

Identification aveugle au second ordre,
ordre et diversité des canaux

Houcem GAZZAH

Département Signal et Image

Institut National des Télécommunications

Candidat au Doctorat

de l'École Nationale Supérieure des Télécommunications

Spécialité : Signal et Images

À mes parents.

Avant-propos

Le présent rapport récapitule les travaux entrepris par l'auteur dans le cadre de sa préparation du Doctorat de l'*École Nationale Supérieure des Télécommunications* commencée en décembre 1997 au sein du *Département Signal et Image* de l'*Institut National des Télécommunications* sous l'encadrement de Mrs Phillip A. Regalia et Jean-Pierre Delmas.

Houcem GAZZAH,
Evry, le 9 Mai 2000.

Remerciements

Je tiens tout d'abord à remercier Monsieur Pierre Duhamel, Professeur à l'ENST pour avoir bien voulu me faire l'honneur de présider mon jury de thèse. Ma gratitude va également à Madame Inbar Fijalkow, Professeur à l'ENSEA, et Monsieur Karim Abed-Meraim, Maître de Conférences à l'ENST, qui ont accepté de donner de leur temps pour rapporter cette thèse. Je remercie Monsieur Dirk Slock, Professeur à Eurecom, et Monsieur Pascal Chevalier, de Thomson-CSF, d'avoir bien voulu accepter d'être membres de ce jury.

Je remercie Monsieur Claude Gimenes, Professeur et Directeur du Département Signal et Image de l'Institut National des Télécommunications pour la qualité d'accueil et l'aide financière qui m'a été accordée tout au long de mon travail de thèse.

J'exprime ma profonde reconnaissance à mes directeurs de thèse, Messieurs Phillip A. Regalia et Jean-Pierre Delmas, Professeurs à l'INT. Monsieur Regalia a toujours su porter un regard critique sur mon travail. Son expérience et sa parfaite connaissance m'ont permis d'orienter et enrichir mes travaux. Je le remercie vivement pour son écoute, sa disponibilité et sa gentillesse. Monsieur Delmas m'a fait profiter d'un suivi intensif, précieux et bienveillant qui m'a permis de progresser rapidement et efficacement dans mes travaux. Je le remercie pour son engagement et la qualité de son encadrement.

Je remercie également la bibliothèque de l'INT et en particulier Madame Corinne Cipièrre pour son aide considérable dans la recherche de documents. Je remercie le personnel du département Signal et Image et en particulier Madame Evelyne Taroni que j'ai dû abondamment solliciter. Mes pensées s'adressent aussi à tous mes amis, Houcine, Naceur, Kofidis, Sorin, . . . Je garderai un agréable souvenir de leur compagnie et de leur humour.

Table des matières

1	Introduction	11
2	Notations	15
3	Égalisation synchrone en communication numérique	17
3.1	Canal de communication	17
3.1.1	Chaîne de communication numérique	17
3.1.2	Effets du canal	18
3.1.3	Équivalent bande de base	19
3.1.4	Interference inter symboles	21
3.2	Égalisation linéaire supervisée	21
3.2.1	Structures et critères d'égalisation	21
3.2.2	Implémentation de l'égalisation linéaire MSE	24
3.3	Égalisation aveugle	25
3.4	Conclusion	27
4	Identification/égalisation spatio-temporelle	29
4.1	Notations et résultats	29
4.1.1	Le canal SIMO	29
4.1.2	Processus multivariables	30
4.1.3	Matrice de filtrage	32
4.2	Égalisation linéaire des canaux SIMO	35
4.2.1	Égalisation ZF	36
4.2.2	Égalisation MMSE	38

4.3	Détection de l'ordre et estimation du bruit	38
4.4	Algorithmes d'identification aveugle	39
4.4.1	L'algorithme sous-espace	41
4.4.2	L'algorithme prédiction linéaire	42
4.4.3	L'algorithme décomposition en produit extérieur	44
4.5	Conclusion	45
5	Identification aveugle au second ordre robuste à la surestimation de l'ordre	47
5.1	Surestimation de l'ordre	47
5.2	Notations	48
5.3	Développement théorique de l'algorithme	49
5.4	Robustesse à la surestimation de l'ordre en statistiques exactes	51
5.5	Cas des statistiques estimées	52
5.6	Robustesse à la surestimation de l'ordre en statistiques estimées	54
5.7	Récapitulatif	56
5.8	Mesure de performances	57
5.9	Simulations	57
5.10	Comparaison avec les algorithmes existants	59
5.11	Conclusion	61
6	Diversité des canaux SIMO	64
6.1	Mesure de la diversité	64
6.2	Notations et Rappel	66
6.3	Matrices blocs de Toeplitz	68
6.3.1	Définitions	68
6.3.2	Distribution asymptotique des valeurs propres	69
6.3.3	Classe particulière des matrices blocs de Toeplitz	72
6.4	Application à l'identification aveugle	74
6.4.1	Application aux matrices de filtrage	74
6.4.2	Application à la diversité des canaux SIMO	75

6.4.3	Cas de la réception fractionnée des signaux à bande limitée	76
6.5	Conclusion	80
7	Modélisation effective de l'ordre	81
7.1	Réponse et ordre effectifs	81
7.1.1	Intérêt	81
7.1.2	Définitions	82
7.1.3	Insuffisances	83
7.2	Modélisation effective en identification aveugle	84
7.2.1	Notations	84
7.2.2	La perturbation de modélisation effective	85
7.3	Approche fonctionnelle de l'identification	85
7.4	Étude de performances en modélisation effective	86
7.4.1	Modèle déterministe	86
7.4.2	Modèle probabiliste	88
7.5	Égalisation en modélisation effective	90
7.5.1	Égalisation ZF	91
7.5.2	Réponse globale	91
7.5.3	Ouverture de l'oeil	91
7.6	Application aux algorithmes	92
7.6.1	L'algorithme sous-espace / moindres carrés	92
7.6.2	L'algorithme prédiction linéaire	94
7.6.3	L'algorithme décomposition en produit extérieur	94
7.6.4	Simulations	94
7.7	Conclusion	105
8	Conclusion	106
8.1	Résultats	106
8.2	Perspectives	107

A Démonstrations de §7	110
A.1 Démonstration de (7.5)	110
A.2 Démonstration de (7.4)	110
A.3 Démonstration de (7.8)	111
A.4 Démonstration de (7.9)	112
A.5 Démonstration de (7.10)	113
A.6 Démonstration de (7.11)	114
A.7 Démonstration de (7.12)	114
A.8 Démonstration de Prop. 7.10	114
A.9 Démonstration de Prop. 7.11	116
B Différentiation des algorithmes d'identification	118
B.1 Estimation du bruit	119
B.2 L'algorithme sous-espace	120
B.3 L'algorithme prédiction linéaire	120
B.4 L'algorithme décomposition en produit extérieur	126
C Démonstration de Prop. 4.1	129
D Démonstration de Prop. 4.11	141
E Théorie de perturbation des matrices	143

Liste des figures

3.1	Chaîne de communication numérique	18
3.2	Schéma numérique équivalent bande de base	20
3.3	Égaliseur linéaire	22
4.1	Canal SIMO bruité	30
4.2	Égalisation d'un canal SIMO	35
4.3	Schéma numérique équivalent	46
5.1	Canal SIMO	48
5.2	Robustesse à la surestimation de l'ordre en statistiques estimées.	55
5.3	Le canal	58
5.4	Comparaison de l'algorithme proposé et des algorithmes SS, LP et OPD avec détection de l'ordre exact, $w = 9$	59
5.5	L'algorithme proposé en sur-estimation de l'ordre du canal. La légende montre l'ordre détecté. $N = 200$. $SNR = 20$	60
5.6	Effet du facteur de lissage. $SNR = 20$	61
5.7	Surestimation de l'ordre	62
6.1	Contre-exemple	65
6.2	$\sigma(\mathcal{T}_l(\mathbf{h}_M)), \mathbf{h}_M^1 = [-0.6804 \ 0.177 \ -0.0902]^T, \mathbf{h}_M^2 = [0.4281 \ -0.2446 \ -0.5043]^T$. La borne (6.17) vaut 0.6747.	75
6.3	Canal avec termes faibles de la tête/queue	76
6.4	Cas de non-identifiabilité absolue	77
6.5	Cas identifiable	78
6.6	Cas de non-identifiabilité absolue	79

7.1	Canal radio hyper-fréquence	82
7.2	Canal type en modélisation effective	84
7.3	Réponses totale et effective : disposition des zeros	95
7.4	Identification de la réponse effective : Modèle déterministe	98
7.5	Modélisation effective : le cas le plus défavorable	99
7.6	Effets séparés du nombre d'observations et de la modélisation effective	103
7.7	Identification de la réponse effective : Modèle probabiliste	103
7.8	Égalisation ZF en modélisation effective	104
7.9	Effet de la diversité.	104
8.1	Exploitation de la connaissance du signal de mise en forme	107

Liste des tableaux

7.1	Diversité des réponses totale et effective	95
7.2	Méthode SS : Biais et EQM	100
7.3	Méthode LP : Biais et EQM	101
7.4	Méthode OP : Biais et EQM	102

Chapitre 1

Introduction

En communication numériques, l'égalisation désigne les techniques de traitement du signal mises en oeuvre pour annuler l'effet dispersif du canal. Le plus souvent sous forme d'un filtre linéaire, l'égaliseur est calculé à la suite d'une phase dite d'*apprentissage* au cours de laquelle l'émetteur envoie une séquence connue à l'avance par le récepteur. En observant la sortie du canal pendant cette phase, l'algorithme (d'égalisation) en déduit l'égaliseur *optimal* qui, à la fin de la phase d'apprentissage, continue d'opérer sur les observations du canal pour récupérer au *mieux* les symboles effectivement transmis, et ce jusqu'à la disponibilité d'une nouvelle séquence d'apprentissage. L'égalisation ainsi réalisée est dite *supervisée* (par la séquence d'apprentissage).

Le problème essentiel provient de la nécessité de reprendre la phase d'apprentissage chaque fois que le canal change (significativement¹) de réponse. Dans des applications où le canal est inévitablement variable dans le temps, notamment les applications de communications avec les mobiles, ceci implique la réservation dans la trame d'une partie dédiée à la séquence d'apprentissage. La perte de bande est souvent conséquente². Dans les communications point à multipoints, lorsque la qualité de transmission se dégrade sur une liaison, la reprise de l'apprentissage s'impose et monopolise le serveur; et toutes les autres liaisons se trouvent suspendues durant [GODR 80]. Ces difficultés ont motivé la recherche de techniques *autodidactes* dites aussi *aveugles*. Il s'agit d'algorithmes d'égalisation qui se passent de la connaissance du signal émis³.

Les premiers algorithmes d'égalisation aveugle utilisent implicitement [GODR 80] ou explicitement [SHAL 90] les statistiques d'ordre supérieur (HOS⁴). Leur efficacité est limitée. D'un côté, une bonne estimation de ces statistiques nécessite un grand nombre de symboles. Ces algorithmes convergent typiquement au bout de quelques milliers de symboles ce qui est excessif vu qu'en pratique la réponse du canal peut changer plus rapidement. D'un autre côté, ces algorithmes faisant intervenir des critères non quadratiques, souffrent de l'existence de minimas locaux.

¹En effet, les algorithmes d'égalisation supervisée ont la capacité de poursuivre des variations limitées de la réponse du canal.

²Pour le standard GSM de communications mobiles, la séquence d'apprentissage est de 26 bits sur les 148 bits de la trame. Elle est de 41 bits sur 87 pour la trame d'accès [LEE 95].

³L'égalisation aveugle trouve son application aussi en dehors du domaine des communications, tel en restauration d'images et en géophysique [TONG 98].

⁴Higher Order Statistics

Récemment, des techniques spatio-temporelles [NAGB 96] furent introduites. Consistant en l'utilisation des réseaux d'antennes à la réception (et éventuellement à l'émission), elles permettent d'améliorer la qualité de la réception supervisée, et par conséquent la capacité des systèmes de communications mobiles. Dans ces conditions de réception, les algorithmes d'égalisation supervisée nécessitent des séquences d'apprentissage plus longues. Cependant, la réception spatio-temporelle a l'avantage de faire apparaître l'information de phase du canal dans ses statistiques du second ordre (SOS⁵) qui n'en contiennent pas lorsque la réception s'effectue par le moyen d'un seul capteur. Ainsi, l'égalisation aveugle est rendu possible au second ordre. Un cadre *équivalent*⁶ à l'utilisation des réseaux d'antennes est le sur-échantillonnage de la sortie du canal à un rythme supérieur au rythme symbole, dit aussi réception *fractionnée*.

L'emploi de l'une et/ou l'autre de ces techniques est désigné par *réception en diversité spatio-temporelle*. Il renvoie à l'identification-égalisation des canaux à une entrée et sorties multiples (SIMO⁷). Ainsi, de nombreux algorithmes d'identification aveugle au second ordre ont été proposés dont [TONG 94, MOUL 95, ABED 97, DINGa97]. Cette thèse est dédiée à l'étude de ce type d'algorithmes. Elle se penche sur deux inconvénients majeurs qui pénalisent leur emploi en pratique. Premièrement, ces algorithmes exigent la connaissance (improbable) de la longueur exacte du canal à identifier. Deuxièmement, il ne leur suffit pas de disposer de la sortie du canal observée en diversité (spatio-temporelle) mais il leur faut une diversité suffisamment forte pour que le problème de l'identification (en statistiques exactes) soit bien conditionné et assurer ainsi de bonnes performances lorsque le canal est observé pendant une durée (forcement) finie (statistiques estimées).

Le rapport aborde ces questions en respectant le plan suivant.

Après avoir précisé les notations (chapitre 2), on rappelle dans le chapitre 3 les techniques d'égalisation (supervisée ou non) synchrones où le canal à égaliser est observé à travers un capteur unique et au rythme des symboles émis.

Ensuite (chapitre 4), les techniques de réception spatio-temporelle sont introduites ainsi que le modèle SIMO. Les notations, ainsi que des résultats asymptotiques, relatives aux SOS d'un SIMO sont présentés. On rappelle les critères de détection de l'ordre. Leur non fiabilité en pratique rend nécessaire pour un algorithme d'identification d'être capable de déduire la réponse impulsionnelle du canal lorsque la longueur de ce dernier est sur-estimée. On désignera cette propriété par *robustesse à la sur-estimation de l'ordre*. L'examen des algorithmes développés à ce jour permet de constater qu'en pratique, aucun ne vérifie cette propriété.

Pour remédier à cet inconvénient, nous proposons dans le chapitre 5 un nouvel algorithme d'identification aveugle au second ordre robuste à la sur-estimation de l'ordre *au sens pratique du terme*, i.e., lorsque l'on dispose d'une valeur sur-estimée de l'ordre du canal et lorsqu'on l'observe pendant une durée finie et en présence de bruit. Outre cet avantage, l'algorithme proposé permet de choisir l'ambiguïté avec laquelle l'identification se fait. La réponse du canal peut être ainsi identifiée à une phase près ou bien à une phase et à une amplitude près. Il montre également de meilleures performances qui de surcroît

⁵Second Order Statistics

⁶En pratique (cf. chapitre 6), les canaux sont à bande limitée et les techniques fractionnées créent une diversité insuffisante pour identifier de tels canaux à partir d'un nombre limité d'observations. La réception en diversité spatiale est souvent citée comme préférable de ce point de vue [Redd 95].

⁷Single Input Multiple Output

peuvent être améliorées sans avoir à augmenter la durée d'observation du canal.

Dans le chapitre 6, on met en évidence la sensibilité des algorithmes d'identification aveugle au second ordre au bon conditionnement de la matrice de corrélation du canal. Cette dernière ayant une structure bloc de Toeplitz, on fait le rapprochement avec le théorème de Szegö qui lie les valeurs propres d'une matrice de Toeplitz Hermitienne à la transformée de Fourier de ses termes. Ceci permet de proposer une définition de la *diversité* d'un canal qui permet d'apprécier son identifiabilité en aveugle indépendamment de l'algorithme en question. Elle est fixée comme étant la limite asymptotique de la plus petite valeur propre non nulle de la matrice de corrélation du canal lorsque le facteur de lissage de cette dernière tend vers l'infini. On propose également une nouvelle démonstration de l'extension du théorème de Szegö aux matrices bloc de Toeplitz, plus simple mais sous des hypothèses moins générales. L'interprétation de ce résultat dans le cadre de l'identification aveugle des canaux SIMO permet de déduire une borne supérieure sur la diversité du canal qui s'exprime en fonction de sa fonction de transfert. Cette borne se prête mieux à l'interprétation physique. Elle permet en particulier de justifier l'inefficacité des algorithmes aveugles à identifier les canaux à bande limitée en diversité temporelle, et rejoint ainsi les conclusions de travaux précédents.

La faible diversité des canaux est souvent le fait de termes faibles à la tête et/ou à la queue de la réponse impulsionnelle. *Liavas et al* introduisent [LIAV 99] la notion de *réponse effective* comme étant la sous partie contiguë de la réponse débarrassée des termes faibles à la tête/queue. Ils proposent [LIAV 99] un algorithme de détection de l'ordre dit effectif de cette réponse et mettent en évidence sa bien meilleure diversité qui la prête mieux à l'identification [LIAV 99, LIAV 00]. Nous commentons la pertinence de cette approche qui en fait ne correspond qu'à certaines situations fréquentes mais pas exhaustives. Dans de telles situations, nous effectuons une étude de performances applicable à tout algorithme du second ordre. Son application à certains algorithmes permet de vérifier la validité de l'étude et de dégager certaines conclusions quant à la sensibilité des algorithmes d'identification à la non prise en compte de termes faibles à la tête et/ou à la queue de la réponse.

Le rapport s'achève par une conclusion qui expose certaines des difficultés persistantes et les perspectives futures.

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

Articles de revues

Houcem Gazzah, Phillip A. Regalia, Jean-Pierre Delmas and Karim Abed-Meraim, “*A Blind Multichannel Identification Algorithm Robust to Order Over Estimation*”, soumis à IEEE Transactions on Signal Processing, octobre 2000.

Houcem Gazzah, Phillip A. Regalia and Jean-Pierre Delmas, “*Asymptotic Eigenvalue Distribution of Block Toeplitz Matrices and Application to Blind SIMO Channels Identification*”, à paraître dans IEEE Transactions on Information Theory, janvier 2001.

Jean-Pierre Delmas, Houcem Gazzah and A. P. Liavas, “*Statistical Analysis of Some Second-Order Methods for Blind Channel Identification/Equalization with Respect to Channel Undermodeling*”, IEEE Transactions on Signal Processing, VOL. 48, NO. 7, pp. 1984-1998, juillet 2000.

Articles de conférences

Houcem Gazzah, Phillip A. Regalia and Jean-Pierre Delmas, “*On Blind (non)Identifiability of Dispersive Bandlimited Channels*”, soumis à IEEE ICASSP, Salt Lake City, Utah, mai 2001.

Houcem Gazzah, Phillip A. Regalia and Jean-Pierre Delmas, “*A Blind Identification Algorithm Robust to Order Over Estimation*”, IEEE ICASSP, Istanbul, juin 2000.

Jean-Pierre Delmas, Houcem Gazzah and A. P. Liavas, “*Statistical Analysis of Second-Order Blind Channel Identification Under Practical Conditions, Application to the Least Squares and Subspace Methods*”, IMA Conference on Mathematics in Communications, Loughborough, UK, 15-17 décembre 1998.

Chapitre 2

Notations

On écrit a (resp. $\mathbf{a}$ et $\mathbf{A}$) un scalaire (resp. un vecteur et une matrice).

Soient $\mathbf{A}$ ($m \times n$) et $\mathbf{B}$ ($p \times q$) deux matrices. On définit (notations de [SCHO 97, §7]) le produit de Kronecker ($\mathbf{A} \otimes \mathbf{B}$) par la matrice $mp \times nq$ telle que le sous-blocs (i, j) est $[a_{ij}\mathbf{B}]$.

On définit l'opérateur linéaire $\text{Vec}(\cdot)$ tel que $\text{Vec}(\mathbf{A}) \triangleq [a_{11}a_{21} \cdots a_{m1}a_{12}a_{22} \cdots a_{mn}]^T$. On désigne par $\|\mathbf{a}\|$ la norme euclidienne du vecteur $\mathbf{a}$ définie par $\|\mathbf{a}\| \triangleq \sqrt{\sum_i |a_i|^2}$. Si $\mathbf{A}$ est une matrice carrée $n \times n$, on désigne par $\|\mathbf{A}\|$ sa norme euclidienne définie par $\|\mathbf{A}\| \triangleq \max_{\|\mathbf{a}\|=1} \|\mathbf{A}\mathbf{a}\|$ et par $|\mathbf{A}_n|$ sa norme de Frobenius définie par $|\mathbf{A}_n| \triangleq \sqrt{\frac{1}{n} \sum_{i=1}^n \sum_{j=1}^n |a_{i,j}|^2} = \frac{1}{\sqrt{n}} \|\text{Vec}(\mathbf{A})\|$.

On note $\sigma_p(\mathbf{A})$ (resp. $\lambda_p(\mathbf{A})$) la $p^{\text{ème}}$ plus grande valeur singulière (resp. propre) de la matrice (resp. de la matrice carrée) $\mathbf{A}$. En particulier, on note $\sigma(\mathbf{A}) \triangleq \sigma_{\text{rg}(\mathbf{A})}(\mathbf{A})$ (resp. $\lambda(\mathbf{A}) \triangleq \lambda_{\text{rg}(\mathbf{A})}(\mathbf{A})$) la plus petite valeur singulière (resp. propre non nulle). $\text{vecp}(\mathbf{A})$ (resp. $\text{VECP}(\mathbf{A})$) désigne le vecteur propre normé¹ associé à la plus petite (resp. la plus grande) valeur propre de la matrice carrée $\mathbf{A}$. $\text{Tr}(\cdot)$ et $\text{rg}(\cdot)$ désignent respectivement la trace et le rang d'une matrice.

$\lfloor x \rfloor$ (resp. $\lceil x \rceil$) d'un réel x désigne le plus grand (resp. le plus petit) entier $\in \mathbf{Z}$ inférieur (resp. supérieur) ou égal à x . $\text{E}(\cdot)$ désigne l'esperance d'une variable aléatoire.

On note les vecteurs et les matrices suivants :

$\mathbf{I}_l$ la matrice identité $l \times l$.

$\mathbf{0}_{k,l}$ la matrice nulle $k \times l$. On note $\mathbf{0}_k \triangleq \mathbf{0}_{k,k}$.

$\mathbf{I}_l^\dagger \triangleq \begin{bmatrix} 0 & \cdots & 0 & 1 \\ \vdots & & 1 & 0 \\ 0 & & \vdots & \\ 1 & 0 & \cdots & 0 \end{bmatrix}$ la matrice d'inversion $l \times l$.

$\mathbf{J}_l \triangleq \begin{bmatrix} 0 & 0 & \cdots & 0 \\ 1 & \ddots & \ddots & \vdots \\ \vdots & \ddots & 0 & 0 \\ 0 & \cdots & 1 & 0 \end{bmatrix}$ la matrice de décalage $l \times l$. On note $\mathbf{J}_l^{-k} \triangleq (\mathbf{J}_l^k)^T$.

¹Unique à une rotation près.

$\mathbf{e}_{L,l}$ le vecteur canonique de dimension L , tel que $(\mathbf{e}_{L,l})_i = \delta_{il}$.

$\mathbf{1}_l = [1, 1, \dots, 1]^T$, l -dim.

$\mathbf{K}_{k,l} \hat{=} \sum_{i=1}^k \sum_{j=1}^l (\mathbf{e}_{k,i} \mathbf{e}_{l,j}^T \otimes \mathbf{e}_{l,j} \mathbf{e}_{k,i}^T)$. On notera $\mathbf{K}_k \hat{=} \mathbf{K}_{k,k}$.

La matrice $\mathbf{K}$ vérifie les propriétés suivantes [SCHO 97, HEND 81]. Soient les matrices $\mathbf{A}$ ($m \times n$), $\mathbf{B}$ ($n \times p$) et $\mathbf{C}$ ($p \times q$) et les vecteurs $\mathbf{x}$ et $\mathbf{y}$.

$$\begin{aligned}
(\mathbf{x} \otimes \mathbf{y}^T) &= \mathbf{x} \mathbf{y}^T \\
\text{Vec}(\mathbf{ABC}) &= (\mathbf{C}^T \otimes \mathbf{A}) \text{Vec}(\mathbf{B}) \\
\text{Vec}(\mathbf{A}^T) &= \mathbf{K}_{m,n} \text{Vec}(\mathbf{A}) \\
(\mathbf{A} \otimes \mathbf{C}) &= \mathbf{K}_{m,p} (\mathbf{C} \otimes \mathbf{A}) \mathbf{K}_{q,n} \\
\mathbf{K}_{l,k} &= \mathbf{K}_{k,l}^T = \mathbf{K}_{k,l}^{-1} \\
\mathbf{K}_{k,1} &= \mathbf{I}_k \\
\mathbf{K}_{a,mn} &= \mathbf{K}_{am,n} \mathbf{K}_{an,m} \\
\text{Si } \mathbf{A} \text{ est } m \times m \text{ et Hermitienne, alors } \text{Vec}(\mathbf{A}^*) &= \mathbf{K}_m \text{Vec}(\mathbf{A})
\end{aligned} \tag{2.1}$$

Chapitre 3

Égalisation synchrone en communication numérique

Les techniques numériques de communications sont en passe de devenir le mode privilégié pour les produits et systèmes de télécommunications. Les communications numériques offrent des possibilités vastes de traitement à tous les niveaux de la chaîne de communication: stockage, compression, cryptage, multiplexage, et surtout lutte contre les erreurs de transmission. Ce but est servi par un nombre de techniques mises en service à la fois à la transmission et à la réception: codage et décodage canal, filtre de mise en forme et réception synchrone, séquence d'apprentissage et réception supervisée. Les progrès enregistrés ont permis d'améliorer la qualité des communications et de proposer de nouveaux services.

Dans ce chapitre, on rappelle des éléments de la théorie des communications nécessaires à l'introduction du problème d'égalisation des canaux de communications. On s'y limite aux récepteurs synchrones fonctionnant au rythme symbole; conçus dans le but de recouvrir la séquence des symboles émis à partir des observations tirées de l'échantillonnage au rythme symbole (à la fréquence de Baud) du signal reçu.

3.1 Canal de communication

3.1.1 Chaîne de communication numérique

Comme montré dans Fig. 3.1, la source d'information numérique subit en premier lieu un codage source qui consiste à compresser (sans ou avec pertes) les données pour minimiser la quantité de symboles à transmettre. Ensuite, un codage canal va injecter de la redondance qui à la réception va permettre de s'assurer si la transmission s'est effectuée sans erreurs et éventuellement de les corriger. Les trains de symboles de plusieurs sources coopératrices peuvent être associés pour partager le même support de transmission (multiplexage). Le modulateur génère un signal qui varie en fonction du multiplex ainsi formé et qui occupe une bande spectrale adaptée aux possibilités de transmission sur le support de communications. Les variations du signal modulé sont interprétées à la réception pour

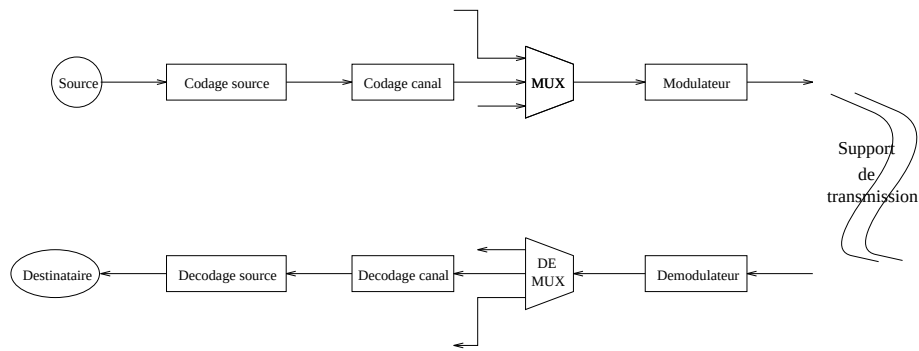


Figure 3.1: Chaîne de communication numérique

déduire les bits transmis, la démodulation. Le modulateur peut éventuellement assurer un étalement en fréquence. Le signal ainsi formé est continu dans le temps et est mis en contact avec le support physique de communication. A ce point, la suite des équipements et opérations engagés constitue l'émetteur.

Le récepteur procède en sens inverse à des tâches similaires. Le signal transmis est reçu à travers un ou plusieurs antennes, conçues et disposées pour capter un maximum de l'énergie utile. Le démodulateur effectue la transposition en fréquence et livre, après synchronisation, des échantillons du signal pris à intervalle de temps régulier. Ces échantillons vont être traités (égalisation) puis comparés à certaines valeurs discrètes d'un alphabet fini désigné par constellation. Décodage canal et source permettent finalement de récupérer au mieux la séquence transmise.

3.1.2 Effets du canal

Tout se serait bien déroulé sans les effets indésirables qui se produisent pendant le parcours du support physique. Le signal arrive affecté par un certain nombre de modifications.

La première est le bruit additif. Il englobe toutes les émissions externes les plus diverses (bruit cosmique, activités électriques de l'atmosphère, autres transmissions, ...) et le bruit généré dans les circuits du récepteur lui même (bruit thermique et impulsif). L'activité de ces sources est supposée être indépendante de la transmission en cours et ne pas évoluer au cours du temps. Elle est souvent modélisée par un processus aléatoire stationnaire gaussien blanc et centré caractérisable par sa seule puissance.

Les performances d'un système de télécommunications doivent se dégrader avec l'aggravation du niveau de bruit. Dans les systèmes multi-utilisateurs, ceci implique une dégradation de la qualité de service à mesure que des utilisateurs se connectent au réseau. Les recherches actuelles s'orientent vers la prise en compte de ces interférences en tant que telles et la prise en considération des propriétés structurales de ces signaux interférant dans la procédure de réception. Ces techniques sont souvent désignées par *détection multi-utilisateurs*. Notre travail s'inscrit en dehors de ce cadre.

Le signal transmis peut être affecté par des distorsions le plus souvent considérées comme linéaires. Ceci se produit lorsque le canal de transmission, dit alors linéaire, vérifie la propriété de superposition

i.e., lorsqu'une combinaison linéaire de signaux se présente à son entrée, il génère la même combinaison linéaire des signaux sortie qui seraient générés par ces signaux entrée se présentant individuellement. Un canal de ce type est le canal à trajet multiples souvent utilisé pour modéliser les conditions de transmission radio dans un environnement où les obstacles se comportent comme des réflecteurs qui font apparaître des versions affaiblies et retardées du signal d'origine.

Ces distorsions linéaires se traduisent alors par la présence à la réception à un instant donné du symbole émis courant mais également de répliques des symboles émis à d'autres instants. Cette interférence inter symboles (ISI¹) est un problème majeur de la communication numérique. Sa suppression, ou du moins sa réduction, est une tâche essentielle du récepteur.

Ce problème s'accroît dans le cas des systèmes de communications avec les mobiles à cause de la diversité (et de l'évolution permanente) du canal physique qui sépare l'émetteur du récepteur, station de base ou terminal selon qu'il s'agit du lien montant ou descendant. En effet, émetteur et récepteur sont rarement en visibilité directe. Les obstacles qui s'interposent sont de deux types. Les obstacles les plus grands (grands bâtiments, montagnes, ...) sont à l'origine chacun d'un trajet. Ces trajets arrivent au niveau du récepteur séparés de plus d'une durée symbole. Ces obstacles sont responsables de l'ISI. Des obstacles plus petits et plus nombreux (personnes, véhicules, arbres, constructions, ...) se trouvent au voisinage immédiat du récepteur. De part leur proximité, les trajets qu'ils génèrent arrivent presque au même instant. Leur effet principal est d'une autre nature puisque à cause de leur mobilité ou celle du récepteur, leur composition varie dans le temps. L'effet de ces obstacles secondaires se traduit par la non stationnarité du canal de transmission, d'autant plus importante que la vitesse de déplacement du mobile augmente.

Le canal peut cependant être non linéaire. C'est le cas par exemple des fibres optiques et guides d'ondes. Entre autres, ces milieux sont le siège de phénomènes de résonances qui génèrent des fréquences harmoniques du signal d'entrée. Le signal est plus riche en fréquences à la sortie qu'à l'entrée. Ce type de distorsions n'est pas adressé dans ce travail, les canaux étant supposés linéaires. Également, la manière avec laquelle le canal altère les signaux transmis peut varier dans le temps (à cause de la mobilité des utilisateurs d'un système de communications avec les mobiles ou bien à cause de l'irruption d'obstacles ou des changements climatiques pour les systèmes de faisceaux hertziens). Ceci se traduit par des chutes brutales du niveau de puissance du signal reçu [KNDY] dits *fading* ou évanouissements rapides.

3.1.3 Équivalent bande de base

Lorsqu'on emploie des modulations linéaires, le plus souvent la modulation en quadrature (QAM²), on montre que la chaîne de communication peut être modélisée comme dans Fig. 3.2.a i.e., équivalente à une transmission en bande de base (i.e., sans modulation) de symboles complexes $s(k)$ à travers un canal $c(t)$ complexe.

La source englobe les symboles à transmettre après codage source et canal et autres manipulations qui n'altèrent pas la nature numérique ou discrète des symboles (multiplexage, entrelacement, brouillage,

¹Inter Symbol Interference

²Quadrature Amplitude Modulations

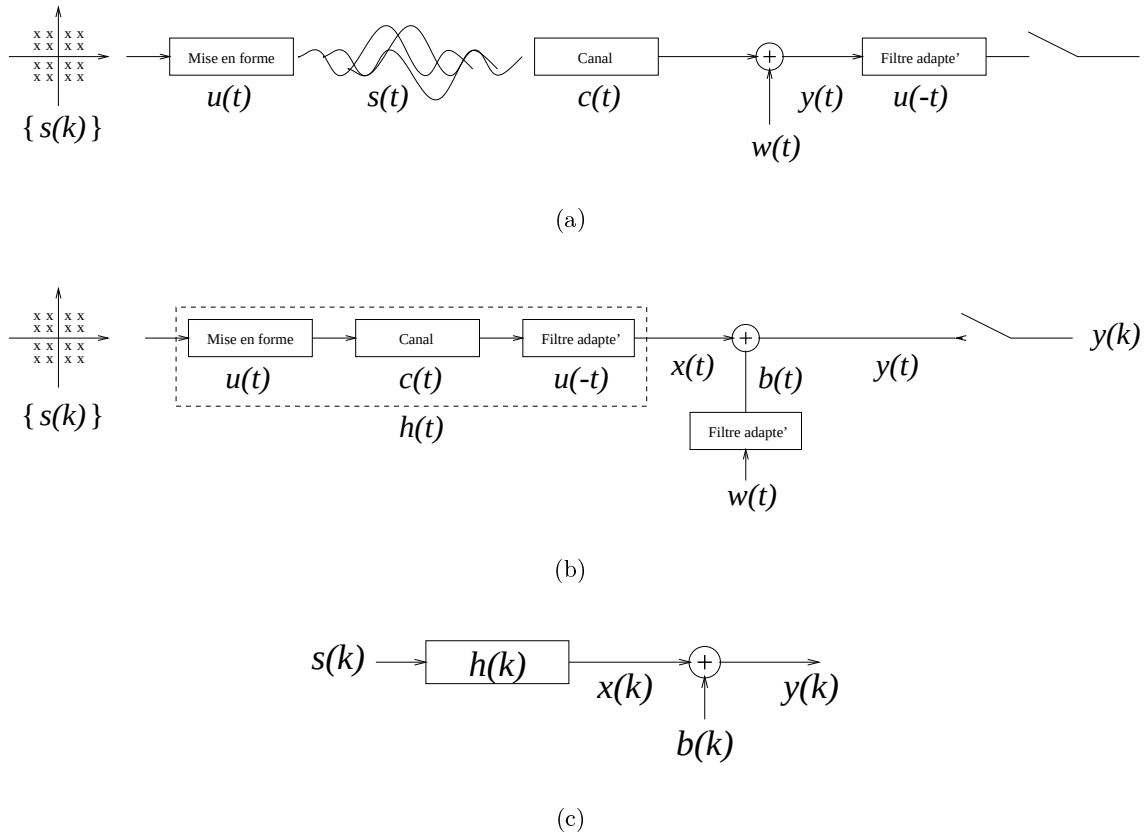


Figure 3.2: Schéma numérique équivalent bande de base

chiffrement, ...). La source est dite à alphabet fini car le nombre de valeur possibles (dits états) que peut prendre un symbole donné est fini. L'ensemble des états forme la constellation. La source est supposée générer des symboles i.i.d. (indépendants et identiquement distribués).

À partir de ces symboles, un filtre d'émission (ou aussi filtre de mise en forme) fabrique un signal continu dans le temps $s(t) = \sum_k s(k)u(t - kT)$. $s(k)$ étant le symbole émis par la source à l'instant kT . $u(t)$ est le signal de mise en forme. Le signal ainsi généré est mis en contact avec le canal.

Le canal $c(t)$ inclut en plus du canal physique, les étages de modulation et de démodulation. Il est souvent modélisé par un filtre linéaire et invariable dans le temps. Son effet se traduit par l'étalement dans le temps du signal émis. Les sources de nuisances sont traitées comme un bruit additif qui entache l'observation de ce canal. On fait l'hypothèse que le bruit discret $b(k)$ est blanc³ malgré, qu'en fait, il est coloré à cause du filtre de réception.

Le récepteur échantillonne le signal reçu à la fréquence symbole à partir d'un instant t_0 . Le filtre de réception, dit aussi filtre adapté blanchisseur a un double rôle. D'une part, il est adapté au filtre d'émission de manière à maximiser le rapport signal sur bruit à l'entrée de l'échantillonneur pour l'instant d'échantillonnage précis. D'autre part, il décorrèle les échantillons du bruit. Ces fonctions ne sont jamais parfaitement achevées du fait que canal et bruit sont inconnus. Puisque l'instant d'échantillonnage intervient dans la conception du filtre de réception, une synchronisation plus ou moins parfaite est nécessaire pour commander l'échantillonneur. Des techniques de verrouil-

³au second ordre i.e., décorrélé.

lage de phase (PLL⁴) sont mises en oeuvre pour mettre en phase l'échantillonneur avec la séquence des symboles émis. Les échantillons ou observations (entachés de bruits) sont par la suite traités pour reconstituer les symboles émis.

Fig. 3.2.a à Fig. 3.2.c montrent que ce schéma peut être ramené à un schéma numérique équivalent en bande de base, celui de Fig. 3.2.c, où la sortie discrète $y(n)$ est une version bruitée de la filtrée des symboles émis par un filtre numérique $h(n)$ qu'on désigne désormais de canal et qu'il convient d'annuler l'effet, ainsi que celui du bruit. Une partie du récepteur, le filtre adapté, est passée dans le canal. Le récepteur numérique se charge de reconstituer les symboles émis. Cette tâche d'égalisation peut être réalisée par un seul bloc, le decodeur de Viterbi, ou plus couramment par un filtre linéaire dont la sortie approxime les symboles émis, suivi d'un bloc de décision qui lève le restant de l'indétermination sur ces symboles en les choisissant dans l'alphabet de la source. Dans ce dernier cas, on appelle égaliseur le seul filtre linéaire.

3.1.4 Interference inter symboles

D'après Fig. 3.2.b, lorsque le signal traverse le canal de réponse impulsionnelle $h(t)$, on récupère à la sortie le signal $y(t) = \sum_k s(k)h(t - kT) + b(t)$ où $b(t)$ représente le bruit d'observation. On a $y(n) \hat{=} y(t_0 + nT) = \sum_k s(k)h(t_0 + (n - k)T) + b(t_0 + nT) = s(n)h(t_0) + \sum_{k \neq n} s(k)h(t_0 + (n - k)T) + b(t_0 + nT)$ donc l'observation à l'instant nT contient, en plus du signal émis à cet instant et du bruit, le terme $\sum_{k \neq n} s(k)h(t_0 + (n - k)T)$ qui traduit l'effet indésirable des symboles voisins appelé interférence intersymboles (ISI).

Pour prévenir l'ISI, le canal $h(t)$ doit remplir la condition dite de Nyquist se traduisant par $h(t_0 + (n - k)T) = 1$, si $n = k$, 0 si non. Une conséquence est que si $[-f_h, f_h]$ est la bande occupée par le canal, alors elle doit vérifier $f_h \geq \frac{1}{2T}$. Une famille populaire de signaux satisfaisant le critère de Nyquist est celle des cosinus sur-élevés. Le respect du critère de Nyquist n'est pas possible en pratique à cause de l'indétermination du canal physique et l'impossibilité d'achever une synchronisation parfaite. Il s'en suit que la présence d'ISI est incontournable et que des techniques, dites d'égalisation, pour la combattre doivent être développées.

On utilise l'ouverture moyenne de l'oeil pour apprécier l'ISI (ou l'ISI résiduelle, après une égalisation) indépendamment des symboles courants. Elle est définie par $\frac{\sum_k |h(t_0 + kT)|^2 - \max_k |h(t_0 + kT)|^2}{\max_k |h(t_0 + kT)|^2}$. Une autre définition courante est $\frac{\sum_k |h(t_0 + kT)| - \max_k |h(t_0 + kT)|}{\max_k |h(t_0 + kT)|}$.

3.2 Égalisation linéaire supervisée

3.2.1 Structures et critères d'égalisation

Un égaliseur appliqué à la sortie de l'échantillonneur va tenter de fabriquer à partir d'une suite de mesures du canal, un symbole (ou bien une séquence de symboles) qui *ressemble* aux mieux au symbole (ou bien à la séquence de symbole) émise, moyennant un délai donné.

⁴Phase Looked Loop

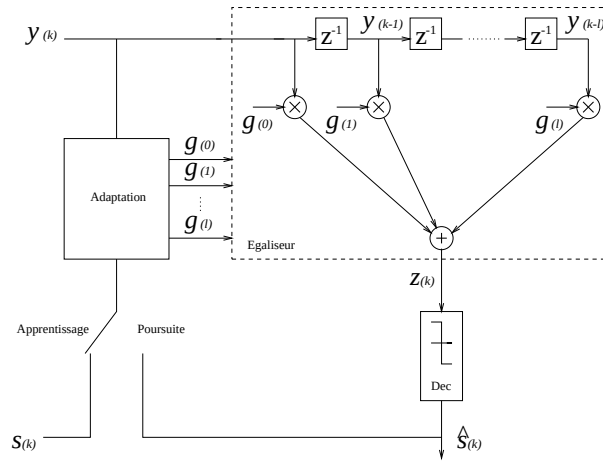


Figure 3.3: Égaliseur linéaire

Le degré de ressemblance est décrit par un critère qui compare les symboles reconstitués aux symboles émis. Le plus pertinent pour les applications de télécommunications est la probabilité d'erreur. Minimiser la probabilité d'erreur est équivalent à maximiser la probabilité à posteriori (MAP⁵). Lorsque les symboles émis sont équiprobables (des techniques de brouillage et d'entrelacement sont utilisées pour ce faire), ce critère est équivalent au critère du maximum de vraisemblance (ML⁶). Lorsque le bruit est blanc et gaussien, l'estimateur optimal au sens ML est l'estimateur de séquence à maximum de vraisemblance (MLSE⁷) lorsque la longueur de la séquence tend vers l'infini. La MLSE est réalisée par l'algorithme de Viterbi. Un tel égaliseur a, cependant, le défaut d'avoir une structure non linéaire et surtout une complexité qui croît exponentiellement avec le nombre d'états. D'autres critères sous-optimaux mais plus simples sont proposés pour la conception des égaliseurs. Il s'agit du critère du Zero-Forcing (ZF) et du critère de l'erreur quadratique moyenne minimale (MMSE⁸) qui permettent d'atteindre des performances proches de l'optimum lorsque les conditions de réception sont favorables (fort rapport signal sur bruit notamment).

Lorsque la réception se fait en présence de bruit, une reconstitution parfaite (ou parfait rime toujours avec *à un délai près*) n'est pas possible sans utiliser des techniques non linéaires. Ou bien l'égaliseur a une structure non linéaire (détecteur de Viterbi); ou bien (cf. Fig. 3.3) il est composé d'un premier bloc linéaire suivi d'un deuxième bloc (de décision) non linéaire (un comparateur ou *slicer* typiquement).

Les systèmes de communications tendent de plus en plus vers l'utilisation de constellation denses, doivent intégrer des utilisateurs nombreux et sont dimensionnés pour affronter des canaux sévères, notamment en termes d'ISI. Les récepteurs de Viterbi sont de ce point de vue mal adaptés. Les structures linéaires sont nettement plus faciles à modéliser et à manipuler, en particulier pour des mises en oeuvre adaptatives. L'algorithme d'égalisation calcule les coefficients du filtre en fonction des observations du canal et éventuellement en connaissance de certains des symboles émis (égalisation supervisée) ou des symboles reconstitués si ces derniers sont crédibles (égalisation à retour de décision ou DFE⁹).

⁵Maximum a Posteriori Probability

⁶Maximum Likelihood

⁷Maximum Likelihood Sequence Estimation

⁸Minimum Mean Square Error

⁹Decision Feedback Equalizer

Cet ajustement des coefficients du filtre égaliseur se fait en fonction du critère de reconstitution : ZF ou MMSE.

Un égaliseur linéaire, d'ordre l , qui permet de retrouver le symbole transmis avec un délai $i \geq 0$ est noté $\mathbf{g} = [g(0), \dots, g(l)]^T$. On note $g(z) = \sum_k g(k)z^{-k}$ sa transformée en z . Pour des considérations pratiques, l'égaliseur est à réponse impulsionnelle finie (FIR¹⁰).

Égaliser un canal $h(z)$ au sens ZF consiste à choisir $g(z) = \frac{z^{-i}}{h(z)}$. Puisque le canal $h(z)$ est souvent un FIR, son inverse est à réponse impulsionnelle infinie (IIR¹¹). $g(z)$ en est une approximation FIR et n'assure donc pas une inversion parfaite du canal. La sortie de l'égaliseur est $s(i) + \sum_k b(k)g(i-k)$. En présence de bruit, certaines composantes du bruit peuvent se trouver significativement amplifiées. En effet, lorsque le canal est sélectif en fréquence, l'égaliseur va présenter des gains importants pour récupérer les fréquences affaiblies par le canal. Le bruit est amplifié par la même occasion. Cette amplification du bruit (*noise enhancement*) amoindrit l'intérêt pour l'égalisation ZF.

L'égalisation MMSE fixe les coefficients de l'égaliseur afin de minimiser l'erreur quadratique moyenne EQM (MSE¹²) entre le symbole reconstitué et le symbole effectivement émis i.e.,

$$\mathbf{g} = \operatorname{argmin} \mathbb{E} \left(|z(n) - s(n-i)|^2 \right)$$

La MSE est minimale lorsque l'erreur $z(n) - s(n-i)$ est décorrélée de $[y(n), \dots, y_{n-l}]$ ce qui mène à avoir

$$\mathbb{E} \left([y(n), \dots, y(n-l)]^H [y(n), \dots, y(n-l)] \right) \mathbf{g} = \mathbb{E} \left([y(n), \dots, y(n-l)]^H s(n-i) \right) \quad (3.1)$$

Lorsque le canal est à réponse impulsionnelle finie, on a

$$\begin{bmatrix} y(n) \\ \vdots \\ y(n-l) \end{bmatrix} = \begin{bmatrix} h(0) & \dots & h(M) & 0 & \dots \\ 0 & h(0) & & h(M) & \ddots \\ \vdots & \ddots & \ddots & \ddots & \\ \dots & 0 & h(0) & \dots & h(M) \end{bmatrix} \begin{bmatrix} s(n) \\ \vdots \\ s(n-M-l) \end{bmatrix} + \begin{bmatrix} b(n) \\ \vdots \\ b(n-l) \end{bmatrix}.$$

Étant donné que le bruit est décorrélé de l'entrée (blanche) du canal, on obtient

$$\begin{aligned} & \mathbb{E} \left(\begin{bmatrix} h(0) & \dots & h(M) & 0 & \dots \\ 0 & h(0) & & h(M) & \ddots \\ \vdots & \ddots & \ddots & \ddots & \\ \dots & 0 & h(0) & \dots & h(M) \end{bmatrix}^* \begin{bmatrix} h(0) & \dots & h(M) & 0 & \dots \\ 0 & h(0) & & h(M) & \ddots \\ \vdots & \ddots & \ddots & \ddots & \\ \dots & 0 & h(0) & \dots & h(M) \end{bmatrix}^T + \sigma_b^2 \mathbf{I}_{l+1} \right) \mathbf{g} \\ &= \begin{bmatrix} h(i) \\ \vdots \\ h(i-l) \end{bmatrix}^* \end{aligned} \quad (3.2)$$

Plus généralement, lorsque le canal est à réponse impulsionnelle infinie, on montre que

$$g(z) = \frac{\sigma_s^2 z^{-i} h^*(z^{-1})}{\sigma_s^2 h(z) h^*(z^{-1}) + \sigma_b^2}$$

¹⁰Finite Impulse Response

¹¹Infinite Impulse Response

¹²Mean Square Error

Son expression traduit le compromis que doit respecter l'égaliseur MMSE entre la correction des distorsions du canal et une amplification limitée du bruit. L'égaliseur MMSE coïncide avec l'égaliseur ZF en l'absence de bruit.

