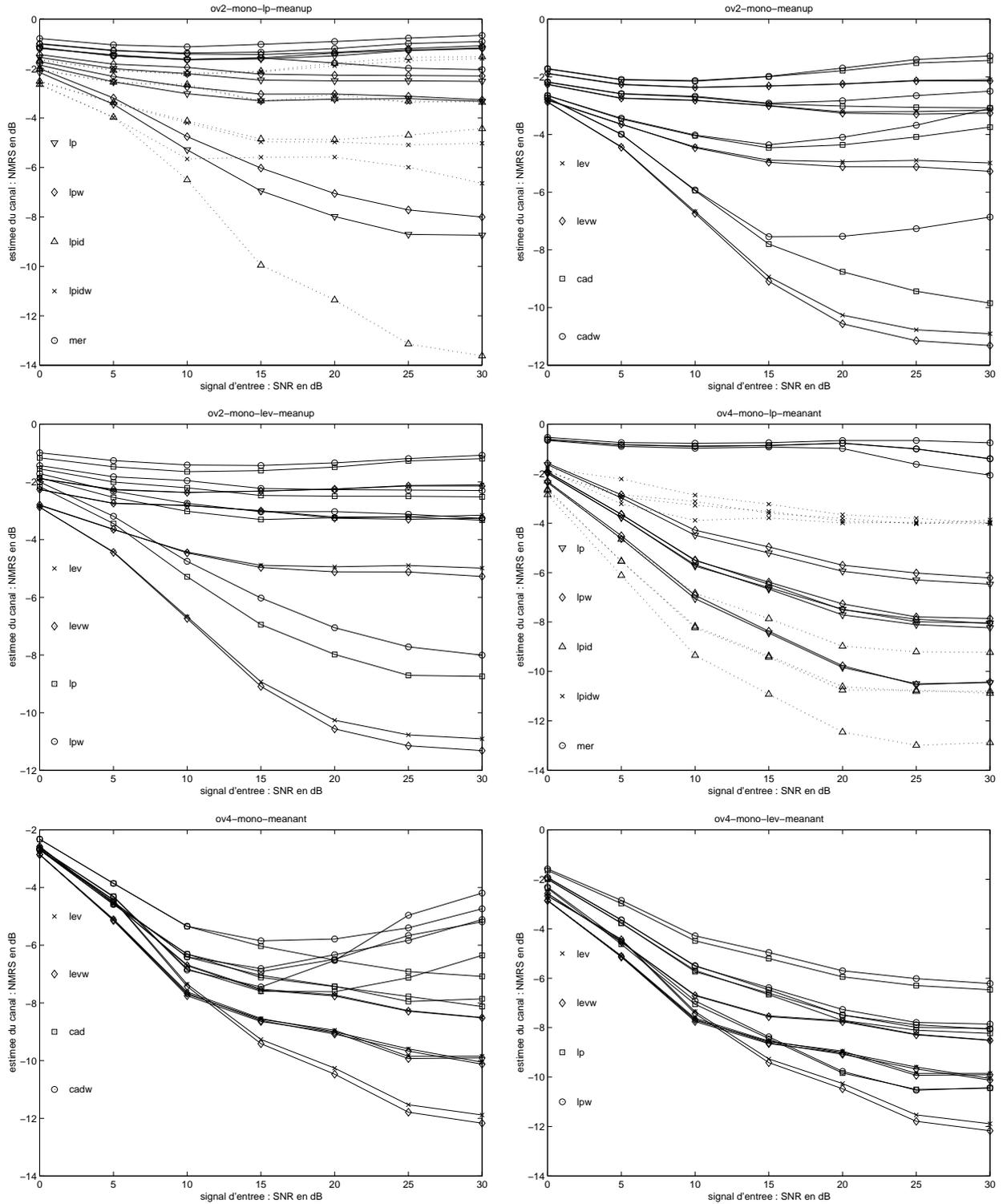


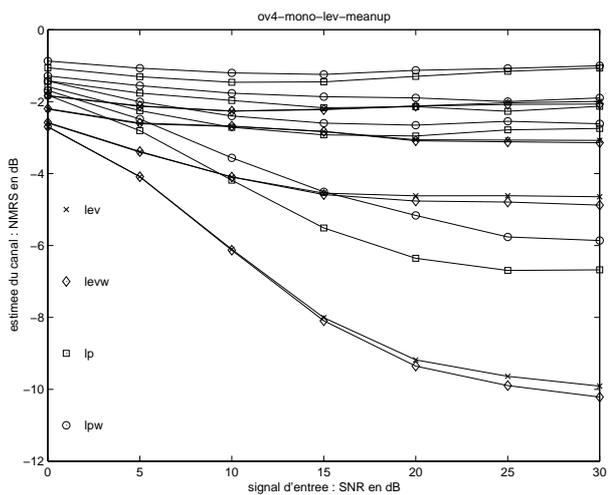
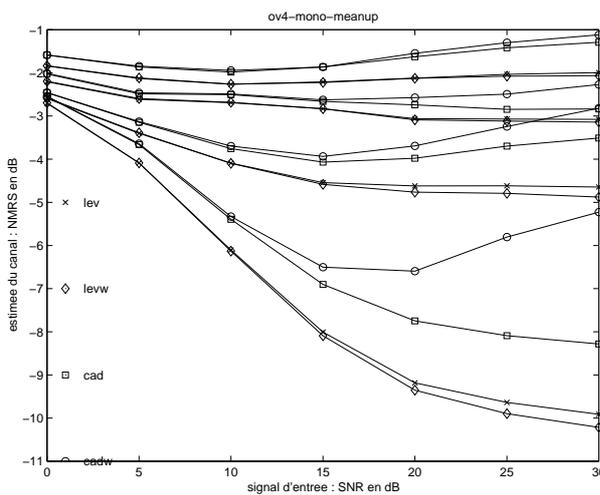
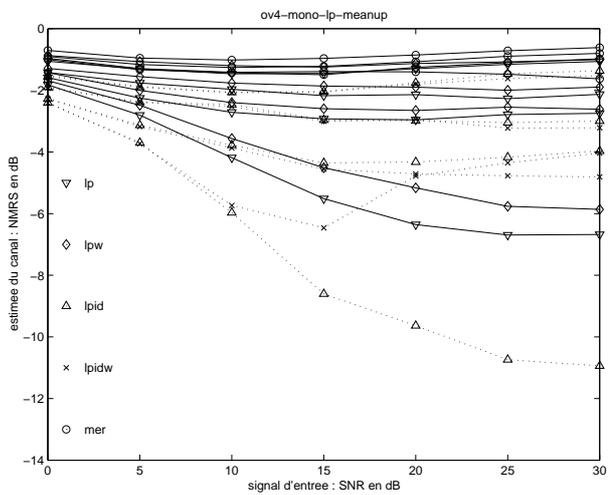
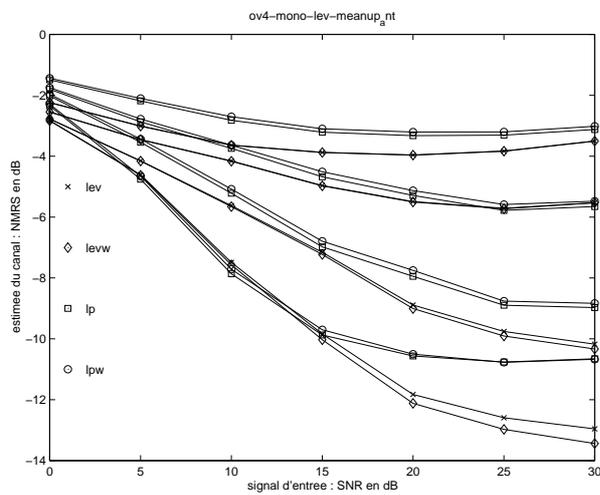
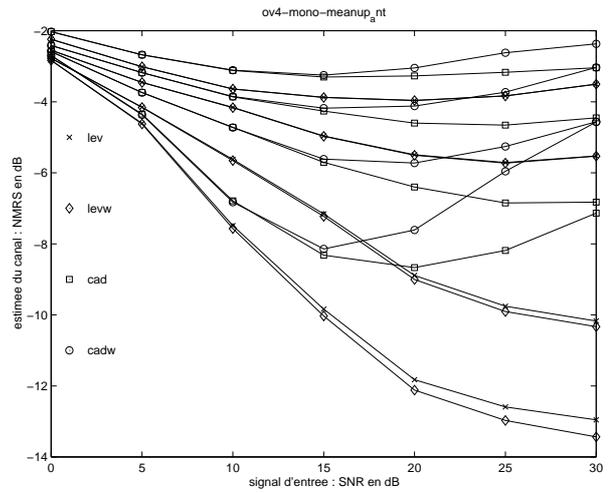
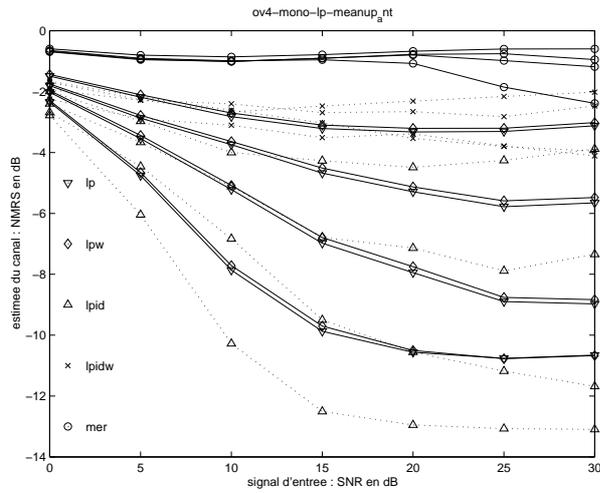
Les résultats des simulations précédentes appellent les commentaires suivants :

- «— La méthode  $lpid$  donne effectivement de meilleurs résultats que la méthode  $lp$ , mais perd sa robustesse à la surestimation de l'ordre (en effet, dans ce cas, la méthode devient une méthode de type déterministe). Les courbes pour 4 éléments d'antenne sont éloquentes à ce niveau, dans la mesure où les courbes pour les largeurs de bandes non-infinies correspondent en fait à une surestimation de la longueur du canal.
- «— La méthode de Cadzow est meilleure que la méthode de Levinson pour une largeur de bande infinie, mais perd en qualité pour les largeurs de bande plus faibles.
- «— La méthode de Levinson, de par l'utilisation de l'estimateur biaisé de la matrice de corrélation, présente un plancher de performances (c'est particulièrement bien visible pour les canaux obtenus par suréchantillonnage). À largeur de bande infinie, sa performance est sensiblement moins bonne que la méthode  $lp[w]$ , mais tend à être meilleure pour des largeurs de bande réalistes.

10.3.2 Résultats avec surestimation de la longueur du canal







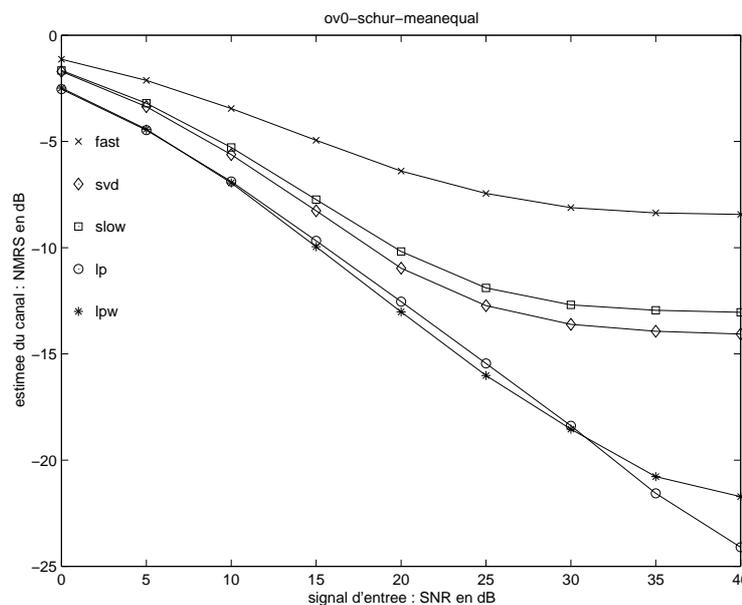
Les résultats des simulations précédentes appellent les commentaires suivants :

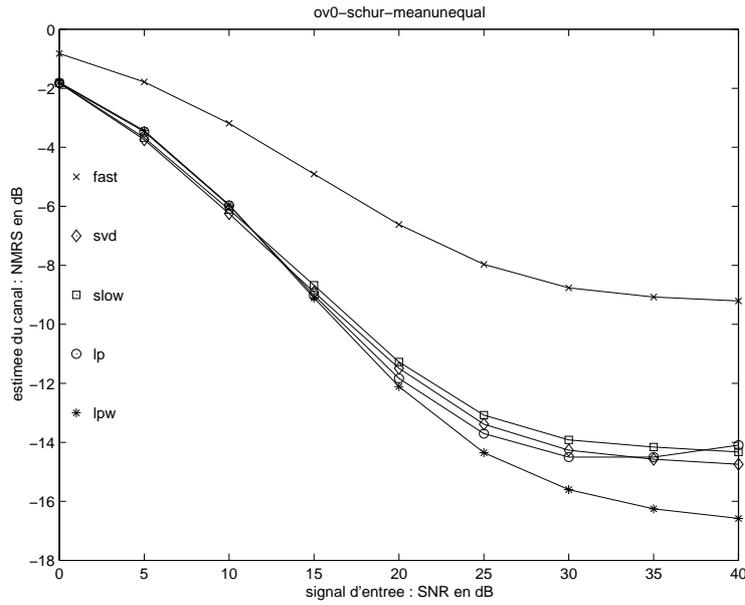
«— La méthode lpid donne toujours les meilleurs résultats, même si, de par la surestimation de l'ordre, on assimile un peu de l'espace bruit dans l'espace signal, ce qui fait échouer la méthode pondérée correspondante.

- «– La méthode de Cadzow est très peu robuste à la surestimation de l'ordre du canal, ce qui semble assez évident au vu de l'équation 6.14. En effet, cette équation suppose la longueur du canal connue, et aucune robustesse ne peut en être déduite.
- «– La méthode de Levinson présente de meilleures propriétés de robustesse que  $lp[w]$ , quoique présentant des performances légèrement inférieures à  $lpid$ . Cependant, on peut remarquer que les performances des méthodes de Levinson sont très peu affectées, comme on pouvait s'y attendre, par la surestimation de la longueur du canal. De ce point de vue, elles semblent être les grandes gagnantes de ces comparaisons.