Les performances (sous-optimales) des égaliseurs linéaires peuvent varier sensiblement selon le choix de la longueur de l'égaliseur et du retard d'égalisation. Alors qu'il va de soi que plus long l'égaliseur, meilleure est l'égalisation; le choix du retard est plus problématique. En pratique, il est le plus souvent constaté (des contre-exemples cependant existent) que les délais extrêmes, minimaux ($i = 0$) et maximaux ($i = M + l$), se montrent défavorables et que les délais moyens offrent de meilleures performances [Touz 98].

3.2.2 Implémentation de l'égalisation linéaire MSE

D'après (3.2), le calcul de l'égaliseur linéaire MMSE $\mathbf{g}$ nécessite la connaissance de la réponse impulsionnelle du canal. Ainsi, une première approche, dite indirecte, passe par l'estimation de la réponse du canal suivie d'une estimation de l'égaliseur MMSE correspondant en utilisant (3.2). Cependant, l'approche suivante, dite directe, lui est préférée. Plus simple, elle se prête mieux à une implémentation adaptative. Ne faisant pas d'hypothèses sur la longueur ni sur la structure du canal, elle est robuste aux erreurs de modélisation de ce dernier. Elle exploite (3.1) pour calculer l'égaliseur MMSE. La connaissance du canal $\mathbf{y}$ est remplacée par la connaissance des bits émis. Ces derniers forment la séquence d'apprentissage¹³. Dans (3.1), les espérances sont remplacées par les moyennes empiriques tirées des observations et symboles courants.

Les situations pratiques correspondent souvent à des canaux variables au court du temps. Pour faire face, l'estimée de l'égaliseur doit être mise à jour continuellement, à l'occasion de toute nouvelle mesure du canal le plus souvent. Un algorithme *adaptatif* ajuste à chaque instant nT les coefficients de l'égaliseur $\mathbf{g}^{(n)}$

$$\mathbf{g}^{(n)} = \mathbf{g}^{(n-1)} + \mu \Delta^{(n)} \left(s(n-i) - [y(n), \dots, y(n-l)] \mathbf{g}^{(n-1)} \right) \quad (3.3)$$

La suite des égaliseurs doit converger, avec le temps, vers la solution de (3.1). (3.3) exprime que les coefficients du filtre sont actualisés à partir des anciennes valeurs, et sont explorés dans la direction $\Delta^{(n)}$, proportionnellement à l'erreur de reconstruction et au pas μ choisi. Le pas μ contrôle la vitesse et la précision (augmenter μ fait augmenter la première et affaiblir la deuxième) avec lesquelles l'algorithme converge vers la solution optimale de (3.1). $\Delta^{(n)}$ est choisi de manière à s'annuler lorsque l'erreur $z(n) - s(n-i)$ entre le symbole égalisé et le symbole émis (de la séquence d'apprentissage) s'annule. Le choix de $\Delta^{(n)}$ dépend de l'algorithme adaptatif. Les plus connus sont l'algorithme du gradient stochastique (LMS¹⁴) et l'algorithme des moindres carrés récursifs (RLS¹⁵).

Quand la phase d'apprentissage du canal est terminée et les symboles transmis ne sont plus connus, l'algorithme adaptatif continue de se dérouler de la même manière, en remplaçant dans (3.3) les symboles émis par leurs estimations $\hat{s}(n)$ prises par le dispositif de décision. Dans ce cas, on dit

¹³training sequence

¹⁴Least Mean Square

¹⁵Recursive Least Square

que l'algorithme est entré dans un mode entraîné (DD¹⁶) qui va lui permettre de suivre des variations pas trop rapide du canal jusqu'à la nouvelle phase d'apprentissage. Le bon fonctionnement du mode DD nécessite que les symboles soient suffisamment bien reconstitués à la sortie de l'égaliseur pour qu'après le bloc de décision, ils correspondent exactement aux symboles effectivement émis. Autrement, une erreur de décision se propage vers les décisions suivantes et provoque la dérive de l'égaliseur. Cette propagation d'erreur limite l'emploi du mode DD aux conditions de transmission favorables i.e., des variations lentes du canal et un bruit faible. Lorsque le canal reste faiblement variable entre deux séquences d'apprentissage successives (c'est le cas, par exemple, du système GSM pour qui les trames sont suffisamment courtes), l'égaliseur n'est plus mis à jour après le passage de la séquence d'apprentissage et le mode DD n'est pas engagé.

L'algorithme LMS est un algorithme de gradient stochastique i.e., qui minimise la fonction coût en explorant la direction opposée du gradient instantané. Il s'en suit que $\Delta^{(n)} = [y(n), \dots, y(n-l)]^H$. Pour l'algorithme RLS, $\Delta^{(n)} = \mathbf{K}(n) [y(n), \dots, y(n-l)]^H$ où $\mathbf{K}(n)$ est une estimée de $\left[\mathbb{E} \left([y(n), \dots, y(n-l)]^H [y(n), \dots, y(n-l)] \right) \right]^{-1}$ i.e., de l'inverse de la conjuguée de la matrice de corrélation des observations. L'algorithme RLS converge plus rapidement vers la solution (3.1) mais souffre d'une complexité calculatoire plus importante¹⁷ et de sa non robustesse à la non blancheur du bruit d'observation [MACC 98].

Étant donné que l'égaliseur achève en quelque sorte l'inversion du canal qui a une réponse finie et souvent courte, l'égaliseur est amené à être plutôt long et de toute les manières bien plus long que la réponse du canal et ainsi, il s'estime moins bien. C'est un inconvénient principal de l'approche indirecte. On verra (cf. §4) que les techniques spatio-temporelles permettent d'éviter cet inconvénient en rendant possible l'égalisation d'un canal FIR par un égaliseur FIR également, et de longueur comparable.

3.3 Égalisation aveugle

L'égalisation aveugle considère la reconstitution des symboles émis sur la base des seules observations bruitées $y(n)$ de la sortie du canal. Les statistiques du second ordre (SOS¹⁸) de $y(n)$ sont données par $S_y(w) \triangleq \sum_k \mathbb{E} (y(n+k)y^*(k)) e^{-jkw} = |h(w)|^2 S_s(w)$. Elles ne contiennent donc pas d'information sur la phase du canal et ne peuvent suffire à elle seules pour l'identifier. L'information de phase apparaît dans les statistiques d'ordre supérieures (HOS¹⁹) du canal rendant possible son égalisation en aveugle. Ceci suppose que les HOS ne sont pas nulles donc que la sortie du canal n'est pas gaussienne, i.e., que les symboles émis ne sont pas gaussiens.

L'information sur la phase du canal peut être ressortie dans les SOS en utilisant des techniques de réception en diversité spatiale (multi-capteurs) et/ou temporelle (réception fractionnée). Ce rapport est consacré à cette approche et l'abordera à partir du chapitre suivant. Pour l'instant, on fait le rappel suivant des techniques d'égalisation aveugle au rythme symbole et en mono-capteur se basant

¹⁶Decision Directed

¹⁷Des versions rapides sont cependant proposés.

¹⁸Second Order Statistics

¹⁹Higer Order Statistics

sur les HOS. Il est à noter que des adaptations des algorithmes HOS aux récepteurs en diversité (spatio-temporelle) ont été proposées, par exemple l'algorithme CMA fractionnel [FIJA 97].

Les algorithmes suivants se basent sur des fonctions de coût, fonction des paramètres de l'égaliseur et dont il est prouvé qu'elles sont minimales lorsque l'égaliseur assure une inversion (au sens ZF) du canal à un retard et une rotation des symboles près. Les implémentations adaptatives utilisent le plus souvent l'algorithme du gradient stochastique.

Le premier des algorithmes aveugles est celui de *Sato* [SATO 75] pour l'égalisation des canaux transmettant des symboles modulés en amplitude (PAM²⁰). Les coefficients de l'égaliseur sont ajustés en fonction du coût suivant

$$J(\mathbf{g}) \triangleq \mathbb{E} \left(\left| \frac{\mathbb{E}(s(n)^2)}{\mathbb{E}(|s(n)|)} \text{Sgn}(z(n)) - z(n) \right|^2 \right)$$

Ensuite, *Godard* a proposé dans [GODR 80] une famille d'algorithmes bâtis autour des fonctions coût

$$J(\mathbf{g}) \triangleq \frac{1}{2r} \mathbb{E} \left(\left[|z(n)|^r - \frac{\mathbb{E}(|s(n)|^{2r})}{\mathbb{E}(|s(n)|^r)} \right]^2 \right)$$

dont le cas $r = 2$ est le plus populaire et est connu sous le nom de CMA²¹. Dans ce dernier cas, *Godard* montre qu'en l'absence de bruit, cette fonction coût est minimale lorsque la sortie de l'égaliseur restitue les données émises à un retard et une rotation près. Ceci est obtenu sous les hypothèses suivantes sur les symboles émis qui doivent être équiprobables, décorrélés, circulaires²² et vérifient la condition dite du Kurtosis $K(s(n)) \triangleq \mathbb{E}(|s(n)|^4) - 2(\mathbb{E}(|s(n)|^2))^2 - |\mathbb{E}(s(n)^2)|^2 = \mathbb{E}(|s(n)|^4) - 2(\mathbb{E}(|s(n)|^2))^2$ négatif. Sous cette contrainte, la séquence émise est dite sous-gaussienne (cette condition est vérifiée par les symboles MAQ). L'implémentation adaptative, utilisant le gradient stochastique, conduit au calcul de $\mathbf{g}^{(n)} = \mathbf{g}^{(n-1)} + \mu [\mathbb{E}(|s(n)|^4) - |z(n)|^2] z(n) [y(n), \dots, y(n-l)]^H$ ce qui traduit que l'algorithme tend à donner un module constant aux symboles reconstitués, d'où l'appellation CMA.

Benevise et al montrent [BENE 80] que l'égalisation est achevée si la densité de probabilité de la sortie de l'égaliseur est égale à celle de l'entrée du canal. *Shalvi et Weinstein* prouvent qu'il est suffisant (et bien sur nécessaire) de récupérer à la sortie de l'égaliseur une variance et un Kurtosis égaux à ceux de l'entrée. Si $\mathbf{r}$ dénote la réponse globale (canal - égaliseur), ceux-ci sont liés par

$$\begin{aligned} \mathbb{E}(|z(n)|^2) &= \mathbb{E}(|s(n)|^2) \|\mathbf{r}\|^2 \\ K(z(n)) &= K(s(n)) \sum_{n=0}^{M+l} |r(n)|^4 \end{aligned}$$

Ainsi, si $\mathbb{E}(|z(n)|^2) = \mathbb{E}(|s(n)|^2)$, alors $|K(z(n))| \leq |K(s(n))|$. L'égalité n'étant atteinte que si $\mathbf{r}$ correspond à un retard et une rotation. Ainsi, le critère proposé revient à maximiser le (module du) Kurtosis de la sortie de l'égaliseur sous la contrainte de conserver la variance :

$$\max_{\mathbb{E}(|z(n)|^2) = \mathbb{E}(|s(n)|^2)} |K(z(n))|$$

²⁰Pulse Amplitude Modulation

²¹Constant Modulus Algorithm

²²i.e., $\mathbb{E}(s(n)^2) = 0$ ce qui est vérifiée lorsque les symboles sont pris dans des constellations symétriques; le cas de toutes les constellations pratiques

Les algorithmes aveugles ci-dessus achèvent l'égalisation à un retard indéterminé qui, en pratique, ne pose pas problème. Étant bâtis autour de fonctions coût, ils sont souvent réalisés sous forme d'algorithmes adaptatifs, du type gradient stochastique le plus souvent²³. Les inconvénients suivants sont partagés par les algorithmes d'égalisation aveugle au rythme Baud :

- Les canaux ont besoin d'être observés pendant des durées importantes (plusieurs centaines, voire quelques milliers de durées symboles) avant que les HOS ne soient suffisamment bien estimées pour obtenir de bonnes performances d'égalisation.
- Les algorithmes aveugles ci-dessus se basent sur la minimisation de fonction coût non quadratiques. Le problème de convergence vers des minima locaux leur est fréquent [DING 91, DING 92] et constitue un handicap majeur à leur exploitation en pratique.
- La blancheur des symboles émis est requise au moments d'ordre supérieurs. Elle est nettement plus difficile à obtenir que pour les moments du second ordre. Elle nécessite la mise en oeuvre de brouilleurs particulièrement élaborés [MACC 98].

3.4 Conclusion

Dans ce chapitre, on est parti du schéma de la communication numérique en bande de base pour établir un modèle discret équivalent qui représente le cadre d'étude et de conception des schémas de communication. On a mis en évidence l'interférence inter symboles comme étant le dégât principal subi par les symboles émis à l'occasion de leur passage à travers le canal. La lutte contre cet effet, désignée par égalisation, est l'une des tâches principales du récepteur. Pour des raisons de complexité, elle est effectuée par une structure le plus souvent linéaire. Le filtre égaliseur cherche à reconstituer au mieux la suite des symboles émis au sens d'un critère : le Zero Forcing ou la MMSE, très comparables lorsque le bruit qui affecte les observations du canal est faible. Outre l'observation du canal, l'égaliseur, implémenté sous la forme d'algorithme adaptatifs, a besoin de connaître les symboles émis (la séquence d'apprentissage) pendant une certaine période avant de pouvoir converger vers une situation où l'ISI résiduelle est devenue acceptable (oeil ouvert). C'est alors qu'un comparateur fini d'achever l'égalisation. Les symboles ainsi estimés sont assez crédibles pour se substituer à ceux de la séquence d'apprentissage dans l'algorithme adaptatif.

Reste que ces techniques d'égalisation supervisée (par la séquence d'apprentissage) deviennent incompatibles avec le développement de systèmes de communications gourmands en bande passante et des modes de communication non coopératifs. L'intégration des techniques de diversité (spatiales et/ou temporelle) dans les systèmes de communications avec les mobiles notamment fait augmenter la longueur des séquences d'apprentissage [LIND 99, NAGB 96] et par conséquent l'intérêt pour les techniques d'égalisation aveugle.

La sortie du canal échantillonnée au rythme symbole ne contient pas d'information sur la phase du canal que dans ses statistiques d'ordre supérieur (à deux). Les algorithmes qui se basent sur celles-ci

²³Pour qu'une telle implémentation soit possible, *Shalvi et Weinstein* proposent une version sans contrainte de leur critère

convergent lentement; vers des points de stabilité indésirables parfois. Un pas important a été franchi par l'introduction des techniques de réception en diversité (spatiales et/ou temporelle) qui permettent de provoquer l'apparition de l'information de phase du canal dans les statistiques d'ordre deux. Le reste du rapport est consacré à ces techniques.

Chapitre 4

Identification/égalisation spatio-temporelle

Le traitement spatio-temporel désigne les techniques de réception (et éventuellement de transmission) qui utilisent un réseau d'antennes (réception multi-capteurs) pour la réception d'un signal et/ou un échantillonnage de ce signal à un rythme supérieur à celui des symboles qu'il transmet (réception fractionnée). Ces techniques, dites aussi réception en diversité spatio-temporelle ou aussi multivariable, offrent la possibilité d'améliorer sensiblement les performances de tâches standards telles que synchronisation, contrôle de puissance, égalisation supervisée mais surtout rendent possible l'égalisation aveugle au second ordre, impossible lorsque le signal est reçu avec une seule antenne et échantillonné à la fréquence symbole. Ce gain en performance se paye en le besoin d'une puissance de calcul importante. Dans le contexte des systèmes de communications avec les mobiles, ces techniques sont destinées en premier lieu aux stations de base, où une telle puissance peut être supposée disponible, ainsi que l'emplacement et le dégagement nécessaires pour les antennes.

4.1 Notations et résultats

4.1.1 Le canal SIMO

On se place dans le contexte de l'égalisation d'un canal de réponse impulsionnelle $f(t)$ excité par une suite de symboles $s(k)$ et affecté par le bruit $b(t)$. Le récepteur spatio-temporel est constitué d'un réseau de c_1 capteurs de réponses impulsionnelles $a^i(t), i = 1, \dots, c_1$. Chacun des capteurs est suivi d'un échantillonneur qui prélève des mesures à la cadence $\frac{c_2}{T}$.

Fig. 4.3.a à Fig. 4.3.d montrent que ce schéma est équivalent à un canal à une entrée et plusieurs sorties (SIMO¹). L'entrée $s(k)$ excite simultanément les $c \triangleq c_1 c_2$ filtres discrets $\mathbf{h}_M^{c'}$ dont le $k^{\text{ème}}$ coefficient est donné par $\mathbf{h}^{c'}(k) \triangleq (f * a^{c'_1}) \left((\frac{c'_2}{c_2} + k)T \right)$ i.e., est la $(c'_2 + 1)^{\text{ème}}$ phase du canal total formé par le canal $f(t)$ et la $(c'_1)^{\text{ème}}$ antenne où $c'_1 \triangleq \lceil \frac{c}{c_2} \rceil$ dans $1, \dots, c_1$ et $c'_2 \triangleq c - 1 - c_2 \lfloor \frac{c-1}{c_2} \rfloor$ dans $0, \dots, c_2 - 1$.

¹Single Input Multiple Output

Le bruit discret $b^{c'}(k)$ qui affecte le filtre $\mathbf{h}_M^{c'}$ est donné par $b^{c'}(k) \triangleq (b * a^{c_1}) \left(\left(\frac{c_2'}{c_2} + k \right) T \right)$ et constitue donc un bruit coloré temporellement mais pas spatialement au sens où $E \left(b^{c'}(k) b^{c''}(l) \right) = 0$ si $c' \neq c''$.

On définit le $m^{\text{ème}}$ coefficient du filtre SIMO $\mathbf{h}_M$ par $\mathbf{h}(m) \triangleq [h_M^1(m), h_M^2(m), \dots, h_M^c(m)]^T$ et son ordre M telque si m_1 est le plus petit entier m telque $\mathbf{h}(m) \neq \mathbf{0}_{c,1}$ et m_2 est le plus grand entier m telque $\mathbf{h}(m) \neq \mathbf{0}_{c,1}$ alors $M \triangleq m_2 - m_1$. On rebaptisera alors $\mathbf{h}(m)$ par $\mathbf{h}(m + m_1)$ pour $m = 0, \dots, M$. Par abus, on note $\mathbf{h}_M^{c'} \triangleq [h_M^{c'}(0), \dots, h_M^{c'}(M)]^T$ les réponses impulsionnelles des différents sous-canaux $\forall c' = 1, \dots, c$ même si elles peuvent avoir des ordres $\leq M$.

Lorsque $f(t)$ est à support fini, M est fini et le canal $\mathbf{h}_M$ est à réponse impulsionnelle finie (FIR). On se place dans les conditions d'un SIMO-FIR.

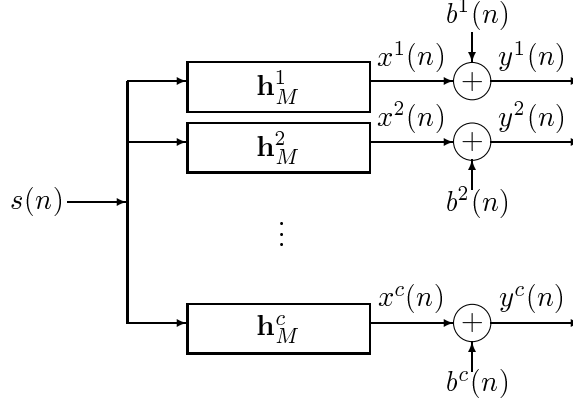


Figure 4.1: Canal SIMO bruité

$\mathbf{h}_M$ est un filtre spatio-temporel. Ses coefficients $h_M^{c'}(m')$ sont indexés doublement par les indices temporels et spatiaux m' et c' , respectivement. Deux rangements sont possibles temps/espace (TS) et espace/temps (ST), donnant lieu aux notations suivantes.

$$\begin{aligned}
 \mathbf{h}_{M,ST} &\triangleq \left[h_M^1(0), \dots, h_M^c(0), h_M^1(1), \dots, h_M^c(1), \dots, h_M^1(M), \dots, h_M^c(M) \right]^T \\
 &= \left[\mathbf{h}^T(0) \dots, \mathbf{h}^T(M) \right]^T \\
 \mathbf{h}_{M,TS} &\triangleq \left[h_M^1(0), \dots, h_M^1(M), h_M^2(0), \dots, h_M^2(M), \dots, h_M^c(0), \dots, h_M^c(M) \right]^T \\
 &= \left[\left(\mathbf{h}_M^1 \right)^T, \dots, \left(\mathbf{h}_M^c \right)^T \right]^T
 \end{aligned}$$

Les deux écritures sont liées par la permutation suivante :

$$\mathbf{h}_{M,TS} = \mathbf{K}_{c,M+1} \mathbf{h}_{M,ST} \quad (4.1)$$

4.1.2 Processus multivariables

Les processus spatio-temporels impliqués dans l'étude des canaux SIMO sont des processus multivariables. Avant d'exploiter la structure particulière aux canaux SIMO, il convient de fixer les notations et présenter certaines propriétés relatives aux processus multi-variables.

Soit le processus vectoriel $\mathbf{x}(n) \triangleq [x^1(n), \dots, x^c(n)]$. n et c' s'interprètent comme des indices dans le temps et dans l'espace, respectivement. Les valeurs prises par le processus entre les instants $n - (D - 1)$ et n ($n \geq D - 1$) peuvent être concaténées de deux manières :

$$\begin{aligned} \mathbf{x}_{D,ST}(n) &\triangleq [x^1(n), \dots, x^c(n), \\ &\quad x^1(n-1), \dots, x^c(n-1), \dots, \\ &\quad x^1(n-(D-1)), \dots, x^c(n-(D-1))] \\ &= [\mathbf{x}(n), \dots, \mathbf{x}(n-D)]^T \\ \text{ou bien } \mathbf{x}_{D,TS}(n) &\triangleq [x^1(n), \dots, x^1(n-(D-1)), \\ &\quad x^2(n), \dots, x^2(n-(D-1)), \dots, \\ &\quad x^c(n), \dots, x^c(n-(D-1))] \end{aligned}$$

Quand le processus est scalaire ($c = 1$), on note $\mathbf{x}_D(n) \triangleq \mathbf{x}_{D,ST}(n) = \mathbf{x}_{D,TS}(n)$. Par analogie à (4.1), on peut facilement montrer que

$$\mathbf{x}_{D,TS}(n) = \mathbf{K}_{c,D} \mathbf{x}_{D,ST}(n) \quad (4.2)$$

Pour les processus stationnaires, on définit $\gamma(k) \triangleq \mathbb{E} [\mathbf{x}(n+k) \mathbf{x}^H(n)] = \gamma^H(-k)$

et $\tau(k) \triangleq \mathbb{E} [\mathbf{x}(n+k) \mathbf{x}^T(n)]$ ainsi que leurs transformées de Fourier respectives $\mathbf{S}(w) \triangleq \sum_k \gamma(k) e^{-jwk} = \mathbf{S}^H(w)$ et $\mathbf{T}(w) \triangleq \sum_k \tau(k) e^{-jwk} = \mathbf{T}^T(-w)$. Si le processus est complexe circulaire, alors τ et $\mathbf{T}$ sont nuls.

On définit la matrice de corrélation spatio-temporelle à D décalages $\mathbf{R}_{D,SPC} \triangleq \mathbb{E} [\mathbf{x}_{D,SPC}(n) \mathbf{x}_{D,SPC}^H(n)]$, $SPC = ST$ ou TS ; qui vérifie

$$\begin{aligned} \mathbf{R}_{D,ST} &= \begin{bmatrix} \gamma(0) & \gamma(1) & \dots & \gamma(D-1) \\ \gamma^H(1) & \gamma(0) & \ddots & \gamma(D-2) \\ \vdots & \ddots & \ddots & \vdots \\ \gamma^H(D-1) & \gamma^H(D-2) & \dots & \gamma(0) \end{bmatrix} \\ \mathbf{R}_{D,TS} &= \mathbf{K}_{c,D} \mathbf{R}_{D,ST} \mathbf{K}_{D,c} \end{aligned} \quad (4.3)$$

On lui associe l'estimateur non biaisé suivant, dit corrélation empirique, calculé à partir de N observations $\hat{\mathbf{R}}_D \triangleq \frac{1}{N-D+1} \sum_{n=D-1}^{N-1} \mathbf{x}_D(n) \mathbf{x}_D^H(n)$ qui vérifie une relation du type (4.3). L'erreur d'estimation de la matrice de corrélation $\text{Vec}(\hat{\mathbf{R}}_D) - \text{Vec}(\mathbf{R}_D)$ vérifie le résultat suivant dont la preuve est établie au §C.

Proposition 4.1 *Soit $\mathbf{x}(n)$ un processus vectoriel stationnaire centré. Soient $\mathbf{R}_D$ sa matrice de corrélation et $\hat{\mathbf{R}}_D$ sa matrice de corrélation empirique. Alors*

$$\begin{aligned} \lim_{N \rightarrow \infty} N \mathbb{E} \left[\text{Vec}(\hat{\mathbf{R}}_{D,ST} - \mathbf{R}_{D,ST}) \text{Vec}(\hat{\mathbf{R}}_{D,ST} - \mathbf{R}_{D,ST})^H \right] &= \mathbf{A} + \mathbf{B} + \mathbf{C} \\ \lim_{N \rightarrow \infty} N \mathbb{E} \left[\text{Vec}(\hat{\mathbf{R}}_{D,ST} - \mathbf{R}_{D,ST}) \text{Vec}(\hat{\mathbf{R}}_{D,ST} - \mathbf{R}_{D,ST})^T \right] &= (\mathbf{A} + \mathbf{B} + \mathbf{C}) \mathbf{K}_{cD} \end{aligned}$$

Avec

$$\mathbf{A} = \frac{1}{2\pi} \int_{w=-\pi}^{\pi} (\mathbf{M}^*(w) \otimes \mathbf{M}(w)) dw$$

$$\begin{aligned}
\mathbf{B} &= \frac{1}{2\pi} \int_{w=-\pi}^{\pi} (\mathbf{N}^*(w) \otimes \mathbf{N}(w)) dw \mathbf{K}_{cD} \\
\mathbf{M}_{ST}(w) &= \left[\left(\mathbf{e}_D(w) \mathbf{e}_D^H(w) \right) \otimes (\mathbf{S}(w)) \right] \\
\mathbf{N}_{ST}(w) &= \left[\left(\mathbf{e}_D(w) \mathbf{e}_D^H(w) \right) \otimes (\mathbf{T}(w)) \right] \\
\mathbf{M}_{TS}(w) &= \left[(\mathbf{S}(w)) \otimes \left(\mathbf{e}_D(w) \mathbf{e}_D^H(w) \right) \right] \\
\mathbf{N}_{TS}(w) &= \left[(\mathbf{T}(w)) \otimes \left(\mathbf{e}_D(w) \mathbf{e}_D^H(w) \right) \right] \\
\mathbf{e}_D &= \left[1, e^{-jw}, \dots, e^{-jw(D-1)} \right]^T
\end{aligned}$$

L'expression de $\mathbf{C}$, complexe, est donnée dans la démonstration (Résultat § C.10).

On remarque que $\mathbf{M}$ est Hermitienne et que $\mathbf{N} = \mathbf{M}$ si $\mathbf{x}(n)$ est un processus réel. Si $\mathbf{x}(n)$ est la somme de deux processus indépendants $\mathbf{x}_a(n)$ et $\mathbf{x}_b(n)$ auxquels sont associées les matrices $\mathbf{A}_a$, $\mathbf{B}_a$, $\mathbf{C}_a$, $\mathbf{A}_b$, $\mathbf{B}_b$ et $\mathbf{C}_b$, alors $\mathbf{A} = \mathbf{A}_a + \mathbf{A}_b$, $\mathbf{B} = \mathbf{B}_a + \mathbf{B}_b$ et $\mathbf{C} = \mathbf{C}_a + \mathbf{C}_b$. $\mathbf{B}$ est nulle si $\mathbf{x}(n)$ est complexe circulaire. $\mathbf{C}$ est nulle si $\mathbf{x}(n)$ est gaussien (complexe ou non; circulaire ou non).

4.1.3 Matrice de filtrage

La relation entrée/sortie du SIMO de Fig. 4.1 est donnée par

$$\mathbf{x}(n) = \mathcal{T}_1(\mathbf{h}_M) \mathbf{s}_{M+1}(n) \triangleq [\mathbf{h}(0) \quad \mathbf{h}(1) \quad \dots \quad \mathbf{h}(M)] \mathbf{s}_{M+1}(n)$$

On s'intéresse au cas où la sortie $\mathbf{y}(n)$ est observée sur l durées symboles. Ces observations étant groupées dans le vecteur $\mathbf{y}_l(n)$, on a alors

$$\mathbf{y}_l(n) = \mathcal{T}_l(\mathbf{h}_M) \mathbf{s}_{M+l}(n) + \mathbf{b}_l(n)$$

où $\mathcal{T}_l(\mathbf{h}_M)$ est la matrice de filtrage du SIMO définie dans le paragraphe suivant.

Définitions et propriétés

$$\mathcal{T}_l(\mathbf{h}_M^{c'}) \triangleq \begin{bmatrix} h_M^{c'}(0) & h_M^{c'}(1) & \dots & h_M^{c'}(M) & \dots & 0 \\ \vdots & \ddots & & & \ddots & \vdots \\ 0 & \dots & h_M^{c'}(0) & h_M^{c'}(1) & \dots & h_M^{c'}(M) \end{bmatrix}, \quad l \times (l + M) \quad (4.4)$$

$$\mathcal{T}_{l,TS}(\mathbf{h}_M) \triangleq \begin{bmatrix} \mathcal{T}_l(\mathbf{h}_M^1) \\ \vdots \\ \mathcal{T}_l(\mathbf{h}_M^c) \end{bmatrix} \quad (4.5)$$

$$\mathcal{T}_{l,ST}(\mathbf{h}_M) \triangleq \begin{bmatrix} \mathbf{h}(0) & \mathbf{h}(1) & \dots & \mathbf{h}(M) & \dots & 0 \\ \vdots & \ddots & & & \ddots & \vdots \\ 0 & \dots & \mathbf{h}(0) & \mathbf{h}(1) & \dots & \mathbf{h}(M) \end{bmatrix} \quad (4.6)$$

$$\mathcal{T}^\circ(\mathbf{h}_M) \triangleq \begin{bmatrix} \mathbf{h}(0) & \mathbf{h}(1) & \cdots & \mathbf{h}(M) \\ \mathbf{h}(1) & \mathbf{h}(2) & & \mathbf{0}_{c,1} \\ \vdots & & \ddots & \vdots \\ \mathbf{h}(M) & \mathbf{0}_{c,1} & \cdots & \mathbf{0}_{c,1} \end{bmatrix} \mathbf{0}_{c(M+1),M-1} \quad (4.7)$$

Comme ça était le cas pour les filtres et processus spatio-temporels, les écritures ST et TS de la matrice de filtrage² sont équivalentes au sens que $\mathcal{T}_{l,TS}(\mathbf{h}_M) = \mathbf{K}_{c,l} \mathcal{T}_{l,ST}(\mathbf{h}_M)$.³

Si $\mathbf{g}_N \in \mathbf{C}^{c(N+1)}$, alors le produit $\mathcal{T}_{M+1}^T(\mathbf{g}_N) \mathbf{h}_M \triangleq \mathcal{T}_{M+1,ST}^T(\mathbf{g}_N) \mathbf{h}_{M,ST} = \mathcal{T}_{M+1,TS}^T(\mathbf{g}_N) \mathbf{h}_{M,TS}$ vérifie la propriété de commutation suivante :

$$\mathcal{T}_{M+1}^T(\mathbf{g}_N) \mathbf{h}_M = \mathcal{T}_{N+1}^T(\mathbf{h}_M) \mathbf{g}_N \quad [\text{MOUL 95, ABED 95}] \quad (4.9)$$

Un résultat important concerne le rang de la matrice de filtrage. $\mathcal{T}_l(\mathbf{h}_M)$ est de rang colonne plein sous les hypothèses suivantes :

Hypothèse H 4.1 • $l \geq \lceil \frac{M}{c-1} \rceil$ [Carv 98, (4)]⁴.

- pas de zéros communs aux polynomes $\left(\sum_{k=0}^M h_M^{c'}(k) z^k \right), c' \in \{1, \dots, c\}$.

La première condition traduit que la matrice doit être sur-déterminée. La dernière condition est dite condition de disparité ou de diversité. Un canal la respectant est dit à minimum de phase [PAPA 95] ou irréductible [Carv 98]. Un autre résultat important stipule qu'une matrice de filtrage de rang colonne plein est entièrement caractérisée par son noyau à gauche :

Proposition 4.2 Noyau [MOUL 95, ABED 95]⁵ Soient $\mathbf{h}_M$ et $\mathbf{g}_M \in \mathbf{C}^{c(M+1)}$ à minimum de phase et $l \geq \lceil \frac{M+1}{c-1} \rceil$. Soit $\mathbf{n}_i \in \mathbf{C}^{cl}, i \in \{1, \dots, (c-1)l - M\}$ une famille libre de vecteurs orthogonaux à gauche de $\mathcal{T}_l(\mathbf{h}_M)$.

$$\mathbf{n}_i^H \mathcal{T}_l(\mathbf{g}_M) = 0 \text{ pour tout } i \text{ ssi il existe un scalaire complexe } a \text{ tel que } \mathbf{h}_M = a \mathbf{g}_M$$

²Dans la littérature, $\mathcal{T}_{l,ST}(\cdot)$ est souvent notée $\mathcal{T}_l(\cdot)$ [ABED 97, DINGa97] alors que $\mathcal{T}_{l,TS}(\cdot)$ est notée $\mathcal{H}_l$ [MOUL 95]. On a choisit ici de leur réserver la même notation vu qu'elles sont identiques à une permutation de lignes près. Par la suite, elles continuent à vérifier les mêmes propriétés (notamment rang et valeurs singulières) et l'indice ST/TS qui les distingue ne sera mentionné que lorsque nécessaire.

³On peut aussi voir que la matrice de filtrage est une fonction linéaire du filtre. Ainsi, $\text{Vec} \left(\mathcal{T}_l^T \left(\mathbf{h}_M^{c'} \right) \right) = \overline{\mathcal{T}}_{l,M} \mathbf{h}_M^{c'}$, $\text{Vec} \left(\mathcal{T}_{l,TS}^T(\mathbf{h}_M) \right) = \overline{\mathcal{T}}_{l,TS,M} \mathbf{h}_{M,TS}$ et $\text{Vec} \left(\mathcal{T}_{l,ST}^T(\mathbf{h}_M) \right) = \overline{\mathcal{T}}_{l,ST,M} \mathbf{h}_{M,ST}$; avec

$$\begin{aligned} \overline{\mathcal{T}}_{l,M} &= \left[\left(\mathbf{1}_{l-1} \otimes \begin{bmatrix} \mathbf{I}_{M+1} \\ \mathbf{0}_{l,M+1} \end{bmatrix} \right) \right] \\ \overline{\mathcal{T}}_{l,TS,M} &= (\mathbf{I}_c \otimes \overline{\mathcal{T}}_{l,M}) \\ \overline{\mathcal{T}}_{l,ST,M} &= (\mathbf{K}_{l,c} \otimes \mathbf{I}_{l+M}) (\mathbf{I}_c \otimes \overline{\mathcal{T}}_{l,M}) \mathbf{K}_{c,M+1} = (\mathbf{K}_{l,c} \otimes \mathbf{I}_{l+M}) \overline{\mathcal{T}}_{l,TS,M} \mathbf{K}_{c,M+1} \end{aligned} \quad (4.8)$$

⁴Rigoureusement, il faut que $l \geq M$ [TONG 94], i.e. que le facteur de lissage soit au moins égal à l'ordre du canal. Cependant, en pratique, la condition $l \geq \lceil \frac{M}{c-1} \rceil$ est (presque) toujours suffisante, ce que nous retenons dans ce rapport.

⁵La démonstration dans [MOUL 95, ABED 95] a été faite pour $\mathcal{T}_{l,TS}(\cdot)$ et sous la condition $l > M$. Cependant, il peut être vérifié que cette condition n'est que suffisante et que le noyau à gauche d'une matrice de filtrage caractérise cette dernière tant qu'il n'est pas vide.

La proposition ci-dessus implique qu'une matrice $\mathcal{T}(\mathbf{h}_M)$ (et par conséquent $\mathbf{h}_M$) est déterminée (à une constante près) par la donnée de son noyau (non vide) à gauche. Le cas particulier $c = 2$ et $l = M + 1$ est d'un intérêt pratique puisque l'unique vecteur normé orthogonal à gauche de $\mathcal{T}_{M+1}(\mathbf{h}_M)$ s'exprime (4.10) en fonction de la réponse $\mathbf{h}_M$. Si on définit $\mathbf{T}_1 \hat{=} \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix}$, $\mathbf{T}_{l,TS} \hat{=} (\mathbf{T}_0 \otimes \mathbf{I}_l)$ et $\mathbf{T}_{l,ST} \hat{=} (\mathbf{I}_l \otimes \mathbf{T}_0)$, on a alors $\mathbf{T}_l^{-1} = \mathbf{T}_l^T = -\mathbf{T}_l$, $\mathbf{T}_{M+1,TS} \mathbf{h}_{M,TS} = \begin{bmatrix} \mathbf{h}_M^2 \\ -\mathbf{h}_M^1 \end{bmatrix}$ et $\mathcal{T}_{l,TS}(\mathbf{T}_{M+1,TS} \mathbf{h}_{M,TS}) = \mathbf{T}_{l,TS} \mathcal{T}_{l,TS}(\mathbf{h}_{M,TS})$. Il s'en déduit que

$$\forall \mathbf{h}_M \in \mathbf{C}^{2(M+1)}, \text{ on a } \quad \mathcal{T}_{M+1}^T(\mathbf{h}_M) \mathbf{T}_{M+1} \mathbf{h}_M = \mathbf{0}_{2M+1,1} \quad (4.10)$$

Statistiques du second ordre

Par défaut, on note $\gamma(k)$ et $\mathbf{R}_D$ les SOS relatifs au processus $\mathbf{y}(n)$. Pour tout autre processus $\mathbf{p}(n)$, elles sont notées $\gamma^p(k)$ et $\mathbf{R}_D^p$.

les SOS du canal sont entièrement contenues dans les matrices $\gamma(k) = \mathbb{E}(\mathbf{y}(n+k) \mathbf{y}^H(n))$ et $\tau(k) = \mathbb{E}(\mathbf{y}(n+k) \mathbf{y}^T(n))$, $k \geq 0$. Les premières sont nulles $\forall |k| \geq M+1$ puisque $\mathbf{x}(n)$ est un MA (Moyenne Ajustée) d'ordre M . Les dernières sont nulles si l'entrée $s(n)$ et le bruit $\mathbf{b}(n)$ sont circulaires⁶. Dans ces conditions, l'information statistique d'ordre deux est entièrement contenue dans la suite $\{\gamma(k), 0 \leq k \leq M\}$ qu'on retrouve dans toute matrice de corrélation $\mathbf{R}_D = \mathbb{E}(\mathbf{y}_D(n+k) \mathbf{y}_D^H(n)) \forall D \geq M+1$.

On suppose que le bruit $\mathbf{b}(n)$ est décorrélé de l'entrée $s(n)$.

Hypothèse H 4.2 *L'entrée $s(n)$ est décorrélée et de variance σ_s^2 .*

$$\begin{aligned} \mathbf{R}_D &= \mathcal{T}_D(\mathbf{h}_M) \mathbf{R}_{D+M}^s \mathcal{T}_D^H(\mathbf{h}_M) + \mathbf{R}_D^b \\ &= \sigma_s^2 \mathcal{T}_D(\mathbf{h}_M) \mathcal{T}_D^H(\mathbf{h}_M) + \mathbf{R}_D^b \quad \text{sous H4.2} \end{aligned}$$

On définit

$$\begin{aligned} \mathbf{R}_D^\circ &\hat{=} \begin{bmatrix} \gamma(0) - \sigma_b^2 \mathbf{I}_c & \gamma(1) & \cdots & \gamma(D-1) \\ \gamma(1) & \gamma(2) & & \gamma(D) \\ \vdots & & & \vdots \\ \gamma(D-1) & 0 & \cdots & \gamma(2D-2) \end{bmatrix} \\ \mathbf{R}_{M+1}^\circ &= \begin{bmatrix} \gamma(0) - \sigma_b^2 \mathbf{I}_c & \gamma(1) & \cdots & \gamma(M) \\ \gamma(1) & \gamma(2) & & \mathbf{0}_c \\ \vdots & & & \vdots \\ \gamma(M) & \mathbf{0}_c & \cdots & \mathbf{0}_c \end{bmatrix} \\ &= \sigma_s^2 \mathcal{T}^\circ(\mathbf{h}_M) \mathcal{T}_{M+1,ST}^H(\mathbf{h}_M) \quad \text{sous H4.2 [DINGa97, (3.10)]} \end{aligned}$$

⁶L'hypothèse entrée circulaire est vérifiée par les constellations les plus fréquentes (MDP et MAQ) et est, par conséquent, retenue dans la suite du rapport.

$$\gamma^x(k) = \mathcal{T}_1(\mathbf{h}_M) \begin{bmatrix} \gamma^s(k) & \gamma^s(k+1) & \cdots & \gamma^s(k+M) \\ \gamma^s(k-1) & \gamma^s(k) & \cdots & \gamma^s(k+M-1) \\ \vdots & \ddots & \ddots & \vdots \\ \gamma^s(k-M) & \cdots & \gamma^s(k) & \end{bmatrix} \mathcal{T}_1^H(\mathbf{h}_M).$$

Sous **H4.2**, le processus $\mathbf{x}(n)$ vérifie $\gamma^x(k) = \sigma_s^2 \mathcal{T}_1(\mathbf{h}_M) \mathbf{J}_{M+1}^k \mathcal{T}_1^H(\mathbf{h}_M) = \sigma_s^2 \sum_{i=0}^M \mathbf{h}(i) \mathbf{h}(i-k)^H, \forall k$, $\mathbf{S}^x(w) = \sigma_s^2 \left(\sum_{k=0}^M \mathbf{h}(k) e^{-jwk} \right) \left(\sum_{k=0}^M \mathbf{h}(k) e^{-jwk} \right)^H$ [PAPA 95, (7.126)] et $\mathbf{R}_D^x = \sigma_s^2 \mathcal{T}_D(\mathbf{h}_M) \mathcal{T}_D^H(\mathbf{h}_M)$ permettant ainsi de simplifier les expressions des matrices **A** et **B** dans Prop. 4.1. Dans ces conditions, la matrice **C** se simplifie comme suit (cf. §D pour la démonstration) :

$$\mathbf{C} = \mathcal{K}_{s_4} \text{Vec}(\mathbf{R}_D^x) \text{Vec}(\mathbf{R}_D^x)^H \quad (4.11)$$

avec $\mathcal{K}_{s_4} \triangleq \text{cum}(s_0, s_0, s_0^*, s_0^*)$. $\mathcal{K}_{s_4} = -2$ dans le cas d'une source binaire équiprobable.

Un résultat moins général que Prop. 4.1 a été donné dans [ABEDa97, Théorème 8] sous les hypothèses supplémentaires d'un processus $\mathbf{x}(n)$ issu du filtrage par un SIMO d'un processus i.i.d., circulaire et affectée d'un bruit circulaire. Une autre expression de **C** y a été donnée mais moins simple que celle de (4.11).

4.2 Égalisation linéaire des canaux SIMO

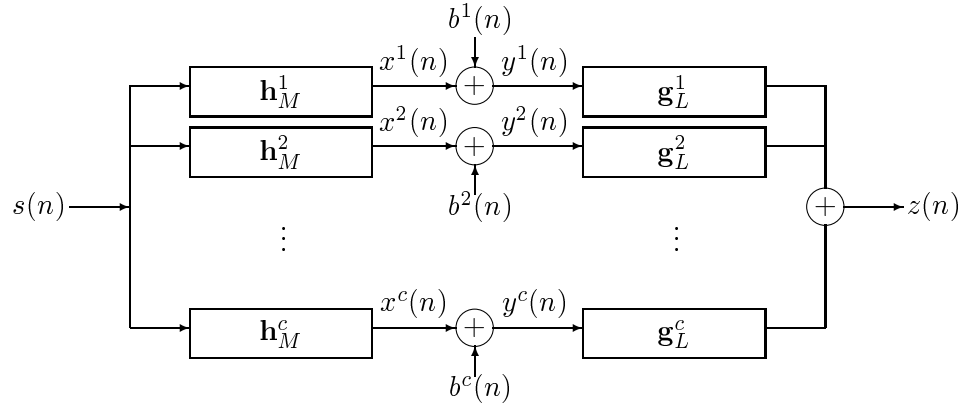


Figure 4.2: Égalisation d'un canal SIMO

On place à la sortie du canal SIMO un filtre linéaire d'ordre l , $\mathbf{g}_l = [\mathbf{g}_l^1, \dots, \mathbf{g}_l^c]^T$. La sortie de l'égaliseur est donnée par $z(n) = \mathbf{y}_{l+1}^T(n) \mathbf{g}_l$. En l'absence de bruit, $z(n) = \mathbf{s}_{M+l+1}^T(n) \mathcal{T}_{l+1}^T(\mathbf{h}_M) \mathbf{g}_l$ et la réponse globale du canal est égale à $\mathbf{r}_{M+l} = \mathcal{T}_{l+1}^T(\mathbf{h}_M) \mathbf{g}_l$. La qualité de la suppression d'ISI est appréciée à travers l'ouverture de l'oeil pour laquelle on retient la définition suivante lorsque le délai de reconstitution est $i - 1$

$$\text{OEM}_i(\mathbf{r}_{M+l}) = \frac{\sum_{k \neq i} |r(k)|^2}{|r(i)|^2}$$

La suppression de l'ISI est possible en appliquant un égaliseur ZF. Cependant, l'ISI n'est pas la seule dégradation que subissent les symboles émis. L'égaliseur doit prendre soin de ne pas aggraver (amplifier) le bruit en tentant d'inverser le canal. La présence de bruit rend impossible une reconstitution

⁷Format *TS*.

parfaite. L'égaliseur doit se contenter de reconstituer au mieux les symboles transmis en comptant sur le bloc de décision pour lever le restant de l'indétermination sur ces symboles. Cette reconstitution se fait au sens du critère de la minimisation de l'erreur quadratique moyenne (MMSE) qui permet de concilier la lutte contre les deux sources d'erreur : l'ISI et le bruit.

De la même manière que les égaliseurs synchrones, le problème du choix d'un délai et d'une longueur pour l'égaliseur se pose. Les remarquent évoqués alors (cf. §3.2.1) restent valables.

4.2.1 Égalisation ZF

Le canal est dit égalisé au sens *Zero forcing* (ZF) quand la réponse globale est un retard pur de $i - 1$ durées symboles, $i \in \{1, M + l + 1\}$. L'égaliseur est alors noté $\mathbf{g}_{l,i}$ et vérifie

$$\mathcal{T}_{l+1}^T(\mathbf{h}_M) \mathbf{g}_{l,i} = \mathcal{T}_{M+1}^T(\mathbf{g}_{l,i}) \mathbf{h}_M = \mathbf{e}_{M+l+1,i} \quad (4.12)$$

Égalisation ZF indirecte

Lorsque la réponse du canal a été identifiée au préalable [MOUL 95, ABED 97, DINGa97], l'égaliseur ZF, d'un ordre et un délai désirés, existe si et seulement si [SCHO 97]

$\left(\mathcal{T}_{l+1}^{\mathfrak{h}}(\mathbf{h}_M) \mathcal{T}_{l+1}(\mathbf{h}_M)\right)^T \mathbf{e}_{M+l+1,i} = \mathbf{e}_{M+l+1,i}$. Puisque on peut se contenter d'une partie des égaliseurs possibles, on se ramène à la recherche de solutions sous la condition suffisante suivante $\mathcal{T}_{l+1}^{\mathfrak{h}}(\mathbf{h}_M) \mathcal{T}_{l+1}(\mathbf{h}_M) = \mathbf{I}_{M+l+1}$ laquelle est satisfaite si $\mathcal{T}_{l+1}(\mathbf{h}_M)$ est de rang colonnes plein, donc si

- $\mathbf{h}_M$ est un canal à minimum de phase i.e., les fonctions de transfert des différents sous-canaux n'ont pas de zéros en commun.
- $l \geq \lceil \frac{M}{c-1} \rceil - 1$.