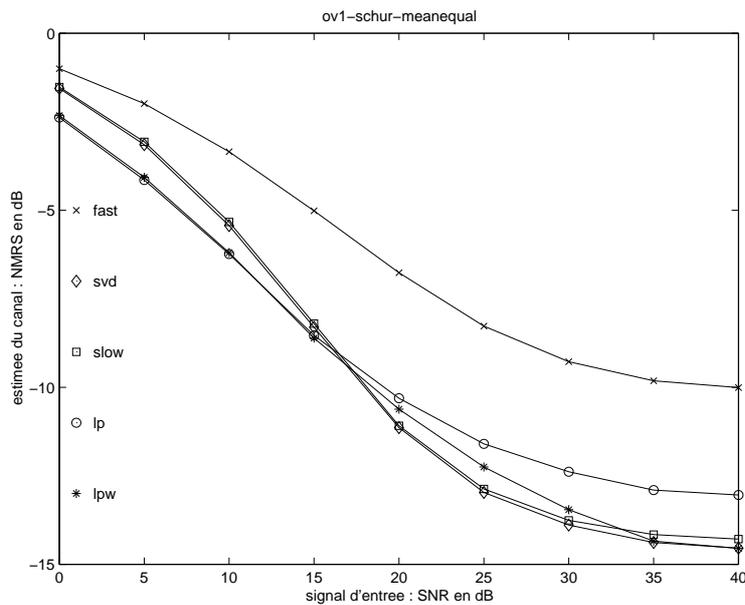
## 10.4 Méthode de Schur

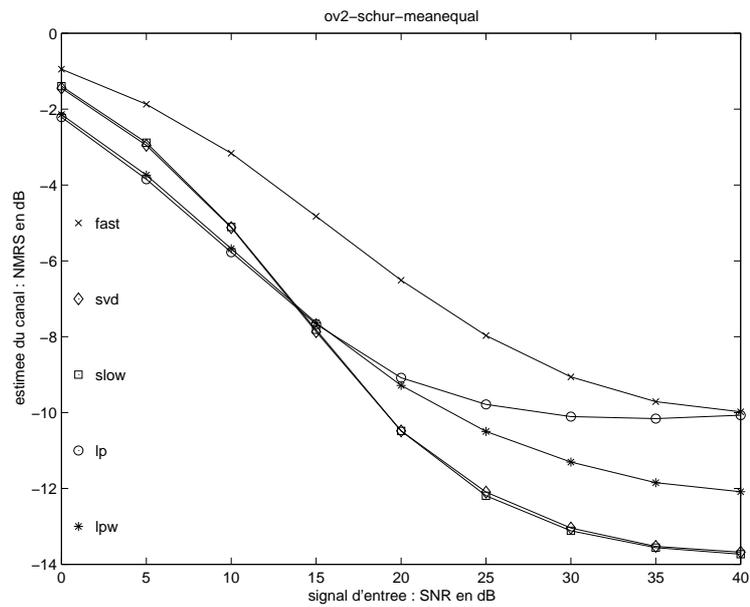
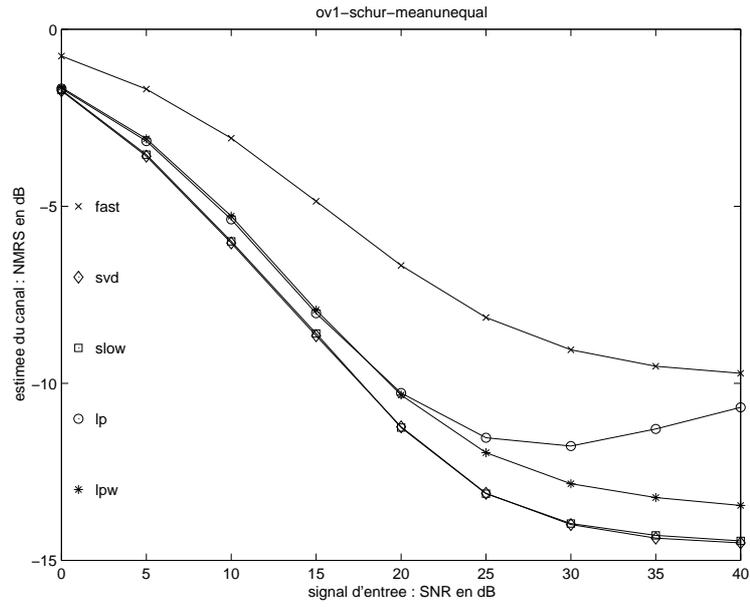
Pour l'identification de canal multi-utilisateurs, nous avons estimé les performances des méthodes de Schur, en les comparant à la méthode de prédiction linéaire, dans sa version  $lp[w]$ , en se basant sur la "sample covariance matrix". Les titres des figures comprennent la quantité par laquelle la longueur de chaque canal est surestimée et une mention indiquant si on utilise deux canaux de longueur égale ou différente. Nous traçons les courbes pour les méthodes  $lp[w]$  et Schur, avec ou sans SVD du générateur à l'étape finale, et par la méthode lente, qui calcule le complément de Schur explicitement.

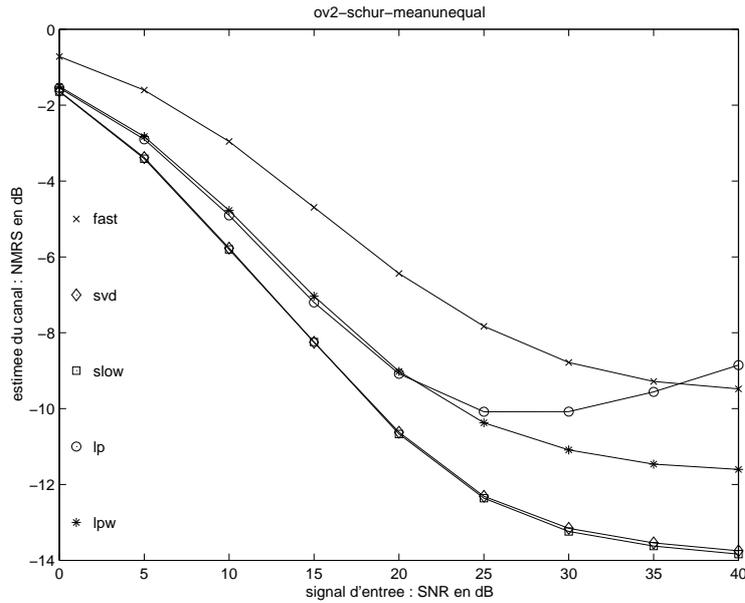




Les performances de la méthode de Schur, si on ne surestime pas la longueur du canal, sont légèrement inférieures à celle de la méthode lpw à un SNR bas, à un SNR élevé, Schur présente un effet de plafonnement des performances, dû à l'utilisation de l'estimation biaisée de la matrice de covariance. Dans le cas de canaux de longueur inégale, la méthode de prédiction linéaire présente également un effet de plafonnement des performances, de par l'estimation imparfaite d'éléments théoriquement nuls, et les deux méthodes se valent.







En moyenne, la méthode de Schur présente d'excellentes propriétés de robustesse à la surestimation de la longueur du canal. Les simulations indiquent que cette robustesse est supérieure à celle des méthodes  $lp[w]$ , ce qui nous encourage effectivement à poursuivre cette piste et à effectuer les analyses de performances théoriques. Cependant, le détail des simulations, que nous ne reproduisons pas ici, sous peine de surcharger le texte, indique des comportements très différents entre les deux familles d'algorithmes, en fonction des réalisations de canal. Dans certains cas, Schur est très médiocre alors que  $lp$  donne de bons résultats, et vice-versa. Ce point mérite certainement une analyse complémentaire.

On remarquera également que l'utilisation de la SVD sur le générateur s'impose et que les performances ainsi obtenues sont quasiment identiques à celles obtenues par la méthode lente.

# Chapitre 11

## Conclusions

### 11.1 Revue des résultats

Dans cette thèse, nous nous sommes attaqués au problème de l'identification de canal aveugle en utilisant les statistiques de second ordre, dans un contexte de communications, soit en présence d'un seul utilisateur, soit en présence de plusieurs utilisateurs partageant la même fréquence, au même moment et sans étalement de spectre.

Dans un premier temps, nous avons exploré les statistiques cycliques du second ordre, débouchant sur deux résultats :

- «– Une famille d'algorithmes dont les performances sont indépendantes de la couleur du bruit, mais malgré tout relativement médiocres [28, 25].
- «– Un nouveau modèle de canal qui permet d'appliquer toutes les méthodes existantes en se basant sur les statistiques cycliques du signal reçu [26, 25]. Ce travail peut être considéré comme étant l'achèvement des travaux de Gladyshev [37].

Ensuite, dans le souci de dériver des algorithmes rapides, nous avons développé une méthode sous-espace sans décomposition en vecteurs/valeurs propres et montré que cette méthode est asymptotiquement équivalente à la méthode initiale [23]. L'incorporation de la matrice de pondération optimale est relativement simple à faire et permet de déduire bon nombre d'équivalences asymptotiques entre diverses méthodes [22].

Toujours dans le souci de dériver des algorithmes rapides, mais également d'explorer la robustesse des méthodes proposées, nous avons exploré divers algorithmes dérivés de la prédiction linéaire [27]. Cette étude a montré que les algorithmes rapides de type Levinson donnent des performances très acceptables et présentent des propriétés de robustesse remarquables, comparées aux autres algorithmes. De plus, ces algorithmes rapides permettent une détection de l'ordre du

processus AR associé au canal et donc une robustesse accrue à la surdétermination de la longueur du canal. Cette détection est une extension des méthodes de détection d'ordre de processus multivariés au cas de processus singuliers.

Nous avons développé une méthode originale d'identification de canal multi-utilisateurs (applicable au cas mono) en se basant sur la décomposition LDU de la matrice de covariance et sur l'algorithme généralisé de Schur. Ce travail étant récent, nous n'avons malheureusement pas pu intégrer l'analyse de performances dans le texte de cette thèse, cependant, nous comptons la faire pour SPAWC'99. Les résultats des simulations sont extrêmement encourageants et laissent penser que cet algorithme est un candidat sérieux pour l'identification de canal aveugle en SDMA. De plus, la complexité de cet algorithme est faible.

Enfin, nous nous sommes intéressé aux méthodes basées sur le principe du Maximum de Vraisemblance. Un premier travail a été présenté dans [24] pour le modèle Gaussien, mais donne des performances médiocres, une version asymptotiquement optimale est en préparation [16]. Nous présentons également un algorithme pour le modèle déterministe [15], qui présente des performances prometteuses, avec bon nombre de garanties théoriques. Cependant, une application brutale de l'algorithme donne lieu à une charge de calcul prohibitive et le développement d'un algorithme rapide devrait être fait pour rendre cette méthode viable, ce qui sort du contexte de cette thèse.

## 11.2 Directions de recherches futures

Dans un premier temps, nous nous attacherons à analyser théoriquement les performances de l'algorithme de Schur, qui nous semble être une bonne méthode d'identification de canal aveugle, donnant de bonnes performances à un prix relativement faible. C'est un bon candidat pour donner une initialisation pour un algorithme itératif plus performant, voire un bon candidat en soi, dans des conditions favorables.

Ensuite, personnellement, j'aimerais explorer la détection multi-utilisateurs en incorporant le codage de canal, que ce soit en SDMA ou en CDMA, ainsi que les questions de théorie de l'information qui y sont associées. En effet, en surfant sur la vague finissante de l'aveugle et sur la vague montante du semi-aveugle, il me semble que l'exploitation jointe des informations apportées par le codage du canal (et donc la nature "alphabet fini" des sources) et par les statistiques du signal reçu devraient nous approcher de la limite théorique atteignable. Ce type de travail a déjà été abordé par un certain nombre de personnes et est très prometteur.

Enfin, la présence prochaine d'une plate-forme de test à Eurécom sera une bonne opportunité pour tester nos algorithmes dans des conditions réelles.

## Annexe A

# Bornes de Cramer-Rao

### A.1 Quelques notions de base

#### A.1.1 Distributions Gaussiennes

Pour des variables à valeurs réelles, la distribution Gaussienne du vecteur  $\mathbf{Y}$  de dimension  $M$ , de matrice de covariance  $C_{YY}$  et de moyenne  $m_Y$  vaut :

$$f(\mathbf{Y}) = \frac{1}{(2\pi \det C_{YY})^{M/2}} \exp \left[ -\frac{1}{2} (\mathbf{Y} - m_Y)^T C_{YY}^{-1} (\mathbf{Y} - m_Y) \right] \quad (\text{A.1})$$

Pour des variables à valeurs complexes, la distribution Gaussienne du vecteur  $\mathbf{Y}$  de dimension  $M$  de matrice de covariance  $C_{YY}$  et de moyenne  $m_Y$  vaut :

$$f(\mathbf{Y}) = \frac{1}{(\pi \det C_{YY})^M} \exp \left[ -(\mathbf{Y} - m_Y)^H C_{YY}^{-1} (\mathbf{Y} - m_Y) \right] \quad (\text{A.2})$$

#### A.1.2 Dérivées par rapport à une variable complexe

La dérivée par rapport à une variable complexe  $\theta = \alpha + j\beta$  est définie par  $\frac{\partial}{\partial \theta} = \frac{1}{2} \left( \frac{\partial}{\partial \alpha} - j \frac{\partial}{\partial \beta} \right)$ .