Cette dernière condition n'est que suffisante pour l'existence de solutions. Il est ainsi prouvé qu'il est possible d'égaliser au sens ZF un canal SIMO-FIR par un FIR également. C'est un avantage immédiat de la réception spatio-temporelle par rapport à la réception Baud pour qui l'inverse d'un canal FIR est un IIR et est donc condamné à la troncature et à des performances limitées.

Sous ces conditions suffisantes, les solutions du système (au nombre infini pour l et i fixés) sont données par $\mathbf{g}_{l,i} = \mathcal{T}_{l+1}^{\mathfrak{h}T}(\mathbf{h}_M) \mathbf{e}_{M+l+1,i} + \left[\mathbf{I}_{c(L+1)} - \left(\mathcal{T}_{l+1}(\mathbf{h}_M) \mathcal{T}_{l+1}^{\mathfrak{h}}(\mathbf{h}_M) \right)^T \right] \mathbf{a}$, $\mathbf{a} \in \mathbf{C}^{c(L+1)}$ quelconque. Parmi les égaliseurs les plus courts, on a $\mathbf{g}_{\lceil \frac{M}{c-1} \rceil - 1, i} = \mathcal{T}_{\lceil \frac{M}{c-1} \rceil}^{\mathfrak{h}T}(\mathbf{h}_M) \mathbf{e}_{\lceil \frac{M}{c-1} \rceil + M, i}$. Les lignes $1, \dots, \lceil \frac{M}{c-1} \rceil + M$ de $\mathcal{T}_{\lceil \frac{M}{c-1} \rceil}^{\mathfrak{h}}(\mathbf{h}_M)$ sont autant d'égaliseurs ZF. Dans le cas pratique $c = 2$, le plus court égaliseur ZF, d'un retard $i - 1 \in \{0, \dots, 2M\}$ fixé, est unique et est donné par

$$\mathbf{g}_{M-1,i} = \mathcal{T}_M^{-T}(\mathbf{h}_M) \mathbf{e}_{2M,i}$$

Égalisation ZF directe

On peut déduire *certaines* égaliseurs ZF directement à partir des (statistiques des) observations du canal. L'algorithme de prédiction linéaire (LP) [Slek 94, ABED 97] (cf. §4.4.2) est un exemple d'algorithmes d'égalisation ZF. Il permet de calculer un égaliseur $\mathbf{g}_{M,1}$. L'algorithme d'identification proposé dans §5 en est un aussi puisqu'il permet de calculer plusieurs égaliseurs ZF de longueur quelconque ($\geq \lceil \frac{M}{c-1} \rceil$ néanmoins) avec un délai maximal ou minimal. Cependant, puisque les égaliseurs ZF à délai maximal et minimal sont connus en pratique pour être les plus défavorables, en terme de MSE, les algorithmes proposés dans [Touz 98, GESB 97] se distinguent en donnant la possibilité de calculer des égaliseurs ZF à un délai fixé à l'avance.

Identification en utilisant des égaliseurs ZF

On s'intéresse à la situation inverse où l'algorithme (LP ou celui proposé dans §5) commence par calculer un égaliseur ZF donné (dont le délai est connu). Il est possible d'en déduire la réponse impulsionnelle du canal comme suit.

Connaissant un égaliseur $\mathbf{g}_{l,i}$, on résout le système $\mathcal{T}_{M+1}^T(\mathbf{g}_{l,i}) \mathbf{h}_M = \mathbf{e}_{M+l+1,i}$; lequel est consistant si $\mathcal{T}_{M+1}(\mathbf{g}_{l,i})$ est de rang colonne plein. Si on suppose que $\mathbf{g}_{l,i}$ respecte la condition minimum de phase, le système admet des solutions dès que $l \leq (c-1)(M+1)$. Dans l'ensemble des solutions $\{\mathcal{T}_{M+1}^{\dagger T}(\mathbf{g}_{l,i}) \mathbf{e}_{M+l+1,i} + [\mathbf{I}_{c(M+1)} - (\mathcal{T}_{M+1}(\mathbf{g}_{l,i}) \mathcal{T}_{M+1}^{\dagger}(\mathbf{g}_{l,i}))^T] \mathbf{a}, \mathbf{a} \in \mathbf{C}^{c(M+1)}\}$, on doit retrouver la réponse exacte $\mathbf{h}_M$. Cette indétermination ne peut être supprimée que dans le cas $\mathcal{T}_{M+1}(\mathbf{g}_{l,i})$ inversible soit $l = (c-1)(M+1)$. Dans ce cas on a $\mathbf{h}_M = \mathcal{T}_{M+1}^{-T}(\mathbf{g}_{(c-1)(M+1),i}) \mathbf{e}_{c(M+1),i}$.

À l'exception de ce cas, l'équation ZF (4.12), i.e., la seule connaissance d'un égaliseur, ne suffit pas à déterminer la réponse du canal. Outre l'expression de l'égaliseur, on aura besoin d'une information statistique (du second ordre) sur la sortie du canal en terme des matrices de corrélation $\gamma(k)$. L'identification s'opère de la manière suivante⁸ :

Étant donné que $\mathbf{y}(n) = [\mathbf{h}(0) \quad \mathbf{h}(1) \quad \dots \quad \mathbf{h}(M)] \mathbf{s}_{M+1}(n) + \mathbf{b}(n)$, le $k^{\text{ème}}$ coefficient du canal se déduit comme $\mathbf{h}(k) = \mathbf{E}(\mathbf{y}(n) s(n-k)^*)$. Les symboles sont retrouvés grâce à l'égaliseur ZF, soit $s(n-k) = \mathbf{x}_{l+1}^T(n-k+i-1) \mathbf{g}_{l,i}$ et par conséquent $\mathbf{h}(k) = \mathbf{E}(\mathbf{y}(n) \mathbf{x}_{l+1}^H(n-k+i-1)) \mathbf{g}_{l,i}^*$.

Étant donné que l'entrée du canal est décorrélée du bruit, on a $\mathbf{h}(k) = \mathbf{E}(\mathbf{x}(n) \mathbf{x}_{l+1}^H(n-k+i-1)) \mathbf{g}_{l,i}^*$. Soit⁹

$$\mathbf{h}_M = \begin{bmatrix} \gamma^x(-i+1) & \gamma^x(-i+2) & \dots & \gamma^x(-i+l+1) \\ \gamma^x(-i+2) & \gamma^x(-i+3) & & \gamma^x(-i+l+2) \\ \vdots & & & \vdots \\ \gamma^x(-i+M+1) & \dots & & \gamma^x(-i+l+M+1) \end{bmatrix} \mathbf{g}_{l,i}^*$$

⁸Cette procédure est la généralisation de celle présentée dans [ABED 97], qui a été faite en réel et pour $l = M$ et $i = 1$.

⁹Format *ST*

$$= \left(\begin{bmatrix} \gamma(-i+1) & \gamma(-i+2) & \cdots & \gamma(-i+l+1) \\ \gamma(-i+2) & \gamma(-i+3) & & \gamma(-i+l+2) \\ \vdots & & & \vdots \\ \gamma(-i+M+1) & \cdots & & \gamma(-i+l+M+1) \end{bmatrix} - \begin{bmatrix} \gamma^b(-i+1) & \gamma^b(-i+2) & \cdots & \gamma^b(-i+l+1) \\ \gamma^b(-i+2) & \gamma^b(-i+3) & & \gamma^b(-i+l+2) \\ \vdots & & & \vdots \\ \gamma^b(-i+M+1) & \cdots & & \gamma^b(-i+l+M+1) \end{bmatrix} \right) \mathbf{g}_{l,i}^*$$

4.2.2 Égalisation MMSE

La reconstitution parfaite des symboles étant impossible, on cherche à l'achever au mieux en fabriquant des symboles aussi proches que possible de ceux émis, et ce au sens de l'erreur quadratique moyenne MSE : $E(|z(n) - s(n-i+1)|^2) = E(|\mathbf{y}_{l+1}^T(n) \mathbf{g}_{l,i} - s(n-i+1)|^2)$, $i \geq 1$. D'une manière analogue au cas des récepteurs synchrones, l'égaliseur MMSE à un délai $i-1$ est donné par

$$\mathbf{R}_{l+1} \mathbf{g}_{l,i}^* = E(\mathbf{y}_{l+1}(n) s^*(n-i+1)) \quad (4.13)$$

Tout comme dans le cas synchrone, le calcul des égaliseurs ZF peut se faire en utilisant un algorithme adaptatif du type LMS ou RLS quand une séquence d'apprentissage est disponible.

(4.13) s'écrit encore

$$\mathcal{T}_{l+1}(\mathbf{h}_M) \mathcal{T}_{l+1}^H(\mathbf{h}_M) \mathbf{g}_{l,i}^* = \mathcal{T}_{l+1}(\mathbf{h}_M) \mathbf{e}_{M+l+1,i} \quad (4.14)$$

4.3 Détection de l'ordre et estimation du bruit

Les algorithmes d'identification exploitent les propriétés de $\mathcal{T}_l(\mathbf{h}_M) \mathcal{T}_l^H(\mathbf{h}_M)$ i.e., la matrice de corrélation des observations non bruitées. Une première étape d'estimation du produit $\mathcal{T}_l(\mathbf{h}_M) \mathcal{T}_l^H(\mathbf{h}_M)$ est alors nécessaire. Elle passe par la détection de l'ordre et l'estimation de la puissance du bruit qui sont ainsi des tâches communes aux algorithmes d'identification aveugles; et ce à partir de la matrice de corrélation du canal. En choisissant un facteur de lissage $l \geq \frac{M+1}{c-1}$, l'ordre du canal peut être déduit de la multiplicité (égale à $(c-1)l - M$) de la plus petite valeur propre de $\mathbf{R}_l$, qui n'est autre que la puissance du bruit. Cependant, les valeurs propres de l'estimée de la matrice de corrélation sont différentes et la distinction d'un groupe de "plus petites valeurs propres" n'est pas toujours évidente.

Wax et Kailath ont montré dans [WAX 85] que des critères de la théorie de l'information, en l'occurrence le critère AIC¹⁰ et le critère MDL¹¹, peuvent s'appliquer au problème de détection de l'ordre pour déduire les estimations suivantes de l'ordre du canal

$$\begin{aligned} M_{\text{AIC}} + 1 &= \min_{k=0, \dots, l-1} Q(k) + k((2c-1)l - k) \\ M_{\text{MDL}} + 1 &= \min_{k=0, \dots, l-1} Q(k) + \frac{1}{2} k((2c-1)l - k) \log(N) \end{aligned}$$

¹⁰Akaike's information theoretic criterion

¹¹Minimum description length criterion

$$\text{où } Q(k) = -\log \left(\frac{\prod_{i=l+k+1}^{cl} \lambda_i(\hat{\mathbf{R}}_l)^{\frac{1}{(c-1)l-k}}}{\frac{1}{(c-1)l-k} \sum_{i=l+k+1}^{cl} \lambda_i(\hat{\mathbf{R}}_l)} \right)^{(p-k)N}$$

Wax et Kailath ont prouvé que le critère MDL est asymptotiquement consistant alors que le critère AIC est asymptotiquement non consistant et conduit à l'estimation d'une valeur supérieure à l'ordre exact du canal. Cette dernière propriété, rajoutée au peu de fiabilité de ces critères dans des conditions réelles d'observations (pendant une durée limitée et en présence de bruit non négligeable), ont motivé la proposition dans §5 d'un algorithme d'identification aveugle au second ordre robuste à une surestimation de l'ordre du canal.

L'ordre du canal étant détecté, la puissance du bruit peut être alors estimée d'une manière consistante [WAX 85] comme étant la valeur moyenne des $(c-1)l - M$ plus petites valeurs propres de $\hat{\mathbf{R}}_l$, soit

$$\sigma_b^2(\hat{\mathbf{R}}_l) \triangleq \frac{\sum_{i=1}^{(c-1)l-M} \lambda_{l+M+i}(\hat{\mathbf{R}}_l)}{(c-1)l - M}$$

La méthode SS se base sur la détermination des sous-espaces bruit et signal et ce en procédant à une recherche des vecteurs orthogonaux à $\mathcal{T}_l(\mathbf{h}_M) \mathcal{T}_l^H(\mathbf{h}_M)$. Ces derniers peuvent également déterminés comme des vecteurs propres de $\mathbf{R}_l$. Ainsi cette étape d'estimation de la puissance du bruit n'est pas prévue pour l'algorithme SS alors qu'elle est indispensable pour les algorithmes LP, OPD et celui de §5.

4.4 Algorithmes d'identification aveugle

Dans ce qui suit, on va introduire les principaux algorithmes¹² d'identification aveugle au second ordre développés à ce jour : l'algorithme sous-espace (SS¹³), l'algorithme prédiction linéaire (LP¹⁴) et l'algorithme décomposition en produit extérieur (OPD¹⁵).

Ces algorithmes se basent sur la connaissance de la matrice de corrélation de la sortie du canal à identifier. Ainsi, on peut adopter une approche fonctionnelle de l'identification (cf. §7.3) et calculer les différentielles correspondantes (cf. §B). Ces algorithmes exploitent les propriétés de (SOS de la) la sortie non bruitée du canal¹⁶. Ils sont formulés donc sous l'hypothèse suivante :

¹²Ces algorithmes seront présentés dans leur version *standard*. On ne tiendra pas compte des différentes versions (sous-espace signal [MOUL 95], sous-espace pondéré [ABEDa97], prédiction linéaire *multi-step* [Gesb 97], ...), d'autant plus qu'elles ne corrigent pas les défauts majeurs (détaillés dans la suite) de ces algorithmes, à savoir la non robustesse à la surestimation de l'ordre du canal et la non robustesse à un mauvais conditionnement de la matrice de corrélation.

¹³Subspace

¹⁴Linear Prediction

¹⁵Outer Product Decomposition

¹⁶Ces algorithmes sont dits *stochastiques* par opposition aux algorithmes dits *déterministes* dont [XU 95, LIU 94, Liu 95, Veena95]. Ces derniers opèrent sur des matrices d'observations. Ils exploitent la relation de filtrage pour factoriser ces matrices en séparant le *vecteur canal* de la *matrice données*. Il s'en suit une estimation de la réponse du canal ou directement des symboles émis. Sous certaines conditions d'identifiabilité, les algorithmes déterministes assurent l'identification de la réponse du canal en l'absence de bruit et à partir d'un nombre limité d'observations. Ceci limite leur intérêt à des situations de fort rapport signal sur bruit et un nombre limité d'observations. Notons qu'ils exigent la connaissance de l'ordre exact du canal. La distinction stochastiques/déterministes est proposée dans [TONG 98] où une classification plus générale des algorithmes d'identification aveugle est faite. Dans ce rapport, on se limite aux algorithmes stochastiques.

Hypothèse H 4.3 *Bruit temporellement et spatialement décorrélé.*

Deux aspects importants de l'identification aveugle sont adressés en particulier, l'ambiguïté et la robustesse à la sur-estimation de l'ordre.

Un algorithme d'identification aveugle au second ordre déduit¹⁷ la réponse impulsionnelle d'un canal SIMO $\mathbf{h}_M$ à minimum de phase d'ordre présumé M à partir de la matrice de corrélation $\mathbf{R}_l = \mathcal{T}_l(\mathbf{h}) \mathcal{T}_l^H(\mathbf{h}) + \sigma_b^2 \mathbf{I}_{cl}$. $\mathbf{R}_l$ est aussi égale à¹⁸ $\mathcal{T}_l \left(e^{j\Phi} \begin{bmatrix} \mathbf{0}_{cM_1,1} \\ \mathbf{h} \\ \mathbf{0}_{cM_2,1} \end{bmatrix} \right) \mathcal{T}_l^H \left(e^{j\Phi} \begin{bmatrix} \mathbf{0}_{cM_1,1} \\ \mathbf{h} \\ \mathbf{0}_{cM_2,1} \end{bmatrix} \right) + \sigma_b^2 \mathbf{I}_{cl}$ pour tout entier M_1 et M_2 positifs et pour tout réel Φ . Il s'en suit que l'ambiguïté de phase est intrinsèque au problème de l'identification et que tout algorithme d'identification se basant (seulement) sur les propriétés algébriques de la matrice de corrélation (exactement connue) doit être (en principe) robuste à la sur-estimation de l'ordre du canal.

C'est ainsi que, comme vérifié dans la suite, les algorithmes LP (cf. §4.4.2), OPD (cf. §4.4.3) et l'algorithme proposé dans §5 sont robustes à la sur-estimation de l'ordre du canal. On qualifie de *théorique* cette robustesse à la sur-estimation de l'ordre puisqu'elle se rapporte au cas (improbable) où les statistiques du canal sont exactement connues. Cette robustesse s'avérera ne pas s'étendre au cas (pratique) où ces statistiques sont estimées à partir d'observations limitées dans le temps; et ce à l'exception de l'algorithme proposé dans §5 pour qui la robustesse à la sur-estimation de l'ordre est qualifiée alors de *réelle*. Notons que l'algorithme sous-espace (cf. §SS) n'est pas robuste (même théoriquement) à la sur-estimation de l'ordre du canal. En effet, cet algorithme se base sur une propriété de la matrice de filtrage (celle énoncée dans Prop. 4.2) qui n'est pas valable dans toute la

$$\text{famille } \left\{ e^{j\Phi} \begin{bmatrix} \mathbf{0}_{cM_1,1} \\ \mathbf{h} \\ \mathbf{0}_{cM_2,1} \end{bmatrix}, M_1, M_2 \text{ entiers positifs et } \Phi \text{ réel quelconques} \right\}.$$

L'algorithme LP et celui proposé dans §5 permettent l'identification avec une ambiguïté de phase. Les algorithmes SS et OPD identifient la réponse du canal dans un sous-espace de dimension 1, donc à une constante près.

L'indétermination de phase, d'amplitude et de l'ordre (sur-estimé) du canal ne sont pas gênant dans le cadre des applications visées, en l'occurrence l'égalisation. L'ambiguïté d'amplitude peut être compensée par un contrôle automatique du gain (CAG). Le codage différentiel permet d'immuniser le récepteur contre une rotation de la constellation, donc contre une ambiguïté de phase. Finalement, un égaliseur $\mathbf{g}_{l,i}$ d'ordre l conçu pour un canal de réponse $e^{j\Phi} \begin{bmatrix} \mathbf{0}_{cM_1,1} \\ \mathbf{h}_M \\ \mathbf{0}_{cM_2,1} \end{bmatrix}$ avec un retard $(i-1) \in \{0, M+l\}$ égalise également le canal $\mathbf{h}_M$ au même sens (ZF ou MMSE) et à un retard inférieur de M_1 durées symboles.

En effet, si $\mathbf{g}_{l,i}$ est un égaliseur ZF, alors d'après (4.12), on a $\mathbf{e}_{M+M_1+M_2+l+1,i} = \mathcal{T}_{l+1}^T \left(\begin{bmatrix} \mathbf{0}_{cM_1,1} \\ \mathbf{h}_M \\ \mathbf{0}_{cM_2,1} \end{bmatrix} \right) \mathbf{g}_{l,i} =$

¹⁷sous certaines conditions d'identifiabilité explicitées plus loin.

¹⁸Dans la suite du paragraphe, les grandeurs sont exprimées dans le format ST .

$$\begin{bmatrix} \mathbf{0}_{M_1, c(M+1)} \\ \mathcal{T}_{l+1}^T(\mathbf{h}_M) \\ \mathbf{0}_{M_2, c(M+1)} \end{bmatrix} \mathbf{g}_{l,i} = \begin{bmatrix} \mathbf{0}_{M_1,1} \\ \mathcal{T}_{l+1}^T(\mathbf{h}_M) \mathbf{g}_{l,i} \\ \mathbf{0}_{M_2,1} \end{bmatrix} \text{ donc } i \in \{M_1 + 1, \dots, M + l + M_1 + 1\} \text{ et } \mathcal{T}_{l+1}^T(\mathbf{h}_M) \mathbf{g}_{l,i} = \mathbf{e}_{M+l+1, i-M_1}.$$

Si $\mathbf{g}_{l,i}$ est un égaliseur MMSE, alors d'après (4.14) on a $\mathcal{T}_{l+1} \left(\begin{bmatrix} \mathbf{0}_{cM_1,1} \\ \mathbf{h}_M \\ \mathbf{0}_{cM_2,1} \end{bmatrix} \right) \mathcal{T}_{l+1}^H \left(\begin{bmatrix} \mathbf{0}_{cM_1,1} \\ \mathbf{h}_M \\ \mathbf{0}_{cM_2,1} \end{bmatrix} \right) \mathbf{g}_{l,i}^* = \mathcal{T}_{l+1} \left(\begin{bmatrix} \mathbf{0}_{cM_1,1} \\ \mathbf{h}_M \\ \mathbf{0}_{cM_2,1} \end{bmatrix} \right) \mathbf{e}_{M+M_1+M_2+l+1,i}$. Donc $i \in \{M_1 + 1, \dots, M + M_1 + l + 1\}$ et $\mathcal{T}_{l+1}(\mathbf{h}_M) \mathcal{T}_{l+1}^H(\mathbf{h}_M) \mathbf{g}_{l,i}^* = \mathcal{T}_{l+1}(\mathbf{h}_M) \mathbf{e}_{M+l+1, i-M_1}$.

4.4.1 L'algorithme sous-espace

L'algorithme SS repose sur la propriété Prop. 4.2 qui permet de caractériser (à une constante complexe près) un canal $\mathbf{h}_M$ à minimum de phase par la donnée du noyau (non vide) à gauche de sa matrice de filtrage $\mathcal{T}_l(\mathbf{h}_M)$ sur-déterminée ($l \geq \lceil \frac{M+1}{c-1} \rceil$). La matrice de corrélation des observations est donnée par $\mathbf{R}_l = \mathcal{T}_l(\mathbf{h}_M) \mathbf{R}_{l+M}^s \mathcal{T}_l^H(\mathbf{h}_M) + \mathbf{R}_l^b$. Si $l \geq \lceil \frac{M+1}{c-1} \rceil$, alors $\mathcal{T}_l(\mathbf{h}_M)$ est de rang colonne plein et $\text{rg}(\mathcal{T}_l(\mathbf{h}_M) \mathbf{R}_{l+M}^s \mathcal{T}_l^H(\mathbf{h}_M)) = \text{rg}(\mathbf{R}_{l+M}^s)$. Si $\mathbf{R}_{l+M}^s$ est inversible, alors $\text{rg}(\mathbf{R}_l - \mathbf{R}_l^b) = l + M$.

Sous H4.3, $\mathbf{R}_l^b = \sigma_b^2 \mathbf{I}_{cl}$ et σ_b^2 est la plus petite valeur propre de $\mathbf{R}_l$ de multiplicité $(c-1)l - M$. Les $(c-1)l - M$ vecteurs propres qui lui sont associés sont orthogonaux à droite de $\mathcal{T}_l(\mathbf{h}_M) \mathbf{R}_{l+M}^s \mathcal{T}_l^H(\mathbf{h}_M)$ donc à gauche de $\mathcal{T}_l(\mathbf{h}_M)$. Ils caractérisent ainsi le canal $\mathbf{h}_M$ à une constante près.

L'algorithme SS suit les étapes suivantes

1. Estimation de $\mathbf{R}_l$, $l \geq \lceil \frac{M+1}{c-1} \rceil$.
2. Détermination des vecteurs propres normés $\mathbf{v}_i \in \mathbf{C}^{cl}$, $i \in \{1, \dots, (c-1)l - M\}$ associés aux $(c-1)l - M$ plus petites valeurs propres de l'estimée de $\mathbf{R}_l$.
3. Minimisation par rapport à $\mathbf{x}_M \in \mathbf{C}^{c(M+1)}$ de la forme quadratique définie positive

$$\begin{aligned} \sum_{i=1}^{(c-1)l-M} \mathbf{v}_i^H \mathcal{T}_l(\mathbf{x}_M) \mathcal{T}_l^H(\mathbf{x}_M) \mathbf{v}_i &= \sum_{i=1}^{(c-1)l-M} \mathbf{x}_M^T \mathcal{T}_{M+1}(\mathbf{v}_i^*) \mathcal{T}_{M+1}^H(\mathbf{v}_i^*) \mathbf{x}_M^* \\ &= \mathbf{x}_M^T \left[\sum_{i=1}^{(c-1)l-M} \mathcal{T}_{M+1}(\mathbf{v}_i^*) \mathcal{T}_{M+1}^H(\mathbf{v}_i^*) \right] \mathbf{x}_M^* \end{aligned}$$

La résolution se fait sous contrainte de manière à éviter la solution triviale $\mathbf{x}_M = \mathbf{0}_{c(M+1),1}$. Sous la contrainte $\|\mathbf{x}_M\| = Cste$, la solution est donnée par le vecteur propre associé à la plus petite valeur propre de la matrice Hermitienne définie positive $\left[\sum_{i=1}^{(c-1)l-M} \mathcal{T}_{M+1}(\mathbf{v}_i^*) \mathcal{T}_{M+1}^H(\mathbf{v}_i^*) \right]$ donc à une rotation près. D'où, compte tenu de la contrainte, la solution $\mathbf{x}_M$ est égale à $\mathbf{h}_M$ à une constante complexe près.

L'algorithme SS peut se contenter d'une estimation de la matrice de corrélation tirée à partir d'un nombre N fini d'observations et en tirer la réponse exacte du canal [QIU 96]. En effet, les statistiques

estimées sont données par $\hat{\mathbf{R}}_l = \frac{1}{(N-l+1)} \mathcal{T}_l(\mathbf{h}_M) \sum_{n=l-1}^{N-1} [\mathbf{s}_{l+M}(n) \mathbf{s}_{l+M}^H(n)] \mathcal{T}_l^H(\mathbf{h}_M)$. $\sum_{n=l-1}^{N-1} [\mathbf{s}_{l+M}(n) \mathbf{s}_{l+M}^H(n)]$ est une somme de matrices de rang 1 chacune et représente l'estimée de la matrice de corrélation de la source. Cette somme atteint un rang complet au bout d'une certaine valeur¹⁹ de N . Dans ce cas, en suivant la même démarche, on déduit que $\hat{\mathbf{R}}_l$ est de rang $l + M$. Ses $(c - 1)l - M$ vecteurs propres associés à 0 sont orthogonaux à gauche de $\mathcal{T}_l(\mathbf{h}_M)$ et suffisent donc à déterminer $\mathbf{h}_M$ à une constante près de la même manière qu'auparavant. Il est intéressant de voir que l'algorithme SS parvient, dans ces conditions, à estimer exactement la réponse du canal; et ce indépendamment de sa diversité, qui peut être arbitrairement faible. Ceci justifie les bonnes performances de l'algorithme SS en faible rapport signal sur bruit. Une autre raison est qu'il est le seul algorithme qui se passe de l'estimation de la puissance du bruit.

Cependant, la bonne estimation de l'ordre exacte M du canal est cruciale pour le bon fonctionnement de l'algorithme. En effet, pour caractériser $\mathbf{h}_M$ (à une constante près), il faut disposer de la totalité des $(c - 1)l - M$ vecteurs du noyau à gauche de $\mathcal{T}_l(\mathbf{h}_M)$.

Le cas particulier $c = 2$ et $l = M$ est d'un intérêt particulier. En effet, le sous-espace bruit est de dimension 1. On note $\mathbf{v}_M$ l'unique $\mathbf{v}_i = \mathbf{v}_1$. D'après (4.10), on a $\mathbf{v}_M = \frac{e^{i\alpha}}{\|\mathbf{h}_M\|} \mathbf{T}_M \mathbf{h}_M^*$, $\alpha \in \mathbf{R}$ quelconque. Dans l'algorithme, l'étape de minimisation de la forme quadratique se réduit au calcul de $\mathbf{T}_M \mathbf{v}_M^*$ qui est égal à $\mathbf{h}_M$ à une constante complexe près. Ce cas ($c = 2$ et $l = M$) présente un autre intérêt puisque [ZENG 96] la solution $(\mathbf{T}_M \mathbf{v}_M^*)$ coïncide avec celle de l'algorithme des moindres carrés (LS²⁰) [XU 95].

4.4.2 L'algorithme prédiction linéaire²¹

La sortie (non bruitée) d'un SIMO d'ordre M est un processus (vectoriel) MA d'ordre M . Slock montre dans [Slick 94] que c'est également un processus auto-régressif (AR) du même ordre i.e., M et dont l'innovation est donnée par $\mathbf{i}(n) = s(n)\mathbf{h}(0)$. Les coefficients du filtre prédicteur $\mathbf{A}_0 = \mathbf{I}_c, \mathbf{A}_1, \dots, \mathbf{A}_M$ ($c \times c$) sont telles que sa sortie (l'innovation) $\mathbf{i}(n) = \sum_{i=0}^M \mathbf{A}_i \mathbf{x}(n)$ soit décorrélée. Seule donc $\gamma^i(0)$ est non nulle et est égale à ([ABED 97, Eqt. 13]²²)

$$\gamma^i(0) = [\mathbf{I}_c \mathbf{A}_1 \dots \mathbf{A}_M] [\gamma(0)\gamma(1) \dots \gamma(M)]^H = \sigma_s^2 \mathbf{h}(0) \mathbf{h}^H(0) \quad (4.15)$$

En l'absence de bruit, l'application du prédicteur à la sortie du canal permet de récupérer l'innovation donc les données $s(n)$ sans retard mais à $\mathbf{h}(0)$ près. Cette ambiguïté peut être levée en identifiant $\mathbf{h}(0)$ comme un vecteur propre de $\gamma^i(0)$ i.e., de $[\mathbf{I}_c \mathbf{A}_1 \dots \mathbf{A}_M] [\gamma(0)\gamma(1) \dots \gamma(M)]^H$. On vérifie que le filtre suivant

$$\frac{1}{\|\mathbf{h}(0)\|^2} [\mathbf{I}_c \mathbf{A}_1 \dots \mathbf{A}_M]^T \mathbf{h}^*(0)$$

est un égaliseur ZF sans délai. On le note $\mathbf{g}_{M,1}$.

¹⁹pratiquement dès que $N - l + 1 \geq cl$.

²⁰Least Square

²¹Dans ce paragraphe, les grandeurs spatio-temporelles sont exprimées dans le format ST .

²²Dans [ABED 97], les résultats sont donnés en réel et en supposant σ_s^2 unitaire. La généralisation à (4.15) est triviale.

La réponse du canal s'en déduit d'après §4.2.1 par :

$$\mathbf{h}_M = \begin{bmatrix} \gamma(0) & \gamma(1) & \cdots & \gamma(M) \\ \gamma(1) & \gamma(2) & & \gamma(M+1) \\ \vdots & & & \vdots \\ \gamma(M) & \cdots & & \gamma(2M) \end{bmatrix} \mathbf{g}_{M,1}^* \quad (4.16)$$

$$\begin{aligned} &= \begin{bmatrix} \gamma(0) & \gamma(1) & \cdots & \gamma(M) \\ \gamma(1) & \gamma(2) & & \mathbf{0}_c \\ \vdots & & & \vdots \\ \gamma(M) & \mathbf{0}_c & \cdots & \mathbf{0}_c \end{bmatrix} \mathbf{g}_{M,1}^* \\ &= \mathbf{R}_{M+1}^\circ \mathbf{g}_{M,1}^* \end{aligned} \quad (4.17)$$

Un processus AR d'ordre M peut être vu comme un AR d'ordre $M' \geq M$ de telle sorte que la démarche ci-dessus reste valable lorsque l'on détecte un ordre du canal $M' \geq M$. La dernière étape (4.16) se

transforme en le calcul de $\begin{bmatrix} \gamma(0) & \gamma(1) & \cdots & \gamma(M) \\ \gamma(1) & \gamma(2) & & \gamma(M+1) \\ \vdots & & & \vdots \\ \gamma(M) & \mathbf{0}_c & \cdots & \mathbf{0}_c \\ \mathbf{0}_c & \cdots & & \\ \vdots & & & \end{bmatrix} \mathbf{g}_{M,1}^* = \begin{bmatrix} \mathbf{h}_M \\ \mathbf{0}_{M'-M,1} \end{bmatrix}$ laquelle est exploitable

pour des besoins d'égalisation. Ainsi, l'algorithme LP suit les étapes suivantes :

1. Détection d'un ordre $M' \geq M$
2. Calcul de $\hat{\mathbf{R}}_{M'+1}$ estimée de $\mathbf{R}_{M'+1}$
3. Le sous-espace bruit est alors de dimension estimée $c(M'+1) - (2M'+1)$ inférieure à sa dimension véridique $c(M'+1) - (M' + M + 1)$. La puissance du bruit est estimée par la moyenne des $c(M'+1) - (2M'+1)$ plus petites valeurs propres de $\hat{\mathbf{R}}_{M'+1}$.

4. Résoudre l'équation de Yule-Walker

$$[\mathbf{A}_1 \cdots \mathbf{A}_{M'}] \left[\hat{\mathbf{R}}_{M'} - \widehat{\sigma_b^2} \mathbf{I}_{cM'} \right] = -[\gamma(1) \cdots \gamma(M')]$$

$$\text{et en déduire } [\mathbf{A}_1 \cdots \mathbf{A}_{M'}] = -[\gamma(1) \cdots \gamma(M')] \left[\hat{\mathbf{R}}_{M'} - \widehat{\sigma_b^2} \mathbf{I}_{cM'} \right]^\dagger$$

5. Calculer $\mathbf{l}$ le vecteur propre normé associé à la plus grande valeur propre λ de $[\mathbf{I}_c \mathbf{A}_1 \cdots \mathbf{A}_{M'}] [\gamma(0)\gamma(1) \cdots \gamma(M')]^H$. On a alors $\mathbf{l} = \frac{1}{\|\mathbf{h}(0)\|} \mathbf{h}(0)$ et $\lambda = \sigma_s^2 \|\mathbf{h}(0)\|^2$. Soit $\mathbf{h}(0) = \frac{\sqrt{\lambda}}{\sigma_s} \mathbf{l}$.

6. Calculer $\mathbf{g}_{M',1} = \frac{1}{\sqrt{\lambda}} [\mathbf{I}_c \mathbf{A}_1 \cdots \mathbf{A}_{M'}]^T \mathbf{l}^*$.

7. Identifier $\mathbf{h}_M = \mathbf{R}_{M'+1}^\circ \mathbf{g}_{M',0}^*$

Le calcul de $\mathbf{l}$ se fait à une rotation près. Cette ambiguïté va se propager jusqu'à l'estimée du canal. On peut, cependant, éditer une version de l'algorithme où l'on tolère une ambiguïté de phase et d'amplitude à la fois en remplaçant l'étape 6 par le calcul de $\mathbf{g}_{M',1} = [\mathbf{I}_c \mathbf{A}_1 \cdots \mathbf{A}_{M'}]^T \mathbf{l}^*$ ²³. En

²³lequel n'est pas un égaliseur au sens suggéré par la notation.

évitant d'estimer λ , on élimine une source d'erreur et obtient de meilleures performances²⁴ au prix d'une ambiguïté supplémentaire.

Lorsque la matrice $\mathbf{R}_{M'+1}$ est estimée, ses $c(M' + 1) - (2M' + 1)$ valeurs propres les plus petites ne sont pas égales mais très proches de la puissance du bruit. L'estimée de cette dernière est proche de chacune d'elles de telle manière que $\hat{\mathbf{R}}_{M'} - \widehat{\sigma_b^2} \mathbf{I}_{cM'}$ ne soit pas déficiente mais très mal-conditionnée pour l'inversion. Il convient d'accomplir cette inversion au sens de la pseudo-inverse et de la manière suivante [LIAV 00] : forcer à 0 les $cM' - 2M'$ plus petites valeurs propres de $\hat{\mathbf{R}}_{M'} - \widehat{\sigma_b^2} \mathbf{I}_{cM'}$ et inverser les autres. Le cas $c = 2$ ne pose pas ce problème étant donné que $\hat{\mathbf{R}}_{M'} - \widehat{\sigma_b^2} \mathbf{I}_{cM'}$ est de facto de rang "numérique" complet.

La sur-estimation de l'ordre du canal proposée dans [ABED 97] comme solution pour contourner l'incertitude sur l'ordre, est numériquement risquée lorsque qu'on dispose de juste une estimation de la matrice de corrélation [LIAV 00, GORO 99]. En effet, le sous-espace bruit ainsi estimé est d'une dimension inférieure à sa dimension véridique. Les valeurs propres de $\hat{\mathbf{R}}_{M'+1}$ de la $(M' + M + 2)^{\text{ème}}$ à la $(2M' + 1)^{\text{ème}}$ plus petite sont proches de σ_b^2 mais seront comptabilisées dans le sous-espace signal. Le calcul de la pseudo-inverse de $\hat{\mathbf{R}}_{M'} - \widehat{\sigma_b^2} \mathbf{I}_{cM'}$ (au sens rappelé ci-dessus) inverse ces valeurs propres diminuées de $\widehat{\sigma_b^2}$ (cf. §5.6 et Fig. 5.2 pour plus de détails). Ainsi l'équation de Yule-Walker est nécessairement mal-conditionnée.

4.4.3 L'algorithme décomposition en produit extérieur²⁵

L'algorithme OPD repose sur le fait que la matrice $(\mathbf{D}_1 - \mathbf{D}_2)$

où $\mathbf{D}_1 \triangleq \frac{1}{\sigma_s^2} \mathbf{R}_{M+1}^{\circ} (\mathbf{R}_{M+1} - \sigma_b^2 \mathbf{I}_{c(M+1)})^{\natural} \mathbf{R}_{M+1}^{\circ H}$ et $\mathbf{D}_2 \triangleq (\mathbf{J}_{M+1}^T \otimes \mathbf{I}_c) \mathbf{D}_1 (\mathbf{J}_{M+1} \otimes \mathbf{I}_c)$ vérifie $(\mathbf{D}_1 - \mathbf{D}_2) = \mathbf{h}_M \mathbf{h}_M^H$. Ainsi, la réponse du canal est donnée (à une constante près) par le vecteur propre associé à la plus grande valeur propre de $(\mathbf{D}_1 - \mathbf{D}_2)$.

On peut voir, qu'à la manière de l'algorithme LP, l'algorithme OPD permet d'identifier la réponse du canal si l'ordre de ce dernier est sur-estimé. En effet, si on suppose $M' \geq M$ l'ordre du canal, on peut

vérifier que $\mathbf{D}'_1 \triangleq \frac{1}{\sigma_s^2} \mathbf{R}_{M'+1}^{\circ} (\mathbf{R}_{M'+1} - \sigma_b^2 \mathbf{I}_{c(M'+1)})^{\natural} \mathbf{R}_{M'+1}^{\circ H} = \begin{bmatrix} \mathbf{D}_1 & \mathbf{0}_{M+1, M'-M} \\ \mathbf{0}_{M'-M, M+1} & \mathbf{0}_{M'-M, M'-M} \end{bmatrix}$

et $\mathbf{D}'_2 \triangleq (\mathbf{J}_{M'+1}^T \otimes \mathbf{I}_c) \mathbf{D}_1 (\mathbf{J}_{M'+1} \otimes \mathbf{I}_c) = \begin{bmatrix} \mathbf{D}_2 & \mathbf{0}_{M+1, M'-M} \\ \mathbf{0}_{M'-M, M+1} & \mathbf{0}_{M'-M, M'-M} \end{bmatrix}$.

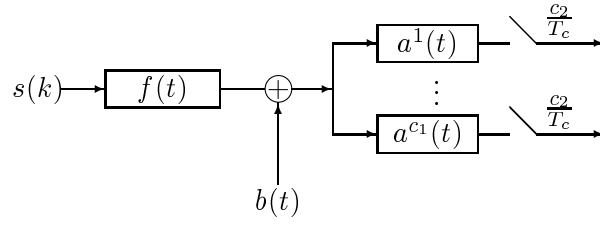
Ainsi $(\mathbf{D}'_1 - \mathbf{D}'_2) = [\mathbf{h}_M \mathbf{0}_{M'-M, 1}] [\mathbf{h}_M \mathbf{0}_{M'-M, 1}]^H$. Cependant, à l'instar de l'algorithme LP, cette robustesse vis à vis de la sur-estimation de l'ordre du canal n'est que théorique. En effet, l'algorithme OPD fait intervenir la pseudo-inverse de $\mathbf{R}_{M'+1} - \sigma_b^2 \mathbf{I}_{c(M'+1)}$ qui, comme dans le cas de l'algorithme LP (cf. §4.4.2), est forcément mal-conditionnée lorsqu'elle est estimée à partir d'un nombre fini d'observations.

²⁴Ceci dépend de la *mesure* de performances adoptée. Pour celle proposée dans §5.8, les deux versions sont équivalentes.

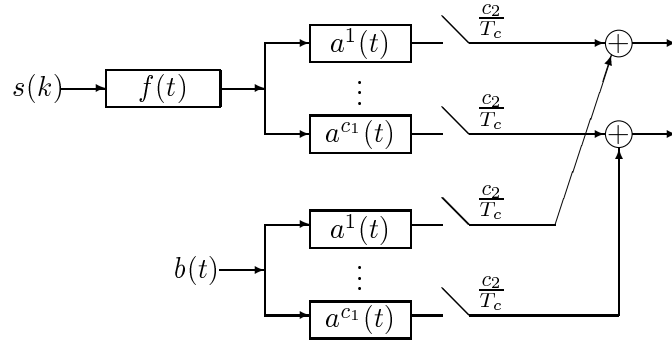
²⁵Dans ce paragraphe, les grandeurs spatio-temporelles sont exprimées dans le format *ST*.

4.5 Conclusion

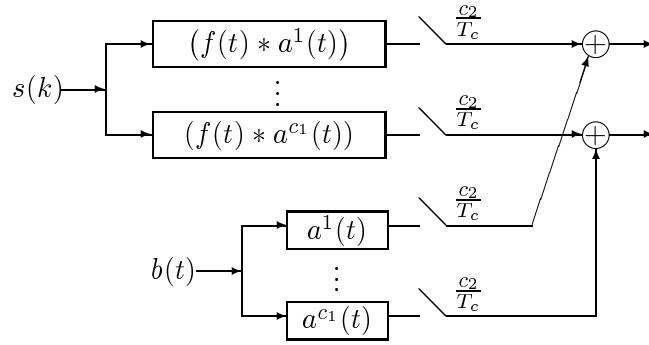
Dans ce chapitre, on a présenté les techniques de réception en diversité spatio-temporelle. L'identification et l'égalisation aveugles à partir des statistiques du second ordre et l'égalisation par des filtres FIR sont rendues possibles par un sur-échantillonnage à un rythme multiple au rythme symbole et/ou une réception à travers un réseau de capteurs. Dans ces conditions, les canaux sont modélisés par des systèmes SIMO. Les algorithmes d'identification des canaux SIMO développés à ce jour présentent une complexité calculatoire importante du fait de leur recours à des décompositions en vecteurs propres/singuliers. Mais leur inconvénient majeur reste leur besoin de connaître la valeur exacte de l'ordre du canal; alors que cette dernière est difficilement détectable à la fois à cause du manque de fiabilité des tests de détection de l'ordre et aussi à cause de l'ambiguïté même de la définition de l'ordre dans le cas des canaux à bande limitée. Dans le chapitre suivant, on se propose d'apporter une réponse à ce problème en développant un algorithme qui se montre robuste à une sur-estimation de l'ordre du canal, en théorie (statistiques exacte) comme en pratique (statistiques estimées).



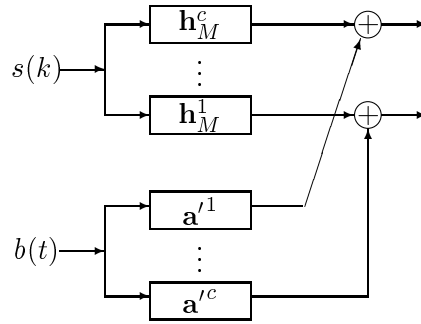
(a)



(b)



(c)



(d)

Figure 4.3: Schéma numérique équivalent

Chapitre 5

Identification aveugle au second ordre robuste à la surestimation de l'ordre

Dans le chapitre précédent, on a présenté l'identification aveugle au second ordre, détaillé certains de ces algorithmes et noté leur non robustesse aux erreurs d'estimation de l'ordre du canal. Dans ce chapitre, on propose un nouvel algorithme d'identification aveugle au second ordre robuste à la surestimation de l'ordre des canaux SIMO [GZH c, Gzh 00]. Cette robustesse démontrée théoriquement d'abord, se confirme en pratique à travers les simulations et constitue le point fort de cet algorithme. Les simulations montrent, d'autre part, que l'algorithme proposé dépasse en performance les algorithmes LP et OPD.

5.1 Surestimation de l'ordre

Un défaut commun aux algorithmes sous-espace (SS), prédiction linéaire (LP) et décomposition en produit extérieur (OPD) est leur inaptitude à bien approximer la réponse du canal sans connaître sa longueur exacte. L'algorithme SS exige dans tous les cas la valeur exacte de l'ordre du canal. LP et OPD sont capables d'estimer la réponse du canal (ou plutôt une version *zero-padded*) s'ils disposent d'une estimation de l'ordre du canal supérieure à sa valeur véridique, situation qu'on désigne désormais par *surestimation de l'ordre*, mais uniquement si la matrice de corrélation des observations est exactement connue. Cette dernière étant toujours estimée, cette robustesse à la surestimation de l'ordre n'est que théorique et le problème reste entièrement posé. On cherche à développer un algorithme qui doit être robuste à la surestimation de l'ordre au vrai sens du terme i.e., lorsque le canal est observé pendant une durée finie et en présence du bruit.

D'autre part, les algorithmes de détection de l'ordre [WAX 85] sont peu fiables, notamment lorsque les mesures (de la sortie du canal) sont bruitées et disponibles en nombre fini. Cependant, ces algorithmes peuvent être biaisés de telle manière à fournir une valeur sûrement surestimée de l'ordre du canal. C'est en particulier le cas du critère AIC (cf. §4.3). On montre dans [WAX 85] que ce critère est non consistant mais asymptotiquement, fournit une valeur surestimée de l'ordre du canal. On peut même disposer d'une valeur surestimée de l'ordre sans avoir à utiliser un test de détection de l'ordre.

En particulier, celle-ci peut être déduite à partir des connaissances à priori sur les conditions de propagation, en terme de dispersion (*delay spread*).

Par conséquent, la robustesse à la surestimation de l'ordre est une propriété utile, recherchée, voire indispensable en pratique pour tout algorithme d'identification aveugle; mais jusqu'ici non vérifiée par les algorithmes proposés.

5.2 Notations

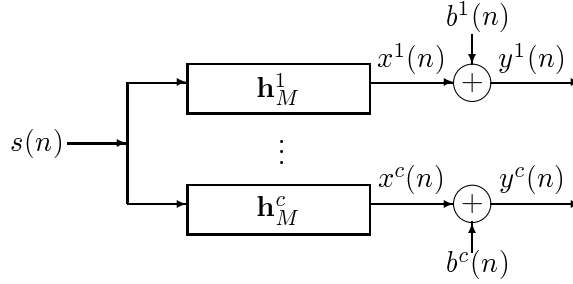


Figure 5.1: Canal SIMO

On reprend le cadre SIMO (fig 5.1). On rappelle la définition de la matrice de corrélation

$$\mathbf{R}_l \triangleq \mathbb{E} [\mathbf{y}_l(n) \mathbf{y}_l^H(n)] = \begin{bmatrix} \gamma(0) & \gamma(1) & \dots & \gamma(l-1) \\ \gamma^H(1) & \gamma(0) & \ddots & \gamma(l-2) \\ \vdots & \ddots & \ddots & \vdots \\ \gamma^H(l-1) & \gamma^H(l-2) & \dots & \gamma(0) \end{bmatrix}$$

et on introduit la matrice de corrélation décalée

$$\mathcal{R}_l \triangleq \mathbb{E} [\mathbf{y}_l(n) \mathbf{y}_l^H(n-1)] = \begin{bmatrix} \gamma(1) & \gamma(2) & \dots & \gamma(l) \\ \gamma(0) & \gamma(1) & \ddots & \gamma(l-1) \\ \vdots & \ddots & \ddots & \vdots \\ \gamma(-l+2) & \gamma(-l+3) & \dots & \gamma(1) \end{bmatrix}$$

Pour tout autre processus¹ $\mathbf{p}(n)$, on notera $\gamma^p(k)$, $\mathbf{R}_l^p$ et $\mathcal{R}_l^p$.

Lorsque les symboles transmis sont décorrélés du bruit, ces matrices sont données par

$$\begin{aligned} \mathbf{R}_l &= \mathcal{T}_l(\mathbf{h}_M) \mathbf{R}_{l+M}^s \mathcal{T}_l^H(\mathbf{h}_M) + \mathbf{R}_l^b \\ \text{et } \mathcal{R}_l &= \mathcal{T}_l(\mathbf{h}_M) \mathcal{R}_{l+M}^s \mathcal{T}_l^H(\mathbf{h}_M) + \mathcal{R}_l^b \end{aligned}$$

Si les symboles $s(n)$ sont décorrélés (H4.2), alors les matrices de corrélation sont données par

$$\begin{aligned} \mathbf{R}_l &= \sigma_s^2 \mathcal{T}_l(\mathbf{h}_M) \mathcal{T}_l^H(\mathbf{h}_M) + \mathbf{R}_l^b \\ \text{et } \mathcal{R}_l &= \sigma_s^2 \mathcal{T}_l(\mathbf{h}_M) \mathbf{J}_{l+M} \mathcal{T}_l^H(\mathbf{h}_M) + \mathcal{R}_l^b \end{aligned}$$

¹éventuellement scalaire.

où $\mathbf{J}_l \triangleq \begin{bmatrix} 0 & 0 & \cdots & 0 \\ 1 & \ddots & \ddots & \vdots \\ \vdots & \ddots & 0 & 0 \\ 0 & \cdots & 1 & 0 \end{bmatrix}$ est la matrice de décalage (vers le bas) $l \times l$. Si, en plus, le bruit est spatialement et temporellement décorrélé (H4.3), alors $\mathbf{R}_l^b = \sigma_b^2 \mathbf{I}_{cl}$ et $\mathcal{R}_l^b = \sigma_b^2 (\mathbf{J}_l \otimes \mathbf{I}_c)$. Ces hypothèses sur les symboles et le bruit sont maintenues dans le reste du chapitre.

5.3 Développement théorique de l'algorithme

L'algorithme proposé se base sur l'exploitation des propriétés structurales de la matrice de corrélation décalée $\mathcal{R}_l$.

La puissance du bruit peut toujours être déterminée comme la plus petite valeur propre de multiplicité $c(l+1) - (l+M+1) = (c-1)(l+1) - M$ de la matrice Hermitienne définie positive $\mathbf{R}_{l+1}$ (cf. §4.3).

La matrice de corrélation de la sortie non bruitée du SIMO est donnée par $\mathcal{R}_l^x = \mathcal{R}_l - \mathcal{R}_l^b = \mathcal{R}_l - \sigma_b^2 (\mathbf{J}_l \otimes \mathbf{I}_c)$. Elle est aussi égale à $\mathcal{R}_l^x = \sigma_s^2 \mathcal{T}_l(\mathbf{h}_M) \mathbf{J}_{l+M} \mathcal{T}_l^H(\mathbf{h}_M)$.

Le canal est supposé à minimum de phase. Sous l'hypothèse H4.1 ($cl \geq l+M$ i.e., $l \geq \lceil \frac{M}{c-1} \rceil$), la matrice $\mathcal{T}_l(\mathbf{h}_M)$ est de rang colonne plein. Par conséquent,

$$\text{rg}(\mathcal{R}_l - \mathcal{R}_l^b) = \text{rg}(\mathcal{T}_l(\mathbf{h}_M) \mathbf{J}_{l+M} \mathcal{T}_l^H(\mathbf{h}_M)) = \text{rg}(\mathcal{T}_l(\mathbf{h}_M) \mathbf{J}_{l+M}) = \text{rg}(\mathbf{J}_{l+M}) = l+M-1.$$

Il s'en suit qu'il existe une famille de vecteurs orthogonaux $\{\mathbf{n}_{l,1}^{(i)}\}_{i=1,\dots,w}$ (resp. $\{\mathbf{n}_{l,2}^{(i)}\}_{i=1,\dots,w}$) orthogonaux à droite (resp. à gauche) de $\mathcal{R}_l - \mathcal{R}_l^b$, où $w \triangleq (c-1)l - M + 1$.

Pour chacun de ces vecteurs, on a

$$\begin{aligned} \mathcal{T}_l(\mathbf{h}_M) \mathbf{J}_{l+M} \mathcal{T}_l^H(\mathbf{h}_M) \mathbf{n}_{l,1}^{(i)} &= \mathbf{0}_{cl,1} \\ \text{et donc } \mathbf{J}_{l+M} \mathcal{T}_l^H(\mathbf{h}_M) \mathbf{n}_{l,1}^{(i)} &= \mathbf{0}_{l+M,1} \end{aligned}$$

Par conséquent, il existe un scalaire complexe $\alpha_1^{(i)}$ telque

$$\mathcal{T}_l^H(\mathbf{h}_M) \mathbf{n}_{l,1}^{(i)} = \alpha_1^{(i)} \mathbf{e}_{l+M,l+M}$$

De la même manière, il existe un scalaire complexe $\alpha_2^{(i)}$ telque

$$\mathcal{T}_l^H(\mathbf{h}_M) \mathbf{n}_{l,2}^{(i)} = \alpha_2^{(i)} \mathbf{e}_{l+M,1}, \quad \forall i = 1, \dots, w.$$

Les inconnues $\alpha_1^{(i)}$ et $\alpha_2^{(i)}$ peuvent être déterminées (à une ambiguïté de phase près) puisque

$$\mathbf{n}_{l,j}^{(i)H} (\mathbf{R}_l - \mathbf{R}_l^b) \mathbf{n}_{l,j}^{(i)} = \sigma_s^2 \mathbf{n}_{l,j}^{(i)H} \mathcal{T}_l(\mathbf{h}_M) \mathcal{T}_l^H(\mathbf{h}_M) \mathbf{n}_{l,j}^{(i)} = \sigma_s^2 \left\| \mathcal{T}_l^H(\mathbf{h}_M) \mathbf{n}_{l,j}^{(i)} \right\|^2 = \sigma_s^2 |\alpha_j^{(i)}|^2, \quad j = 1, 2$$

Ainsi, les vecteurs

$$\mathbf{g}_{l-1,l+M}^{(i)} \triangleq \frac{1}{\sigma_s \sqrt{\mathbf{n}_{l,1}^{(i)H} (\mathbf{R}_l - \mathbf{R}_l^b) \mathbf{n}_{l,1}^{(i)}}} \mathbf{n}_{l,1}^{(i)*} \quad (5.1)$$

$$\text{et } \mathbf{g}_{l-1,1}^{(i)} \triangleq \frac{1}{\sigma_s \sqrt{\mathbf{n}_{l,2}^{(i)H} (\mathbf{R}_l - \mathbf{R}_l^b) \mathbf{n}_{l,2}^{(i)}}} \mathbf{n}_{l,2}^{(i)*} \quad (5.2)$$

vérifient respectivement

$$\begin{aligned}\mathcal{T}_l^T(\mathbf{h}_M)\mathbf{g}_{l-1,l+M}^{(i)} &= \frac{\alpha_1^{(i)*}}{|\alpha_1^{(i)}|}\mathbf{e}_{l+M,l+M} \\ \text{et } \mathcal{T}_l^T(\mathbf{h}_M)\mathbf{g}_{l-1,1}^{(i)} &= \frac{\alpha_2^{(i)*}}{|\alpha_2^{(i)}|}\mathbf{e}_{l+M,1}\end{aligned}$$

Ils correspondent donc à des égaliseurs ZF d'ordre $(l-1)$ avec un délai maximal et minimal respectivement; i.e., qu'en l'absence de bruit, ils permettent une restitution parfaite des symboles transmis avec l'amplitude exacte et à une rotation près.

Chacun des vecteurs $\mathbf{n}_{l,1}^{(i)}$ et $\mathbf{n}_{l,2}^{(i)}$ permet de calculer un égaliseur ZF. On dispose ainsi d'une famille de $2w$ égaliseurs ZF pour un choix donné des familles $\{\mathbf{n}_{l,1}^{(i)}\}_{i=1,\dots,w}$ et $\{\mathbf{n}_{l,2}^{(i)}\}_{i=1,\dots,w}$. Sur la base de ces égaliseurs ZF, on peut déterminer les coefficients de la réponse impulsionnelle $\mathbf{h}_M$ en suivant la procédure décrite dans §4.16, ce qui donne :

Étant donné $\mathbf{g}_{l-1,1}^{(i)}$, la réponse du canal est

$$\mathbf{h}_M = \begin{bmatrix} \gamma^x(0) & \gamma^x(1) & \cdots & \gamma^x(l-1) \\ \gamma^x(1) & \gamma^x(2) & & \gamma^x(l) \\ \vdots & & & \vdots \\ \gamma^x(M) & \cdots & & \gamma^x(l+M-1) \end{bmatrix} \mathbf{g}_{l-1,1}^{(i)*} \quad (5.3)$$

Étant donné $\mathbf{g}_{l-1,l+M}^{(i)}$, la réponse du canal est

$$\mathbf{h}_M = \begin{bmatrix} \gamma^x(-l-M+1) & \gamma^x(-l-M+2) & \cdots & \gamma^x(-M) \\ \gamma^x(-l-M+2) & \gamma^x(-l-M+3) & & \gamma^x(-M+1) \\ \vdots & & & \vdots \\ \gamma^x(-l+1) & \cdots & & \gamma^x(0) \end{bmatrix} \mathbf{g}_{l-1,l+M}^{(i)*} \quad (5.4)$$

Puisque les égaliseurs $\mathbf{g}_{l-1,1}^{(i)}$ et $\mathbf{g}_{l-1,l+M}^{(i)}$ sont calculés à une rotation près, la réponse du canal est déterminée à une constante complexe unitaire près.

Les systèmes d'équation (5.3) et (5.4) font intervenir les termes $\gamma(l), \dots, \gamma(M)$ qui manquent à la matrice $\mathcal{R}_l$. Pour achever l'identification de $\mathbf{h}_M$, on renforce la contrainte **H4.1** sur le facteur de lissage l qui doit dépasser l'ordre du canal, soit

Hypothèse H 5.1 $l \geq M$.

Étant donné **H5.1** et le fait que le processus $\mathbf{x}(n)$ est M -MA, on réécrit (5.3) comme suit

$$\mathbf{h}_M = \begin{bmatrix} \gamma^x(0) & \cdots & \gamma^x(M-1) & \gamma^x(M) & \mathbf{0} & \cdots \\ \gamma^x(1) & & \gamma^x(M) & \mathbf{0} & \cdots & \\ \vdots & & \vdots & & & \\ \gamma^x(M) & \mathbf{0} & \cdots & & & \end{bmatrix} \mathbf{g}_{l-1,1}^{(i)*}, \text{ si } l \geq M+1$$

$$= \begin{bmatrix} \gamma^x(0) & \gamma^x(1) & \cdots & \gamma^x(M-1) \\ \gamma^x(1) & \gamma^x(2) & \cdots & \gamma^x(M) \\ \vdots & & & \mathbf{0} \\ & & & \vdots \\ \gamma^x(M) & \mathbf{0} & \cdots & \mathbf{0} \end{bmatrix} \mathbf{g}_{M-1,1}^{(i)*}$$

et on réécrit (5.4) comme suit

$$\begin{aligned} \mathbf{h}_M &= \begin{bmatrix} & & \cdots & \mathbf{0} & \gamma^x(-M) \\ & & & \vdots & \\ & \cdots & \mathbf{0} & \gamma^x(-M) & \cdots & \gamma^x(-1) \\ \cdots & \mathbf{0} & \gamma^x(-M) & \gamma^x(-M+1) & \cdots & \gamma^x(0) \end{bmatrix} \mathbf{g}_{l-1,l+M}^{(i)*}, \text{ si } l \geq M+1 \\ &= \begin{bmatrix} \mathbf{0} & \cdots & \mathbf{0} & \gamma^x(-M) \\ \vdots & & & \vdots \\ \mathbf{0} & & & \\ \gamma^x(-M) & \cdots & & \gamma^x(-1) \\ \gamma^x(-M+1) & \cdots & & \gamma^x(0) \end{bmatrix} \mathbf{g}_{M-1,2M}^{(i)*} \end{aligned}$$

où $\mathbf{0}$ représente ici des matrices nulles $c \times c$.

5.4 Robustesse à la surestimation de l'ordre en statistiques exactes

On donne ici la preuve théorique que l'algorithme proposé est robuste à la surestimation de l'ordre du canal. Lorsque l'algorithme est tourné en presumant un ordre M' du canal supérieur à l'ordre exact M , la réponse estimée (d'ordre M') est, à des termes nuls limitrophes près, égale à la réponse exacte du canal, lorsque les statistiques (la matrice de corrélation) sont exactement connues.

Si une erreur de détection de l'ordre du canal $\mathbf{h}_M$ conduit à lui supposer un ordre $M' > M$, le facteur de lissage l doit être choisi de manière à remplir **H5.1** donc $l \geq M' > M$. Ainsi choisi, l est praticable pour achever l'identification suivant la procédure ci-dessus. Le rang de $\mathcal{R}_l$ est estimé à $l + M' + 1$ i.e., supérieur à $l + M + 1$ qui est la valeur exacte du rang. Ainsi, on calcule moins de vecteurs $\mathbf{n}_{l,1}^{(i)}$ et $\mathbf{n}_{l,2}^{(i)}$ qu'il en existe vraiment, soit $i = 1, \dots, w' = (c-1)l - M' + 1$ au lieu de $i = 1, \dots, (c-1)l - M + 1$. L'algorithme va pouvoir aboutir car il lui suffit de disposer de l'un des vecteurs $\mathbf{n}_{l,1}^{(i)}$ et $\mathbf{n}_{l,2}^{(i)}$.

En poursuivant la procédure dans §5.3, l'algorithme calcule les vecteurs $\mathbf{g}_{l-1,l+M'}^{(i)}$ et $\mathbf{g}_{l-1,1}^{(i)}$, et identifie les réponses impulsionnelles suivantes :

En utilisant (5.3)

$$\begin{aligned} \begin{bmatrix} \gamma^x(0) & \gamma^x(1) & \cdots & \gamma^x(l-1) \\ \gamma^x(1) & \gamma^x(2) & & \gamma^x(l) \\ \vdots & & & \vdots \\ \gamma^x(M) & \cdots & & \gamma^x(l+M-1) \\ \vdots & & & \vdots \\ \gamma^x(M') & \cdots & & \gamma^x(l+M'-1) \end{bmatrix} \mathbf{g}_{l-1,1}^{(i)*} &= \begin{bmatrix} & & \mathbf{h}_M \\ \left[\begin{array}{ccc} \gamma^x(M+1) & \cdots & \gamma^x(l+M) \\ \vdots & & \vdots \\ \gamma^x(M') & \cdots & \gamma^x(l+M'-1) \end{array} \right] \mathbf{g}_{l-1,1}^{(i)*} \end{bmatrix} \\ &= \begin{bmatrix} \mathbf{h}_M \\ \mathbf{0}_{c(M'-M),1} \end{bmatrix} \end{aligned}$$

ou bien en utilisant (5.4)

$$\begin{aligned}
& \begin{bmatrix} \gamma^x(-l-M'+1) & \gamma^x(-l-M'+2) & \cdots & \gamma^x(-M') \\ \vdots & \vdots & & \vdots \\ \gamma^x(-l-M) & \gamma^x(-l-M+1) & \cdots & \gamma^x(-M-1) \\ \gamma^x(-l-M+1) & \gamma^x(-l-M+2) & \cdots & \gamma^x(-M) \\ \gamma^x(-l-M+2) & \gamma^x(-l-M+3) & & \gamma^x(-M+1) \\ \vdots & & & \vdots \\ \gamma^x(-l+1) & \cdots & & \gamma^x(0) \end{bmatrix} \mathbf{g}_{l-1,l+M}^{(i)*} \\
&= \begin{bmatrix} \begin{bmatrix} \gamma^x(-l-M'+1) & \gamma^x(-l-M'+2) & \cdots & \gamma^x(-M') \\ \vdots & \vdots & & \vdots \\ \gamma^x(-l-M) & \gamma^x(-l-M+1) & \cdots & \gamma^x(-M-1) \end{bmatrix} \mathbf{g}_{l-1,l+M}^{(i)*} \\ \mathbf{h}_M \end{bmatrix} \\
&= \begin{bmatrix} \mathbf{0}_{c(M'-M),1} \\ \mathbf{h}_M \end{bmatrix}
\end{aligned}$$

où on a exploité le fait que puisque le processus $\mathbf{x}(n)$ est un M -MA, $\gamma^x(k) = \mathbf{0}_{c,c}$, $\forall |k| > M$.

Ainsi, on obtient à chaque fois une version *zero-padded* de la vraie réponse impulsionnelle. Donc l'estimée obtenue peut être utilisée pour les besoins de l'égalisation (cf. §4.4).

5.5 Cas des statistiques estimées

En pratique, l'algorithme d'identification ne dispose que d'une estimée des statistiques du canal, ici la matrice de corrélation $\mathcal{R}_l$. L'estimée de $\mathcal{R}_l$ ne doit pas être déficiente et les vecteurs $\mathbf{n}_{l,1}^{(i)}$ et $\mathbf{n}_{l,2}^{(i)}$ tels que définis dans §5.3 ne doivent pas exister. L'algorithme doit être réécrit en conséquence.

Le vecteur $\mathbf{n}_{l,1}^{(i)}$ (resp. $\mathbf{n}_{l,2}^{(i)}$) doit être choisi comme étant le vecteur singulier à droite (resp. à gauche) de l'estimée $\hat{\mathcal{R}}_l$ de $\mathcal{R}_l$ associé à sa $i^{\text{ème}}$ plus petite valeur singulière. Chacun de ces vecteurs permet, par le même enchaînement décrit dans §5.3, de calculer une estimée d'un égaliseur ZF à délai maximal (resp. minimal) qui à son tour permet de calculer une estimée de la réponse du canal (ou bien une version *zero padded* si la valeur détectée de l'ordre du canal est supérieure à l'ordre exact). Ces estimées n'ayant aucune raison d'être les mêmes, il faut définir un critère pour retenir le *meilleur* et réécrire l'algorithme en conséquence.

Tout d'abord, nous annonçons le résultat suivant qui nous sera utile par la suite.

Proposition 5.1 *Soient $\mathbf{x}$ et $\mathbf{y}$ deux vecteurs complexes de même taille, alors*

$$\min_{\gamma \in \mathbf{C}} \|\mathbf{x} - \gamma \mathbf{y}\|^2 = \|\mathbf{x}\|^2 - \left(\frac{|\mathbf{x}^H \mathbf{y}|}{\|\mathbf{y}\|} \right)^2$$

Démonstration:

On a $\|\mathbf{x} - \gamma \mathbf{y}\|^2 = \|\mathbf{x}\|^2 + |\gamma|^2 \|\mathbf{y}\|^2 - \gamma \mathbf{x}^H \mathbf{y} - \gamma^* \mathbf{y}^H \mathbf{x}$. En écrivant $\gamma = \alpha + i\beta$ où $\alpha, \beta \in \mathbf{R}$, le

minimum est atteint lorsque $2\alpha \|\mathbf{y}\|^2 - (\mathbf{x}^H \mathbf{y} + \mathbf{y}^H \mathbf{x}) = 0$ et $2\beta \|\mathbf{y}\|^2 - i(\mathbf{x}^H \mathbf{y} - \mathbf{y}^H \mathbf{x}) = 0$; donc lorsque

$$\begin{aligned} \gamma &= \frac{\mathbf{x}^H \mathbf{y} + \mathbf{y}^H \mathbf{x} - (\mathbf{x}^H \mathbf{y} - \mathbf{y}^H \mathbf{x})}{2 \|\mathbf{y}\|^2} \\ &= \frac{\mathbf{x}^H \mathbf{y} + (\mathbf{x}^H \mathbf{y})^* - (\mathbf{x}^H \mathbf{y} - (\mathbf{x}^H \mathbf{y})^*)}{2 \|\mathbf{y}\|^2} \\ &= \frac{\operatorname{Re}(\mathbf{x}^H \mathbf{y}) - i(\operatorname{Im}(\mathbf{x}^H \mathbf{y}))}{\|\mathbf{y}\|^2} \\ &= \frac{\mathbf{y}^H \mathbf{x}}{\|\mathbf{y}\|^2} \end{aligned}$$

Donc

$$\begin{aligned} \|\mathbf{y}\|^4 \min_{\gamma \in \mathbb{C}} \|\mathbf{x} - \gamma \mathbf{y}\|^2 &= [\|\mathbf{y}\|^2 \mathbf{x}^H - (\mathbf{x}^H \mathbf{y}) \mathbf{y}^H] [\|\mathbf{y}\|^2 \mathbf{x} - (\mathbf{y}^H \mathbf{x}) \mathbf{y}] \\ &= \|\mathbf{y}\|^4 \|\mathbf{x}\|^2 (\mathbf{y}^H \mathbf{x})(\mathbf{y}^H \mathbf{x})^* \|\mathbf{y}\|^2 - \|\mathbf{y}\|^2 (\mathbf{x}^H \mathbf{y}) \mathbf{y}^H \mathbf{x} - \|\mathbf{y}\|^2 (\mathbf{y}^H \mathbf{x}) \mathbf{x}^H \mathbf{y} \\ &= \|\mathbf{y}\|^4 \|\mathbf{x}\|^2 - \|\mathbf{y}\|^2 |\mathbf{x}^H \mathbf{y}|^2 \end{aligned}$$

d'où le résultat. ■

L'algorithme conduit au calcul de $2w$ estimées de la réponse impulsionnelle du canal émanant des w vecteurs $\mathbf{n}_{l,1}^{(i)}$ et des w vecteurs $\mathbf{n}_{l,2}^{(i)}$. On note $\{\hat{\mathbf{h}}_M^{(i)}, i = 1, \dots, 2w\}$ l'ensemble de ces estimées. Pour comparer ces estimées à la vraie réponse du canal, on ne dispose que de l'information $\hat{\mathbf{R}}_l$ i.e., l'estimée de la matrice de corrélation de sa sortie. On se propose d'évaluer la différence entre $\mathcal{T}_l(\hat{\mathbf{h}}_M^{(i)}) \mathcal{T}_l^H(\hat{\mathbf{h}}_M^{(i)})$ et $\hat{\mathbf{R}}_l - \widehat{\sigma_b^2} \mathbf{I}_{cl}$ par le biais d'une norme matricielle. Il s'en suit le critère suivant : on retient l'estimée $\hat{\mathbf{h}}_M$ telle que

Critère 5.1

$$\hat{\mathbf{h}}_M = \operatorname{argmin}_i \left(\min_{\beta} \left\| \hat{\mathbf{R}}_l - \widehat{\sigma_b^2} \mathbf{I}_{cl} - \beta \mathcal{T}_l(\hat{\mathbf{h}}_M^{(i)}) \mathcal{T}_l^H(\hat{\mathbf{h}}_M^{(i)}) \right\|^2 \right) \quad (5.5)$$

où $\|\cdot\|$ désigne une norme matricielle.

On peut choisir $\beta = 1$ ou β complexe quelconque selon que l'on veut résoudre le problème de l'identification avec une ambiguïté de phase² ou une ambiguïté de phase et d'amplitude, respectivement i.e., si on veut déterminer la réponse du canal à une constante complexe de norme unitaire ou de norme quelconque, respectivement.

On rappelle que la norme de Frobenius d'une matrice $\mathbf{A}$ $n \times n$ est donnée par $\|\mathbf{A}\| \triangleq \frac{1}{\sqrt{n}} \|\operatorname{Vec}(\mathbf{A})\|$. Pour cette norme, et pour une ambiguïté de phase et d'amplitude (β complexe quelconque), (5.5) se simplifie comme suit :

$$\begin{aligned} \hat{\mathbf{h}}_M &= \operatorname{argmin}_i \left(\min_{\beta} \left\| \hat{\mathbf{R}}_l - \widehat{\sigma_b^2} \mathbf{I}_{cl} - \beta \mathcal{T}_l(\hat{\mathbf{h}}_M^{(i)}) \mathcal{T}_l^H(\hat{\mathbf{h}}_M^{(i)}) \right\|^2 \right) \\ &= \operatorname{argmin}_i \left(\min_{\beta} \left\| \operatorname{Vec}(\hat{\mathbf{R}}_l - \widehat{\sigma_b^2} \mathbf{I}_{cl}) - \beta \operatorname{Vec}(\mathcal{T}_l(\hat{\mathbf{h}}_M^{(i)}) \mathcal{T}_l^H(\hat{\mathbf{h}}_M^{(i)})) \right\|^2 \right) \end{aligned}$$

²Comme discuté dans §4.4, l'ambiguïté de phase est inhérente au problème de l'identification

$$\begin{aligned}
&= \operatorname{argmin}_i \left(\left\| \operatorname{Vec} \left(\hat{\mathbf{R}}_l - \widehat{\sigma_b^2} \mathbf{I}_{cl} \right) \right\|^2 - \frac{\left| \operatorname{Vec} \left(\hat{\mathbf{R}}_l - \widehat{\sigma_b^2} \mathbf{I}_{cl} \right)^H \operatorname{Vec} \left(\mathcal{T}_l \left(\hat{\mathbf{h}}_M^{(i)} \right) \mathcal{T}_l^H \left(\hat{\mathbf{h}}_M^{(i)} \right) \right) \right|^2}{\left\| \operatorname{Vec} \left(\mathcal{T}_l \left(\hat{\mathbf{h}}_M^{(i)} \right) \mathcal{T}_l^H \left(\hat{\mathbf{h}}_M^{(i)} \right) \right\|^2} \right) \\
&= \operatorname{argmin}_i \left(cl \left| \hat{\mathbf{R}}_l - \widehat{\sigma_b^2} \mathbf{I}_{cl} \right|^2 - \frac{\left| \operatorname{Vec} \left(\hat{\mathbf{R}}_l - \widehat{\sigma_b^2} \mathbf{I}_{cl} \right)^H \operatorname{Vec} \left(\mathcal{T}_l \left(\hat{\mathbf{h}}_M^{(i)} \right) \mathcal{T}_l^H \left(\hat{\mathbf{h}}_M^{(i)} \right) \right) \right|^2}{cl \left| \mathcal{T}_l \left(\hat{\mathbf{h}}_M^{(i)} \right) \mathcal{T}_l^H \left(\hat{\mathbf{h}}_M^{(i)} \right) \right|^2} \right) \\
\text{Soit, } \hat{\mathbf{h}}_M &= \operatorname{argmax}_i \frac{\left| \operatorname{Vec} \left(\hat{\mathbf{R}}_l - \widehat{\sigma_b^2} \mathbf{I}_{cl} \right)^H \operatorname{Vec} \left(\mathcal{T}_l \left(\hat{\mathbf{h}}_M^{(i)} \right) \mathcal{T}_l^H \left(\hat{\mathbf{h}}_M^{(i)} \right) \right) \right|}{\left| \mathcal{T}_l \left(\hat{\mathbf{h}}_M^{(i)} \right) \mathcal{T}_l^H \left(\hat{\mathbf{h}}_M^{(i)} \right) \right|}
\end{aligned}$$

Toujours dans le cas d'une ambiguïté de phase et d'amplitude (i.e., β réel quelconque), il n'est plus nécessaire de calculer les égaliseurs ZF comme décrit dans §5.3. On modifie (5.1) et (5.2) pour calculer $\mathbf{g}_{l-1,l+M}^{(i)} \hat{=} \mathbf{n}_{l,1}^{(i)*}$ et/ou $\mathbf{g}_{l-1,1}^{(i)} \hat{=} \mathbf{n}_{l,2}^{(i)*}$. L'identification s'achève comme auparavant.

5.6 Robustesse à la surestimation de l'ordre en statistiques estimées

Étant données les considérations "statistiques estimées" ci-dessus (cf. §5.5), on montre dans ce qui suit que la propriété de robustesse à la surestimation de l'ordre, prouvée dans le cas des "statistiques exactes" (cf. §5.4), se maintient dans le cas statistiques estimées. En particulier, on fera le parallèle avec les algorithmes Prédiction Linéaire (LP) et Décomposition en Produit Extérieur (OPD) (cf. 5.2).

L'algorithme proposé (resp. les algorithmes LP et OPD) exploitent les propriétés de la matrice de corrélation décalée (resp. standard) débruitée $\mathcal{R}_l - \mathcal{R}_l^b$ (resp. $\mathbf{R}_l - \mathbf{R}_l^b$). Son estimée $\hat{\mathcal{R}}_l - \hat{\mathcal{R}}_l^b$ (resp. $\hat{\mathbf{R}}_l - \hat{\mathbf{R}}_l^b$) approxime donc une matrice de rang déficient. Ses plus petites valeurs singulières (resp. propres), les $(c-1)l - M + 1$ (resp. $(c-1)l - M$) plus petites, correspondant au sous-espace bruit, sont proches de 0.

Les algorithmes LP et OPD font intervenir la pseudo-inverse de la matrice $\hat{\mathcal{R}}_l - \hat{\mathcal{R}}_l^b$. Pour l'inverser proprement (forcer à zero les valeurs propres du sous-espace bruit et inverser celles du sous-espace signal), l'algorithme doit connaître la dimension exacte du sous-espace bruit (donc l'ordre exact du canal). Lorsque l'ordre du canal est surestimé, la dimension du sous-espace bruit est sousestimée et certaines de ses valeurs propres sont comptabilisées dans le sous-espace signal, donc seront inversées pour calculer la pseudo-inverse. Ainsi, dans le cas des statistiques estimées, la surestimation de l'ordre conduit au mauvais conditionnement de la matrice de corrélation (standard) débruitée pour l'inversion, et par conséquent à l'échec des algorithmes LP et OPD.

De la même manière que pour la matrice de corrélation standard débruitée, la surestimation de l'ordre conduit à sousestimer la dimension du sous-espace bruit de la matrice de corrélation décalée débruitée. Cependant, contrairement à ce qui se produit avec les algorithmes LP et OPD, l'algorithme proposé ne procède pas à une inversion de la matrice. Ainsi, son mauvais conditionnement (pour l'inversion) n'a pas d'incidence sur les performances de l'algorithme. La surestimation de l'ordre aura, comme seul effet, une baisse de performance de l'algorithme. En effet, dans ces conditions, on estime seulement une partie du sous-espace bruit. Chacun des vecteurs $\mathbf{n}_{l,j}^{(i)}, i = 1, \dots, \hat{w} \hat{=} (c-1)l - \hat{M} + 1 \leq$

$$\begin{array}{c}
\hat{\mathbf{R}}_l - \hat{\mathbf{R}}_l^b \\
\text{estimée} \\
\text{débruitée}
\end{array}
\begin{bmatrix}
X & & & & \\
& \ddots & & & \\
& & X & & \\
& & & \epsilon & \\
& & & & \ddots \\
& & & & & \epsilon
\end{bmatrix}
\begin{array}{c}
\text{Ordre exact} \\
1/(\dots) \quad 0
\end{array}$$

$$\begin{array}{c}
\hat{\mathcal{R}}_l - \hat{\mathcal{R}}_l^b \\
\text{estimée} \\
\text{débruitée}
\end{array}
\begin{bmatrix}
X & & & & \\
& \ddots & & & \\
& & X & & \\
& & & \epsilon & \\
& & & & \ddots \\
& & & & & \ddots \\
& & & & & & \epsilon
\end{bmatrix}
\begin{array}{c}
\text{Ordre surestimé} \\
1/(\dots) \quad 0
\end{array}$$

$$\begin{array}{c}
\mathbf{n}^{(cl)} \quad \dots \quad \mathbf{n}^{(w)} \dots \mathbf{n}^{(\hat{w})} \dots \mathbf{n}^{(1)} \\
\vdots \quad \vdots \\
\hat{\mathbf{h}}_M^{(\hat{w})} \quad \hat{\mathbf{h}}_M^{(1)} \\
\hline
\hat{\mathbf{h}}_M
\end{array}
\begin{array}{c}
\text{Ordre surestimé}
\end{array}$$

Figure 5.2: Robustesse à la surestimation de l'ordre en statistiques estimées.

$(c-1)l - M + 1, j = 1, 2$, conduit à une estimée de la réponse impulsionnelle du canal. On récupère ainsi moins d'estimées que ce que l'on aurait récupéré si l'on avait connu la valeur exacte de l'ordre. La sélection de la *meilleure estimée* se fera dans un ensemble plus petit et l'erreur de performance se détériore par conséquent.

Ainsi, l'algorithme proposé peut se contenter d'une seule direction dans le sous-espace bruit (de la matrice de corrélation décalée) pour retrouver une estimée de la réponse du canal; alors que l'algorithme sous-espace (SS) a besoin de détecter la totalité du sous-espace bruit (de la matrice de corrélation standard) pour caractériser la réponse du canal. Tout comme les algorithmes LP et OPD, l'algorithme proposé opère sur une matrice mal conditionnée (pour l'inversion). Mais puisqu'il ne nécessite pas l'inversion de cette matrice, il continue d'être performant. Ainsi, vis-à-vis de la surestimation de l'ordre, l'algorithme SS est *fondamentalement* non robuste (même en statistiques exactes). La robustesse de LP et OPD est *théorique* (seulement en statistiques exactes). Celle de l'algorithme proposé est *réelle*.

Toutefois, la surestimation de l'ordre a pour effet la baisse de performance de l'algorithme proposé

à cause, comme expliqué ci-dessus, d'une restriction du nombre des estimées de la réponse du canal. Ceci peut être compensé en augmentant le facteur de lissage.

On peut également autoriser l'algorithme à scruter tous les vecteurs singuliers $\mathbf{n}_{l,j}^{(i)}, j = 1, 2$, ceux du sous-espace bruit ($i = 1, \dots, (c-1)l - M + 1$) et ceux du sous-espace signal ($i = (c-1)l - M + 2, \dots, cl$). Ces derniers vont naturellement aboutir à de *mauvaises* estimées de la réponse du canal. La prise en compte de ces estimées est sans danger si le critère de sélection est capable de les écarter. Celui proposé (critère 5.1), sous-optimal par construction, s'est révélé, à travers les simulations, pouvant être induit en erreur par une *mauvaise* estimée; au sens où une estimée $\hat{\mathbf{h}}_M^{(i)}$ issue d'un vecteur du sous-espace signal peut avoir un coût $\min_{\beta} \left\| \hat{\mathbf{R}}_l - \widehat{\sigma_b^2} \mathbf{I}_{cl} - \beta \mathcal{T}_l \left(\hat{\mathbf{h}}_M^{(i)} \right) \mathcal{T}_l^H \left(\hat{\mathbf{h}}_M^{(i)} \right) \right\|^2$ inférieur à ceux atteints par les estimées issues des vecteurs du sous-espace bruit. Ainsi, l'exploitation de cette possibilité reste suspendue à la résolution, d'une manière optimale, du choix de la *meilleure* estimée dans l'ensemble $\{\hat{\mathbf{h}}_M^{(i)}\}$.

5.7 Récapitulatif

Étant données les modifications ci-dessus, l'algorithme se réécrit comme suit :

1. Estimer une valeur M supérieure à l'ordre exact du canal.
2. Choisir un facteur de lissage $l \geq M$.
3. Calculer $\hat{\mathcal{R}}_l$ estimée de $\mathcal{R}_l$.
4. Estimer la puissance du bruit $\widehat{\sigma_b^2}$ comme la valeur moyenne des $(c-1)(l+1) - M$ plus petites valeurs propres de $\hat{\mathbf{R}}_{l+1}$.
5. Calculer les vecteurs cl -dim $\mathbf{n}_{l,1}^{(i)}$ et $\mathbf{n}_{l,2}^{(i)}$ singuliers à gauche et à droite respectivement, associés aux $w \hat{=}(c-1)l - M + 1$ plus petites valeurs singulières de $\hat{\mathcal{R}}_l - \widehat{\sigma_b^2} (\mathbf{J}_l \otimes \mathbf{I}_c)$.
Constituer l'ensemble $\{\mathbf{n}_l^{(i)}, i = 1, \dots, 2w\} = \{\mathbf{n}_{l,1}^{(i)}, i = 1, \dots, w\} \cup \{\mathbf{n}_{l,2}^{(i)}, i = 1, \dots, w\}$.
6. Pour chaque $\mathbf{n}_l^{(i)}$, estimer l'égaliseur ZF

$$\mathbf{g}_{l-1}^{(i)} \hat{=} \frac{1}{\sigma_s \sqrt{\mathbf{n}_l^{(i)H} \left(\hat{\mathbf{R}}_l - \widehat{\sigma_b^2} \mathbf{I}_{cl} \right) \mathbf{n}_l^{(i)}}} \mathbf{n}_l^{(i)*} \quad (\text{ambiguïté de phase})$$

ou bien $\mathbf{g}_{l-1}^{(i)} \hat{=} \mathbf{n}_l^{(i)*} \quad (\text{ambiguïté de phase et d'amplitude})$

7. Pour chaque $\mathbf{g}_{l-1}^{(i)}$, calculer l'estimée $\hat{\mathbf{h}}_M^{(i)}$ correspondante en utilisant (5.3) ou (5.4) selon que $\mathbf{n}_l^{(i)}$ est un vecteur singulier à gauche ou à droite, respectivement.
8. Choisir

$$\hat{\mathbf{h}}_M = \underset{i}{\operatorname{argmin}} \left(\left\| \hat{\mathbf{R}}_l - \widehat{\sigma_b^2} \mathbf{I}_{cl} - \mathcal{T}_l \left(\hat{\mathbf{h}}_M^{(i)} \right) \mathcal{T}_l^H \left(\hat{\mathbf{h}}_M^{(i)} \right) \right\| \right)$$

(ambiguïté de phase)

$$\text{ou bien } \hat{\mathbf{h}}_M = \underset{i}{\operatorname{argmax}} \frac{\left| \operatorname{Vec} \left(\hat{\mathbf{R}}_l - \widehat{\sigma_b^2} \mathbf{I}_{cl} \right)^H \operatorname{Vec} \left(\mathcal{T}_l \left(\hat{\mathbf{h}}_M^{(i)} \right) \mathcal{T}_l^H \left(\hat{\mathbf{h}}_M^{(i)} \right) \right) \right|}{\left| \mathcal{T}_l \left(\hat{\mathbf{h}}_M^{(i)} \right) \mathcal{T}_l^H \left(\hat{\mathbf{h}}_M^{(i)} \right) \right|}$$

(ambiguïté de phase et d'amplitude)

5.8 Mesure de performances

Dans ce qui suit, on propose une mesure de performance $\mathcal{E}(\hat{\mathbf{h}}_M)$ inspirée de [MORG 98]. Pour les applications ciblées, l'erreur doit être nulle lorsque l'estimée est égale à la réponse exacte du canal à une constante complexe quelconque et à des coefficients nuls à la tête et ou à la queue près. Autrement dit, l'erreur $\mathcal{E}(\hat{\mathbf{h}}_{M'})$ sur $\mathbf{h}_M$ doit vérifier³

$$\mathcal{E} \left(\beta \begin{bmatrix} \mathbf{0}_{cM_1,1} \\ \mathbf{h}_M \\ \mathbf{0}_{cM_2,1} \end{bmatrix} \right) = 0, \forall \beta \in \mathbf{C}, \text{ et } \forall M_1 \text{ et } M_2 \text{ entiers positifs.}$$

Ainsi, pour mesurer les performances de l'algorithme, on propose l'erreur suivante

$$\mathcal{E}(\hat{\mathbf{h}}_{M'}) \triangleq \min_{\beta, M_1 + M_2 = M' - M} \left(\frac{\left\| \begin{bmatrix} \mathbf{0}_{cM_1,1} \\ \mathbf{h}_M \\ \mathbf{0}_{cM_2,1} \end{bmatrix} - \beta \hat{\mathbf{h}}_{M'} \right\|}{\|\mathbf{h}_M\|} \right)^2$$

Cependant, cette erreur peut se simplifier pour notre algorithme en particulier. En effet, une estimée d'ordre $M' > M$ est prévue pour *ressembler* à $\begin{bmatrix} \mathbf{0}_{c(M'-M),1} \\ \mathbf{h}_M \end{bmatrix}$ ou bien à $\begin{bmatrix} \mathbf{h}_M \\ \mathbf{0}_{c(M'-M),1} \end{bmatrix}$. Par conséquent, l'erreur se réécrit comme

$$\mathcal{E}(\hat{\mathbf{h}}_{M'}) \triangleq \left[\min_{\beta \in \mathbf{C}} \left\| \begin{bmatrix} \mathbf{0}_{c(M'-M),1} \\ \mathbf{h}_M \end{bmatrix} - \beta \hat{\mathbf{h}}_{M'} \right\|^2, \min_{\beta \in \mathbf{C}} \left\| \begin{bmatrix} \mathbf{h}_M \\ \mathbf{0}_{c(M'-M),1} \end{bmatrix} - \beta \hat{\mathbf{h}}_{M'} \right\|^2 \right]$$

On peut montrer, en utilisant Prop. 5.1, que ceci est égale à

$$\begin{aligned} \mathcal{E}(\hat{\mathbf{h}}_{M'}) &= \min \left[1 - \left(\frac{\left| \begin{bmatrix} \mathbf{0}_{1,c(M'-M)} \mathbf{h}_M^H \end{bmatrix} \hat{\mathbf{h}}_{M'} \right|}{\|\mathbf{h}_M\| \|\hat{\mathbf{h}}_{M'}\|} \right)^2, 1 - \left(\frac{\left| \begin{bmatrix} \mathbf{h}_M^H \mathbf{0}_{1,c(M'-M)} \end{bmatrix} \hat{\mathbf{h}}_{M'} \right|}{\|\mathbf{h}_M\| \|\hat{\mathbf{h}}_{M'}\|} \right)^2 \right] \\ &= 1 - \left(\max \left(\left| \begin{bmatrix} \mathbf{0}_{1,c(M'-M)} \mathbf{h}_M^H \end{bmatrix} \hat{\mathbf{h}}_{M'} \right|, \left| \begin{bmatrix} \mathbf{h}_M^H \mathbf{0}_{1,c(M'-M)} \end{bmatrix} \hat{\mathbf{h}}_{M'} \right| \right) \right)^2 \\ &\quad \frac{1}{\|\mathbf{h}_M\| \|\hat{\mathbf{h}}_{M'}\|} \end{aligned}$$

5.9 Simulations

Une série de simulations a été conduite pour tester l'algorithme, et en particulier sa robustesse quant à la surestimation de l'ordre, et pour le comparer en terme d'erreur d'estimation avec d'autres algorithmes, en l'occurrence les algorithmes SS, LP et OPD. On a testé les performances de l'algorithme

³en format ST.

en jouant sur les conditions de réception (la puissance du bruit et la durée d'observation) et sur le facteur de lissage. Les tests en surestimation de l'ordre ont été conduit uniquement sur l'algorithme proposé étant donné que, dans cette situation, les autres algorithmes échouent systématiquement.

Le but de ces simulations étant de vérifier les développements théoriques précédents, le canal n'a pas été choisi avec le souci de correspondre à des situations de propagation réalistes⁴. Le canal testé (cf. 5.3.a) est repris de [MOUL 95]. Il s'agit d'un SIMO à 4 sous-canaux et d'ordre $M = 4$. Ses coefficients sont donnés par

$$\begin{aligned} \mathbf{h}(0) &= \begin{bmatrix} -0.049 + 0.359\sqrt{-1} & 0.443 - 0.0364\sqrt{-1} & -0.211 - 0.322\sqrt{-1} & 0.417 + 0.030\sqrt{-1} \end{bmatrix}^T, \\ \mathbf{h}(1) &= \begin{bmatrix} 0.482 - 0.569\sqrt{-1} & 1 & -0.199 + 0.918\sqrt{-1} & 1 \end{bmatrix}^T, \\ \mathbf{h}(2) &= \begin{bmatrix} -0.556 + 0.587\sqrt{-1} & 0.921 - 0.194\sqrt{-1} & 1 & 0.873 + 0.145\sqrt{-1} \end{bmatrix}^T, \\ \mathbf{h}(3) &= \begin{bmatrix} 1 & 0.189 - 0.208\sqrt{-1} & -0.284 - 0.524\sqrt{-1} & 0.285 + 0.309\sqrt{-1} \end{bmatrix}^T \text{ et} \\ \mathbf{h}(4) &= \begin{bmatrix} -0.171 + 0.061\sqrt{-1} & -0.087 - 0.054\sqrt{-1} & 0.136 - 0.190\sqrt{-1} & -0.049 + 0.161\sqrt{-1} \end{bmatrix}^T. \end{aligned}$$

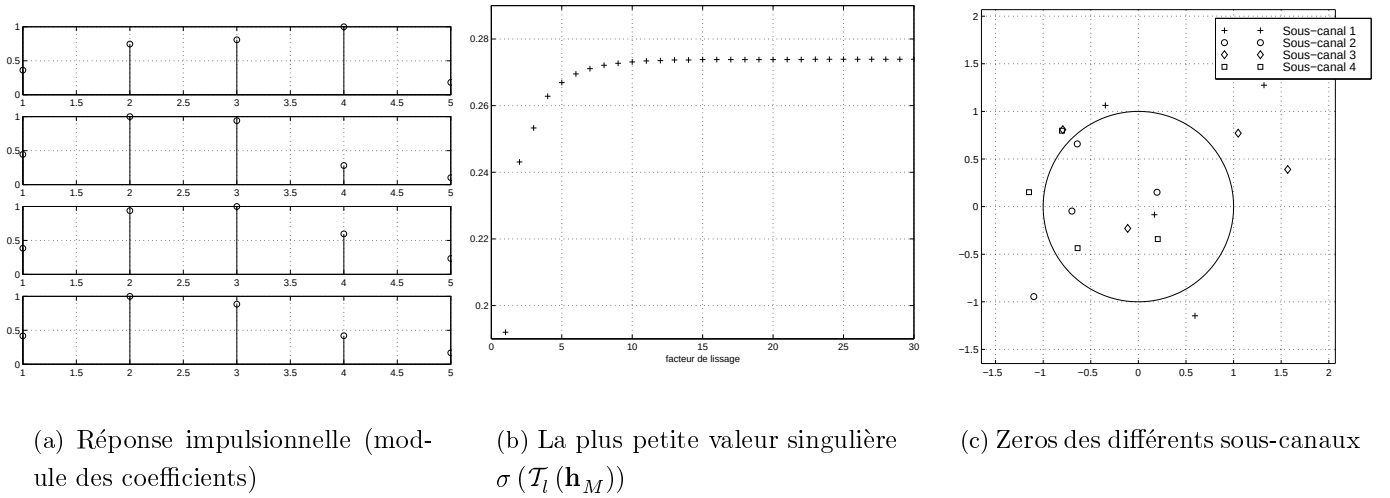


Figure 5.3: Le canal

Le canal montre une bonne diversité d'après la plus petite valeur singulière de sa matrice de filtrage (cf. 5.3.b) et la disposition des zeros de ses différents sous-canaux (cf. 5.3.c). Le canal est excité par une source de symboles QAM-4, i.i.d. et de variance unité. Le bruit d'observation est blanc et gaussien (AWGN). Le rapport signal sur bruit est défini par⁵ $SNR \triangleq \frac{\mathbb{E}(\|\mathbf{x}(n)\|^2)}{\mathbb{E}(\|\mathbf{b}(n)\|^2)} = \frac{\sigma_s^2 \|\mathbf{h}_M\|^2}{c \sigma_b^2}$.

L'erreur $\mathcal{E}$ est moyennée sur des réalisations Monte Carlo au nombre de 100.