### A.2 CRBs pour des variables circulaires complexes

Nous faisons l'hypothèse que nous sommes en présence de variables complexes circulaires, ce qui est d'habitude le cas. Nous définissons alors :

$$J_{\phi\psi} = \mathbf{E}_{\mathbf{Y}|\theta} \left( \frac{\partial \ln f(\mathbf{Y}|\theta)}{\partial \phi^*} \right) \left( \frac{\partial \ln f(\mathbf{Y}|\theta)}{\partial \psi^*} \right)^H \quad (\text{A.3})$$

En considérant le vecteur réel  $\theta_R = [\Re(\theta) \ \Im(\theta)]$ , on peut relier  $J_{\theta_R \theta_R}$  à  $J_{\theta\theta}$  et  $J_{\theta\theta^*}$  par :

$$J_{\theta_R\theta_R} = 2 \begin{bmatrix} \Re(J_{\theta\theta}) & -\Im(J_{\theta\theta}) \\ \Im(J_{\theta\theta}) & \Re(J_{\theta\theta}) \end{bmatrix} + 2 \begin{bmatrix} \Re(J_{\theta\theta^*}) & \Im(J_{\theta\theta^*}) \\ \Im(J_{\theta\theta^*}) & -\Re(J_{\theta\theta^*}) \end{bmatrix} \quad (\text{A.4})$$

# Bibliographie

- [1] K. Abed-Meraim, J.-F. Cardoso, A. Y. Gorokhov, P. Loubaton, and E. Moulines. «On Subspace Methods for Blind Identification of Single-Input Multiple-Output FIR Systems». *IEEE Trans. on Signal Processing*, 45(1):42–55, January 1997.
- [2] K. Abed-Meraim, E. Moulines, and P. Loubaton. «Prediction error method for second-order blind identification». *IEEE Trans. on Signal Processing*, 45(3), March 1997.
- [3] Karim Abed-Meraim, Wanzhi Qiu, and Yingbo Hua. «Blind System Identification». *Proceedings of the IEEE*, 85(8):1310–1322, August 1997.
- [4] Alexei Gorokhov. «*Séparation autodidacte des mélanges convolutifs : méthodes du second ordre*». PhD thesis, Ecole Nationale Supérieure des Télécommunications, 1997.
- [5] Jahouar Ayadi, Elisabeth de Carvalho, and Dirk Slock. «Blind and Semi-Blind Maximum Likelihood Methods for FIR multichannel identification,». In *ICASSP*, Seattle, U.S.A., May 1998.
- [6] Jaouhar Ayadi, Luc Deneire, and Dirk T.M. Slock. «Channel-based Blind Multichannel Identification without Order Overestimation Problems,». In *13<sup>th</sup> international Conference on Digital Signal Processing*, Santorini, Greece, July 1997.
- [7] Luiz A. Baccala and Sumit Roy. «A New Blind Time-Domain Channel Identification Method Based on Cyclostationarity,». «*IEEE Signal Processing Letters*», 1(6):89–91, june 1994.
- [8] R.R. Bitmead, S.Y. Kung, B.D.O. Anderson, and T. Kailath. «Greatest Common Divisors via Generalized Sylvester and Bezout Matrices,». *IEEE Trans. on Automatic Control*, 23(6):1043–1047, December 1978.
- [9] James A. Cadzow. «Spectral Estimation: An Overdetermined Rational Model Equation Approach,». *Proceedings of the IEEE*, 70(9):907–939, september 1982.

- [10] Changhua Chen, Richard A. Davis, and Peter J. Brockwell. «Order Determination for Multivariate Autoregressive Processes Using Resampling Methods». *Journal of Multivariate Analysis*, 57(2):175–190, May 1996.
- [11] Pascal Chevalier. «Méthodes aveugles de filtrage d’antennes». *Revue d’Electricité*, (3):48–58, Septembre 1995.
- [12] P. COMON. Structured matrices and inverses. In A. Bojanczyk and G. Cybenko, editors, *Linear Algebra for Signal Processing*, volume 69 of *IMA Volumes in Mathematics and its Applications*, pages 1–16. Springer Verlag, 1995.
- [13] Pierre Comon. Separation of sources using high-order cumulants. In *SPIE Conf. Adv. Alg. Archi. Sig. Proc.*, pages 170–181, San Diego, Ca, August 8-10 1989. vol. Real-time signal processing XII.
- [14] Pierre Comon. «Tensor Diagonalization, A useful Tool in Signal Processing,». In *IFAC-SYSID, 10th IFAC Symposium on System Identification*, volume 1, Copenhagen, Denmark, July 4–6 1994. M. Blanke and T. Soderstrom, editors. invited session.
- [15] Elisabeth de Carvalho, Luc Deneire, and Dirk Slock. «Blind and Semi-Blind Maximum Likelihood Techniques for Multiuser Multichannel identification». In *European Association for Signal Processing EUSIPCO-98*, Island of Rhodes, Greece, September, 8–11 1998.
- [16] Elisabeth de Carvalho, Luc Deneire, and Dirk Slock. «A Fast Gaussian Maximum Likelihood Method for Blind Multiuser Multichannel Estimation». In *Signal Processing Advances in Wireless Communications workshop*, 1999.
- [17] Elisabeth de Carvalho and Dirk Slock. «Identifiability Conditions for Blind and Semi-Blind Multichannel Estimation». In *European Association for Signal Processing EUSIPCO-98*, Island of Rhodes, Greece, September, 8–11 1998.
- [18] Elisabeth de Carvalho and Dirk T.M. Slock. «Cramer-Rao Bounds for Semi-Blind, Blind and Training Sequence Based Channel Estimation». In *IEEE SP Workshop on Signal Processing Advances for Wireless Communications*, pages 129–133, Paris, France, April 16–18 1997. IEEE.
- [19] Elisabeth de Carvalho and Dirk T.M. Slock. «Maximum-Likelihood Blind FIR Multi-Channel Estimation with Gaussian Prior for the Symbols». In *Proc. ICASSP*, Munich, Germany, April 1997.