L'algorithme proposé ne peut être comparé aux algorithmes SS, LP et OPD que dans le cas où l'ordre du canal est exactement connu. Dans ces conditions, il ressort de Fig. 5.4 que l'algorithme proposé

⁴Il ne correspond pas, par exemple, à une onde en cosinus sur-élevé transmises à travers un canal à trajets multiples

⁵En effet, $SNR = \frac{\sum_{c'=1}^c \mathbb{E}(\|x^{c'}(n)\|^2)}{\sum_{c'=1}^c \mathbb{E}(\|b^{c'}(n)\|^2)} = \frac{\sum_{c'=1}^c \|\mathbf{h}_M^{c'}\|^2 \mathbb{E}(\|s(n)\|^2)}{c \mathbb{E}(\|b^1(n)\|^2)} = \frac{\mathbb{E}(\|s(n)\|^2) \sum_{c'=1}^c \|\mathbf{h}_M^{c'}\|^2}{c \sigma_b^2}$

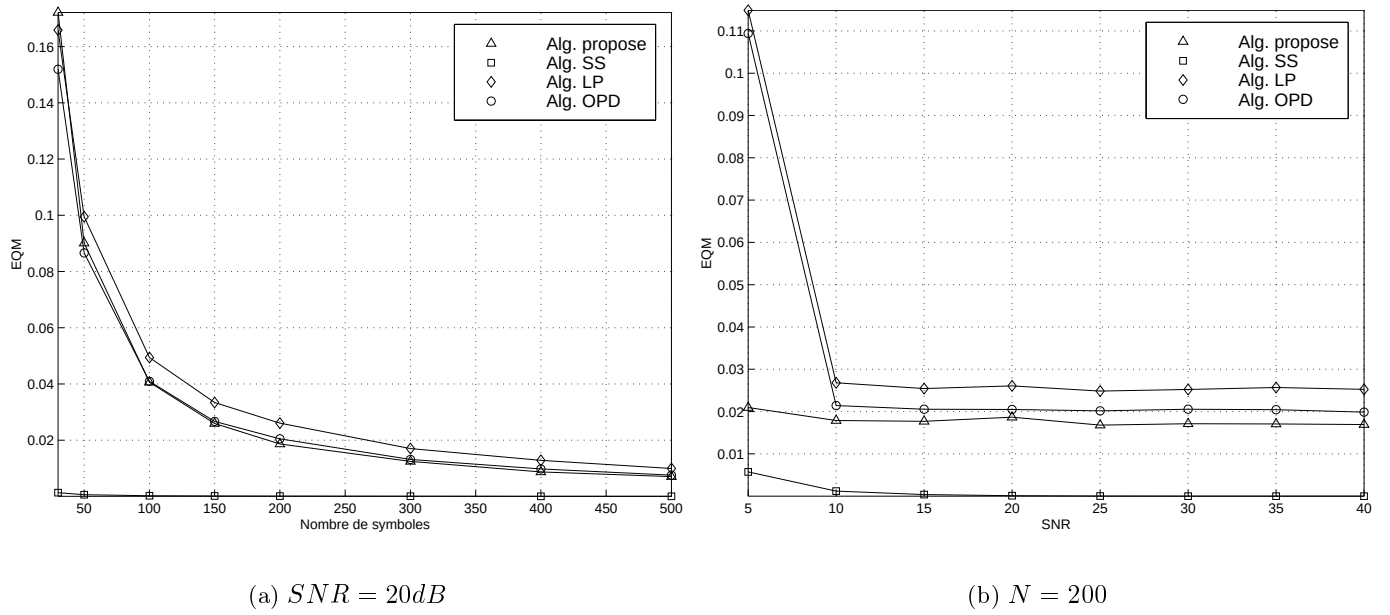


Figure 5.4: Comparaison de l'algorithme proposé et des algorithmes SS, LP et OPD avec détection de l'ordre exact, $w = 9$

dépasse en performances les algorithmes LP et OPD (pour d'autres canaux, il a été constaté un écart plus important [Gzh 00]). L'algorithme SS continue de montrer les meilleures performances. Il est intéressant de voir que les performances de l'algorithme proposé se montrent peu sensibles à une détérioration du rapport signal sur bruit, comparé aux algorithmes LP et OPD.

Lorsque l'on surestime l'ordre du canal (Fig. 5.5), l'algorithme proposé se montre capable de *bien* estimer la réponse du canal, une situation où les autres algorithmes aboutissent à des erreurs d'estimation très importantes, non reportées sur la figure. L'erreur d'estimation dans le cas d'une surestimation de l'ordre du canal peut être diminuée en augmentant le paramètre w (grâce à l'augmentation du facteur de lissage) i.e., le nombre des estimées *candidates*.

5.10 Comparaison avec les algorithmes existants

Dans ce qui suit, on justifie qualitativement les performances de l'algorithme proposé comparées à celles des algorithmes SS, LP et OPD. La robustesse à la surestimation est une qualité vérifiée uniquement par l'algorithme proposé. Elle trouve sa justification dans ce qui suit. La surestimation de l'ordre du canal se traduit par un mauvais conditionnement pour l'inversion des estimées des matrices de corrélation débruitées standard ($\mathbf{R}_l - \mathbf{R}_l^b$) et décalée ($\mathcal{R}_l - \mathcal{R}_l^b$). Leur inversion nécessite la connaissance exacte du *rang théorique* de la matrice, i.e., la connaissance exacte du canal. Puisque les algorithmes LP et OPD font intervenir la pseudo-inverse de $\mathbf{R}_l - \mathbf{R}_l^b$, leur échec est prévisible en surestimation de l'ordre du canal. L'algorithme proposé ne procède pas à une telle inversion et sa robustesse à la surestimation de l'ordre se maintient dans des conditions pratiques d'observation (limitée dans le temps et bruitée).

L'algorithme proposé présente un autre avantage qui lui est propre. Pour les autres algorithmes, le

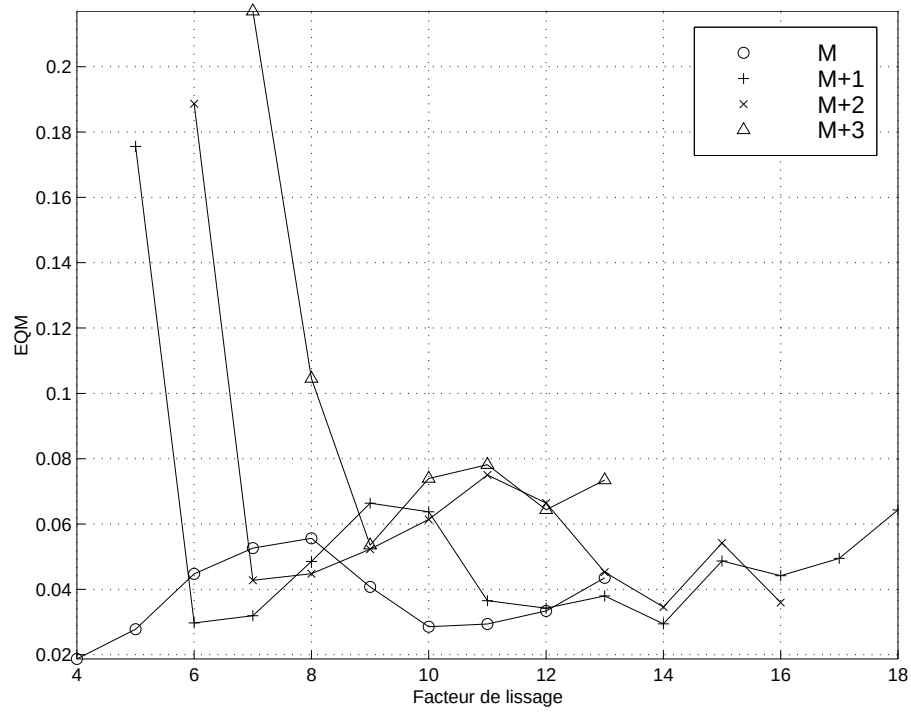


Figure 5.5: L'algorithme proposé en sur-estimation de l'ordre du canal. La légende montre l'ordre détecté. $N = 200$. $SNR = 20$.

seul moyen pour améliorer les performances d'estimation est d'affiner l'estimation de la matrice de corrélation en observant plus longtemps le canal. Ceci n'est pas toujours possible car en pratique les canaux sont variables dans le temps et l'algorithme doit achever l'identification pendant la durée de stationnarité du canal. L'algorithme proposé dispose d'un autre moyen pour améliorer les performances en augmentant le facteur de lissage i.e., la taille de la matrice de corrélation. Cette option ne gaspille pas de ressources physiques mais induit une augmentation du temps de traitement, moins préoccupante cependant⁶. L'augmentation du facteur de lissage est particulièrement utile lorsque l'algorithme tourne dans des mauvaises conditions, notamment en sur-estimation de l'ordre du canal (Fig. 5.5 et Fig. 5.6).

Lorsque l'ordre est exactement connu, l'algorithme proposé montre des performances intermédiaires entre l'approche sous-espace et l'approche prédiction linéaire (l'approche OPD étant très comparable à LP).

Contrairement à l'algorithme SS, l'algorithme proposé passe par l'estimation de la puissance du bruit du canal, ce qui est une source d'erreur supplémentaire. Ceci justifie que l'algorithme SS continue d'exhiber les meilleures performances. Rajouter à cela qu'il (SS) est le seul capable, en l'absence de bruit, d'estimer la réponse exacte du canal même si seulement un nombre limité d'observations est disponible.

De la même manière que l'algorithme LP, l'algorithme proposé passe par le calcul d'égaliseurs ZF⁷ avant de déduire la réponse impulsionnelle du canal. Cependant, il calcule toute une famille d'égaliseurs

⁶On note en particulier l'existence d'algorithmes rapides de recherche de vecteurs propres/singuliers.

⁷Ceux-là sont à délai minimum et maximum et donc les moins favorables en présence de bruit [Touz 98].

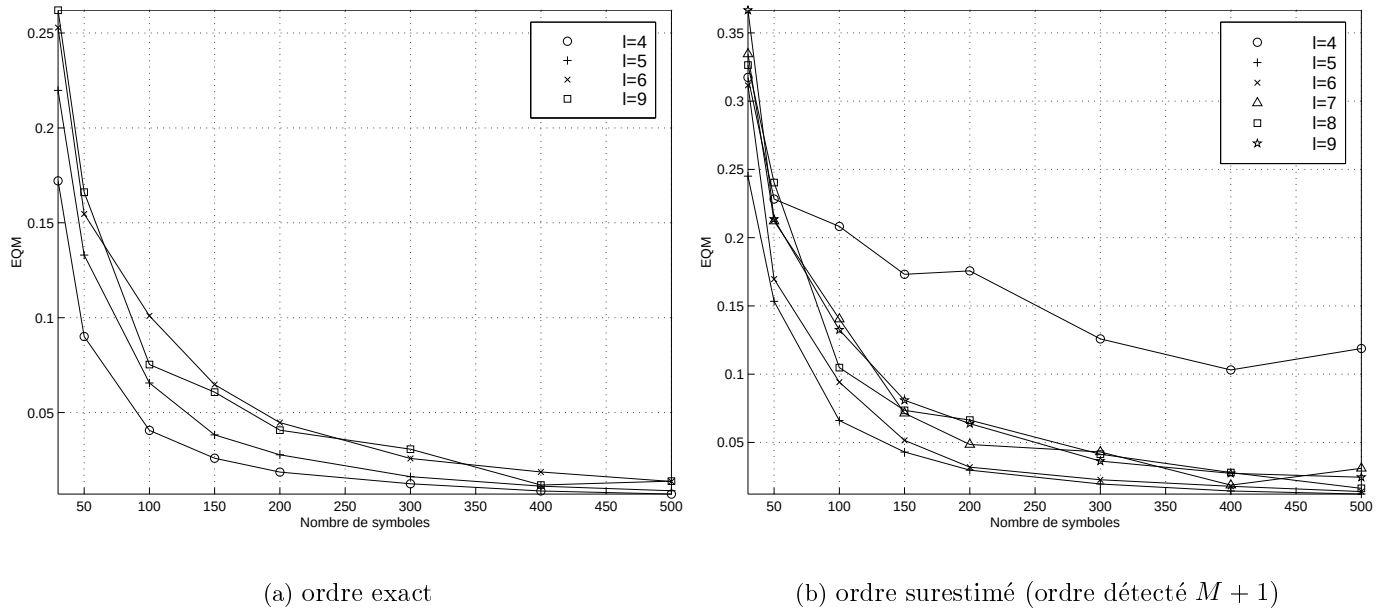


Figure 5.6: Effet du facteur de lissage. $SNR = 20$

(donc d'estimées *candidates*) et dispose ainsi d'une meilleure chance d'aboutir à une erreur d'estimation faible.

Notons que, tant que l'égalisation ZF est concernée, l'algorithme proposé offre plusieurs égaliseurs ZF, de délais maximaux et minimaux et d'ordres $\geq \lceil \frac{M}{c-1} \rceil - 1$. Au contraire, l'algorithme LP n'en donne qu'un seul avec un délai minimal et d'ordre M .

Les meilleures performances de l'algorithme SS sont à apprécier par rapport à une ambiguïté plus importante (de phase et d'amplitude). L'algorithme LP achève l'identification avec la plus faible ambiguïté possible (de phase). L'algorithme proposé, quant à lui, offre la possibilité de choisir le degré d'ambiguïté, ce qui lui rend d'une plus large utilité.

5.11 Conclusion

On a proposé un nouvel algorithme d'identification aveugle au second ordre des canaux SIMO. Il corrige un défaut essentiel de la plupart des algorithmes développés à ce jour : la nécessité de connaître l'ordre exact du canal. L'algorithme proposé est robuste à la surestimation de l'ordre du canal; et ce au sens pratique du terme : la réponse du canal peut être arbitrairement approchée si on connaît une valeur supérieure à la valeur exacte de l'ordre et si le canal est observé pendant une durée finie et en présence de bruit. Ceci est qualifié de *vraie robustesse* en opposition aux algorithmes LP et OPD dont la robustesse à la surestimation de l'ordre est purement théorique, car fondée sur les statistiques exactes.

En ordre exact, l'algorithme proposé montre des performances meilleures que celles des algorithmes LP et OPD car il aboutit au calcul de plusieurs estimées de la réponse du canal. Un choix qui permet une meilleure estimation de cette dernière et constitue, par le moyen de sa taille ajustable, un moyen d'améliorer davantage l'erreur d'estimation autre que l'augmentation de la durée d'observation.

L'ambiguïté, avec laquelle la réponse du canal est estimée, peut être fixée en fonction de l'application.

Le critère proposé dans §5.5 a été choisi intuitivement. C'est là une possibilité d'amélioration éventuelle de l'algorithme. En particulier, on se demande si la sélection peut se faire plutôt dans la procédure d'identification en choisissant un égaliseur sur l'ensemble des égaliseurs calculés et continuer l'identification avec l'égaliseur retenu. Ceci permettra un gain considérable en temps de calcul, en particulier si le facteur de lissage est important, ce qui est le cas lorsqu'on se place dans le cas d'une surestimation de l'ordre du canal, lequel cas constitue l'intérêt de cet algorithme.

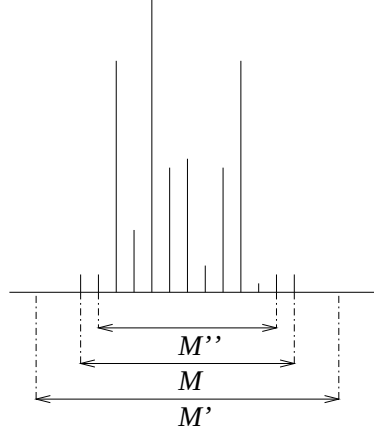


Figure 5.7: Surestimation de l'ordre

Finalement, on revient sur la définition donnée à la surestimation de l'ordre. En se référant à Fig. 5.7, une situation de surestimation correspond à la détection d'un ordre M' supérieur ou égale à l'ordre M de la réponse *totale* (au sens strict) i.e., y compris éventuellement les termes faibles de la tête/queue. L'estimation d'un ordre M'' qui correspond à la prise en compte de seulement une partie des termes faibles de la tête/queue n'est pas, relativement à notre définition, une situation de surestimation. Avec un ordre détecté $M' \geq M$, le canal présente une matrice de corrélation mal conditionnée pour l'inversion (il est alors dit *mal diversifié*) à cause de la présence de termes faibles de la tête/queue (cf. §7.1.1 et [LIAV 99]). L'algorithme proposé, qui tente à l'identifier en opérant une SVD de l'estimée de cette dernière, échoue à bien approximer (une version *zero-padded* de) la réponse d'ordre M (cf. §6).

On a choisi cette définition de l'ordre (à surestimer) de manière à ce que robustesse à la surestimation de l'ordre et robustesse à la mauvaise diversité du canal (cf. §6) soient deux problèmes indépendants qui puissent être adressés séparément. Plus précisément, ainsi définie, la surestimation de l'ordre n'affecte pas le conditionnement/diversité du problème. En particulier, cette définition évite qu'une surestimation de l'ordre (au sens "prendre en compte des termes faibles à la tête ou à la queue de la réponse") En effet, une situation du type M'' peut être interprétée comme un cas de surestimation d'un *ordre pratique* ou comme un cas de mauvaise diversité. Pour nous, il s'agira de ce dernier cas et l'algorithme proposé (d'après nous robuste à la surestimation de l'ordre) montrera alors de mauvaises performances à cause d'une mauvaise diversité de la réponse recherchée. Cette définition de l'ordre (et de la surestimation) vise aussi à éviter l'ambiguïté de la définition, souvent *subjective*, d'un ordre *pratique* ou d'une manière équivalente de présence de termes *faibles*.

Le problème de la (non) robustesse à la mauvaise diversité du canal reste non résolu par l'algorithme

proposé ici et ce pour les mêmes raisons discutées dans §6.1, à savoir que la recherche de vecteurs singuliers est sensible au conditionnement de la matrice de corrélation (débruitée). Le reste du rapport se consacre désormais à ce problème.

Chapitre 6

Diversité des canaux SIMO

L'information sur la réponse impulsionnelle d'un canal SIMO est contenue dans les statistiques d'ordre deux de sa sortie. Le vecteur $\mathbf{h}_M$ apparaît dans l'expression de la matrice de corrélation. Pour l'en extraire, les algorithmes d'identification doivent inverser explicitement ou implicitement la matrice de corrélation. Les performances de l'algorithme d'identification aveugle, disposant d'un nombre fini d'observations, s'en trouvent sensibles au conditionnement de la matrice de corrélation.

L'identifiabilité du canal est souvent citée [FIJA 97] comme fonction de sa diversité, ou encore sa disparité, sans qu'une mesure stricte n'en soit donnée. Ce constat incite à proposer une définition stricte de la diversité d'un canal se basant sur la plus petite valeur propre non nulle de la matrice de corrélation.

À l'exception des algorithmes LP et OPD, le facteur de lissage n'est pas fixe. En particulier, pour l'algorithme §5, l'erreur d'identification diminue lorsque ce paramètre augmente. Ainsi, le comportement asymptotique (pour tailles grandissantes des matrices de corrélation) de la plus petite valeur propre non nulle mérite d'être étudié. Ces matrices de corrélation ont des structures blocs de Toeplitz. Ceci rappelle les résultats établis par Szegö sur le comportement asymptotique des valeurs propres des matrices de Toeplitz Hermitiennes, et invite à généraliser ces résultats aux matrices blocs de Toeplitz.

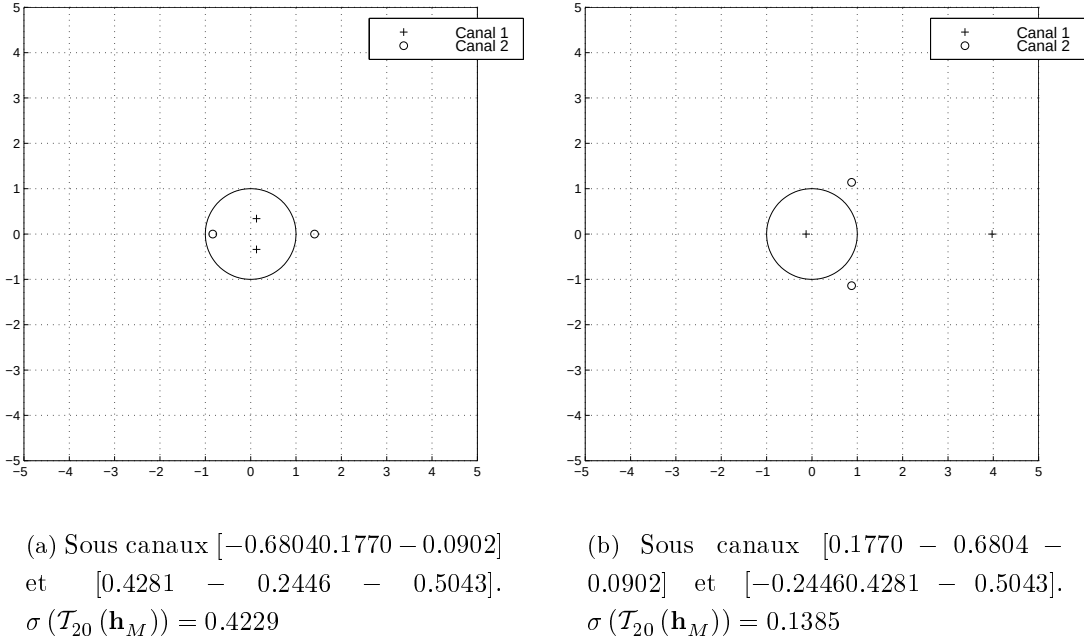
Dans ce chapitre, on arrête une définition de la diversité (§6.1). On généralise le théorème de Szegö (§6.2) au cas des matrices de Toeplitz (§6.3). On en déduit une borne supérieure sur la diversité des canaux SIMO. Il s'en suit une discussion sur les implications de ce résultat sur l'identifiabilité des canaux dans des conditions pratiques d'observations. On s'intéresse en particulier au cas des canaux à bande limitée.

6.1 Mesure de la diversité

Les algorithmes du second ordre d'identification aveugle sont capables de déduire la réponse impulsionnelle d'un canal SIMO à partir des statistiques du second ordre de sa sortie non bruitée si le canal remplit la condition de diversité (dite aussi disparité ou minimum de phase). Par diversité on entend que les fonctions de transfert des réponses impulsionnelles des différents sous-canaux ne possèdent pas de zéros en commun. Dans ce cas, le canal est dit à minimum de phase, par extension aux canaux

SIMO de la notion de minimum de phase déjà définie pour les canaux synchrones. En pratique, les statistiques ne sont pas exactement connues. Elles sont estimées sur la base d'un nombre fini d'observations. La qualité de l'estimation dépend alors de la *qualité de la diversité* du canal ou aussi le *degré de diversité* duquel dépend *l'identifiabilité du canal observé pendant une durée limitée*.

Une interprétation *géométrique* de la diversité a été donnée dans [FIJA 97, TOUZ 98]. Plus les zéros des sous-canaux sont distants, meilleure est la diversité. Dans le cas limite où les sous-canaux possèdent des zéros en communs, le canal est non identifiable même en connaissance des statistiques exactes de sa sortie. Des contre-exemples (Fig. 6.1) attestent que cette interprétation n'est pas toujours valable.



Le canal de (a) présente des zéros plus rapprochés que pour le canal (b) mais, à l'examen de la plus petite valeur singulière de la matrice de filtrage (avec un facteur de lissage choisi arbitrairement), est mieux diversifié.

Figure 6.1: Contre-exemple

Une interprétation *algébrique* de la diversité est donnée par *Liavas et al* dans [LIAV 99, LIAV 99, LIAV 00]¹. La diversité du canal dépend du conditionnement de sa matrice de filtrage (sur-déterminée). Plus cette dernière est loin de la singularité, meilleure est la diversité. *Liavas et al* ont proposé la plus petite valeur singulière de la matrice de filtrage $\mathcal{T}_l(\mathbf{h}_M)$ (l étant le facteur de lissage employé par l'algorithme) comme une mesure de la diversité. Cette interprétation a été discutée dans [Veen 95] également.

Cette interprétation est plus pertinente que l'interprétation géométrique dans [FIJA 97, TOUZ 98] et permet surtout de proposer une mesure de la diversité qui ne soit pas seulement visuelle. Le carré de la plus petite valeur singulière de la matrice de filtrage (de rang colonne plein) est égale à la plus petite valeur propre non nulle de la matrice de corrélation (non bruitée) du SIMO à identifier. Or, les

¹L'étude *Liavas et al* a été faite dans le cas de deux sous-canaux (H7.1) mais intuitivement la mesure de diversité qu'ils proposent s'étend facilement au cas d'un nombre de phases ≥ 2 quelconque.

algorithmes d'identification inversent explicitement² ou implicitement³ (par une recherche de vecteurs propres ou singuliers) la matrice de corrélation. La stabilité de la solution du problème (la réponse du canal) par rapport à une perturbation de la matrice de corrélation (due à une erreur d'estimation inévitable) est très dépendante du conditionnement de cette matrice. D'ailleurs, sa pseudo-inverse ressort dans toutes les expressions des différentielles de ces algorithmes (cf. § B.2, B.3 et B.4).

Cette interprétation a besoin d'être affinée pour ne plus dépendre du facteur de lissage. En remarquant que la plus petite valeur singulière de la matrice de filtrage d'un SIMO à minimum de phase est convergente⁴ lorsque le facteur de lissage tend vers l'infini⁵, une *mesure de la diversité* peut être fixée comme étant⁶

$$\frac{1}{\|\mathbf{h}_M\|} \lim_{l \rightarrow \infty} \sigma(\mathcal{T}_l(\mathbf{h}_M)) \quad (6.1)$$

Une conséquence de cette interprétation de la diversité est que la diversité augmente lorsqu'on rajoute

un sous-canal supplémentaire i.e.,⁷ $\begin{bmatrix} \mathbf{h}_M^1 \\ \vdots \\ \mathbf{h}_M^c \\ \mathbf{h}_M^{c+1} \end{bmatrix}$ est mieux diversifié⁸ que $\begin{bmatrix} \mathbf{h}_M^1 \\ \vdots \\ \mathbf{h}_M^c \end{bmatrix} \forall \mathbf{h}_M^{c+1} \in \mathbf{C}^l$ et l ; donc

indépendamment de la réponse supplémentaire. Cependant, cette solution n'est pas intéressante car le coût que ça implique (augmentation de la taille du réseau de capteurs) ne justifie souvent pas le gain en diversité. Pour cette raison, la valeur $c = 2$ est souvent retenue en pratique.

6.2 Notations et Rappel

Un résultat célèbre de Szegö [GREN 84] stipule que le comportement asymptotique des valeurs propres d'une matrice de Toeplitz Hermitienne est décrit par la transformée de Fourier de ses termes. Le théorème de Szegö énonce que les valeurs propres d'une suite matricielle de Toeplitz Hermitiennes sont distribuées comme les échantillons de la transformée de Fourier de leurs termes. La plus petite (resp. plus grande) est décroissante (resp. croissante) et converge vers le minima (resp. maxima) de cette transformée de Fourier. Dans ce qui suit, on rappelle ce résultat :

Soit $\{t_k\}_{k=-\infty, -1, 0, 1, \dots}$ une suite infinie de nombres complexes absolument sommable (i.e., $\sum_k |t_k| < \infty$) de telle manière que la transformée de Fourier associée $t(w) \triangleq \sum_k t_k e^{-ikw}$ soit bien définie. On défini

²tel LP.

³tels SS, OPD et l'algorithme proposé dans §5.

⁴Cette convergence a été systématiquement vérifiée en pratique mais une preuve stricte n'a pu être établie.

⁵On verra également que cette limite approxime bien la plus petite valeur singulière pour des valeurs pratiques du facteur de lissage

⁶On normalise la réponse du canal pour ne pas avoir, par exemple, le résultat absurde qu'est $2\mathbf{h}_M$ mieux diversifié que $\mathbf{h}_M$. Ce qui revient à dire que le rapport des valeurs singulières minimale et maximale est le facteur déterminant pour le conditionnement par rapport à l'inversion de la matrice.

⁷Format *TS*

⁸Ceci peut être démontré en utilisant [HORN91, Corollaire 3.1.3].

la suite matricielle $\{\mathbf{T}_n(t)\}_{n \geq 1}$ où $\mathbf{T}_n(t)$ est une matrice de Toeplitz $n \times n$ donnée par

$$\mathbf{T}_n(t) \triangleq \begin{bmatrix} t_0 & t_{-1} & \cdots & t_{-(n-1)} \\ t_1 & \ddots & \ddots & t_{-(n-2)} \\ \vdots & \ddots & & \vdots \\ t_{n-1} & t_{n-2} & \cdots & t_0 \end{bmatrix}. \quad (6.2)$$

Si $\mathbf{A}$ est une matrice $n \times n$, on continue de noter $\|\mathbf{A}_n\| \triangleq \sqrt{\max_{\|\mathbf{x}\|=1} \mathbf{x}^H \mathbf{A}_n^H \mathbf{A}_n \mathbf{x}}$ la norme euclidienne de $\mathbf{A}$ qu'on dira aussi *norme forte*; et $|\mathbf{A}_n| \triangleq \sqrt{\frac{1}{n} \sum_{i=1}^n \sum_{j=1}^n |a_{i,j}|^2}$ la norme de Frobenius de $\mathbf{A}$ qu'on dira aussi *norme faible* [GRAY 72].

On introduit la notion d'équivalence asymptotique entre suites matricielles carrées :

Définition 6.1 Équivalence Asymptotique [GRAY 72]

Deux suites matricielles $\{\mathbf{A}_n\}$ et $\{\mathbf{B}_n\}$, $n = 1, 2, \dots$ sont dites *asymptotiquement équivalentes* et notées $\{\mathbf{A}_n\} \sim \{\mathbf{B}_n\}$ si

$$\exists M < \infty \text{ tel que } \forall n, \|\mathbf{A}_n\| \leq M \text{ and } \|\mathbf{B}_n\| \leq M \quad (6.3)$$

$$\lim_{n \rightarrow \infty} |\mathbf{A}_n - \mathbf{B}_n| = 0. \quad (6.4)$$

Deux suites matricielles asymptotiquement équivalentes vérifient la propriété suivante quant à leurs valeurs propres.

Theorem 6.1 [GRAY 72]

Si $\{\mathbf{A}_n\} \sim \{\mathbf{B}_n\}$ et si $\lim_{n \rightarrow \infty} \frac{1}{n} \sum_{k=1}^n \lambda_k^s(\mathbf{A}_n)$ existe et est finie pour un entier s donné,

$$\lim_{n \rightarrow \infty} \frac{1}{n} \sum_{k=1}^n \lambda_k^s(\mathbf{A}_n) = \lim_{n \rightarrow \infty} \frac{1}{n} \sum_{k=1}^n \lambda_k^s(\mathbf{B}_n)$$

À toute suite de matrices de Toeplitz, on peut associer une suite de matrices circulantes asymptotiquement équivalentes :

Theorem 6.2 [GRAY 72]

Pour toute suite $\{t_k\}_{k=\dots, -1, 0, 1, \dots}$ absolument sommable, il existe une suite de matrices circulantes $\{\mathbf{C}_n(t)\}$ asymptotiquement équivalentes à la suite de matrices $\{\mathbf{T}_n(t)\}$ et est donnée par $\mathbf{C}_n(t) = \mathbf{U}_n^H \mathbf{D}_n(t) \mathbf{U}_n$ où

- $\mathbf{D}_n(t)$ est une matrice diagonale dont le $k^{\text{ème}}$ terme est donné par $(\mathbf{D}_n(t))_{k,k} = t\left(\frac{2\pi(k-1)}{n}\right)$.
- $\mathbf{U}_n$ est la matrice DFT unitaire $(\mathbf{U}_n)_{k,l} = \frac{1}{\sqrt{n}} e^{-i2\pi \frac{(k-1)(l-1)}{n}}$.

Les expressions de $\mathbf{C}_n(t)$ et de $\mathbf{U}_n$ (cf. [GRAY 98]) ne sont cependant pas nécessaires par la suite. Dans notre démonstration de Th. 6.7, on exploite le fait que les vecteurs propres d'une matrice circulante ne dépendent pas de ses termes, i.e., de la suite $\{t_k\}$, et que ses valeurs propres sont les échantillons équidistants de la transformée de Fourier $t(w)$.

On rappelle finalement le théorème de Szegö :

Theorem 6.3 Théorème de Szegö [GREN 84]

Pour toute suite $\{t_k\}_{k=\dots,-1,0,1,\dots}$ absolument sommable, si $\mathbf{T}_n(t)$ est Hermitienne, alors pour toute fonction F continue sur $[\min_w t(w), \max_w t(w)]$, on a

$$\lim_{n \rightarrow \infty} \frac{1}{n} \sum_{k=1}^n F(\lambda_k(\mathbf{T}_n(t))) = \frac{1}{2\pi} \int_{w=-\pi}^{\pi} F(t(w)) dw.$$

Le théorème précédent décrit le comportement de groupe des valeurs asymptotiques. Il s'en déduit le résultat suivant sur les valeurs extrêmes.

Theorem 6.4 [GREN 84, GRAY 98]

Pour toute suite $\{t_k\}_{k=\dots,-1,0,1,\dots}$ absolument sommable, si $\mathbf{T}_n(t)$ est Hermitienne, alors pour tout l , la $l^{\text{ème}}$ plus petite (resp. plus grande) valeur propre de $\mathbf{T}_n(t)$ $\lambda_k(\mathbf{T}_n(t))$ (resp. $\lambda_{n-k}(\mathbf{T}_n(t))$) est décroissante (resp. croissante) en n et converge vers $\min_w t(w)$ (resp. $\max_w t(w)$).

6.3 Matrices blocs de Toeplitz

L'application du théorème de Szegö est directe pour les processus scalaires stationnaires (au second ordre), pour lesquels les matrices de corrélation sont de Toeplitz. Plusieurs extensions du théorème de Szegö ont été faites [TILL 99]. La plus importante concerne les matrices blocs de Toeplitz avec blocs non Toeplitz où le nombre de blocs tend vers l'infini [TILL 98]. Cependant, la démonstration qui y est donnée fait appel à des outils mathématiques sophistiqués. On se propose de donner une nouvelle démonstration plus simple que celle dans [TILL 98] de l'extension du théorème de Szegö aux matrices blocs de Toeplitz. La démonstration exploite l'équivalence asymptotique entre séquences matricielles et plus particulièrement le résultat établi par Gray dans [GRAY 72] sur l'équivalence asymptotique entre séquences de matrices de Toeplitz et séquences de matrices circulantes (Th. 6.2).

6.3.1 Définitions

On considère la matrice blocs de Toeplitz avec des blocs non Toeplitz (BTNTB⁹)

$$\mathcal{T}'_n(\{t^{u,v}\}) = \begin{bmatrix} \mathbf{T}_0 & \mathbf{T}_{-1} & \cdots & \mathbf{T}_{-(n-1)} \\ \mathbf{T}_1 & \ddots & \ddots & \mathbf{T}_{-(n-2)} \\ \vdots & \ddots & & \vdots \\ \mathbf{T}_{n-1} & \mathbf{T}_{n-2} & \cdots & \mathbf{T}_0 \end{bmatrix}, \quad (6.5)$$

où $(\mathbf{T}_k)_{k=-(n-1),\dots,n-1}$ sont des matrices $c \times c$ qui ne sont pas forcément de Toeplitz. Leurs termes sont données par $t_k^{u,v} \hat{=} (\mathbf{T}_k)_{u,v}$, $u, v = 1, \dots, c$. On lui associe la matrice non blocs de Toeplitz avec

⁹Block Toeplitz with Non Toeplitz Blocks

des blocs de Toeplitz (NBTTB¹⁰)

$$\mathcal{T}_n(\{t^{u,v}\}) \hat{=} \begin{bmatrix} \mathbf{T}_n(t^{1,1}) & \mathbf{T}_n(t^{1,2}) & \cdots & \mathbf{T}_n(t^{1,c}) \\ \mathbf{T}_n(t^{2,1}) & \mathbf{T}_n(t^{2,2}) & \cdots & \mathbf{T}_n(t^{2,c}) \\ \vdots & & & \vdots \\ \mathbf{T}_n(t^{c,1}) & \mathbf{T}_n(t^{c,2}) & \cdots & \mathbf{T}_n(t^{c,c}) \end{bmatrix} \quad (6.6)$$

où , conformément à la notation (6.2), $\mathbf{T}_n(t^{u,v})$, $u, v = 1, \dots, c$ sont des matrices de Toeplitz $n \times n$ données par

$$\begin{bmatrix} t_0^{u,v} & t_{-1}^{u,v} & \cdots & t_{-(n-1)}^{u,v} \\ t_1^{u,v} & \ddots & \ddots & t_{-(n-2)}^{u,v} \\ \vdots & \ddots & & \vdots \\ t_{n-1}^{u,v} & t_{n-2}^{u,v} & \cdots & t_0^{u,v} \end{bmatrix}.$$

On suppose que chacune des suites $t_k^{u,v}$, $u, v = 1, \dots, c$ est absolument sommable. Ainsi, on associe à chacune des suites la transformée de Fourier $t^{u,v}(w) \hat{=} \sum_k t_k^{u,v} e^{-ikw}$. Puisque les matrices BTNTB et NBTTB, $\mathcal{T}'_n(\{t^{u,v}\})$ et $\mathcal{T}_n(\{t^{u,v}\})$, ainsi définies sont les mêmes à une permutation de lignes et de colonnes près (cf. (6.7)), elles sont équivalentes du point de vue des valeurs propres. Cependant, on préfère la formulation NBTTB car elle permet de manipuler directement des blocs de Toeplitz pour lesquels les résultats déjà établis (cf. §6.2) peuvent être utilisés.

L'une et l'autre des formulations apparaissent lorsqu'il s'agit de la matrice de corrélation d'un processus spatio-temporel stationnaire, selon que les données sont arrangées temps par temps (BTNTB) ou phase par phase¹¹ (NTBBT).

On peut vérifier que

$$\mathcal{T}'_n(\{t^{u,v}\}) = \mathbf{K}_{c,n} \mathcal{T}_n(\{t^{u,v}\}) \mathbf{K}_{n,c} \quad (6.7)$$

Notons finalement que $\mathcal{T}_n(\{t^{u,v}\})$ est Hermitienne ssi $\mathbf{T}_n(t^{v,u}) = \mathbf{T}_n^H(t^{u,v})$, $\forall u, v = 1, \dots, c$ ou d'une manière équivalente $t^{v,u}(w) = (t^{u,v}(w))^*$, $\forall u, v = 1, \dots, c$.

Dans la suite de §6.3, on donne une démonstration de l'extension du théorème de Szegö aux matrices blocs de Toeplitz. Elle est plus simple que celle dans [TILL 98]. Cependant dans [GREN 84, TILL 98], des hypothèses moins restrictives sont données sur les séquences $\{t_k^{u,v}\}_{k=\dots, -1, 0, 1, \dots}$ où elles ont été supposées carré sommables uniquement. Dans ce cas, une définition plus générale de la transformée de Fourier est applicable et les démonstrations du théorème de Szegö et de ses extensions sont plus difficiles à établir.

6.3.2 Distribution asymptotique des valeurs propres

En exploitant le fait (Th. 6.2) que les blocs Toeplitz de la matrice $\mathcal{T}_n(\{t^{u,v}\})$ sont asymptotiquement équivalents à des matrices circulantes qui elles sont diagonalisables via le même changement de base, on étend Th. 6.2 au cas NBTTB comme suit :

¹⁰Non block Toeplitz with Toeplitz blocks

¹¹ou bien aussi capteur par capteur

Theorem 6.5 Si les suites $\{t_k^{u,v}\}_{k=\dots,-1,0,1,\dots}$ sont absolument sommable, il existe une suite matricielle $\{\mathcal{C}_n(\{t^{u,v}\})\}$ asymptotiquement équivalente à $\{\mathcal{T}_n(\{t^{u,v}\})\}$ et donnée par $\mathcal{C}_n(\{t^{u,v}\}) = \mathcal{U}_n^H \mathcal{D}_n(\{t^{u,v}\}) \mathcal{U}_n$ avec $\mathcal{U}_n$ est une matrice unitaire $nc \times nc$ indépendante de $\mathcal{T}_n(\{t^{u,v}\})$ et où $\mathcal{D}_n(\{t^{u,v}\})$ est la matrice

$$\mathcal{D}_n(\{t^{u,v}\}) \hat{=} \begin{bmatrix} \mathbf{D}_n(t^{1,1}) & \mathbf{D}_n(t^{1,2}) & \dots & \mathbf{D}_n(t^{1,c}) \\ \mathbf{D}_n(t^{2,1}) & \mathbf{D}_n(t^{2,2}) & \dots & \mathbf{D}_n(t^{2,c}) \\ \vdots & & & \vdots \\ \mathbf{D}_n(t^{c,1}) & \mathbf{D}_n(t^{c,2}) & \dots & \mathbf{D}_n(t^{c,c}) \end{bmatrix} \quad (6.8)$$

avec $\mathbf{D}_n(t^{u,v})$ est une matrice diagonale définie comme dans Th. 6.2.

Démonstration:

D'après Th. 6.2, $\forall u$ et v , il existe une suite de matrices circulantes $\mathbf{C}_n(t^{u,v})$ asymptotiquement équivalentes à $\mathbf{T}_n(t^{u,v})$ et donnée par $\mathbf{C}_n(t^{u,v}) = \mathbf{U}_n^H \mathbf{D}_n(t^{u,v}) \mathbf{U}_n$ où $\mathbf{U}_n$ est indépendant de $t^{u,v}(w)$ et la matrice diagonale $\mathbf{D}_n(t^{u,v})$ est définie par ses termes $(\mathbf{D}_n(t^{u,v}))_{k,k} = t^{u,v} \left(\frac{2\pi(k-1)}{n} \right)$.

Soit

$$\mathcal{C}_n(\{t^{u,v}\}) \hat{=} \begin{bmatrix} \mathbf{C}_n(t^{1,1}) & \mathbf{C}_n(t^{1,2}) & \dots & \mathbf{C}_n(t^{1,c}) \\ \mathbf{C}_n(t^{2,1}) & \mathbf{C}_n(t^{2,2}) & \dots & \mathbf{C}_n(t^{2,c}) \\ \vdots & & & \vdots \\ \mathbf{C}_n(t^{c,1}) & \mathbf{C}_n(t^{c,2}) & \dots & \mathbf{C}_n(t^{c,c}) \end{bmatrix}.$$

On peut vérifier que $\mathcal{C}_n(\{t^{u,v}\}) = \mathcal{U}_n^H \mathcal{D}_n(t^{u,v}) \mathcal{U}_n$ avec $\mathcal{U}_n \hat{=} (\mathbf{I}_c \otimes \mathbf{U}_n)$.

Reste à démontrer que les suites $\{\mathcal{T}_n(\{t^{u,v}\})\}$ et $\{\mathcal{C}_n(\{t^{u,v}\})\}$ sont asymptotiquement équivalentes. Vérifions alors les deux conditions de Déf. 6.1.

Considérons le vecteur normé $\mathbf{x} \in \mathbf{C}^{cn}$ tel que $\|\mathcal{T}_n(\{t^{u,v}\})\mathbf{x}\| = \|\mathcal{T}_n(\{t^{u,v}\})\|$. On décompose $\mathbf{x} = [\mathbf{x}^{1^T}, \dots, \mathbf{x}^{c^T}]^T$, où $\mathbf{x}^k \in \mathbf{C}^n$, $k = 1, \dots, c$, et donc $\|\mathbf{x}^k\| \leq 1$.

$$\begin{aligned} \|\mathcal{T}_n(t^{u,v})\|^2 &= \left\| \begin{bmatrix} \sum_{v=1}^c \mathbf{T}_n(t^{1,v}) \mathbf{x}^v \\ \vdots \\ \sum_{v=1}^c \mathbf{T}_n(t^{c,v}) \mathbf{x}^v \end{bmatrix} \right\|^2 \\ &= \sum_{u=1}^c \left\| \sum_{v=1}^c \mathbf{T}_n(t^{u,v}) \mathbf{x}^v \right\|^2 \\ &\leq \sum_{u=1}^c \left(\sum_{v=1}^c \|\mathbf{T}_n(t^{u,v}) \mathbf{x}^v\| \right)^2 \\ &\leq \sum_{u=1}^c \left(\sum_{v=1}^c \|\mathbf{T}_n(t^{u,v})\| \right)^2. \end{aligned}$$

De la même manière, on montre que $\|\mathcal{C}_n(\{t^{u,v}\})\|^2 \leq \sum_{u=1}^c (\sum_{v=1}^c \|\mathbf{C}_n(t^{u,v})\|)^2$.

Puisque $\forall u$ et v , les suites $\mathbf{T}_n(t^{u,v})$ et $\mathbf{C}_n(t^{u,v})$ sont asymptotiquement équivalent, alors les deux suites sont bornées au sens de (6.3), et donc il existe $M^{u,v}$ tel que $\|\mathcal{T}_n(\{t^{u,v}\})\|^2 \leq \sum_{u,v=1}^c (M^{u,v})^2$ et $\|\mathcal{C}_n(\{t^{u,v}\})\|^2 \leq \sum_{u,v=1}^c (M^{u,v})^2$. Ainsi, les suites $\mathcal{T}_n(\{t^{u,v}\})$ et $\mathcal{C}_n(\{t^{u,v}\})$ sont absolument bornées au sens de (6.3).

Finalement, il est facile de voir que $\lim_{n \rightarrow \infty} |\mathcal{T}_n(\{t^{u,v}\}) - \mathcal{C}_n(\{t^{u,v}\})| = 0$ car $|\mathcal{T}_n(\{t^{u,v}\}) - \mathcal{C}_n(\{t^{u,v}\})|^2 = \frac{1}{c} \sum_{u,v=1}^c |\mathbf{T}_n(t^{u,v}) - \mathbf{C}_n(t^{u,v})|^2$. ■

Notons que $\mathcal{C}_n(\{t^{u,v}\})$ n'est pas une matrice circulante. $\mathcal{D}_n(\{t^{u,v}\})$ n'est pas diagonale non plus.

Il s'en déduit le résultat suivant concernant les moments asymptotiques des valeurs propres des matrices blocs de Toeplitz :

Theorem 6.6

Pour tout entier $s \geq 1$,

$$\lim_{n \rightarrow \infty} \frac{1}{n} \sum_{k=1}^{cn} \lambda_k^s(\mathcal{T}_n(\{t^{u,v}\})) = \frac{1}{2\pi} \int_{w=-\pi}^{\pi} \sum_{1 \leq k_1, \dots, k_s \leq c} t^{k_1, k_2}(w) t^{k_2, k_3}(w) \dots t^{k_s, k_1}(w) dw \quad (6.9)$$

Démonstration:

Puisque les suites $\mathcal{T}_n(\{t^{u,v}\})$ et $\mathcal{C}_n(\{t^{u,v}\})$ sont asymptotiquement équivalentes, on a d'après Th. 6.1,

$$\lim_{n \rightarrow \infty} \frac{1}{cn} \sum_{k=1}^{cn} \lambda_k^s(\mathcal{T}_n(\{t^{u,v}\})) = \lim_{n \rightarrow \infty} \frac{1}{cn} \sum_{k=1}^{cn} \lambda_k^s(\mathcal{C}_n(\{t^{u,v}\})).$$

Étant donné que $\lambda_k^s(\mathcal{C}_n(\{t^{u,v}\}))$ sont toutes les valeurs propres de $\mathcal{C}_n^s(\{t^{u,v}\})$, cette dernière sommation n'est autre que la trace de $\mathcal{C}_n^s(\{t^{u,v}\})$. $\mathcal{C}_n^s(\{t^{u,v}\})$ étant similaire à $\mathcal{D}_n^s(\{t^{u,v}\})$, ceci est également la trace de $\mathcal{D}_n^s(\{t^{u,v}\})$.

$\mathcal{D}_n^s(\{t^{u,v}\})$ a la même forme que (6.8) i.e., formée de c^2 blocs diagonaux $\Delta_n^{u,v}$, $u, v = 1, \dots, c$ où la matrice $\Delta_n^{u,v}$ est définie par ses termes

$$\begin{aligned} (\Delta_n^{u,u})_{k,k} &= \sum_{1 \leq k_1, \dots, k_{s-1} \leq c} \left(\mathbf{D}_n(t^{u, k_1}) \right)_{k,k} \left(\mathbf{D}_n(t^{k_1, k_2}) \right)_{k,k} \dots \left(\mathbf{D}_n(t^{k_{s-1}, u}) \right)_{k,k} \\ &= \sum_{1 \leq k_1, \dots, k_{s-1} \leq c} t^{u, k_1} \left(\frac{2\pi k}{n} \right) t^{k_1, k_2} \left(\frac{2\pi k}{n} \right) \dots t^{k_{s-1}, u} \left(\frac{2\pi k}{n} \right) \end{aligned}$$

Ainsi

$$\begin{aligned} \text{Tr}(\mathcal{D}_n^s(\{t^{u,v}\})) &= \sum_{u=1}^c \sum_{1 \leq k_1, \dots, k_{s-1} \leq c} \sum_{k=1}^n t^{u, k_1} \left(\frac{2\pi k}{n} \right) t^{k_1, k_2} \left(\frac{2\pi k}{n} \right) \dots t^{k_{s-1}, u} \left(\frac{2\pi k}{n} \right) \\ &= \sum_{1 \leq k_1, \dots, k_s \leq c} \sum_{k=1}^n t^{k_1, k_2} \left(\frac{2\pi k}{n} \right) t^{k_2, k_3} \left(\frac{2\pi k}{n} \right) \dots t^{k_s, k_1} \left(\frac{2\pi k}{n} \right) \end{aligned}$$

Donc

$$\begin{aligned} &\lim_{n \rightarrow \infty} \frac{1}{n} \sum_{k=1}^{cn} \lambda_k^s(\mathcal{T}_n(\{t^{u,v}\})) \\ &= \lim_{n \rightarrow \infty} \frac{1}{n} \sum_{1 \leq k_1, \dots, k_s \leq c} \sum_{k=1}^n t^{k_1, k_2} \left(\frac{2\pi k}{n} \right) t^{k_2, k_3} \left(\frac{2\pi k}{n} \right) \dots t^{k_s, k_1} \left(\frac{2\pi k}{n} \right) \\ &= \sum_{1 \leq k_1, \dots, k_s \leq c} \lim_{n \rightarrow \infty} \frac{1}{n} \sum_{k=1}^n t^{k_1, k_2} \left(\frac{2\pi k}{n} \right) t^{k_2, k_3} \left(\frac{2\pi k}{n} \right) \dots t^{k_s, k_1} \left(\frac{2\pi k}{n} \right). \end{aligned}$$

Reste à y voir la définition de l'intégrale de Riemann. ■

Nous posons $\mathbf{T}(w) \hat{=} \begin{bmatrix} t^{1,1}(w) & \dots & t^{1,c}(w) \\ \vdots & & \vdots \\ t^{c,1}(w) & \dots & t^{c,c}(w) \end{bmatrix}$. Si $\mathcal{T}_n(\{t^{u,v}\})$ est Hermitienne, alors $\mathbf{T}(w)$ est Hermitienne pour tout w . La sommation $\sum_{1 \leq k_1, \dots, k_s \leq c} t^{k_1, k_2}(w) t^{k_2, k_3}(w) \dots t^{k_s, k_1}(w)$ n'est autre que la trace de $(\mathbf{T}(w))^s$ dont les valeurs propres sont celles de $\mathbf{T}(w)$ élevées à la puissance s . Par conséquent, (6.9) est équivalente à

$$\lim_{n \rightarrow \infty} \frac{1}{n} \sum_{k=1}^{cn} \lambda_k^s(\mathcal{T}_n(\{t^{u,v}\})) = \frac{1}{2\pi} \int_{w=-\pi}^{\pi} \sum_{u=1}^c \lambda_u^s(\mathbf{T}(w)) dw.$$

Il en découle que pour tout polynôme P , on a

$$\lim_{n \rightarrow \infty} \frac{1}{n} \sum_{k=1}^{cn} P[\lambda_k(\mathcal{T}_n(\{t^{u,v}\}))] = \frac{1}{2\pi} \int_{w=-\pi}^{\pi} \sum_{u=1}^c P[\lambda_u(\mathbf{T}(w))] dw. \quad (6.10)$$

Lorsque $\mathbf{T}(w)$ est Hermitienne $\forall w$, et en utilisant le théorème de Stone-Weierstrass (rappelé dans [GRAY 72]), (6.10) s'étend à toute fonction F continue sur $[\min_w \lambda_c(\mathbf{T}(w)), \max_w \lambda_1(\mathbf{T}(w))]$. On obtient ainsi l'extension suivante du théorème de Szegő's aux matrices blocs de Toeplitz :

Theorem 6.7 *Supposons que $\mathcal{T}_n(\{t^{u,v}\})$ est Hermitienne; alors pour tout fonction F continue, on a :*

$$\lim_{n \rightarrow \infty} \frac{1}{n} \sum_{k=1}^{cn} F[\lambda_k(\mathcal{T}_n(\{t^{u,v}\}))] = \frac{1}{2\pi} \int_{w=-\pi}^{\pi} \sum_{u=1}^c F[\lambda_u(\mathbf{T}(w))] dw$$

Il convient de noter que ce résultat est plus approprié que celui établi dans [VOOI 96, BOSE 98, TILL 97, SERR 98]. Les résultats asymptotiques dans [VOOI 96, BOSE 98, TILL 97, SERR 98] sont énoncés pour des matrices blocs de Toeplitz avec des blocs de Toeplitz où à la fois la taille et le nombre des blocs tend vers l'infini; alors que dans notre travail, seule la taille n des blocs tend vers l'infini. Ceci est plus approprié dans le contexte des matrices de corrélation des processus multi-variables où n désigne l'intervalle d'observation du processus et c désigne la taille du réseau d'antennes et/ou le facteur de sur-échantillonnage qui naturellement ne sont pas prévus pour prendre des valeurs importantes. Cependant, les résultats dans [VOOI 96, BOSE 98, TILL 97, SERR 98] restent exploitables dans l'étude des processus bi-dimensionnels; le cas par exemple des applications de traitement d'images.

Conjugué au fait que le spectre associé à la matrice Hermitienne $\mathcal{T}_n(\{t^{u,v}\})$ est contenu dans $[\min_w \lambda_c(\mathbf{T}(w)), \max_w \lambda_1(\mathbf{T}(w))]$ [TILL, Th. 3.1], Th. 6.7 implique que (cf. [TILL, MIRA 96] pour les détails) pour tout entier l , la $l^{\text{ème}}$ plus petite (resp., plus grande) valeur propre est convergente en n , et que

$$\lim_{l \rightarrow \infty} \lambda_{cn-l+1}(\mathcal{T}_n(\{t^{u,v}\})) = \min_w \lambda_c(\mathbf{T}(w)) \quad (6.11)$$

$$\lim_{l \rightarrow \infty} \lambda_l(\mathcal{T}_n(\{t^{u,v}\})) = \max_w \lambda_1(\mathbf{T}(w)) \quad (6.12)$$

6.3.3 Classe particulière des matrices blocs de Toeplitz

Dans ce paragraphe, on s'intéresse au cas particulier où la séquence de matrices blocs de Toeplitz vérifie la condition suivante :

Hypothèse H 6.1 $\mathbf{T}(w)$ est de rang 1, $\forall w$.

Ceci est équivalent à avoir pour tout w

$$\mathbf{T}(w) = [t^1(w), \dots, t^c(w)]^T |\text{Tr}(\mathbf{T}(w))| [t^1(w), \dots, t^c(w)]$$

où $[t^1(w), \dots, t^c(w)]^T$ et $[t^1(w), \dots, t^c(w)]^T$ sont des vecteurs unitaires singuliers à gauche et à droite respectivement, associés à la seule valeur singulière ($|\text{Tr}(\mathbf{T}(w))|$ non nulle) de $\mathbf{T}(w)$. Cette valeur singulière étant de multiplicité unité, elle, ainsi que les vecteurs singuliers associés, sont des fonctions continues de $\mathbf{T}(w)$ [BJRK 96, Th. 1.2.8, p.14]¹² qui est, à son tour, fonction continue de w par construction. Ainsi, en redéfinissant $t^u(w)$ et $t'^u(w)$ comme étant respectivement $\sqrt{|\text{Tr}(\mathbf{T}(w))|} t^u(w)$ et $\sqrt{|\text{Tr}(\mathbf{T}(w))|} t'^u(w)$, **H6.1** est équivalente au suivant : il existe des fonctions $t^u(w)$ et $t'^u(w)$ continues tels que $t^{u,v}(w) = t^u(w)t'^v(w)$, pour tout $u, v = 1, \dots, c$.

$t^u(w)$ et $t'^u(w)$ étant continues, les séquences $\{t_k^u\}_{k=\dots, -1, 0, 1, \dots}$ et $\{t_k'^u\}_{k=\dots, -1, 0, 1, \dots}$, obtenues par la transformée de Fourier inverse de $t^u(w)$ et $t'^u(w)$ respectivement, $u = 1, \dots, c$, sont carré sommables et non identiquement nulles. L'hypothèse **H6.1** est équivalente à avoir $t_i^{u,v} = t_i^u * t_i'^v \triangleq \sum_k t_k^u t_{i-k}'^v$, $i = \dots, -1, 0, 1, \dots$, i.e., $t_i^{u,v} = [\dots, t_{-1}^u, t_0^u, t_1^u, \dots] [\dots, t_{i+1}'^v, t_i'^v, t_{i-1}'^v, \dots]^T$. Par conséquent, L'hypothèse **H6.1** est équivalente à la factorisation suivante:

$$\mathcal{T}_n(\{t^{u,v}\}) = \begin{bmatrix} \mathbf{T}_{(n)}(t^1) \\ \vdots \\ \mathbf{T}_{(n)}(t^c) \end{bmatrix} \begin{bmatrix} \mathbf{T}_{(n)}(t'^1) \\ \vdots \\ \mathbf{T}_{(n)}(t'^c) \end{bmatrix}^T, \quad (6.13)$$

où $\mathbf{T}_{(n)}(t^u)$ (resp., $\mathbf{T}_{(n)}(t'^u)$), $u = 1, \dots, c$, désignent la matrice de Toeplitz à n lignes dont la première (ligne) est $[\dots, t_{-1}^u, t_0^u, t_1^u, \dots]$ (resp., $[\dots, t_{-1}'^u, t_0'^u, t_1'^u, \dots]$). En plus, si $\mathcal{T}_n(\{t^{u,v}\})$ est Hermitienne et définie positive, **H6.1** est remplie ssi $\forall u = 1, \dots, c$ on a $t_k'^u = (t_{-k}^u)^*$, ou bien aussi $\mathbf{T}_{(n)}(t'^u) = \mathbf{T}_{(n)}^*(t^u)$, i.e.,

$$\mathcal{T}_n(\{t^{u,v}\}) = \begin{bmatrix} \mathbf{T}_{(n)}(t^1) \\ \vdots \\ \mathbf{T}_{(n)}(t^c) \end{bmatrix} \begin{bmatrix} \mathbf{T}_{(n)}(t^1) \\ \vdots \\ \mathbf{T}_{(n)}(t^c) \end{bmatrix}^H. \quad (6.14)$$

Cette condition est fréquemment vérifiée dans les applications de traitement du signal. En particulier, (6.14) représente la matrice de corrélation d'un processus multi-variables stationnaire de dimension c obtenu par filtrage SIMO d'un processus scalaire stationnaire. Remarquons, cependant, que cette factorisation, et donc **H6.1**, n'est pas le cas des matrices de corrélation des processus multi-variables stationnaires en général¹³. De la même manière, (6.13) représente la matrice d'inter-corrélation de deux processus multi-variables stationnaires issus du filtrage SIMO d'un même processus scalaire stationnaire.

Sous l'hypothèse **H6.1**, le théorème suivant est immédiatement déduit :

¹²Le vecteur singulier unitaire à gauche/droite n'est pas unique et est donné à une rotation près. Cette rotation est mieux exprimée à travers la paramétrisation de Givens. La restriction de ces paramètres à un domaine donné garantit l'unicité du vecteur singulier. Étant ainsi unique, l'application de [BJRK 96, Th. 1.2.8, p.14] conduit à la continuité de chacune des composantes du vecteur singulier.

¹³Ce n'est, par exemple, pas le cas des sortie des canaux MIMO.

Theorem 6.8 Si $\mathcal{T}_n(\{t^{u,v}\})$ est Hermitienne et vérifie **H6.1**, alors pour toute fonction F continue, on a :

$$\lim_{n \rightarrow \infty} \frac{1}{n} \sum_{k=1}^{cn} F(\lambda_k(\mathcal{T}_n(\{t^{u,v}\}))) = (c-1)F(0) + \frac{1}{2\pi} \int_{w=-\pi}^{\pi} F\left(\sum_{u=1}^c t^{u,u}(w)\right) dw$$

6.4 Application à l'identification aveugle

6.4.1 Application aux matrices de filtrage

La matrice de corrélation de la sortie d'un canal SIMO à minimum de phase $\mathbf{h}_M$ est une matrice blocs de Toeplitz. Formellement, en respectant les notations de (6.5) et (6.6), on a $\mathcal{T}_{l,ST}(\mathbf{h}_M) \mathcal{T}_{l,ST}^H(\mathbf{h}_M) = \mathcal{T}'_n(\{h^u h^{v*}\})$ et $\mathcal{T}_{l,TS}(\mathbf{h}_M) \mathcal{T}_{l,TS}^H(\mathbf{h}_M) = \mathcal{T}_n(\{h^u h^{v*}\})$ donc vérifient bien l'hypothèse **H6.1**. L'application de Th. 6.8 implique que pour toute fonction F continue, on a

$$\lim_{l \rightarrow \infty} \frac{1}{l} \sum_{k=1}^{cl} F(\lambda_k(\mathcal{T}_l(\mathbf{h}_M) \mathcal{T}_l^H(\mathbf{h}_M))) = (c-1)F(0) + \frac{1}{2\pi} \int_{w=-\pi}^{\pi} F\left(\sum_{k=1}^c |h^k(w)|^2\right) dw \quad (6.15)$$

Si $l + M \leq cl$, on a $\lambda_k(\mathcal{T}_l(\mathbf{h}_M) \mathcal{T}_l^H(\mathbf{h}_M)) = (\sigma_k(\mathcal{T}_l(\mathbf{h}_M)))^2$ pour $k = 1, \dots, l + M$ et $\lambda_k(\mathcal{T}_l(\mathbf{h}_M) \mathcal{T}_l^H(\mathbf{h}_M)) = 0$ pour $k > l + M$. Ainsi

$$(c-1)F(0) + \lim_{l \rightarrow \infty} \frac{1}{l} \sum_{k=1}^{l+M} F(\sigma_k^2(\mathcal{T}_l(\mathbf{h}_M))) = (c-1)F(0) + \frac{1}{2\pi} \int_{w=-\pi}^{\pi} F\left(\sum_{k=1}^c |h^k(w)|^2\right) dw$$

Par conséquent, on a

Theorem 6.9 Pour tout fonction F continue

$$\lim_{l \rightarrow \infty} \frac{1}{l} \sum_{k=1}^{l+M} F(\sigma_k(\mathcal{T}_l(\mathbf{h}_M))) = \frac{1}{2\pi} \int_{w=-\pi}^{\pi} F\left(\sqrt{\sum_{k=1}^c |h^k(w)|^2}\right) dw. \quad (6.16)$$

Ceci décrit le comportement de groupe des valeurs singulières asymptotiques des matrices de filtrage. Il est important de caractériser le comportement asymptotique de la plus petite valeur singulière $\sigma(\mathcal{T}_l(\mathbf{h}_M)) \hat{=} \sigma_{(l+M)}(\mathcal{T}_l(\mathbf{h}_M))$. Cependant, elle ne pas être écrite sous la forme de $\lambda_{l'}(\mathcal{T}_l(\mathbf{h}_M) \mathcal{T}_l^H(\mathbf{h}_M))$ or $\lambda_{lc-l'+1}(\mathcal{T}_l(\mathbf{h}_M) \mathcal{T}_l^H(\mathbf{h}_M))$, pour un l' fixe. On ne peut donc mettre à profit (6.11) ni (6.12). Seul le résultat suivant a pu être démontré :

Theorem 6.10 Si $\sigma(\mathcal{T}_l(\mathbf{h}_M))$ converge en l , alors

$$\lim_{l \rightarrow \infty} \sigma(\mathcal{T}_l(\mathbf{h}_M)) \leq \min_w \left(\sqrt{\sum_{u=1}^c |h^u(w)|^2} \right) \quad (6.17)$$

Démonstration:

La démonstration s'inspire de celle de [TILL, Corollaire 3.9]. On se donne la fonction réelle positive $t(w) \hat{=} \sqrt{\sum_{u=1}^c |h^u(w)|^2}$ et on pose $m_t \hat{=} \min_w t(w)$. On suppose que la suite

$\sigma(\mathcal{T}_l(\mathbf{h}_M)), l \geq 0$ est convergente vers la limite $\mathcal{L}$ lorsque $l \rightarrow \infty$. On suppose que $\mathcal{L} > m_t$. Il existe alors a and b tels que $m_t < a < b < \mathcal{L}$. On défini la fonction $F(x) = 1$ si $x < a$, $F(x) = 0$ si $x > b$. Si $x \in [a, b]$, $F(x)$ est choisi de manière à ce que F soit continue et positive.

Il existe un entier L tel que pour tout entier $l > L$, $\sigma_k(\mathcal{T}_l(\mathbf{h}_M)) \geq \sigma(\mathcal{T}_l(\mathbf{h}_M)) > b$ et par conséquent, $F(\sigma_k(\mathcal{T}_l(\mathbf{h}_M))) = 0$. Il s'en suit que le terme de gauche de (6.16) est nul. Le terme de droite, cependant, est égale à

$$\begin{aligned} \frac{1}{2\pi} \int_{w=-\pi}^{\pi} F(t(w)) dw &= \frac{1}{2\pi} \int_{w \in [-\pi, +\pi] \text{ and } t(w) < b} F(t(w)) dw \\ &> \frac{1}{2\pi} \int_{w \in [-\pi, +\pi] \text{ and } t(w) < a} F(t(w)) dw \\ &= \frac{1}{2\pi} \int_{w \in [-\pi, +\pi] \text{ and } t(w) < a} dw \\ &> 0 \end{aligned}$$

ce qui est absurde. Ainsi, on ne peut qu'avoir $\mathcal{L} \leq m_t$. ■

Une caractérisation plus fine du comportement asymptotique de $\sigma(\mathcal{T}_l(\mathbf{h}_M))$ s'est avérée difficile à établir. En effet, les exemples numériques montrent qu'elle est bien convergente mais, par contre, non monotone (cf. Fig. 6.2).

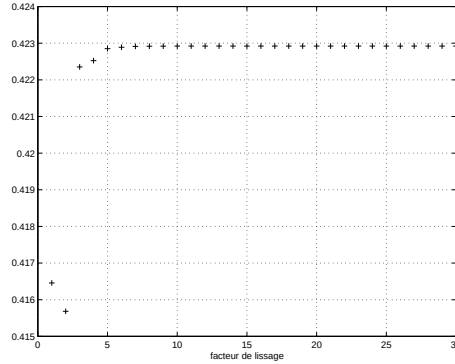


Figure 6.2: $\sigma(\mathcal{T}_l(\mathbf{h}_M)), \mathbf{h}_M^1 = [-0.6804 \ 0.177 \ -0.0902]^T, \mathbf{h}_M^2 = [0.4281 \ -0.2446 \ -0.5043]^T$. La borne (6.17) vaut 0.6747.

6.4.2 Application à la diversité des canaux SIMO

Le canal SIMO étant supposé à minimum de phase, la matrice de corrélation de sa sortie est asymptotiquement de rang colonne plein de telle manière que (6.17) représente une borne supérieure sur la diversité, telle que définie dans §6.1. Cette borne peut être utilisée pour évaluer l'identifiabilité en aveugle et sur la base d'un nombre fini d'observations. En particulier, au cas où cette borne est faible, le canal est forcément mal conditionné et est non identifiable en pratique.

Un aspect important de cette borne est que, lorsque faible, elle induit, non seulement la mauvaise diversité de la réponse totale, mais également celle de la réponse effective, i.e., celle ne tenant pas

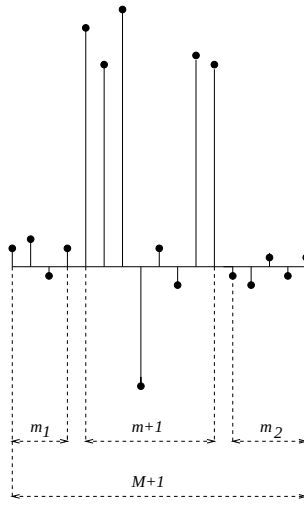


Figure 6.3: Canal avec termes faibles de la tête/queue

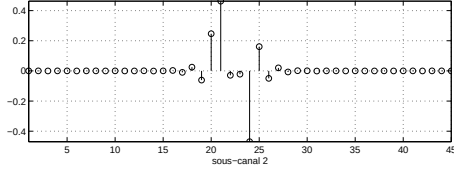
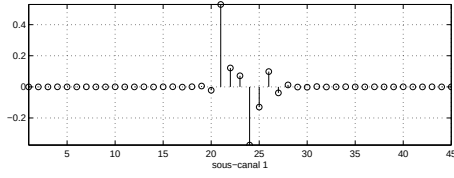
compte des termes faibles de la tête et/ou queue (Fig. 6.3)¹⁴. En effet, lorsque de tels termes existent, la réponse totale $\mathbf{h}_M$ peut se scinder en une réponse effective $\mathbf{h}_m$ et une partie tête/queue (notations de §7.2.1). La borne (6.17) est presque la même pour $\mathbf{h}_M$ et $\mathbf{h}_m$ au sens que $\sum_{u=1}^c |h_M^u(w)|^2 \simeq \sum_{u=1}^c |h_m^u(w)|^2$. Ainsi la mauvaise diversité de la réponse totale *au sens de la borne (6.17)* implique la mauvaise diversité de la réponse effective. Lorsque la borne (6.17) évaluée pour $\mathbf{h}_m$ n'est pas faible, elle est également pas faible pour $\mathbf{h}_M$. Dans ce cas, la borne (6.17) n'est pas serrée car la diversité de $\mathbf{h}_M$ est faible à cause des termes faibles de tête/queue (cf. §7.1.1 et [LIAV 99]). Plus généralement, la borne (6.17) s'avère non serrée (cf. Fig. 6.2) et ne permet donc pas de détecter toutes les situations de mauvaise diversité. Cependant, lorsque elle est faible, elle indique *une non-identifiabilité absolue* au sens que l'identification n'est pas possible quelque soit l'ordre présumé du canal; car ou bien on aura pris en compte des termes faibles de la tête et/ou queue qui alors provoquent le mauvais conditionnement/diversité; ou bien on aura négligé des termes significatifs. Et même si on néglige exactement les termes faibles de la tête/queue, la réponse recherchée est, d'après (6.17), de mauvaise diversité. Un exemple est donné dans le cas d'une réception fractionnée d'un canal à bande limitée (cf. §6.4.3).

La borne ici établie se montre un outil simple pour l'évaluation de l'identifiabilité en aveugle des canaux SIMO. Elle permet d'avoir une interprétation spectrale de la diversité. Elle est mieux interprétable que l'étude [Cibl 98] qui elle se base sur des notions mathématiques difficiles (*spheroidal wave sequences*). Elle est également indépendante de l'algorithme d'identification aveugle en question alors que l'étude [Cibl 98] concerne l'algorithme sous-espace.

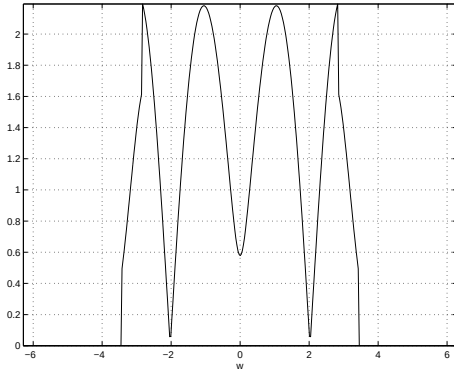
6.4.3 Cas de la réception fractionnée des signaux à bande limitée

On s'intéresse à la réception fractionnée des signaux à bande limitée. Si les sous-canaux $h^k(w)$, $k = 1, \dots, c$ sont issus du sur-échantillonnage d'un signal $h(t)$, alors $h^k(w)$ est la périodisation de la transformée de Fourier du signal $h\left(t - \frac{k-1}{c}\right)$ soit $h^k(w) = \sum_l h(w - 2l\pi) e^{-j(w - 2l\pi)\frac{k-1}{c}}$. Si $h(t)$ est à bande

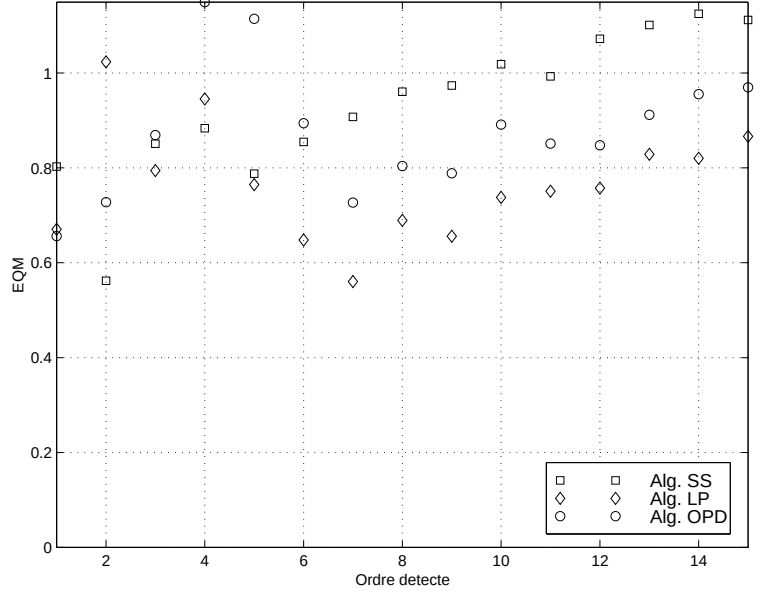
¹⁴cf. définition de la réponse effective §7



(a) Réponse impulsionnelle



(b) Réponse fréquentielle



(c) Identification avec 300 observations et $SNR = 20\text{dB}$

Rolloff de 0.3. Multi-trajets avec les retards respectifs 0, 1.0341 T , 1.1826 T , 1.3605 T , 1.8780 T , 1.9039 T , 2.1689 T , 2.2460 T , 3.3337 T , 3.4301 T , 3.7009 T et 4 T et les atténuations respectives 1, 0.9868, -0.3123 , -0.3661 , -0.6635 , 0.6918, -0.4045 , 0.5188, -0.4929 , -0.7333 , -0.0459 et .6. La borne (6.17) vaut 0.0851.

Figure 6.4: Cas de non-identifiabilité absolue

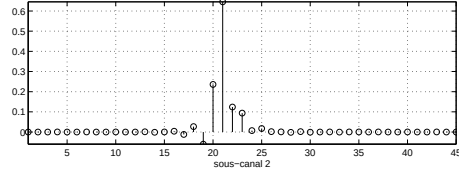
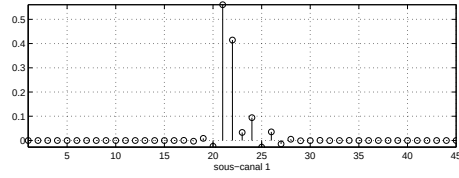
limitée (à $[-\frac{1}{T}, \frac{1}{T}]$), alors

$$\begin{aligned}
 h(w) &= 0 \text{ si } w \notin [-2\pi, 2\pi] \\
 h^k(w) &= h(w)e^{-jw\frac{k-1}{c}} + h(w-2\pi)e^{-j(w-2\pi)\frac{k-1}{c}} \text{ si } w \in [0, 2\pi] \\
 |h^k(w)|^2 &= |h(w)|^2 + |h(w-2\pi)|^2 + 2 \operatorname{Re} \left(h(w)h^*(w-2\pi)e^{2j\pi\frac{k-1}{c}} \right) \\
 \sum_{u=1}^c |h^u(w)|^2 &= c \left(|h(w)|^2 + |h(w-2\pi)|^2 \right) + 2 \operatorname{Re} \left(h(w)h^*(w-2\pi) \sum_{k=0}^{c-1} e^{2j\pi\frac{k}{c}} \right)
 \end{aligned}$$

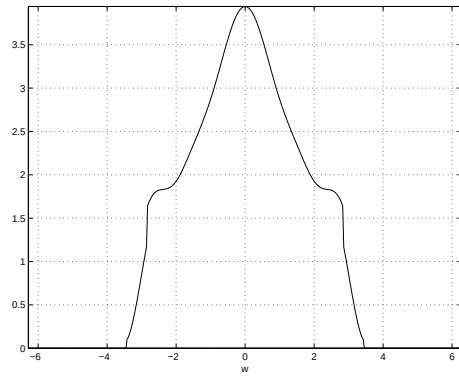
Ainsi, (6.17) se simplifie comme suit

$$\lim_{l \rightarrow \infty} \sigma(\mathcal{T}_l(\mathbf{h}_M)) \leq \sqrt{c} \min_w \left(\sqrt{|h(w)|^2 + |h(w-2\pi)|^2} \right)$$

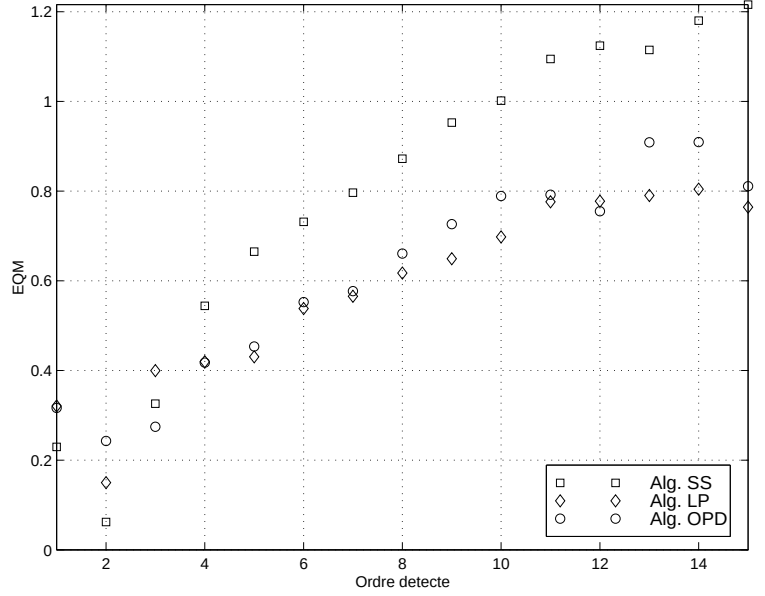
En pratique $h(w)$ est la mise en cascade d'une onde de mise en forme à bande limitée (du type cosinus sur-élevé le plus souvent) et de la réponse d'un canal à trajets multiples. Les trajets multiples se



(a) Réponse impulsionnelle



(b) Réponse fréquentielle



(c) Identification avec 300 observations et $SNR = 20dB$

Rolloff de 0.3. Multi-trajets avec les retards respectifs 0, 0.0470 T , 0.5461 T , 1.0093 T , 1.0858 T , 2.9492 T , 3.5756 et 4 T et les atténuations respectives 1, 0.8381, 0.4418, 0.3689, 0.8672, 0.6775, -0.6518 et 0.4. La borne (6.17) vaut 1.0690.

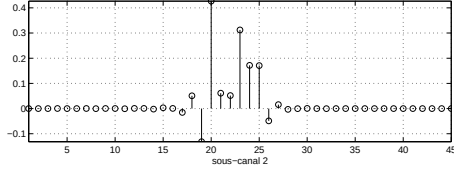
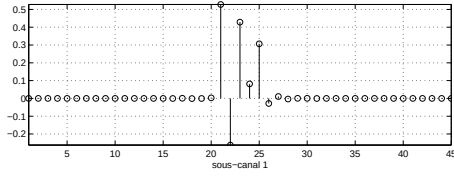
Figure 6.5: Cas identifiable

traduisent souvent par une sélectivité en fréquence. Lorsqu'une fréquence est sévèrement atténuée, la borne peut prendre alors des valeurs faibles. Ceci explique les mauvaises performances des algorithmes d'identification aveugle pour les canaux à bande limitée et rejoint les remarques faites dans [Cibl 98, Veen 95].

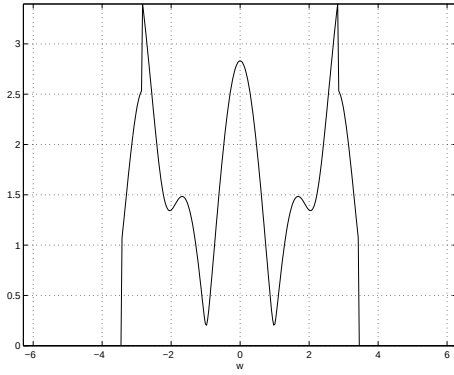
Les simulations confirment les commentaires avancés dans le paragraphe précédent (§6.4.2). Un tirage aléatoire de canaux à multi-trajets¹⁵ a été établi. Les canaux pour qui la borne (6.17) est faible¹⁶ ont

¹⁵Le trajet le plus retardé a été fixé avec un retard de 4 T (T étant la durée symbole) et un affaiblissement de 0.4, par rapport à un trajet principal non retardé et non atténué. Le nombre de trajets intermédiaires était à chaque fois aléatoire, ainsi que les retards et les atténuations. Le signal de mise en forme est un cosinus sur-élevé de rolloff 0.3 pour une première série de tirages, et de rolloff 0.6 pour une deuxième série. Sa réponse impulsionnelle est tronquée sur 40 T . Les symboles transmis étant binaires, indépendants et équiprobables. Le nombre de canaux tirés est 1000 par série.

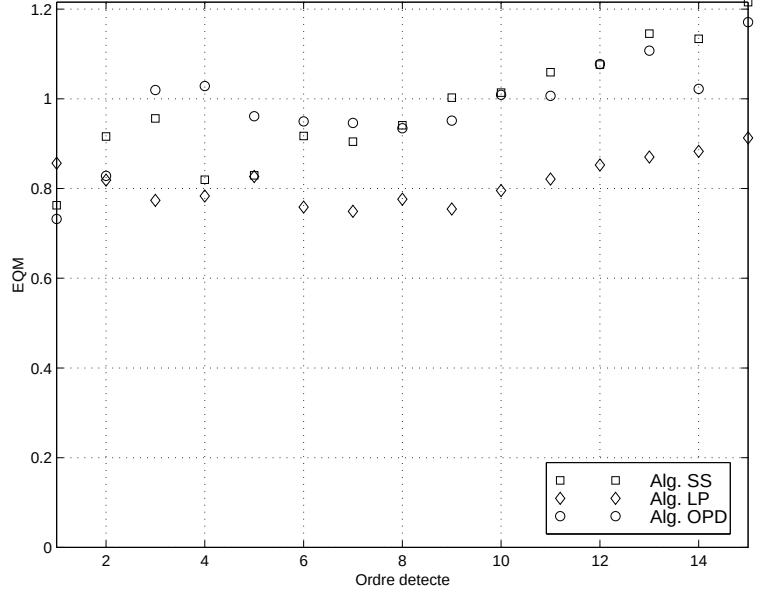
¹⁶La borne était considérée faible lorsque ≤ 0.1 .



(a) Réponse impulsionnelle



(b) Réponse fréquentielle



(c) Identification avec 300 observations et $SNR = 20\text{dB}$

Rolloff de 0.3. Multi-trajets avec les retards respectifs 0, $0.0611 T$, $0.7949 T$, $0.7953 T$, $1.0888 T$, $1.6746 T$, 1.7804 , 1.8640 , 2.4152 , 2.9871 , 3.7273 et $4 T$ et les atténuations respectives 1, 0.4952, 0.6583, -0.8267 , -0.9669 , 0.8591, 0.8911, -0.6676 , 0.5272, -0.0201 , 0.4824 et 0.4. La borne (6.17) vaut 0.3033.

Figure 6.6: Cas de non-identifiabilité absolue

systématiquement été non identifiables¹⁷ en réception fractionnée¹⁸ à n'importe quel ordre considéré¹⁹. Ces exemples (dont un est illustré dans Fig. 6.4) confirment l'interprétation de *non-identifiabilité absolue* lorsque la borne (6.17) est faible. Dans le cas contraire, cependant, on ne peut pas conclure sur l'existence (Fig. 6.5) ou non (Fig. 6.6) d'un ordre avec qui l'identification aboutit à l'approximation d'une sous-réponse (effective) du canal en question. Il a été constaté que lorsque des trajets significatifs se manifestent à des délais supérieurs à la durée symbole, le canal n'est pas identifiable, quelque soit l'ordre présumé et quelque soit la valeur que prend la borne (6.17).

¹⁷L'identification a été tentée par l'algorithme sous-espace, à partir de 300 observations et avec un SNR de 20 dB. Si $\mathbf{h}_M$ est la réponse totale du canal multi-trajets, et si m est l'ordre de l'estimée $\hat{\mathbf{h}}_m$, alors l'erreur quadratique moyenne est calculée comme étant (la réponse du canal étant réelle)

$$\min_{m_1=0, \dots, M-m} \left\{ \left\| \frac{\mathbf{h}_M}{\|\mathbf{h}_M\|} - \begin{bmatrix} \mathbf{0}_{1,cm_1} & \frac{\hat{\mathbf{h}}_m^T}{\|\hat{\mathbf{h}}_m\|} & \mathbf{0}_{1,c(M-m-m_1)} \end{bmatrix}^T \right\|, \left\| \frac{\mathbf{h}_M}{\|\mathbf{h}_M\|} + \begin{bmatrix} \mathbf{0}_{1,cm_1} & \frac{\hat{\mathbf{h}}_m^T}{\|\hat{\mathbf{h}}_m\|} & \mathbf{0}_{1,c(M-m-m_1)} \end{bmatrix}^T \right\| \right\},$$
 (format ST). Lorsque cette dernière est ≥ 0.1 , la réponse était considérée non identifiable.

¹⁸Facteur de sur-échantillonnage égal à 2.

¹⁹Les ordres tentés ont varié à chaque fois de 1 à 20.

6.5 Conclusion

Dans ce chapitre, on a mis en évidence que les performances des algorithmes d'identification dépendent du bon conditionnement de la matrice de filtrage. En pratique, il ne suffit pas à un canal d'être à minimum de phase pour être identifiable en aveugle mais il lui faut également être *bien diversifié*. Une mesure de la diversité a été proposée et est exprimée en fonction de la limite asymptotique de la plus petite valeur singulière de la matrice de filtrage du canal lorsque le facteur de lissage de cette dernière tend vers l'infini.

En remarquant que cette plus petite valeur singulière est aussi une valeur propre de la matrice de corrélation de la sortie du canal et que cette matrice possède une structure en blocs de Toeplitz, le lien avec le théorème de Szegő, dont une démonstration plus simple est proposée, est établi et exploité pour lier la diversité du canal à sa réponse fréquentielle. Cette interprétation spectrale de la diversité permet en particulier d'évaluer la non identifiabilité dans des conditions pratiques d'observations des canaux à bande limitée.

Chapitre 7

Modélisation effective de l'ordre

Le but du chapitre est de définir une méthodologie d'analyse de performance applicable aux algorithmes d'identification du second ordre en modélisation effective de l'ordre du canal (et autre conditions pratiques). On sous-entend par modélisation effective de l'ordre du canal, l'idée de ne prendre en compte dans la réponse impulsionnelle du canal que les termes les plus significatifs. Comment ces termes sont définis et comment cet ordre est déterminé font l'objet du §7.1.

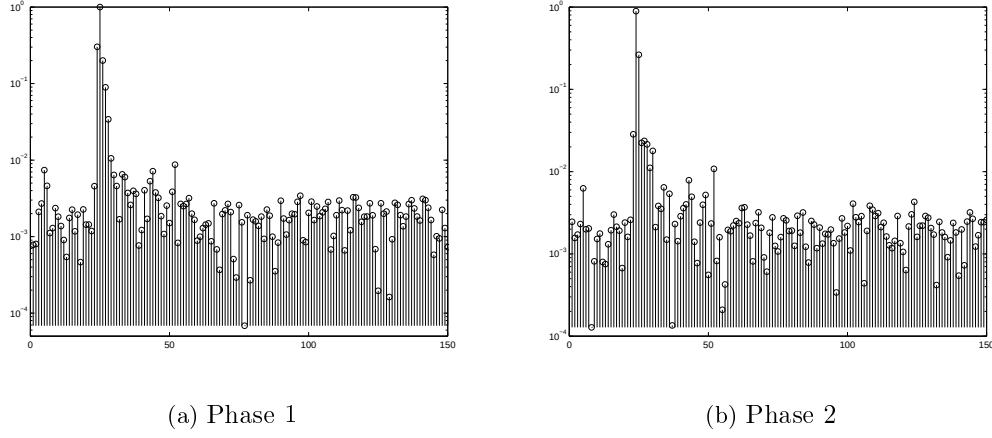
L'idée fut introduite par *Liavas, Regalia* et *Delmas*. Dans [LIAV 99], ils ont présenté un algorithme (qu'on désigne désormais par LRD) de détection de l'*ordre effectif* d'un canal, puis ont démontré que dans les situations pratiques, cet ordre effectif permet d'identifier l'essentiel de la réponse du canal et de l'égaliser. L'étude de performances dans ces conditions a été testée avec les algorithmes sous-espace SS/LS [LIAV 99] et prédiction linéaire LP [LIAV 00]. En particulier, des bornes (rappelées dans §7.6.4) pour majorer l'erreur (quadratique moyenne) d'estimation y ont été développées.

Le présent travail reprend l'idée de la modélisation effective et mène une analyse de performance plus générale que dans [LIAV 99, LIAV 00]. Elle est appliquée aux algorithmes SS, LS, LP et OPD (§7.6) et peut être appliquée à n'importe quel autre algorithme du second ordre. Utilisant les méthodes de perturbation au premier ordre, elle donnera des expressions analytiques du biais et de la variance asymptotiques qui s'avéreront être plus fines que les bornes données dans [LIAV 99, LIAV 00]. Cette étude s'intéresse également à d'autres paramètres dérivés de la réponse impulsionnelle du canal; en l'occurrence les égaliseurs ZF, la réponse combinée canal/égaliseur et l'ouverture moyenne de l'oeil. Elle se fera en supposant un nombre fini d'échantillons reçus, contrairement à [LIAV 99, LIAV 00] où les statistiques (du second ordre) des observations ont été supposées exactement connues.

7.1 Réponse et ordre effectifs

7.1.1 Intérêt

La notion de partie effective fut introduite dans [LIAV 99]. L'examen de réponses impulsionnelles de canaux de communications permet de constater que le plus souvent les réponses impulsionnelles (Fig. 7.1) présentent une majorité de termes faibles et seulement quelques termes prépondérants. Les tests



Réponse impulsionnelle^a, ramenée en bande de base, calculée à partir de mesures faites sur une liaison radio à 30 Mbits/s. La réponse impulsionnelle du canal est sur-échantillonnée d'un facteur de 2. Sont représentés les modules des termes de la réponse. D'autres exemples se trouvent à l'adresse URL <http://spib.rice.edu/spib/microwave.html>.

^aSource : Signal Processing Information Base (SPIB). Canal 1.

Figure 7.1: Canal radio hyper-fréquence

de détection de l'ordre total du canal [WAX 85] conduisent à la prise en considération d'aux moins certains de ces termes faibles de la tête/queue, responsables alors de la mauvaise diversité (au sens précisé dans §6.1) de la réponse recherchée.

En effet, on peut décomposer (cf. §7.2.1) la réponse totale $\mathbf{h}_M$ en une partie effective $\mathbf{h}_m$ ($m \leq M$) et une tête/queue qui joue le rôle d'un terme de perturbation faible. $\mathcal{T}_l(\mathbf{h}_M) \mathcal{T}_l^H(\mathbf{h}_M)$ s'exprime alors comme une perturbation de $\mathcal{T}_l(\mathbf{h}_m) \mathcal{T}_l^H(\mathbf{h}_m)$ et, par conséquent, $\lambda_{l+M}(\mathcal{T}_l(\mathbf{h}_M) \mathcal{T}_l^H(\mathbf{h}_M))$ est très proche de $\lambda_{l+M}(\mathcal{T}_l(\mathbf{h}_m) \mathcal{T}_l^H(\mathbf{h}_m)) = 0$. Ainsi l'échec des algorithmes d'identification est prévisible. D'un autre côté, il est évident que la partie effective ainsi définie est d'une meilleure diversité que la réponse totale puisque $\sigma(\mathcal{T}_l(\mathbf{h}_m)) = \sigma_{l+m}(\mathcal{T}_l(\mathbf{h}_m)) > 0 = \sigma_{l+M}(\mathcal{T}_l(\mathbf{h}_m)) \simeq \sigma_{l+M}(\mathcal{T}_l(\mathbf{h}_M)) = \sigma(\mathcal{T}_l(\mathbf{h}_M))$.

7.1.2 Définitions

La définition de la partie effective est partie de l'examen des canaux du type Fig. 7.1. La définition de la partie effective a besoin d'être généralisée. On appellera m *ordre effectif* d'un canal $\mathbf{h}_M$ celui qui émane de l'application de l'algorithme LRD à la matrice de corrélation exacte de sa sortie. À partir de l'estimée $\hat{\mathbf{R}}_l$ de la matrice de corrélation du canal $\mathbf{h}_M$ (le facteur de lissage l est choisi arbitrairement "grand"), l'ordre m est choisi comme étant

$$m \hat{=} -l + \operatorname{argmin}_{q=1, \dots, cl} r(q) \quad \text{où} \quad r(q) \hat{=} \frac{\lambda_{q+1}(\hat{\mathbf{R}}_l)}{\lambda_q(\hat{\mathbf{R}}_l) - 2} \quad \text{si} \quad \lambda_{q+1}(\hat{\mathbf{R}}_l) \leq \frac{\lambda_q(\hat{\mathbf{R}}_l)}{3}, \quad 1 \quad \text{si non} \quad (7.1)$$

Liavas et al montrent que lorsqu'il y a un *saut* ou *gap* important entre $\lambda_{[\text{argmin}_{q=1,\dots,cl} r(q)]}(\hat{\mathbf{R}}_l)$ et $\lambda_{[\text{argmin}_{q=1,\dots,cl} r(q)]+1}(\hat{\mathbf{R}}_l)$, alors le canal en question se partitionne bien en une partie effective et une tête/queue faible à la manière de Fig. 7.1. Autrement, une telle partition n'est pas possible et on ne peut donner une interprétation physique à la valeur qui émane de (7.1). Ils avancent que cette estimée est robuste par rapport au nombre des observations et au rapport signal sur bruit; et ce en contraste avec les critères AIC et MDL (cf. §4.3).

À partir de l'ordre effectif, on définit la *partie effective* comme étant la sous-partie $\mathbf{h}_m$ d'ordre m qui concentre le plus d'énergie, soit¹ :

$$\mathbf{h}_m = [\mathbf{h}_{m_1+1} \cdots \mathbf{h}_{m_1+m+1}]^T$$

$$\text{où } m_1 = \text{argmin}_{m'} \left(\left\| \mathbf{h}_M - \left[\underbrace{\mathbf{0}_{c,1}^T, \dots, \mathbf{0}_{c,1}^T}_{m'}, \mathbf{h}_{m'+1}^T \cdots \mathbf{h}_{m'+m+1}^T, \underbrace{\mathbf{0}_{c,1}^T, \dots, \mathbf{0}_{c,1}^T}_{M-m-m'} \right]^T \right\| \right)$$

On notera $m_2 = M - m - m_1$

Il convient de confronter cette définition mathématique de la partie effective à la définition *visuelle*. Ceci est discuté dans le paragraphe suivant.

7.1.3 Insuffisances

L'algorithme LRD fournit, en plus de l'ordre m , une indication sur la pertinence de cette valeur. L'algorithme effectue un test de *gap* duquel il déduit la valeur m . Ce *gap* peut, cependant, être faible et dans ce cas, *Liavas et al* reconnaissent que la valeur de m ne peut pas être interprétée comme correspondant à une partition partie effective et tête/queue. Inversement, lorsque une partie effective prépondérante existe *visuellement*, l'algorithme LRD ne permet la détection de sa longueur que si elle a une bonne diversité. Autrement, il détecte une longueur plus courte.

À cause de l'emploi des signaux à bande limitée, des termes faibles existent toujours à la tête/queue et seront pris en compte par les tests de détection (AIC et MDL). La détection de l'ordre effectif est une alternative qui peut cependant échouer lorsque la partie effective (visuelle) est de mauvaise diversité. En pratique, ceci se produit lorsque le filtre d'émission (en cosinus sur-élevé par exemple) traverse un canal en trajets multiples avec des retards importants (supérieurs à une durée symbole). Cette situation est plus difficile puisque l'identification échouera quelque soit l'ordre de la partie qu'on cherchera à identifier (cf. Fig. 6.4 et Fig. 6.6).

En résumé, l'approche *ordre/partie effectifs* adresse une situation particulière, celle des canaux non identifiables en aveugle à cause d'une mauvaise diversité due (uniquement) à la présence de termes faibles au début et/ou à la fin de la réponse. Elle permet de s'intéresser à (identifier en aveugle) l'essentiel de la réponse, bien diversifié cette fois. Un égaliseur optimisé pour cette partie (effective) du canal le sera presque parfaitement pour la réponse totale du canal.

¹Format *ST*

7.2 Modélisation effective en identification aveugle

Il convient de redéfinir le vocabulaire sous/sur modélisation pour tenir compte de la notion d'ordre/réponse effectifs. On parlera de modélisation totale (resp. effective) si on détecte l'ordre total M (resp. effectif m). On parlera de sous/sur modélisation totale (resp. effective) si on détecte un ordre inférieur/supérieur à l'ordre total (resp. l'ordre effectif).