- 
- [20] Elisabeth de Carvalho and Dirk T.M. Slock. «A Pseudo-Quadratic Maximum-Likelihood method for Semi-Blind and Blind FIR Multichannel Estimation». *In preparation for IEEE Trans. on Signal Processing*, 1998.
- [21] Elisabeth de Carvalho and Dirk T.M. Slock. «Cramer-Rao Bounds for Semi-Blind, Blind and Training Sequence FIR Multichannel Estimation». *In preparation for IEEE Trans. on Signal Processing*, 1998.
- [22] Luc Deneire, Jahouar Ayadi, and Dirk T.M. Slock. «Weighted and Unweighted Subspace Fitting without Eigendecomposition». In *5th Symposium on Communications and Vehicular Technology in the Benelux*, Twente, Netherlands, october 1997.
- [23] Luc Deneire, Jaouhar Ayadi, and Dirk T.M. Slock. «Subspace Fitting Without Eigendecomposition». In «*13<sup>th</sup> international Conference on Digital Signal Processing*», Santorini, Greece, July 1997.
- [24] Luc Deneire and Dirk Slock. «Gaussian Maximum Likelihood Blind Multichannel Multiuser Identification». In *48th annual Vehicular Technology Conference*, Ottawa, Canada, May, 18–21 1998.
- [25] Luc Deneire and Dirk T.M. Slock. «Blind channel identification based on cyclic statistics». *IEE Proceedings - Radar, Sonar and Navigation, Special issue on Antenna array processing techniques*, 145(1):58–62, February 1988.
- [26] Luc Deneire and Dirk T.M. Slock. «Linear Prediction And Subspace Fitting Blind Channel Identification Based On Cyclic Statistics.». In «*13<sup>th</sup> international Conference on Digital Signal Processing*», Santorini, Greece, July 1997.
- [27] Luc Deneire and Dirk T.M. Slock. «On robustness of Linear Prediction based blind identification». In *5th Symposium on Communications and Vehicular Technology in the Benelux*, Twente, Netherlands, october 1997.
- [28] Luc Deneire and Dirk T.M. Slock. «Second-order Cyclic Statistics based blind channel identification and equalization». In *Proc. 1st IEEE SP Workshop on Signal Processing Advances for Wireless Communications*, pages 21–24. IEEE, April 1997.
- [29] Z. Ding. «Characteristics of band-limited channels unidentifiable from second-order cyclostationary statistics,». *IEEE Signal Processing letters*, 3:pp. 150–152, May 1996.
- [30] ETSI. «European digital cellular telecommunications system (Phase 2); Radio transmission and reception (GSM 05.05),». ETSI, December 1995. fifth edition.

- [31] Jean-Jacques Fuchs. «Rectangular Pisarenko Method Applied to Source Localization». *IEEE Transactions on Signal Processing*, 44(10):2377–2383, October 1996.
- [32] K.A Gallivan, S. Thirumalai, P. Van Dooren, and V. Vermaut. «High Performance Algorithms for Toeplitz and block Toeplitz matrices». *Linear Algebra and its Applications*, 1994.
- [33] William A. Gardner. «Cyclic Wiener Filtering : Theory and Method». «*IEEE Transactions on Communications* », 41(1):151–163, january 1993.
- [34] William A. Gardner, editor. «*Cyclostationarity in Communications and Signal Processing* ». IEEE Press, New York, 1994.
- [35] Irfan Ghauri and Dirk T.M. Slock. «Blind and Semi-blind Single User Receiver Techniques for Asynchronous CDMA in Multipath Channels». In *IEEE Globecom*, Sydney, Australia, November 1998.
- [36] Georgios B. Giannakis. «Linear Cyclic Correlation Approaches for Blind Identification of FIR Channels». In *29<sup>th</sup> Asilomar Conference on Signals, Systems and Computers*, pages 420–424, 1995.
- [37] E.G. Gladyshev. «Periodically random sequences». *Soviet Mathematics*, 2, 1961.
- [38] G.A. Glentis and N. Kalouptsidis. «Fast Adaptive Algorithms for Multichannel Filtering and System Identification». *IEEE Transactions on Signal Processing*, 40(10):2433–2457, october 1992.
- [39] A. Gorokhov and Ph. Loubaton. «Subspace based techniques for second order blind separation of convolutive mixtures with temporally correlated sources». *IEEE Trans. on Circuits and Systems*, July 1997.
- [40] M. Gurelli and C.L. Nikias. «A New Eigenvector-Based Algorithm for Multichannel Blind Deconvolution of Input Colored Signals». In *Proc. ICASSP*, pages 448–451, 1993.
- [41] Steven D. Halford and Georgios B. Giannakis. «Channel order Determination based on Sample Cyclic Correlations». In *Twenty-Eight Asilomar Conference on Signal, Systems & Computers*, October 1994.
- [42] G. Harikumar and Y. Bresler. «Analysis and Comparative Evaluation of Techniques for Multichannel Blind Deconvolution,». In *Proc. Workshop Statistical Signal and Array Proc.*, pages 332–335, Corfu, June 1996.

- 
- [43] Gopal Harikumar and Yoram Bresler. «FIR Perfect Signal Reconstruction from Multiple Convolutions: Minimum Deconvolver Orders». *IEEE Trans. on Signal Processing*, 46(1):215–218, January 1998.
- [44] H.Liu and G.Xu. «A deterministic approach to blind symbol estimation». *spl*, 1:205–207, December 1994.
- [45] Bertrand Hochwald and Arye Nehorai. «On Identifiability and Information-Regularity in Parameterized Normal Distributions». *Circuits, Systems, and Signal Processing*, 16(1), 1997.
- [46] Yingbo Hua. «Fast Maximum Likelihood for Blind Identification of Multiple FIR Channels». *IEEE Trans. on Signal Processing*, 44(3):661–672, March 1996.
- [47] Yingbo Hua and Mati Wax. «Strict Identifiability of Multiple FIR Channels Driven by an Unknown Arbitrary Sequence». *IEEE Trans. on Signal Processing*, 44(3):756–759, March 1996.
- [48] J. Chun. «*Fast Array Algorithms for Structured Matrices*». PhD thesis, Information Systems Laboratory, Stanford University, June 1989.
- [49] Thomas Kailath. «*Linear Systems*». Prentice Hall, 1980.
- [50] Martin Kristensson, Björn Ottersten, and Dirk Slock. «Blind Subspace Identification of a BPSK Communication Channel». In *Proc. Asilomar 96, Asilomar, USA*, November 1996.
- [51] Hui Liu and Guanghan Xu. «Smart Antennas in Wireless Systems: Uplink Multiuser Blind Channel and Sequence Detection».
- [52] Ph. Loubaton. «Quelques aspects méthodologiques concernant la séparation des mélanges convolutifs». «Ecole des Techniques avancées Signal Image Parole, Grenoble», septembre 1996.
- [53] Karim Abed Meraim. «*Identification/Egalisation autodidacte au second ordre: Algorithmes et Performances*». PhD thesis, ENST, Paris, 1995.
- [54] A.G. Miamee. «Periodically Correlated Processes and their Stationary Dilations». *SIAM Journal of Applied Mathematics*, 50(4):1194–1199, August 1990.
- [55] A.G. Miamee. «Explicit Formula for the Best Linear Predictor of Periodically Correlated Sequences». *SIAM Journal of Mathematic Analysis*, 24(3):703–711, May 1993.