7.2.1 Notations

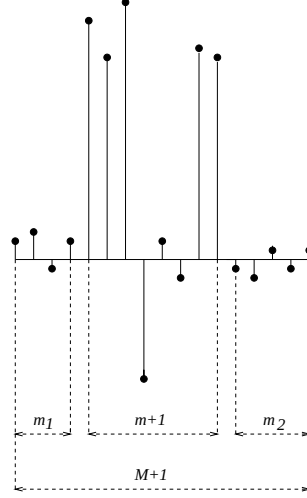


Figure 7.2: Canal type en modélisation effective

On se donne une réponse impulsionnelle $\mathbf{h}_M$ d'ordre (total) M du type Fig. 7.2 de manière à pouvoir y distinguer une partie effective $\mathbf{h}_m$ d'ordre (effectif) $m \leq M$ et m_1 termes faibles à la tête et m_2 termes faibles à la queue de la réponse, regroupés dans le vecteur $\mathbf{d}_{m_1, m_2}$. La réponse totale peut alors s'écrire

$$\mathbf{h}_M = \mathbf{h}_m^z + \mathbf{d}_{m_1, m_2}^z \quad (7.2)$$

avec²

$$\begin{aligned} \mathbf{h}_m^z &= \left[\underbrace{\mathbf{0}_{c,1}^T, \dots, \mathbf{0}_{c,1}^T}_{m_1}, \mathbf{h}_m^T, \underbrace{\mathbf{0}_{c,1}^T, \dots, \mathbf{0}_{c,1}^T}_{m_2} \right]^T \\ \mathbf{d}_{m_1, m_2}^z &= \left[\underbrace{\mathbf{h}^T(0), \dots, \mathbf{h}^T(m_1 - 1)}_{m_1}, \mathbf{0}_{c(m+1)}^T, \underbrace{\mathbf{h}^T(m_1 + m + 1), \dots, \mathbf{h}^T(m_1 + m + m_2)}_{m_2} \right]^T \\ \mathbf{h}_m &= [\mathbf{h}(m_1), \dots, \mathbf{h}(m_1 + m)]^T \\ \mathbf{d}_{m_1, m_2} &= [\mathbf{h}(0), \dots, \mathbf{h}(m_1 - 1), \mathbf{h}(m_1 + m + 1), \dots, \mathbf{h}(m_1 + m + m_2)]^T \end{aligned}$$

$\mathbf{d}_{m_1, m_2}^z$ est une version *zero-padded* de $\mathbf{d}_{m_1, m_2}$ et est lui relié par l'intermédiaire d'une matrice de

²Format ST

sélection³.

7.2.2 La perturbation de modélisation effective

Dans la suite du chapitre, on note $\mathcal{T}(\cdot) \triangleq \mathcal{T}_{m+1}(\cdot)$, $\mathbf{R} \triangleq \mathbf{R}_{m+1}$ et $\mathbf{R}_{h_m} \triangleq \mathcal{T}_{m+1}(\mathbf{h}_m) \mathcal{T}_{m+1}^H(\mathbf{h}_m) + \mathbf{R}_{m+1}^b$.⁴

La décomposition (7.2) se répercute sur la matrice de filtrage du canal $\mathcal{T}(\mathbf{h}_M) = \mathcal{T}(\mathbf{h}_m^z) + \mathcal{T}(\mathbf{d}_{m_1, m_2}^z)$ et sur la matrice de corrélation $\mathbf{R}$ des observations

$$\begin{aligned} \mathbf{R} &\triangleq \mathcal{T}(\mathbf{h}_M) \mathcal{T}^H(\mathbf{h}_M) + \mathbf{R}_{m+1}^b \\ &= \mathcal{T}(\mathbf{h}_m^z) \mathcal{T}^H(\mathbf{h}_m^z) + \mathcal{T}(\mathbf{d}_{m_1, m_2}^z) \mathcal{T}^H(\mathbf{d}_{m_1, m_2}^z) + \mathcal{T}(\mathbf{h}_m^z) \mathcal{T}^H(\mathbf{d}_{m_1, m_2}^z) + \mathcal{T}(\mathbf{d}_{m_1, m_2}^z) \mathcal{T}^H(\mathbf{h}_m^z) + \mathbf{R}_{m+1}^b \\ &= \mathbf{R}_{h_m} + \delta \mathbf{R}^d + \mathcal{O}\left(\|\mathbf{d}_{m_1, m_2}\|^2\right) \end{aligned}$$

où $\delta \mathbf{R}^d \triangleq \mathcal{T}(\mathbf{h}_m^z) \mathcal{T}^H(\mathbf{d}_{m_1, m_2}^z) + \mathcal{T}(\mathbf{d}_{m_1, m_2}^z) \mathcal{T}^H(\mathbf{h}_m^z)$ est la matrice de perturbation due à la non modélisation des termes de la tête/queue. Elle est linéaire en $\mathbf{d}_{m_1, m_2}$ au sens où il existe une matrice $\mathbf{C}$ telle que⁵

$$\text{Vec}(\delta \mathbf{R}^d) = \mathbf{C} \mathbf{d}_{m_1, m_2}^* + \mathbf{K}_{c(m+1)} \mathbf{C} \mathbf{d}_{m_1, m_2} \quad (7.5)$$

On définit également la matrice de corrélation empirique calculée à partir de N observations

$$\hat{\mathbf{R}} \triangleq \frac{1}{N-m} \sum_{n=m}^{N-1} \mathbf{y}_{m+1}(n) \mathbf{y}_{m+1}^H(n)$$

7.3 Approche fonctionnelle de l'identification

Un algorithme d'identification (aveugle), fonctionnant avec un ordre M supposé du canal, peut être formulé comme une fonction matricielle $\text{alg} : \mathbf{C}^{c(M+1)} \times \mathbf{C}^{c(M+1)} \longrightarrow \mathbf{C}^{c(M+1)}$. On note $\mathcal{M}_M \triangleq \{\mathcal{T}_{M+1}(\mathbf{x}_M) \mathcal{T}_{M+1}^H(\mathbf{x}_M) + x \mathbf{I}_{c(M+1)}, x \in \mathbf{R}^+, \mathbf{x}_M \in \mathbf{C}^{c(M+1)} \text{ à minimum de phase}\}$. La restriction de alg à $\mathcal{M}_M$ est :

$$\begin{aligned} \text{alg}_{/\mathcal{M}_M} : \mathcal{M}_M &\longrightarrow \mathbf{C}^{c(M+1)} \\ \mathcal{T}_{M+1}(\mathbf{x}) \mathcal{T}_{M+1}^H(\mathbf{x}) + x \mathbf{I}_{c(M+1)} &\longrightarrow a \mathbf{x}_M \end{aligned}$$

a traduit une ambiguïté *tolérable*⁶.

$$\mathbf{d}_{m_1, m_2}^z = \mathbf{I}^z \mathbf{d}_{m_1, m_2} \quad (7.3)$$

où $\mathbf{I}_{ST}^z \triangleq (\mathbf{I}_{m_1, m_2}^z \otimes \mathbf{I}_c)$, $\mathbf{I}_{TS}^z = (\mathbf{I}_c \otimes \mathbf{I}_{m_1, m_2}^z)$ et $\mathbf{I}_{m_1, m_2}^z \triangleq \begin{bmatrix} \mathbf{I}_{m_1} & \mathbf{0}_{m_1, m_2} \\ \mathbf{0}_{m+1, m_1} & \mathbf{0}_{m+1, m_2} \\ \mathbf{0}_{m_2, m_1} & \mathbf{I}_{m_2} \end{bmatrix}$.

⁴matrice de corrélation de la sortie du canal fictif $\mathbf{h}_m$ affecté par le même bruit.

⁵L'expression de $\mathbf{C}$ est donnée dans la démonstration (cf. §A.1). Dans le cas réel H7.2, il existe (cf. §A.2) une matrice $\mathbf{C}^r$ telle que

$$\text{Vec}(\delta \mathbf{R}^d) = \mathbf{C}^r \mathbf{d}_{m_1, m_2} \quad (7.4)$$

⁶de phase et d'amplitude (a complexe quelconque), ou bien de phase uniquement (a complexe de module 1).

Hypothèse H 7.1 *Le nombre de phases $c = 2$.*

On montrera (cf §B), pour chacun des algorithmes $\text{alg} \in \{SS/LS, LP, OPD\}$ dans le cas de 2 phases (H7.1), que la différentielle⁷ de alg au voisinage de $\mathbf{X}$ peut être définie comme

$$\text{alg}(\mathbf{X} + \mathbf{E}) = \text{alg}(\mathbf{X}) + \mathbf{A}_{\mathbf{X}}^1 \text{Vec}(\mathbf{E}) + \mathbf{A}_{\mathbf{X}}^2 \text{Vec}(\mathbf{E})^* + O(\|\mathbf{E}\|^2).$$

Quand $\mathbf{E}$ est Hermitienne, et en utilisant (2.1), on a $\text{alg}(\mathbf{X} + \mathbf{E}) = \text{alg}(\mathbf{X}) + \mathbf{A}_{\mathbf{X}} \text{Vec}(\mathbf{E}) + O(\|\mathbf{E}\|^2)$ où $\mathbf{A}_{\mathbf{X}} \triangleq \mathbf{A}_{\mathbf{X}}^1 + \mathbf{A}_{\mathbf{X}}^2 \mathbf{K}_{2(M+1)}$.

Hypothèse H 7.2 Cas réel *Données, canal et bruit sont réels.*

7.4 Étude de performances en modélisation effective

L'étude suppose la détection de l'ordre effectif exacte d'un canal SIMO excité par une entrée i.i.d. et centrée et observé pendant une durée finie et en présence d'un bruit additif, centré et décorrélié de l'entrée. Elle se fera selon deux modèles statistiques de la tête/queue i.e., du vecteur $\mathbf{d}_{m_1, m_2}$:

- Un modèle déterministe où la tête/queue est considérée comme constante.
- Un modèle probabiliste où les termes de la tête/queue sont modélisés par des variables aléatoires i.i.d. et centrées.

7.4.1 Modèle déterministe

Dans ce modèle, on s'intéresse au biais et l'EQM qu'introduit la présence d'une tête/queue déterministe sur l'estimation de la réponse effective du canal. On écrit $\hat{\mathbf{R}} = \mathbf{R} + \delta\mathbf{R}^N$, avec $\delta\mathbf{R}^N \triangleq \hat{\mathbf{R}} - \mathbf{R}$ pouvant être perçue comme une perturbation (Hermitienne) due à l'estimation à partir d'un nombre fini N d'observations. Elle vérifie $[\text{Vec}(\delta\mathbf{R}^N)] [\text{Vec}(\delta\mathbf{R}^N)]^H = \frac{1}{N} \mathcal{C}_{\mathbf{R}} + o\left(\frac{1}{N}\right)$ où $\mathcal{C}_{\mathbf{R}}$ est défini d'après Prop. 4.1. Ainsi [HANN 70], $\sqrt{N} \delta\mathbf{R}^N \xrightarrow{L} \mathcal{N}(\mathbf{0}_{c(m+1), 1}, \mathcal{C}_{\mathbf{R}}, \mathcal{C}_{\mathbf{R}} \mathbf{K}_{c(D+1)})$.

$$\begin{aligned} \hat{\mathbf{h}}_m &= \text{alg}(\hat{\mathbf{R}}) \\ &= \text{alg}(\mathbf{R} + \delta\mathbf{R}^N) \\ &= \text{alg}(\mathbf{R}) + \mathbf{A}_{\mathbf{R}} \text{Vec}(\delta\mathbf{R}^N) + O(\|\delta\mathbf{R}^N\|^2) \\ &= \text{alg}(\mathbf{R}) + \mathbf{A}_{\mathbf{R}} \text{Vec}(\delta\mathbf{R}^N) + O\left(\frac{1}{N}\right) \\ \mathbb{E}(\hat{\mathbf{h}}_m) &= \text{alg}(\mathbf{R}) + O\left(\frac{1}{N}\right) \\ \hat{\mathbf{h}}_m &= \mathbb{E}(\hat{\mathbf{h}}_m) + \mathbf{A}_{\mathbf{R}} \text{Vec}(\delta\mathbf{R}^N) + O\left(\frac{1}{N}\right) \end{aligned}$$

⁷Il convient de noter que la fonction alg n'est pas dérivable en $\mathbf{X}$ au vrai sens du terme puisque $\nexists \mathbf{A}_{\mathbf{X}}$ telle que $\text{alg}(\mathbf{X} + \mathbf{E}) = \text{alg}(\mathbf{X}) + \mathbf{A}_{\mathbf{X}} \text{Vec}(\mathbf{E}) + O(\|\mathbf{E}\|^2)$, $\forall \mathbf{E} \in \mathbf{C}^{2(M+1)} \times \mathbf{C}^{2(M+1)}$. Dans le cas réel (H7.2), alg est dérivable en $\mathbf{X}$. Sa différentielle est donnée par la seule matrice $\mathbf{A}_{\mathbf{X}} = \mathbf{A}_{\mathbf{X}}^1$. $\mathbf{A}_{\mathbf{X}}^2$ étant indéfinie.

$$\text{Cov}(\hat{\mathbf{h}}_m) = \mathbf{A}_R \mathbb{E} \left([\text{Vec}(\delta \mathbf{R}^N)] [\text{Vec}(\delta \mathbf{R}^N)]^H \right) \mathbf{A}_R^H + \mathcal{O}\left(\frac{1}{N\sqrt{N}}\right)$$

Proposition 7.1 Covariance asymptotique

$$\text{Cov}(\hat{\mathbf{h}}_m) = \frac{1}{N} \mathbf{A}_R \mathcal{C}_R \mathbf{A}_R^H + \mathcal{O}\left(\frac{1}{N}\right) \quad \text{et} \quad \lim_{N \rightarrow \infty} N \text{Cov}(\hat{\mathbf{h}}_m) = \mathbf{A}_R \mathcal{C}_R \mathbf{A}_R^H$$

$\mathbf{R}$ est à son tour vue comme une matrice $\mathbf{R}_{h_m}$ perturbée, soit

$$\begin{aligned} \mathbf{R} &= \mathbf{R}_{h_m} + \delta \mathbf{R}^d + \mathcal{O}\left(\|\mathbf{d}_{m_1, m_2}\|^2\right) \\ \text{alg}(\mathbf{R}) &= \text{alg}(\mathbf{R}_{h_m}) + \mathbf{A}_{\mathbf{R}_{h_m}} \text{Vec}(\delta \mathbf{R}^d) + \mathcal{O}\left(\|\delta \mathbf{R}^d\|^2\right) + \mathcal{O}\left(\|\mathbf{d}_{m_1, m_2}\|^2\right) \\ &= \text{alg}(\mathbf{R}_{h_m}) + \mathbf{A}_{\mathbf{R}_{h_m}} \left(\mathbf{C} \mathbf{d}_{m_1, m_2}^* + \mathbf{K}_{c(m+1)} \mathbf{C} \mathbf{d}_{m_1, m_2} \right) + \mathcal{O}\left(\|\mathbf{d}_{m_1, m_2}\|^2\right) \\ &= \mathbf{h}_m + \mathbf{B} \mathbf{d}_{m_1, m_2} + \mathbf{B}^\circ \mathbf{d}_{m_1, m_2}^* + \mathcal{O}\left(\|\mathbf{d}_{m_1, m_2}\|^2\right) \\ \text{avec } \mathbf{B} &\triangleq \mathbf{A}_{\mathbf{R}_{h_m}} \mathbf{K}_{c(m+1)} \mathbf{C} \\ \text{et } \mathbf{B}^\circ &\triangleq \mathbf{A}_{\mathbf{R}_{h_m}} \mathbf{C} \end{aligned} \tag{7.6}$$

$$\tag{7.7}$$

On déduit

Proposition 7.2 Biais

$$\begin{aligned} \mathbb{E}(\hat{\mathbf{h}}_m) - \mathbf{h}_m &= \mathbf{B} \mathbf{d}_{m_1, m_2} + \mathbf{B}^\circ \mathbf{d}_{m_1, m_2}^* + \mathcal{O}\left(\|\mathbf{d}_{m_1, m_2}\|^2\right) + \mathcal{O}\left(\frac{1}{N}\right) \\ \lim_{N \rightarrow \infty} \left(\mathbb{E}(\hat{\mathbf{h}}_m) - \mathbf{h}_m \right) &= \mathbf{B} \mathbf{d}_{m_1, m_2} + \mathbf{B}^\circ \mathbf{d}_{m_1, m_2}^* + \mathcal{O}\left(\|\mathbf{d}_{m_1, m_2}\|^2\right) \end{aligned}$$

La non modélisation des termes de la tête/queue (regroupés dans le vecteur $\mathbf{d}_{m_1, m_2}$) introduit un biais d'estimation qui est asymptotiquement (presque⁸) proportionnel⁹ à $\mathbf{d}_{m_1, m_2}$. Ce dernier peut être majoré par

Proposition 7.3 Majoration du biais asymptotique

$$\lim_{N \rightarrow \infty} \left\| \mathbb{E}(\hat{\mathbf{h}}_m) - \mathbf{h}_m \right\| \leq (\sigma_1(\mathbf{B}) + \sigma_1(\mathbf{B}^\circ)) \left\| \mathbf{d}_{m_1, m_2} \right\| + \mathcal{O}\left(\|\mathbf{d}_{m_1, m_2}\|^2\right)$$

Cette borne est atteinte si l'une des matrices $\mathbf{B}$ ou $\mathbf{B}^\circ$ est nulle. Le vecteur $\mathbf{d}_{m_1, m_2}$ qui occasionne le biais le plus important (en norme) est alors colinéaire au vecteur singulier à droite de celle des matrices $\mathbf{B}$ ou $\mathbf{B}^\circ$ qui est non nulle. Dans ce cas, $\sigma_1(\mathbf{B})$ ou $\sigma_1(\mathbf{B}^\circ)$ peut être interprétée comme une *mesure de la sensibilité* de l'algorithme en question à l'existence d'une tête/queue de termes non modélisés et ce pour une partie effective donnée. En particulier, le cas $\mathbf{B}$ nulle est celui de l'algorithme *SS* (§B.2) et aussi celui du cas réel ($\mathbf{B}^\circ$ est nulle par définition).

Proposition 7.4 Erreur asymptotique

$$\sqrt{N} \left(\hat{\mathbf{h}}_m - \mathbb{E}(\hat{\mathbf{h}}_m) \right) \xrightarrow{L} \mathcal{N} \left(\mathbf{0}_{c(m+1), 1}, \mathbf{A}_R \mathcal{C}_R \mathbf{A}_R^H, \mathbf{A}_R \mathcal{C}_R \mathbf{K}_{c(m+1)} \mathbf{A}_R^H \right)$$

⁸à $\mathcal{O}\left(\|\mathbf{d}_{m_1, m_2}\|^2\right)$ près.

⁹au sens où le biais est fonction bilinéaire de $\mathbf{d}_{m_1, m_2}$ et $\mathbf{d}_{m_1, m_2}^*$.

7.4.2 Modèle probabiliste

Les termes de la tête et queue sont nettement moins stables que ceux de la partie effective. La non stationnarité tend à affecter davantage les termes faibles que les termes significatifs. On les modélise ici par des variable aléatoires centrées et i.i.d.. On suppose qu'on a une idée sur leur nombre $(M - m)$, positions $(m_1$ et $m_2)$ et variance $(\frac{E(\|\mathbf{d}_{m_1, m_2}\|^2)}{M - m})$.

$\hat{\mathbf{R}}$ peut être considérée comme une perturbation de la situation idéale où le canal ne présente pas de tête/queue faible et les statistiques du second degré de sa sortie sont exactement connues. On montre (cf §A.3) le résultat suivant :

$$\hat{\mathbf{R}} = \mathbf{R}_{h_m} + \delta \mathbf{R}_{h_m}^N + \delta \mathbf{R}^d + O\left(E\left(\|\mathbf{d}_{m_1, m_2}\|^2\right)\right) + O\left(\frac{1}{N}\right) + O\left(\frac{E\left(\|\mathbf{d}_{m_1, m_2}\|^2\right)}{N}\right) \quad (7.8)$$

La matrice $\delta \mathbf{R}_{h_m}^N$ (définie dans la démonstration A.3) traduit l'erreur d'estimation, due à un nombre fini d'observations, de la matrice de corrélation du canal fictif $\mathbf{h}_m$. On lui associe (Prop. 4.1) la matrice $\mathcal{C}_{\mathbf{R}_{h_m}}$ telle que [HANN 70] $\sqrt{N} \delta \mathbf{R}_{h_m}^N \xrightarrow{L} \mathcal{N}(\mathbf{0}_{c(m+1), 1}, \mathcal{C}_{\mathbf{R}_{h_m}}, \mathcal{C}_{\mathbf{R}_{h_m}} \mathbf{K}_{c(m+1)})$.

$$\begin{aligned} \hat{\mathbf{h}}_m &= \text{alg}(\hat{\mathbf{R}}) \\ &= \text{alg}\left(\mathbf{R}_{h_m} + \left[\delta \mathbf{R}_{h_m}^N + \delta \mathbf{R}^d + O\left(E\left(\|\mathbf{d}_{m_1, m_2}\|^2\right)\right) + O\left(\frac{1}{N}\right) + O\left(\frac{E\left(\|\mathbf{d}_{m_1, m_2}\|^2\right)}{N}\right)\right]\right) \\ &= \mathbf{h}_m + \mathbf{A}_{\mathbf{R}_{h_m}} \text{Vec}(\delta \mathbf{R}_{h_m}^N) + \mathbf{A}_{\mathbf{R}_{h_m}} \text{Vec}(\delta \mathbf{R}^d) \\ &\quad + O\left(\frac{E\left(\|\mathbf{d}_{m_1, m_2}\|^2\right)}{\sqrt{N}}\right) + O\left(\frac{1}{N\sqrt{N}}\right) + O\left(\frac{E\left(\|\mathbf{d}_{m_1, m_2}\|^2\right)}{N\sqrt{N}}\right) \\ &\quad + O\left(\left\|\delta \mathbf{R}_{h_m}^N + \delta \mathbf{R}^d + O\left(E\left(\|\mathbf{d}_{m_1, m_2}\|^2\right)\right) + O\left(\frac{1}{N}\right)\right\|^2\right) \end{aligned}$$

$O\left(\left\|\delta \mathbf{R}_{h_m}^N\right\|^2\right) = O\left(\frac{1}{N}\right)$ et $O\left(\left\|\delta \mathbf{R}^d\right\|^2\right) = O\left(E\left(\|\mathbf{d}_{m_1, m_2}\|^2\right)\right)$. En rappelant (7.6) et (7.7), on obtient

Proposition 7.5

$$\begin{aligned} \hat{\mathbf{h}}_m &= \mathbf{h}_m + \mathbf{A}_{\mathbf{R}_{h_m}} \text{Vec}(\delta \mathbf{R}_{h_m}^N) + \mathbf{B} \mathbf{d}_{m_1, m_2} + \mathbf{B}^\circ \mathbf{d}_{m_1, m_2}^* \\ &\quad + O\left(E\left(\|\mathbf{d}_{m_1, m_2}\|^2\right)\right) + O\left(\frac{1}{N}\right) + O_1\left(\frac{1}{N}, E\left(\|\mathbf{d}_{m_1, m_2}\|\right)\right) \\ \text{avec} \\ O_1\left(\frac{1}{N}, E\left(\|\mathbf{d}_{m_1, m_2}\|\right)\right) &\triangleq O\left(\frac{E\left(\|\mathbf{d}_{m_1, m_2}\|\right)}{\sqrt{N}}\right) + O\left(\frac{E\left(\|\mathbf{d}_{m_1, m_2}\|\right)}{N}\right) \\ &\quad + O\left(\frac{E\left(\|\mathbf{d}_{m_1, m_2}\|^2\right)}{\sqrt{N}}\right) + O\left(\frac{E\left(\|\mathbf{d}_{m_1, m_2}\|^2\right)}{N}\right) \end{aligned}$$

$$\begin{aligned}
& +O\left(\frac{E\left(\left\|\mathbf{d}_{m_1,m_2}\right\|^2\right)}{N\sqrt{N}}\right) + O\left(\frac{E\left(\left\|\mathbf{d}_{m_1,m_2}\right\|^2\right)}{N^2}\right) \\
& +O\left(\frac{E\left(\left\|\mathbf{d}_{m_1,m_2}\right\|^3\right)}{N}\right) + O\left(\frac{E\left(\left\|\mathbf{d}_{m_1,m_2}\right\|^4\right)}{N}\right) + O\left(\frac{E\left(\left\|\mathbf{d}_{m_1,m_2}\right\|^4\right)}{N^2}\right)
\end{aligned}$$

Proposition 7.6 Erreur d'estimation

$$\begin{aligned}
E\left(\hat{\mathbf{h}}_m\right) - \mathbf{h}_m &= O\left(E\left(\left\|\mathbf{d}_{m_1,m_2}\right\|\right)^2\right) + O\left(\frac{1}{N}\right) + O_1\left(\frac{1}{N}, E\left(\left\|\mathbf{d}_{m_1,m_2}\right\|\right)\right) \\
\lim_{N \rightarrow \infty} \left(E\left(\hat{\mathbf{h}}_m\right) - \mathbf{h}_m\right) &= O\left(E\left(\left\|\mathbf{d}_{m_1,m_2}\right\|\right)^2\right)
\end{aligned}$$

Ainsi, l'estimateur de $\mathbf{h}_m$ est asymptotiquement non biaisé lorsque $E\left(\left\|\mathbf{d}_{m_1,m_2}\right\|\right) \rightarrow 0$. En effet, le moyennage sur les différentes réalisations de $\mathbf{d}_{m_1,m_2}$ fait disparaître le biais qu'elles introduisent (voir modèle déterministe). L'effet de la modélisation effective passe dans l'EQM via la covariance (voire Prop. 7.7) de l'estimateur et ce en terme de $\frac{E\left(\left\|\mathbf{d}_{m_1,m_2}\right\|^2\right)}{c(M-m)}$ qui exprime la variance *individuelle* d'un terme non modélisé de la tête/queue. On définit par rapport à $\mathbf{h}_m$ la covariance asymptotique $\text{Cov}\left(\hat{\mathbf{h}}_m\right) \triangleq E\left(\left[\hat{\mathbf{h}}_m - \mathbf{h}_m\right]\left[\hat{\mathbf{h}}_m - \mathbf{h}_m\right]^H\right)$.

Hypothèse H 7.3 Les termes de $\mathbf{d}_{m_1,m_2}$ sont i.i.d centrés, de même variance, indépendants entre eux, indépendants de $\delta\mathbf{R}_{h_m}^N$ et circulaires si complexes (s'ils sont réels, tenir compte de § 7.1).

Proposition 7.7 Covariance asymptotique

Sous H7.3

$$\begin{aligned}
\text{Cov}\left(\hat{\mathbf{h}}_m\right) &= \frac{1}{N}\mathbf{A}_{\mathbf{R}_{h_m}}\mathbf{C}_{\mathbf{R}_{h_m}}\mathbf{A}_{\mathbf{R}_{h_m}}^H + \frac{E\left(\left\|\mathbf{d}_{m_1,m_2}\right\|^2\right)}{c(M-m)}\mathbf{B}\mathbf{B}^H + \frac{E\left(\left\|\mathbf{d}_{m_1,m_2}\right\|^2\right)}{c(M-m)}\mathbf{B}^\circ\mathbf{B}^{\circ H} \\
&+ O\left(E\left(\left\|\mathbf{d}_{m_1,m_2}\right\|\right)^3\right) + O\left(\frac{1}{N}\right) + O_2\left(\frac{1}{N}, E\left(\left\|\mathbf{d}_{m_1,m_2}\right\|\right)\right)
\end{aligned}$$

avec

$$\begin{aligned}
O_2\left(\frac{1}{N}, E\left(\left\|\mathbf{d}_{m_1,m_2}\right\|\right)\right) &\triangleq O\left(\frac{E\left(\left\|\mathbf{d}_{m_1,m_2}\right\|\right)}{\sqrt{N}}\right) + O\left(\frac{E\left(\left\|\mathbf{d}_{m_1,m_2}\right\|\right)}{N}\right) \\
&+ O\left(\frac{E\left(\left\|\mathbf{d}_{m_1,m_2}\right\|^2\right)}{\sqrt{N}}\right) + O\left(\frac{E\left(\left\|\mathbf{d}_{m_1,m_2}\right\|^2\right)}{N}\right) \\
&+ O_1\left(\frac{1}{N}, E\left(\left\|\mathbf{d}_{m_1,m_2}\right\|\right)\right) O\left(E\left(\left\|\mathbf{d}_{m_1,m_2}\right\|\right)\right) \\
&+ O_1\left(\frac{1}{N}, E\left(\left\|\mathbf{d}_{m_1,m_2}\right\|\right)\right) O\left(E\left(\left\|\mathbf{d}_{m_1,m_2}\right\|^2\right)\right) \\
&+ O_1\left(\frac{1}{N}, E\left(\left\|\mathbf{d}_{m_1,m_2}\right\|\right)\right) O\left(\frac{1}{\sqrt{N}}\right) + O_1\left(\frac{1}{N}, E\left(\left\|\mathbf{d}_{m_1,m_2}\right\|\right)\right) O\left(\frac{1}{N}\right) \\
&+ O_1\left(\frac{1}{N}, E\left(\left\|\mathbf{d}_{m_1,m_2}\right\|\right)\right) O_1\left(\frac{1}{N}, E\left(\left\|\mathbf{d}_{m_1,m_2}\right\|\right)\right)
\end{aligned}$$

L'EQM induite par la non modélisation des termes faibles de la tête/queue n'est pas fonction du nombre de ceux-ci mais de leur variance individuelle.

Les deux sources d'erreur ($\delta \mathbf{R}^d$: modélisation effective de l'ordre et $\delta \mathbf{R}_{h_m}^N$: nombre fini d'observations) peuvent être étudiées d'une manière similaire. On se ramène à deux cas : *Statistiques exactes* et *Modélisation totale* avec des résultats similaires énoncés sous forme de théorèmes "central limite".

Proposition 7.8 Modélisation totale ($\mathbf{d}_{m_1, m_2} = \mathbf{0}_{M-m, 1}$)

Lorsque $N \rightarrow \infty$

$$\sqrt{N} (\hat{\mathbf{h}}_m - \mathbf{h}_m) \xrightarrow{L} \mathcal{N} \left(\mathbf{0}_{c(m+1), 1}, \mathbf{A}_{\mathbf{R}_{h_m}} \mathbf{C}_{\mathbf{R}_{h_m}} \mathbf{A}_{\mathbf{R}_{h_m}}^H, \mathbf{A}_{\mathbf{R}_{h_m}} \mathbf{C}_{\mathbf{R}_{h_m}} \mathbf{K}_{c(m+1)} \mathbf{A}_{\mathbf{R}_{h_m}}^H \right)$$

Proposition 7.9 Statistiques exactes ($N \rightarrow \infty$) - **Cas complexe**

Si $\mathbf{d}_{m_1, m_2}$ est un vecteur complexe, centré, gaussien et circulaire¹⁰ et lorsque $E \left(\left\| \mathbf{d}_{m_1, m_2} \right\|^2 \right) \rightarrow 0$

$$\sqrt{\frac{c(M-m)}{E \left(\left\| \mathbf{d}_{m_1, m_2} \right\|^2 \right)}} (\hat{\mathbf{h}}_m - \mathbf{h}_m) \xrightarrow{L} \mathcal{N} \left(\mathbf{0}_{c(m+1), 1}, \mathbf{B} \mathbf{B}^H + \mathbf{B}^\circ \mathbf{B}^{\circ H}, \mathbf{B} \mathbf{B}^{\circ T} + \mathbf{B}^\circ \mathbf{B}^T \right)$$

Remarque 7.1 Le cas réel Quand $\mathbf{d}_{m_1, m_2}$ est réel (H7.2), la matrice $\mathbf{B}^\circ$ n'est pas définie. Les résultats de l'analyse (modèles déterministe et probabiliste) restent valides en redéfinissant $\mathbf{B} \triangleq \mathbf{A}_{\mathbf{R}_m} \mathbf{C}^T$ et $\mathbf{B}^\circ \triangleq \mathbf{0}_{c(m+1), c(M-m)}$.

7.5 Égalisation en modélisation effective

Dans ce paragraphe, on s'intéresse à d'autres paramètres dérivés de la partie effective. Il s'agit des égaliseurs ZF, de la réponse combinée canal/égaliseur et de l'ouverture moyenne de l'oeil. On étendra les résultats (sur le biais et l'EQM) à ces paramètres et ce dans chacun des modèles déterministe et probabiliste et en supposant un nombre de sous-canaux $c = 2$.

$$\begin{array}{rcl}
 \hat{\mathbf{R}} & \text{matrice de corrélation empirique} & \\
 \downarrow & & \\
 \hat{\mathbf{h}}_m = \text{alg}(\hat{\mathbf{R}}) & \text{estimée de la partie effective} & \\
 \downarrow & & \\
 \hat{\mathbf{g}}_{m-1, i} = \mathcal{T}_m^{-T} (\hat{\mathbf{h}}_m) \mathbf{e}_{2m, i} & \text{égaliseur ZF d'ordre minimal} & \\
 \downarrow & & \\
 \mathbf{r}_{M+m-1, i} = \mathcal{T}_m^T (\mathbf{h}_M) \hat{\mathbf{g}}_{m-1, i} & \text{réponse combinée canal/égaliseur} & \\
 \downarrow & & \\
 \text{OEM}_{m_1+i} (\mathbf{r}_{M+m-1, i}) = \frac{\sum_{k \neq m_1+i} |r(k)|^2}{|r(m_1+i)|^2} & \text{ouverture moyenne de l'oeil} &
 \end{array}$$

¹⁰En réel (H7.2), il suffit à $\mathbf{d}_{m_1, m_2}$ d'être centré et gaussien.

7.5.1 Égalisation ZF

Comme décrit au §4.2.1, on peut égaliser un canal $\mathbf{h}_m$ par les $2m$ filtres $\mathbf{g}_{m-1,i} = \mathcal{T}_m^{-T}(\mathbf{h}_m) \mathbf{e}_{2m,i}$, avec des délais $i-1, i=1, \dots, 2m$. À partir d'une estimée de $\mathbf{h}_m$, on construit la famille de filtres suivants :

$$\begin{aligned}
\hat{\mathbf{g}}_{m-1,i} &= \mathcal{T}_m^{-T}(\hat{\mathbf{h}}_m) \mathbf{e}_{2m,i} \\
&= \left[\mathcal{T}_m^T(\mathbf{h}_m) + \mathcal{T}_m^T(\hat{\mathbf{h}}_m - \mathbf{h}_m) \right]^{-1} \mathbf{e}_{2m,i} \\
&= \left[\mathcal{T}_m^{-T}(\mathbf{h}_m) - \mathcal{T}_m^{-T}(\mathbf{h}_m) \mathcal{T}_m^T(\hat{\mathbf{h}}_m - \mathbf{h}_m) \mathcal{T}_m^{-T}(\mathbf{h}_m) + O\left(\left\| \mathcal{T}_m^T(\hat{\mathbf{h}}_m - \mathbf{h}_m) \right\|^2\right) \right] \mathbf{e}_{2m,i} \\
&= \mathbf{g}_{m-1,i} - \mathcal{T}_m^{-T}(\mathbf{h}_m) \mathcal{T}_m^T(\hat{\mathbf{h}}_m - \mathbf{h}_m) \mathbf{g}_{m-1,i} + O\left(\left\| \hat{\mathbf{h}}_m - \mathbf{h}_m \right\|^2\right) \\
&= \mathbf{g}_{m-1,i} - \mathcal{T}_m^{-T}(\mathbf{h}_m) \mathcal{T}^T(\mathbf{g}_{m-1,i}) (\hat{\mathbf{h}}_m - \mathbf{h}_m) + O\left(\left\| \hat{\mathbf{h}}_m - \mathbf{h}_m \right\|^2\right)
\end{aligned}$$

$\hat{\mathbf{g}}_{m-1,i}$ est alors une estimation de $\mathbf{g}_{m-1,i}$ avec l'erreur :

$$\hat{\mathbf{g}}_{m-1,i} - \mathbf{g}_{m-1,i} = \mathcal{T}_m^{-T}(\mathbf{h}_m) \mathcal{T}^T(\mathbf{g}_{m-1,i}) (\hat{\mathbf{h}}_m - \mathbf{h}_m) + O\left(\left\| \hat{\mathbf{h}}_m - \mathbf{h}_m \right\|^2\right)$$

7.5.2 Réponse globale

La réponse combinée canal/égaliseur est donnée par $\mathbf{r}_{M+m-1,i} = \mathcal{T}_m^T(\mathbf{h}_M) \hat{\mathbf{g}}_{m-1,i}$. On montre (cf. §A.4) qu'elle approxime un retard pur de $m_1 + i - 1$, avec

$$\begin{aligned}
\mathbf{r}_{M+m-1,i} - \mathbf{e}_{M+m,m_1+i} &= \mathcal{T}_{M+1}^T(\mathbf{g}_{m-1,i}) \mathbf{d}_{m_1,m_2}^z - \mathcal{T}^T(\mathbf{g}_{m-1,i}^z) \\
&\quad + O\left(\left\| \hat{\mathbf{h}}_m - \mathbf{h}_m \right\|^2\right) + O\left((\hat{\mathbf{h}}_m - \mathbf{h}_m) \mathbf{d}_{m_1,m_2}\right)
\end{aligned} \tag{7.9}$$

où $\mathbf{g}_{m-1,i}^z$ est une version *zero-padded* de $\mathbf{g}_{m-1,i}$ i.e., augmentée de m_1 termes nuls à la tête et m_2 termes nuls à la queue¹¹.

7.5.3 Ouverture de l'oeil

On se place dans le cas réel (H7.2). La i^{eme} réponse globale $\mathbf{r}_{M+m-1,i}$ présente un retard de $m_1 + i - 1$ par rapport à la réponse idéale.

L'ouverture de l'oeil est alors mesurée par $\text{OEM}_{m_1+i}(\mathbf{r}_{M+m-1,i}) = \frac{\sum_{k \neq m_1+i} r(k)^2}{r(m_1+i)^2}$. On montre (cf. §A.5) qu'elle est donnée par

$$\text{OEM}_{m_1+i} = \mathbf{r}_{M+m-1,i}^{\dagger T} \mathbf{r}_{M+m-1,i}^{\dagger} + O\left(\left\| \mathbf{r}_{M+m-1,i} - \mathbf{e}_{M+m,m_1+i} \right\|^3\right) \tag{7.10}$$

avec $\mathbf{r}_{M+m-1,i}^{\dagger} = [r(1) \cdots r(m_1 + i - 1), 0, r(m_1 + i + 1) \cdots r(M + m)]^T = \ddot{\mathbf{I}}_i (\mathbf{r}_{M+m-1,i} - \mathbf{e}_{M+m,m_1+i})$ où $\ddot{\mathbf{I}}_i \triangleq \mathbf{I}_{M+m} - \mathbf{e}_{M+m,m_1+i} \mathbf{e}_{M+m,m_1+i}^T$.

¹¹ $\mathbf{g}_{m-1,i}^z = [\mathbf{0}_{1,2m_1}, \mathbf{g}_{m-1,i}^T, \mathbf{0}_{1,2m_2}]^T$, formal *ST*.

Une réponse globale différente d'un retard pur se traduit par une ouverture moyenne de l'oeil égale à

$$\begin{aligned}
E(\text{OEM}_{m_1+i}) &= E\left((\mathbf{r}_{M+m-1,i} - \mathbf{e}_{M+m,m_1+i})^T \ddot{\mathbf{I}}_i (\mathbf{r}_{M+m-1,i} - \mathbf{e}_{M+m,m_1+i})\right) \\
&= \text{Tr}\left(E\left((\mathbf{r}_{M+m-1,i} - \mathbf{e}_{M+m,m_1+i})^T \ddot{\mathbf{I}}_i (\mathbf{r}_{M+m-1,i} - \mathbf{e}_{M+m,m_1+i})\right)\right) \\
&= E\left(\text{Tr}\left((\mathbf{r}_{M+m-1,i} - \mathbf{e}_{M+m,m_1+i})(\mathbf{r}_{M+m-1,i} - \mathbf{e}_{M+m,m_1+i})^T \ddot{\mathbf{I}}_i\right)\right) \\
&= \text{Tr}\left(E\left((\mathbf{r}_{M+m-1,i} - \mathbf{e}_{M+m,m_1+i})(\mathbf{r}_{M+m-1,i} - \mathbf{e}_{M+m,m_1+i})^T\right) \ddot{\mathbf{I}}_i\right)
\end{aligned}$$

7.6 Application aux algorithmes

Dans ce paragraphe, on applique l'étude de performance décrite dans §7.4 à trois principaux algorithmes du second ordre d'identification aveugle : l'algorithme sous-espace (SS), l'algorithme prédiction linéaire (LP) et l'algorithme décomposition en produit extérieur (OPD). L'application aux algorithmes se fera sous les hypothèses supplémentaires suivantes : bruit temporellement et spatialement décorrélé (H4.3) et canal SIMO à deux sous-canaux (H7.1).

Les matrices $\mathbf{A}_{\mathbf{R}}$ et $\mathbf{A}_{\mathbf{R}_{h_m}}$ sont les inconnues desquelles se déduisent toutes les expressions et résultats de l'analyse. Dans une première étape (§7.6.1, §7.6.2 et §7.6.3), on calcule ces matrices pour chacun des algorithmes. Ensuite, on exprimera les performances en terme de biais et d'EQM et comparera les valeurs théoriques développées dans §7.4 et celles obtenues par simulations au §7.6.4.

7.6.1 L'algorithme sous-espace / moindres carrés

D'après Prop. B.2, on a

$$\begin{aligned}
\mathbf{A}_{\mathbf{R}} &= -\mathbf{T}_{m+1} \left(\left(\mathbf{R}^T - \lambda(\mathbf{R}) \mathbf{I}_{2(m+1)} \right)^{\natural} \otimes \text{vecp}(\mathbf{R})^H \right) \\
\mathbf{A}_{\mathbf{R}_{h_m}} &= -\mathbf{T}_{m+1} \left(\left[\mathcal{T}^*(\mathbf{h}_m) \mathcal{T}^T(\mathbf{h}_m) \right]^{\natural} \otimes [\mathbf{h}_m^T \mathbf{T}_{m+1}] \right), \text{ en utilisant (4.10)}
\end{aligned}$$

Le comportement de l'algorithme en statistiques approchées est entièrement décrit par la matrice $\mathbf{A}_{\mathbf{R}}$ ou aussi $\mathbf{B}$. On a pu montré, dans le cas de l'algorithme SS, les propriétés suivantes, qui ont pu être constatées par calcul numérique pour les algorithmes LP et OPD.

Les performances de l'algorithme (en terme d'EQM) ne dépendent pas des statistiques d'ordres supérieurs de l'entrée. En effet, celles-ci sont décrites par le terme $\mathbf{A}_{\mathbf{R}} \mathcal{C}_{\mathbf{R}} \mathbf{A}_{\mathbf{R}}^H$ pour le modèle déterministe (cf. Prop. 7.1 et Prop. 7.4). Pour le modèle probabiliste, il s'agit du même terme, en remplaçant $\mathbf{R}$ par $\mathbf{R}_{h_m}$ (cf. Prop. 7.7). Dans l'expression de $\mathcal{C}_{\mathbf{R}}$ (cf. Prop. 4.1), seul le terme (4.11), soit $\text{cum}(s_0, s_0, s_0^*, s_0^*) \text{Vec}(\mathbf{R}) \text{Vec}(\mathbf{R})^H$, dépend des statistiques d'ordres supérieurs de l'entrée. On montre (cf. §A.6) que¹²

$$\mathbf{A}_{\mathbf{R}} \text{Vec}(\mathbf{R}) = \mathbf{0}_{2(m+1),1} \quad (7.11)$$

Ceci justifie l'affirmation ci-dessus.

¹²Ceci est valable pour le cas particulier $\mathbf{R} = \mathbf{R}_{h_m}$.

Également, on montre (cf. §A.7) que l'expression de $\mathbf{B}$ se simplifie comme suit :

$$\mathbf{B} = \mathbf{T}_{m+1} \left[\mathcal{T}^* (\mathbf{h}_m) \mathcal{T}^T (\mathbf{h}_m) \right]^{\mathfrak{h}} \mathcal{T} (\mathbf{h}_m^{z*}) \mathcal{T}_M^T (\mathbf{T}_{m+1} \mathbf{h}_m) \mathbf{I}^z \quad (7.12)$$

Il s'en suit les résultats de Prop. 7.10 et Prop. 7.11 (démontrés dans §A.8 et §A.9 respectivement).

Proposition 7.10

$$\mathbf{B}\mathbf{B}^H = \Phi(\mathbf{h}_m, m_1, m_2)$$

où $\Phi(\mathbf{h}_m, m_1, m_2)$ est une matrice vérifiant :

$$\Phi(\mathbf{h}_m, m_1, m_2) = \Phi(\mathbf{h}_m, m, m_2) \quad \forall m_1 \geq m \quad (7.13)$$

$$= \Phi(\mathbf{h}_m, m_1, m) \quad \forall m_2 \geq m \quad (7.14)$$

$$= \Phi(\mathbf{h}_m, m, m) = Cst \quad \forall m_1, m_2 \geq m \quad (7.15)$$

Proposition 7.11 Si $m_1 \geq m$ et $m_2 \geq m$, alors $\sigma_1(\mathbf{B})$ n'est pas fonction ni de m_1 ni de m_2 et le

vecteur singulier normé associé est de la forme¹³ $\begin{bmatrix} \mathbf{0}_{2(m_1-m),1} \\ \mathbf{u}_m \\ \mathbf{v}_m \\ \mathbf{0}_{2(m_2-m),1} \end{bmatrix}$ avec $\mathbf{u}_m, \mathbf{v}_m \in \mathbf{C}^{2m}$ ne sont fonction ni de m_1 ni de m_2 .

Prop. 7.11 renseigne sur le “cas le plus défavorable”, celui où le biais asymptotique occasionné par une tête/queue $\mathbf{d}_{m_1, m_2}$, de norme donnée et fixe, est maximal. Ce biais, qui vaut alors $\|\mathbf{d}_{m_1, m_2}\| \sigma_1(\mathbf{B})$, est atteint lorsque $\mathbf{d}_{m_1, m_2}$ est dans la direction du vecteur singulier à droite de $\mathbf{B}$ associé à sa plus grande valeur singulière. Dans le cas particulier $m_1 \geq m$ et $m_2 \geq m$, qui est pratiquement toujours le cas, on vérifie

- Pour une norme fixée de $\mathbf{d}_{m_1, m_2}$, cette borne n'est pas fonction ni de m_1 ni de m_2 .
- La tête/queue $\mathbf{d}_{m_1, m_2}$ qui occasionne le biais le plus important (en norme) est de la forme¹⁴ $\mathbf{d}_{m_1, m_2} = \|\mathbf{d}_{m_1, m_2}\| [\mathbf{0}_{1, 2(m_1-m)} \mathbf{u}_1^T \cdots \mathbf{u}_m^T \mathbf{v}_1^T \cdots \mathbf{v}_m^T \mathbf{0}_{1, 2(m_2-m)}]^T$. Les $4m$ termes de $\mathbf{u}$ et $\mathbf{v}$ ne sont fonction ni de m_1 ni de m_2 .

On peut déduire que les performances asymptotique en termes de biais (modèle déterministe) ou d'EQM (modèle probabiliste) dans le cas le plus défavorable ne se dégradent pas davantage lorsque le nombre de termes non modélisés augmente tant que leur nombre (à la fois à la tête et à la queue) reste supérieur à l'ordre effectif, i.e., presque toujours. Ceci sous la condition que reste fixée la norme de $\mathbf{d}_{m_1, m_2}$ (modèle déterministe) ou la variance individuelle des termes non modélisés (modèle probabiliste). Pour le modèle déterministe, le cas le plus défavorable est celui des mêmes $2m$ termes ($\forall m_1$ et m_2) concentrés autour de la réponse effective au nombre de m termes à droite et m termes à gauche.

¹³Format ST

¹⁴Format ST . Ou bien aussi en format TS

$$\mathbf{d}_{m_1, m_2} = \|\mathbf{d}_{m_1, m_2}\| \begin{bmatrix} \mathbf{0}_{1, m_1-m} & u_1^1 \cdots u_m^1 v_1^1 \cdots v_m^1 & \mathbf{0}_{1, m_2-m} \\ \mathbf{0}_{1, m_1-m} & u_1^2 \cdots u_m^2 v_1^2 \cdots v_m^2 & \mathbf{0}_{1, m_2-m} \end{bmatrix}^T$$

7.6.2 L'algorithme prédiction linéaire

On déduit de Prop. B.5 et Prop. B.6 :

$$\begin{aligned}\mathbf{A}_{\mathbf{R}} &= \left(\Delta_6 - \left[\left(\mathbf{g}^H \right) \otimes \left(\mathbf{Y}^\circ \begin{bmatrix} \mathbf{0}_c & \mathbf{0}_{c,cm} \\ \mathbf{0}_{cm,c} & \mathbf{Z}^{-1} \end{bmatrix} \right) \right] \right) \overline{\mathcal{G}}(\mathbf{R}) \\ \mathbf{A}_{\mathbf{R}_{h_m}} &= \left\{ \left(\mathbf{g}^H \otimes \mathbf{I}_{2(m+1)} \right) \overline{\mathcal{S}}_{2,m} - \frac{1}{2} \left[\left(\mathbf{g}^H \right) \otimes \left(\mathbf{h}_m \mathbf{g}^T \right) \right] \right. \\ &\quad \left. - \sigma_s^2 \left[\left(\mathbf{g}^H \right) \otimes \left(\mathcal{T}^\circ(\mathbf{h}_m) \mathcal{T}^H(\mathbf{h}_m) \begin{bmatrix} \mathbf{0}_2 & \mathbf{0}_{2,2m} \\ \mathbf{0}_{2m,2} & \mathbf{Z}^{-1} \end{bmatrix} \right) \right] \right\} \overline{\mathcal{G}}(\mathbf{R}_{h_m})\end{aligned}$$

Notons que dans les expressions de $\mathbf{A}_{\mathbf{R}}$ et $\mathbf{A}_{\mathbf{R}_{h_m}}$, $\mathbf{Y}$, $\mathbf{Z}$, $\mathbf{g}$, ... sont déduits de $\mathbf{R}$ et $\mathbf{R}_{h_m}$ comme décrit dans §B.3.