- [56] E. Moulines, P. Duhamel, J.-F. Cardoso, and S. Mayrargue. «Subspace Methods for the Blind Identification of Multichannel FIR filters». *IEEE Trans. on Signal Processing*, 43(2):516–525, February 1995.
- [57] B. Ottersten, M. Viberg, P. Stoica, and A. Nehorai. *Radar Array Processing*, chapter Exact and Large Sample ML Techniques for Parameter Estimation and Detection in Array Processing. S. Haykin, J. Litva, T.J. Shepherd, eds., Springer-Verlag, 1993.
- [58] Björn Ottersten, David Aztély, Marin Kristensson, and Stefan Parkvall. «A Statistical Approach to Subspace Based Estimation with Applications in Telecommunications». In *tlssix*, 1996.
- [59] Constantin Papadias. «*Méthodes d'égalisation et d'identification aveugle pour les canaux linéaires*». PhD thesis, ENST, Paris, 1995.
- [60] Constantinos Papadias and D.T.M. Slock. Fractionally-spaced equalization of linear poly-phase channels and related blind techniques based on multichannel linear prediction. *IEEE Trans. on Signal Processing*, 47(3), March 1999.
- [61] Boaz Porat. «*Contributions to the Theory and Applications of Lattice Filters*». PhD thesis, Stanford University, August 1982.
- [62] John G. Proakis. *Digital Communications*. McGraw-Hill, 2nd edition, 1989.
- [63] Vellenki U. Reddy, Constantinos B. Papadias, and Arogyaswami J. Paulraj. «Blind Identifiability of Certain Classes of Multipath Channels from Second-Order Statistics Using Antenna Arrays». *IEEE Signal Processing Letters*, 4(5):138–141, May 1997.
- [64] Gregory C. Reinsel. «*Elements of Multivariate Time Series Analysis*». Springer Series in Statistics. Springer-Verlag, 1993.
- [65] Ali H. Sayed and Thomas Kailath. «A Look-ahead Block Schur Algorithm for Toeplitz-like Matrices». *SIAM J. Matrix Anal. Appl.*, 16(2):pp. 388–414, April 1995.
- [66] Louis L. Scharf. «*Statistical Signal Processing, Detection, Estimation, and Time Series Analysis*». Addison-Wesley, 1991.
- [67] Dirk T.M. Slock. «Blind Joint Equalization of Multiple Synchronous Mobile Users Using Oversampling and/or Multiple Antennas». In *Twenty-Eight Asilomar Conference on Signal, Systems & Computers*, October 1994.

- 
- [68] D.T.M. Slock. «Blind Fractionally-Spaced Equalization, Perfect-Reconstruction Filter Banks and Multichannel Linear Prediction». In *Proc. ICASSP Conf.*, Adelaide, Australia, April 1994.
- [69] 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, August 1994.
- [70] 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.
- [71] Lang Tong. «Blind Channel Identification using Cyclic Spectra». In *CISS94*, 1994.
- [72] Lang Tong, Guanghan Xu, B. Hassibi, and T. Kailath. «Blind Channel Identification Based on Second-Order Statistics: A Frequency-Domain Approach». «*IEEE Transactions on Information Technology*», 41(1):329–334, January 1995.
- [73] A. Touzni and I. Fijalkow. «Robustness of Blind Fractionally-Spaced Identification/Equalization to Loss of Channel Disparity». In *Proc. ICASSP*, volume 5, pages 3937–3940, Munich, Germany, April 1997.
- [74] H. Trigui and D.T.M. Slock. «Cochannel Interference Cancellation within the current GSM Standard». *To appear in a special issue of the European Journal Signal Processing*, 1999.
- [75] D.W. Tufts and R. Kumaresan. «Estimation of frequencies of multiple sinusoids: making linear prediction perform like maximum likelihood». *Proceedings of the IEEE*, 70:975–989, 1982.
- [76] Jitendra K. Tugnait. «On Blind Identifiability of Multipath Channels Using Fractional Sampling and Second-Order Cyclostationary Statistics». «*IEEE Transactions on Information Technology*», 41(1):308–311, January 1995.
- [77] M. Wax and T. Kailath. «Detection of signals by information theoretic criteria». *IEEE Trans. Acoust., Speech, Signal Processing*, 33:387–392, 1985.
- [78] M. Wax and I. Ziskind. «Detection of the number of coherent signals by the MDL principle». *IEEE Trans. Acoust., Speech, Signal Processing*, 37:1190–1196, 1989.
- [79] G. Xu, H. Liu, L. Tong, and T. Kailath. «A Least-Squares Approach to Blind Channel Identification». *IEEE Trans. on Signal Processing*, 43(12):2982–2993, december 1995.

- [80] Hanks H. Zeng and Lang Tong. «Blind Channel Estimation Using the Second-Order Statistics : Asymptotic Performance and Limitations». *submitted to IEEE Transactions on Signal Processing*, January 1997.