7.6.3 L'algorithme décomposition en produit extérieur

On déduit de Prop. B.7 et Prop. B.8 :

$$\begin{aligned}\mathbf{A}_{\mathbf{R}} &= -\frac{1}{\sigma_s^2} \left([VECP(\mathbf{D}_1 - \mathbf{D}_2)]^T \otimes (\mathbf{D}_1 - \mathbf{D}_2 - \lambda \mathbf{I}_{2(m+1)})^{\natural} \right) \\ &\quad \left(\mathbf{I}_{4(m+1)^2} - ((\mathbf{J}_{m+1} \otimes \mathbf{I}_2) \otimes (\mathbf{J}_{m+1} \otimes \mathbf{I}_2)) \right) \\ &\quad (\Delta_2 + \Delta_3 \mathbf{K}_{2(m+1)}) \overline{\mathcal{G}}(\mathbf{R}) \\ \mathbf{A}_{\mathbf{R}_{h_m}} &= -\frac{1}{\sigma_s^2} \left(\mathbf{h}_m^T \otimes (\mathbf{h}_m \mathbf{h}_m^H - \mathbf{I}_{2(m+1)})^{\natural} \right) \\ &\quad \left(\mathbf{I}_{4(m+1)^2} - ((\mathbf{J}_{m+1} \otimes \mathbf{I}_2) \otimes (\mathbf{J}_{m+1} \otimes \mathbf{I}_2)) \right) \\ &\quad (\Delta_2 + \Delta_3 \mathbf{K}_{2(m+1)}) \overline{\mathcal{G}}(\mathbf{R}_{h_m})\end{aligned}$$

$\mathbf{Y}$, $\mathbf{Z}$, $\mathbf{g}$, ... sont déduites de $\mathbf{R}$ et $\mathbf{R}_{h_m}$ comme décrit dans §B.4.

7.6.4 Simulations

Paramètres des simulations

Les canaux simulés sont d'ordre total $M = 12$. Une partie effective d'ordre $m = 2$ est entourée d'une tête et d'une queue longues respectivement de $m_1 = 6$ et $m_2 = 4$ termes. Pour étudier l'effet de la diversité des parties effectives sur les performances des algorithmes, on s'est donné deux réponses¹⁵ $\mathbf{h}_m$ de module 1 :

$$\begin{aligned}\mathbf{h}_m^a &= [-0.6804 \quad 0.177 \quad -0.0902 \quad 0.4281 \quad -0.2446 \quad -0.5043]^T \\ \mathbf{h}_m^b &= [-0.6804 \quad 0.177 \quad -0.0902 \quad 0.5902 \quad -0.2556 \quad -0.2803]^T\end{aligned}$$

D'après Tab. 7.1, la réponse effective $\mathbf{h}_m^b$ ainsi choisie est moins bien diversifiée que $\mathbf{h}_m^a$. Pour les deux, néanmoins, la diversité est suffisante pour que l'algorithme [LIAV 99] soit capable de détecter

¹⁵Format *TS*

l'ordre effectif dans les pages de SNR considérées. On se donne le vecteur $\mathbf{d}_{m_1, m_2}$ suivant

$$\begin{aligned}\mathbf{d}_{m_1, m_2}^1 &= [0.0100 \quad 0.0029 \quad -0.0123 \quad 0.0167 \quad 0.0036 \quad 0.0184 \quad 0.0142 \quad -0.0085 \quad -0.0161 \quad -0.0009]^T \\ \mathbf{d}_{m_1, m_2}^2 &= [-0.0041 \quad -0.0105 \quad 0.0062 \quad 0.0185 \quad -0.0068 \\ &\quad -0.0105 \quad -0.0031 \quad 0.0082 \quad -0.0106 \quad -0.0149]^T \\ \|\mathbf{d}_{m_1, m_2}\| &= 0.05\end{aligned}$$

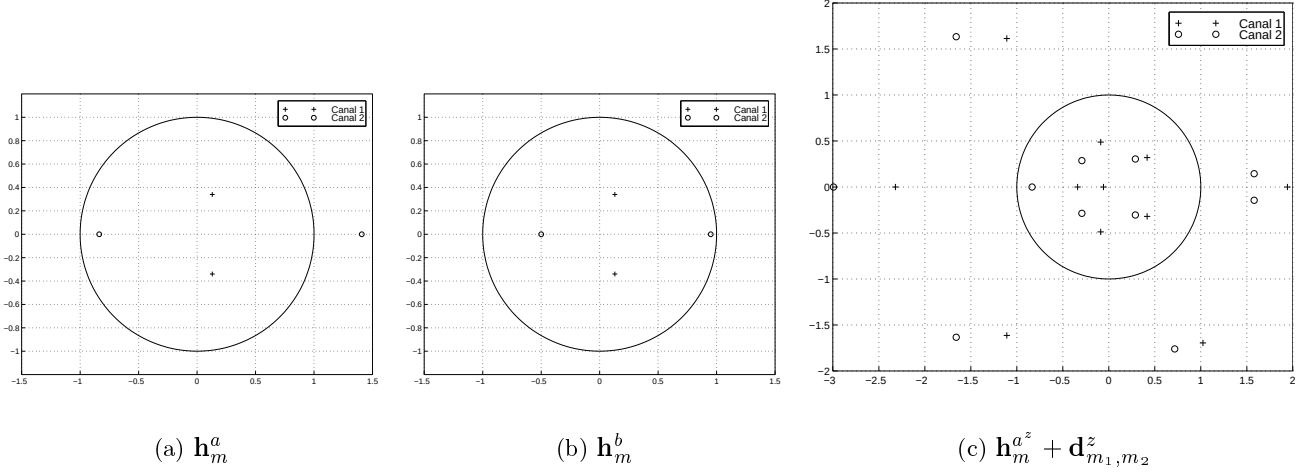


Figure 7.3: Réponses totale et effective : disposition des zeros

	$\mathbf{h}_m^a$	$\mathbf{h}_m^b$
$\sigma_{2m+1}(\mathcal{T}_m(\mathbf{h}_m))$	0.4157	0.2369
$\sigma_{2m}(\mathcal{T}_{m-1}(\mathbf{h}_m))$	0.4165	0.2354
$\sigma_{2m+1}\left(\mathcal{T}_m\left(\mathbf{h}_m + \mathbf{d}_{m_1, m_2}\right)\right)$	0.4195	0.2429
$\sigma_{2m}\left(\mathcal{T}_{m-1}\left(\mathbf{h}_m + \mathbf{d}_{m_1, m_2}\right)\right)$	0.4220	0.2424
$\sigma_{2M+1}\left(\mathcal{T}_M\left(\mathbf{h}_m + \mathbf{d}_{m_1, m_2}\right)\right)$	0.0003	0.0004
$\sigma_{2M}\left(\mathcal{T}_{M-1}\left(\mathbf{h}_m + \mathbf{d}_{m_1, m_2}\right)\right)$	0.0001	0.0001

Tableau 7.1: Diversité des réponses totale et effective

Remarque 7.2 Sur les figures, les courbes se rapportant à une réponse effective $\mathbf{h}_m^a$ (resp. $\mathbf{h}_m^b$) sont représentés en continu (resp. en interrompu).

Le canal SIMO est excité par une source binaire à symboles indépendants et équiprobables ($\sigma_s^2 = 1$). Le bruit est gaussien, spatialement et temporellement décorrélé. On retient, pour le rapport signal sur bruit, la même définition que celle donne dans §5.9, soit $SNR \triangleq \frac{\mathbb{E}(\|\mathbf{x}(n)\|^2)}{\mathbb{E}(\|\mathbf{b}(n)\|^2)} = \frac{\sigma_s^2 \|\mathbf{h}_M\|^2}{c \sigma_b^2} \simeq \frac{1}{2 \sigma_b^2}$.

L'erreur quadratique moyenne (EQM) est calculée par $\|\mathbf{h}_m - \hat{\mathbf{h}}_m\|^2$ moyennée sur l'ensemble des réalisations de Monte Carlo, au nombre de 1000.

Liavas et al ont réalisé une étude de performances des algorithmes sous-espace [LIAV 99] et prédiction linéaire [LIAV 00] en modélisation effective de l'ordre. Par rapport à l'étude de §7.4, ils ont fait les hypothèses supplémentaires suivantes

- La tête/queue $\mathbf{d}_{m_1, m_2}$ est déterministe.
- La matrice de corrélation des observations $\mathbf{R}$ est exactement connue.
- La sortie du canal est non bruitée.

L'estimée de la réponse effective dans ces conditions correspond à l'estimée asymptotique (lorsque le nombre d'observations tend vers l'infini), en effet :

Par continuité de alg on a $\lim_{N \rightarrow \infty} \text{alg}(\hat{\mathbf{R}}) = \text{alg}(\lim_{N \rightarrow \infty} \hat{\mathbf{R}})$.

Par ergodicité des observations, on a $\lim_{N \rightarrow \infty} \hat{\mathbf{R}} = \mathbf{R}$.

D'où $\lim_{N \rightarrow \infty} \hat{\mathbf{h}}_m = \text{alg}_m(\mathbf{R})$.

À chaque fois, *Liavas et al* ont donné une borne supérieure de l'erreur d'estimation :

$$\left\| \lim_{N \rightarrow \infty} \hat{\mathbf{h}}_m - \mathbf{h}_m \right\| \leq 2\sqrt{2(m+1)} \frac{\|\mathbf{d}_{m_1, m_2}\|}{\sigma_{2m+1}(\mathcal{T}_m(\mathbf{h}_m))} + \|\mathbf{d}_{m_1, m_2}\|^2$$

Sous-espace

(7.16)

$$\left\| \lim_{N \rightarrow \infty} \hat{\mathbf{h}}_m - \mathbf{h}_m \right\| \leq \sqrt{m+1} \frac{\|\mathbf{d}_{m_1, m_2}\|}{\sigma_{2m}(\mathcal{T}_{m-1}(\mathbf{h}_m))} + (2m+3) \frac{\|\mathbf{d}_{m_1, m_2}\|}{\|\mathbf{h}(p)\|} \sqrt{1 + \frac{1}{[\sigma_{2m}(\mathcal{T}_{m-1}(\mathbf{h}_m))]^2}}$$

Prédiction linéaire

(7.17)

$\|\mathbf{d}_{m_1, m_2}\|$ exprime l'importance des termes non modélisés alors que σ est interprétée comme une mesure de la diversité du canal. Le rapport $\frac{\|\mathbf{d}_{m_1, m_2}\|}{\sigma}$ traduit alors le compromis

une partie effective de bonne diversité vs une partie non modélisée de faible énergie
capacité de l'algorithme (à identifier $\mathbf{h}_m$) vs validité du modèle ($\mathbf{d}_{m_1, m_2}$ est un terme de perturbation)

Ces bornes sont à comparer aux résultats du modèle déterministe :

- Tab. 7.2, 7.3 et 7.4 font le récapitulatif des résultats des simulations, des valeurs asymptotiques et des bornes de *Liavas et al*.
- Fig. 7.4 compare l'EQM obtenu par simulation à l'EQM asymptotique exprimé lui par

$$\mathbb{E} \left(\|\hat{\mathbf{h}}_m - \mathbf{h}_m\|^2 \right) = \|\mathbf{h}_m - \mathbb{E}(\hat{\mathbf{h}}_m)\|^2 + \text{Cov}(\hat{\mathbf{h}}_m) \sim \|\mathbf{B}\mathbf{d}_{m_1, m_2}\|^2 + \frac{1}{N} \mathbf{A}_R \mathbf{C}_R \mathbf{A}_R^H$$

- Fig. 7.5 calcule le vecteur singulier à droite de $\mathbf{B}$ qui correspondent à la tête/queue la plus défavorable (les calculs ont été faits en réel (H7.2)).

Ils ressort de ces tableaux et figures les conclusions suivantes

- La validité des formules asymptotiques du biais et de l'EQM est confirmée sur une large plage de SNR et de nombre d'observations N .
- Pour les algorithmes LP et OPD, en comparant les tableaux (a) et (c), on constate l'amélioration (baisse de la norme) du biais. Elle cache une détérioration plus importante de la variance de l'estimateur. En définitif, l'EQM se trouve aggravée.
- Les bornes (7.16) et (7.17) de *Liavas et al* s'avèrent pessimistes comparées à celle de la Prop. 7.3 qui elle est atteinte (puisque les simulations sont faites en réel).
- Une mauvaise diversité du canal pénalise plus l'algorithme SS que les algorithmes LP et OPD. Une explication qualitative peut être trouvée dans l'expression de $\mathbf{A}$. Dans le cas des algorithmes LP et OPD, les matrices $\mathbf{A}$ font apparaître des produits de la la matrice de corrélation ¹⁶ et de sa pseudo-inverse. La baisse de diversité provoque une baisse des termes de la première compensée par une augmentation de ceux de la deuxième. D'où une matrice $\mathbf{A}$ qui est moins sensible à la diversité.
- Les résultats confirment la supériorité de l'algorithme SS par rapport aux algorithmes LP et OPD. Cette supériorité tend toutefois à diminuer quand l'identification se fait dans de mauvaises conditions (bruit important, diversité faible, la partie non modélisée $\mathbf{d}_{m_1, m_2}$ est importante ou nombre de symboles insuffisant).
- La tête/queue théorique qui (lorsque non modélisée) détériore le plus (en terme de biais) les performances de l'algorithme d'identification (cf. Fig. (7.5)) a toujours la forme prévue par Prop. 7.11. Cette forme ressort systématiquement pour d'autres choix de $\mathbf{h}_m$, m_1 et m_2 laissant penser que Prop. 7.11 peut être généralisée à LP et OPD également (dans le cas réel).

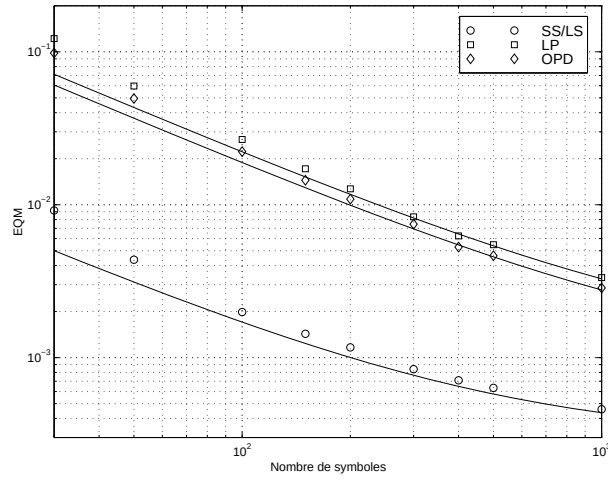
Modèle probabiliste

Les termes de $\mathbf{d}_{m_1, m_2}$ sont générés comme étant des variables aléatoires gaussiennes, indépendantes, centrées et de variance $\frac{E(\|\mathbf{d}_{m_1, m_2}\|^2)}{2(M-m)}$. On continue d'envisager les deux parties effectives $\mathbf{h}_m^a$ et $\mathbf{h}_m^b$.

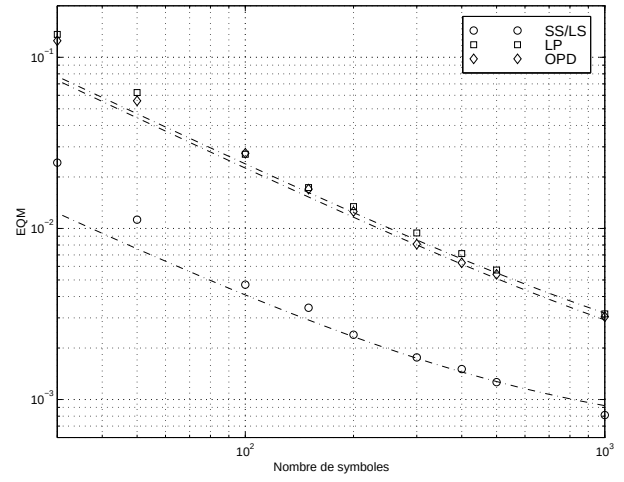
Les résultats obtenus (cf. Fig. 7.6, 7.7, 7.8 et 7.9) confirment la validité des expressions asymptotiques de l'EQM sur de larges plages de nombre de symboles N , rapport signal sur bruit SNR et variance des termes non modélisés $E(\|\mathbf{d}_{m_1, m_2}\|^2)$. Ils permettent d'étendre ces conclusions à l'égaliseur ZF d'ordre $m-1$, la réponse combinée canal-égaliseur et l'ouverture de l'oeil. Pour ces derniers, le délai choisi est le délai moyen (ici 2 durées symbole) qui habituellement [Touz 98] donne les meilleures performances.

Fig. 7.6 permet de séparer les effets des deux sources d'erreur, soit la modélisation effective de l'ordre et l'insuffisance de l'information statistique. Fig 7.6.a confirme la supériorité de l'algorithme SS. Mais SS s'avère plus sensible à la non modélisation des termes de la tête/queue. Ceci explique l'allure des courbes dans Fig. 7.7 ou l'écart entre SS et LP/OPD s'amincit à fur et à mesure que la partie non modélisée est plus importante.

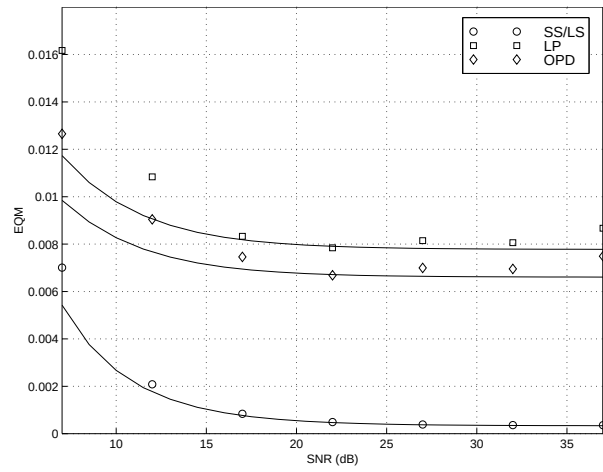
¹⁶ou des matrices qui lui sont proportionnelles telle que $\mathbf{R}_m^o$



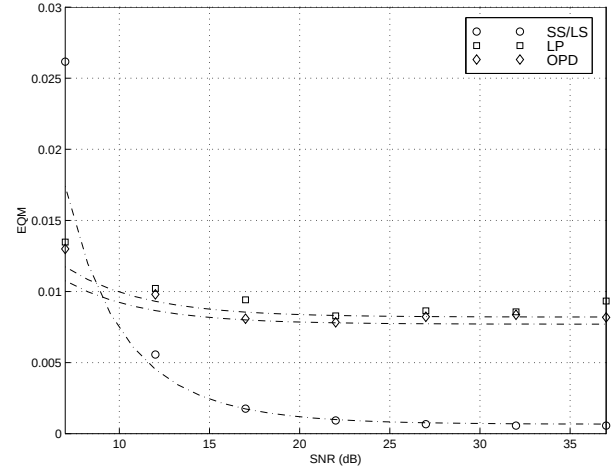
(a) $\mathbf{h}_M = \mathbf{h}_m^{a^z} + \mathbf{d}_{m_1, m_2}^z, SNR = 17dB$



(b) $\mathbf{h}_M = \mathbf{h}_m^{b^z} + \mathbf{d}_{m_1, m_2}^z, SNR = 17dB$



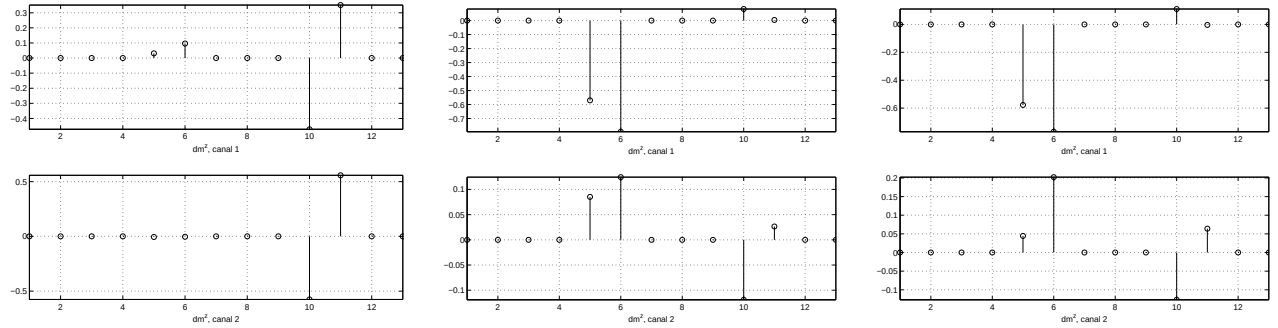
(c) $\mathbf{h}_M = \mathbf{h}_m^{a^z} + \mathbf{d}_{m_1, m_2}^z, N = 300$



(d) $\mathbf{h}_M = \mathbf{h}_m^{b^z} + \mathbf{d}_{m_1, m_2}^z, N = 300$

Évolution de l'erreur d'identification en fonction du bruit, du nombre d'observation et de la diversité de la réponse effective

Figure 7.4: Identification de la réponse effective : Modèle déterministe



(a) SS

(b) LP

(c) OPD

La tête/queue $\mathbf{d}_{m_1, m_2}^z$ qui introduit le biais le plus important sur l'estimation de la réponse effective $\mathbf{h}_m^a$

Figure 7.5: Modélisation effective : le cas le plus défavorable

	<i>Estimé (N=300)</i>	<i>Estimé (N=1000)</i>	<i>Asymptotique</i>		
Biais	0.0007	0.0004	0.0006		
	-0.0083	-0.0080	-0.0086		
	-0.0002	0.0000	0.0010		
	-0.0088	-0.0087	-0.0079		
	-0.0121	-0.0122	-0.0114		
	-0.0048	-0.0046	-0.0052	Borne Prop. 7.3	Borne (7.16)
Norme du Biais	0.0178	0.0176	0.0172	0.1141	0.5917
EQM	0.0008	0.0005	0.0008		

$$(a) \mathbf{h}_M = \mathbf{h}_m^{a^z} + \mathbf{d}_{m_1, m_2}^z$$

	<i>Estimé (N=300)</i>	<i>Asymptotique</i>		
Biais	0.0008	0.0013		
	-0.0158	-0.0172		
	-0.0018	0.0020		
	-0.0195	-0.0157		
	-0.0259	-0.0229		
	-0.0083	-0.0104	Borne Prop. 7.3	Borne (7.16)
Norme du Biais	0.0370	0.0343	0.2282	1.1885
EQM	0.0021	0.0018		

$$(b) \mathbf{h}_M = \mathbf{h}_m^{a^z} + 2\mathbf{d}_{m_1, m_2}^z$$

	<i>Estimé (N=300)</i>	<i>Asymptotique</i>		
Biais	-0.0042	-0.0046		
	-0.0176	-0.0184		
	0.0028	0.0055		
	-0.0038	-0.0028		
	0.0032	0.0044		
	-0.0102	-0.0123	Borne Prop. 7.3	Borne (7.16)
Norme du Biais	0.0215?	0.0239	0.2236	1.0362
EQM	0.0018	0.0017		

$$(c) \mathbf{h}_M = \mathbf{h}_m^{b^z} + \mathbf{d}_{m_1, m_2}^z$$

Tableau 7.2: Méthode SS : Biais et EQM

	<i>Estimé (N=300)</i>	<i>Estimé (N=1000)</i>	<i>Asymptotique</i>		
Biais	0.0142	0.0074	-0.0123		
	0.0074	0.0098	-0.0086		
	-0.0019	-0.0030	0.0019		
	-0.0165	-0.0128	0.0152		
	-0.0253	-0.0281	0.0266		
	-0.0016	-0.0032	-0.0020	Borne Prop. 7.3	Borne (7.17)
Norme du Biais	0.0343	0.0335	0.0342	0.0972	1.3403
EQM	0.0083	0.0033	0.0082		

$$(a) \mathbf{h}_M = \mathbf{h}_m^{a^z} + \mathbf{d}_{m_1, m_2}^z$$

	<i>Estimé (N=300)</i>	<i>Asymptotique</i>		
Biais	0.0195	-0.0247		
	0.0192	-0.0171		
	-0.0039	0.0037		
	-0.0297	0.0304		
	-0.0549	0.0532		
	-0.0080	-0.0040	Borne Prop. 7.3	Borne (7.17)
Norme du Biais	0.0688	0.0685	0.1944	2.6809
EQM	0.0127	0.0121		

$$(b) \mathbf{h}_M = \mathbf{h}_m^{a^z} + 2\mathbf{d}_{m_1, m_2}^z$$

	<i>Estimé (N=300)</i>	<i>Asymptotique</i>		
Biais	0.0139	-0.0115		
	0.0096	-0.0086		
	-0.0050	0.0024		
	-0.0196	0.0168		
	-0.0216	0.0202		
	0.0017	-0.0014	Borne Prop. 7.3	Borne (7.17)
Norme du Biais	0.0341	0.0301	0.0796	2.0633
EQM	0.0094	0.0086		

$$(c) \mathbf{h}_M = \mathbf{h}_m^{b^z} + \mathbf{d}_{m_1, m_2}^z$$

Tableau 7.3: Méthode LP : Biais et EQM

	<i>Estimé (N=300)</i>	<i>Estimé (N=1000)</i>	<i>Asymptotique</i>	
Biais	0.0093	0.0074	-0.0067	Borne Prop. 7.3
	0.0122	0.0119	-0.0119	
	-0.0066	-0.0035	0.0032	
	-0.0126	-0.0113	0.0099	
	-0.0257	-0.0254	0.0261	
	0.0019	0.0003	0.0001	
Norme du Biais	0.0332	0.0314	0.0313	0.0912
EQM	0.0075	0.0029	0.0069	

$$(a) \mathbf{h}_M = \mathbf{h}_m^{a^z} + \mathbf{d}_{m_1, m_2}^z$$

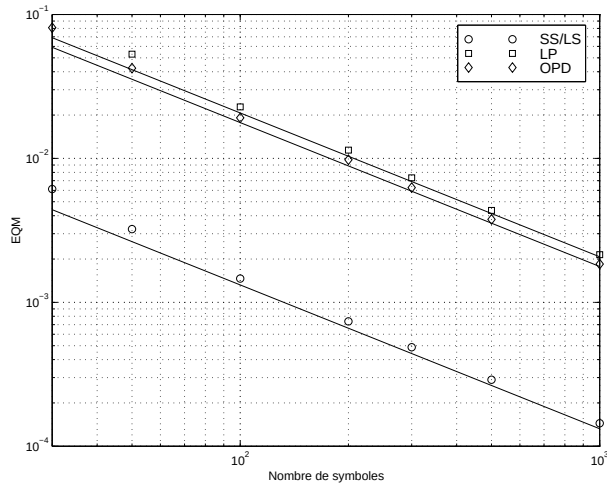
	<i>Estimé (N=300)</i>	<i>Asymptotique</i>	
Biais	0.0173	-0.0134	Borne Prop. 7.3
	0.0245	-0.0237	
	-0.0070	0.0065	
	-0.0260	0.0199	
	-0.0507	0.0522	
	0.0002	0.0002	
Norme du Biais	0.0647	0.0625	0.1825
EQM	0.0114	0.0101	

$$(b) \mathbf{h}_M = \mathbf{h}_m^{a^z} + 2\mathbf{d}_{m_1, m_2}^z$$

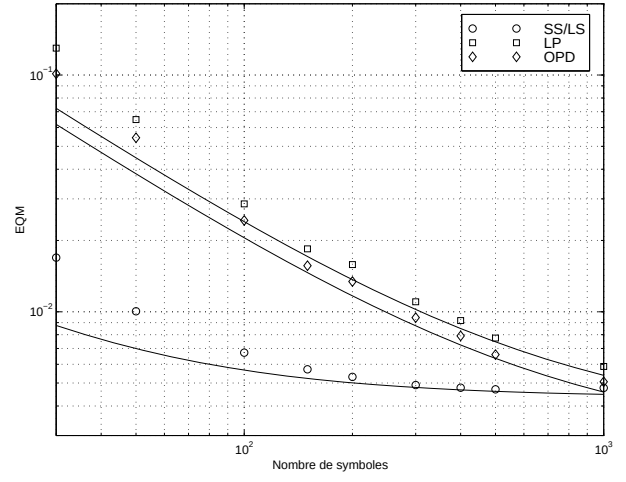
	<i>Estimé (N=300)</i>	<i>Asymptotique</i>	
Biais	0.0073	-0.0043	Borne Prop. 7.3
	0.0109	-0.0115	
	-0.0028	0.0033	
	-0.0124	0.0092	
	-0.0220	0.0223	
	-0.0011	0.0004	
Norme du Biais	0.0286	0.0272	0.0761
EQM	0.0081	0.0080	

$$(c) \mathbf{h}_M = \mathbf{h}_m^{b^z} + \mathbf{d}_{m_1, m_2}^z$$

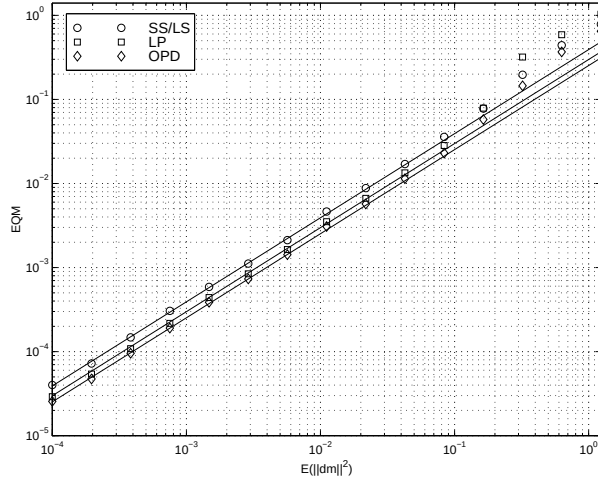
Tableau 7.4: Méthode OP : Biais et EQM



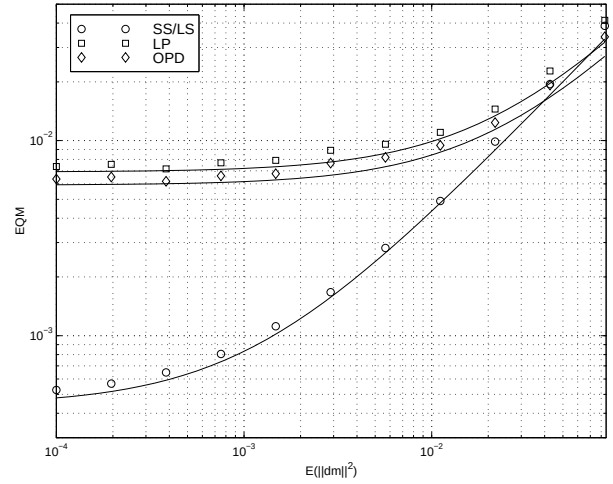
(a) Effet du nombre d'observations en modélisation exacte.



(a) Effet du nombre des symboles.
 $E(\|\mathbf{d}_{m_1, m_2}\|^2) = 0.011, SNR = 17dB$



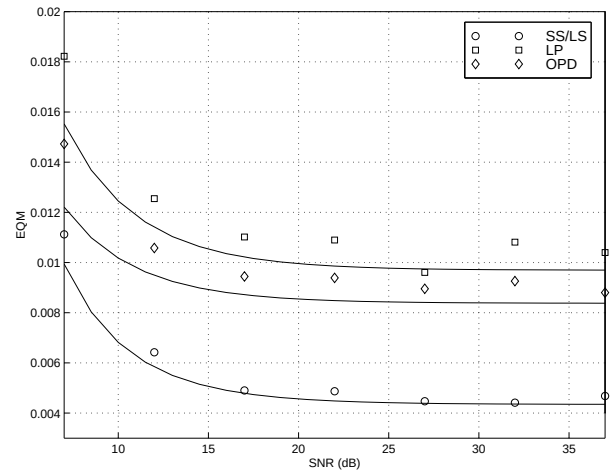
(b) Effet de la modélisation effective en statistiques exactes.



(b) Effet de la modélisation effective de l'ordre.
 $N = 300, SNR = 17dB$

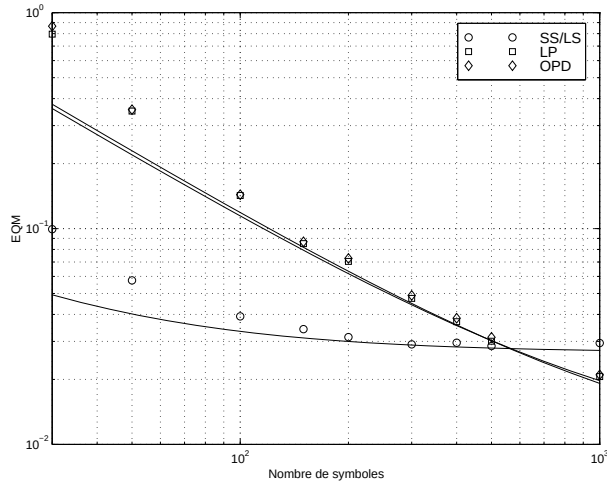
$$SNR = 17dB$$

Figure 7.6: Effets séparés du nombre d'observations et de la modélisation effective

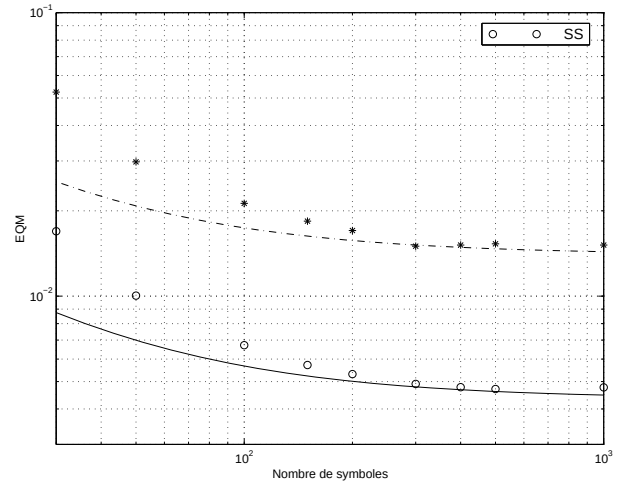


(c) Effet du bruit. $N = 300, E(\|\mathbf{d}_{m_1, m_2}\|^2) = 0.011$

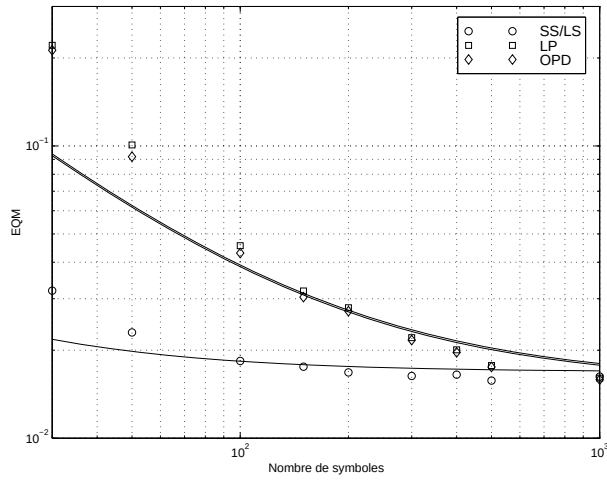
Figure 7.7: Identification de la réponse effective :
Modèle probabiliste



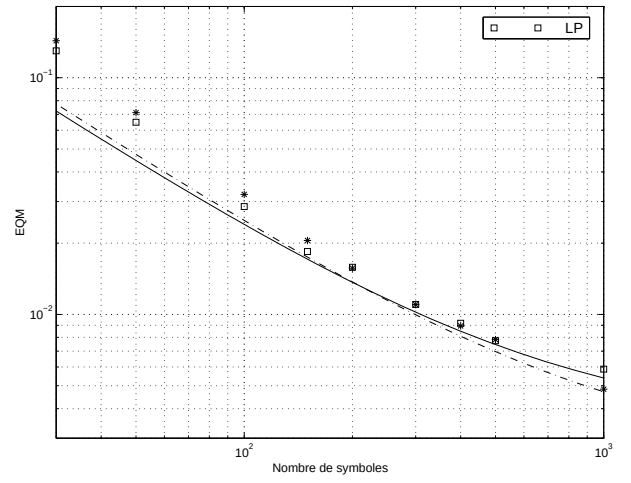
(a) Égaliseur



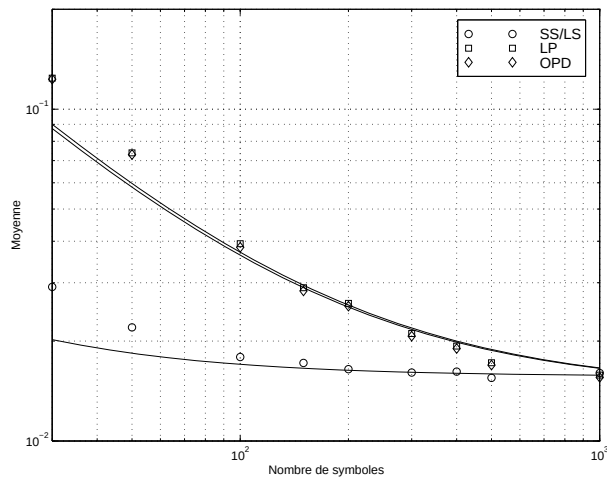
(a)



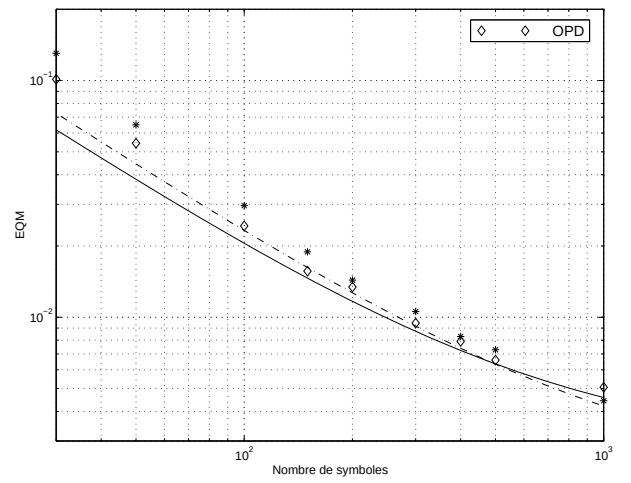
(b) Réponse globale



(b)



(c) Ouverture de l'oeil



(c)

$$E \left(\|\mathbf{d}_{m_1, m_2}\|^2 \right) = 0.011. \text{ SNR} = 17dB. \text{ retard}=1$$

$$E \left(\|\mathbf{d}_{m_1, m_2}\|^2 \right) = 0.011. \text{ SNR} = 17dB$$

Figure 7.8: Égalisation ZF en modélisation effective

Figure 7.9: Effet de la diversité.

7.7 Conclusion

La modélisation effective de l'ordre du canal est une approche pratique de la réponse impulsionnelle. En s'affranchissant de certains termes trop faibles et souvent très nombreux situés à la tête et à la queue de la réponse du canal, on se ramène à des réponses plus courtes et mieux diversifiées.

On a établi une méthodologie d'analyse de performances en modélisation effective applicable à tout algorithme du second ordre. Les termes non modélisés de la tête/queue ont été considérés comme connus (modèle déterministe) puis comme des variables aléatoires i.i.d. et centrés (modèle probabiliste).

Cette méthodologie a permis de mener une étude comparative des algorithmes sous-espace, prédiction linéaire et décomposition en produit extérieur qui confirme la supériorité du premier; les deux autres étant très proches. Comparée à l'étude faite par *Liavas et al*, elle repose sur des hypothèses plus pratiques et moins restrictives et a donné des résultats plus précis qui ont porté aussi sur d'autres paramètres dérivés de la réponse impulsionnelle (égaliseur ZF, réponse combinée canal/égaliseur et ouverture de l'oeil).

On a proposé une mesure (Prop. 7.3) de la sensibilité de l'algorithme à la non modélisation des termes de la tête/queue et donné la forme (Prop. 7.11) de la tête/queue qui occasionne l'erreur d'estimation la plus importante. Sous certaines hypothèses (cf. remarques à la fin de §7.6.1), les performances de l'algorithme en modélisation effective ne dépendent pas du nombre des termes non modélisés de la tête/queue.

La notion d'ordre effectif reste très liée à la classe de canaux du type Fig. 7.1. Elle reste une solution adaptée à une famille de canaux large mais pas exhaustive. En particulier, l'examen des canaux issus de conditions sévères de propagation (trajets multiples avec des retards importants) met en évidence les limites de son applicabilité (cf. §7.1.3).

Chapitre 8

Conclusion

8.1 Résultats

Depuis la démonstration par *Tong et al* [TONG 94] que l'identification aveugle est possible au second ordre, les travaux dédiés à ce types d'algorithmes se sont multipliés, motivés par l'importance des applications potentielles et les promesses en termes de performances de ces algorithmes. Ainsi, de nombreux algorithmes ont été proposés depuis. Ce rapport a été consacré aux algorithmes, dits stochastiques, qui estiment la réponse du canal à partir d'une estimation de la matrice de corrélation; et dont les plus courants sont les algorithmes sous-espace, prédiction linéaire et décomposition en produit extérieur. Leur examen dans des conditions pratiques de réception (bruit de mesure, temps d'observation limité, canaux à bande limitée) soulève deux types de non robustesse. Premièrement, la connaissance (en pratique improbable) de l'ordre exact du canal est requise. Deuxièmement, les performances dépendent d'une bonne diversité du canal.

Dans ce rapport, on pallie à la première insuffisance en proposant (§5) un algorithme robuste capable de *bien* estimer la réponse du canal à partir d'une estimée de la matrice de corrélation lorsque l'on dispose d'une valeur surestimée de l'ordre du canal; des situations où les autres algorithmes sont incapables d'opérer. Cet algorithme montre également des propriétés intéressantes en termes du choix de l'ambiguïté et en termes de performances.

Quant à la diversité des canaux, le rapport y apporte une meilleure compréhension. Elle se base sur l'extension du théorème de Szegö aux matrices blocs de Toeplitz; dont une nouvelle démonstration est donnée. L'interprétation de ces résultats pour les matrices de filtrage des canaux SIMO permet de fixer une définition de la diversité comme étant sa plus petite valeur singulière lorsque la taille de cette matrice (ou encore le facteur de lissage) tend vers l'infini. Vu que les plus petites/grandes valeurs singulières des matrices de Toeplitz convergent¹ à la vitesse $\frac{1}{n^2}$, il se confirme que cette limite approxime bien la plus petite valeur singulière d'une matrice de filtrage d'une taille *pratique*. Ainsi définie, la diversité permet d'apprécier l'*identification en aveugle et dans des conditions pratique d'observation* des canaux SIMO indépendamment de l'algorithme utilisé. L'application du théorème de Szegö permet

¹lorsque la taille n de ces matrices tend vers l'infini. Ce résultat est prouvé théoriquement dans [GREN 84, Théorème p. 72] pour les matrices de Toeplitz. Fait constaté numériquement pour les matrices blocs de Toeplitz.

de déduire une borne supérieure sur cette mesure qui s'exprime en termes des fonctions de transfert des différents sous-canaux. Son évaluation dans le cas de canaux à bande limitée sur-échantillonnés à la réception montre de faibles valeurs et prouve ainsi l'incapacité du sur-échantillonnage à créer suffisamment de diversité pour identifier en aveugle.

Le rapport s'est intéressé à la proposition de *Liavas et al* de ne chercher à estimer que *l'ordre effectif* du canal, sur la base duquel identification et égalisation seront effectuées. L'étude de performances faite dans ce cadre aboutit à la proposition de *mesure* pour apprécier la sensibilité d'un algorithme à la non prise en compte de termes faibles à la tête/queue de la réponse. Elle montre l'insensibilité (sous certaines conditions) des algorithmes au nombre de ces termes là. De part l'inspiration même de cette approche de l'ordre, elle n'adresse que des situations particulières (fut elles fréquentes); celles où la mauvaise diversité (de la réponse globale) est le fait de la présence de termes faibles à la tête/queue. Elle ne résout pas les problèmes de robustesse (à l'ordre et à la diversité) qui se trouvent transposés à la réponse effective. Ainsi, le problème du comportement des algorithmes d'identification dans le cas de canaux mal diversifiés reste entièrement posé.

8.2 Perspectives

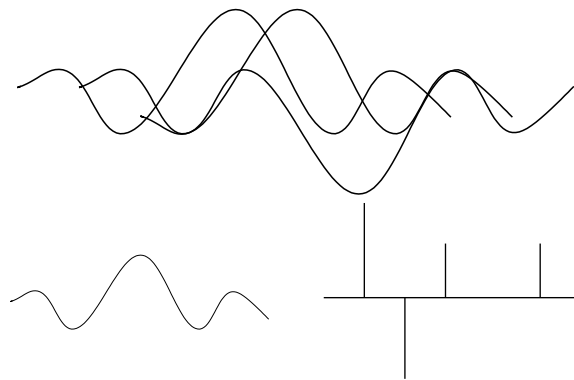


Figure 8.1: Exploitation de la connaissance du signal de mise en forme

Ce travail a proposé une solution au problème de la surestimation de l'ordre d'un canal. Il a apporté une meilleure compréhension du problème de la (mauvaise) diversité des canaux; mais n'a pas proposé de solution. Le cas particulier des signaux à bande limitée (du type cosinus sur-élevé) se propageant dans des canaux à trajets multiples constitue un cadre réaliste pour tester la *vraie* robustesse des algorithmes d'identification proposés. La prise en compte des connaissances sur le signal de mise en forme (cf. Fig. 8.1) permet non seulement de réduire la dimension du problème mais surtout de séparer l'aspect propagation (multi-trajets) de l'aspect bande limitée (lié au signal de mise en forme). Ce dernier est pressenti comme étant la *cause principale* de la mauvaise diversité du canal. Certains travaux [Ding 95, DING 97, SCHL 94] ont proposé de prendre en compte les connaissances sur le signal de mise en forme pour améliorer les performances des algorithmes d'identification. La matrice de filtrage du canal se factorise en un terme mal conditionné mais connu (fonction du signal de mise en forme) et d'un terme court et inconnu (le multi-trajets)². Dans [Ding 95, DING 97] (resp. [SCHL 94, DING 97]), cette factorisation est exploitée dans une approche sous-espace (SS) (resp.

²On préférera cette modélisation à celle, dite spéculaire, proposée dans [Veen 98] puis dans [Perr 99] où le canal

moindres-carrés (LS)). Les solutions proposées ont ainsi hérité la non robustesse des algorithmes SS et LS à la surestimation de l'ordre du canal. Nous pensons qu'il serait utile de mettre à profit cette information sur la structure du canal mais dans une démarche semblable à celle de l'algorithme proposé dans §5. Ainsi défini, le problème est mieux (et probablement bien) conditionné puisque l'inconnue du problème n'est pas à bande limitée et serait en même temps robuste à une surestimation de l'ordre à la manière de l'algorithme dans §5. Le peu de capacité de la réception fractionnée à créer suffisamment de diversité incite à explorer cette approche dans le cadre d'une réception multi-capteurs.

Également concernant l'algorithme proposé dans §5, une amélioration possible est d'introduire un critère autre que le critère 5.1 qui lui a été introduit d'une manière intuitive et donc probablement non optimale. En particulier, la définition d'un critère sur les égaliseurs permet de réduire significativement la complexité de l'algorithme, notamment lorsque le facteur de lissage est important, ce qui est le cas lorsqu'une surestimation de l'ordre du canal est envisagée.

Une piste prometteuse consiste à calculer directement des égaliseurs à partir des statistiques du second ordre du canal, sans passer par l'estimation de la réponse au préalable. L'égalisation directe est d'autant plus intéressante qu'elle est possible avec des filtres de longueurs limitées, contrairement au cas de la réception synchrone. Comparée à l'approche *identifier puis égaliser*, elle présente les deux avantages suivants. Premièrement, elle est intrinsèquement robuste à la surestimation de l'ordre (cf. §4.4)³ Deuxièmement, on peut être conduit à penser que l'égalisation directe est immunisée contre un mauvais conditionnement de la matrice de corrélation, donc robuste à une mauvaise diversité du canal. En effet, l'identification de la réponse du canal implique toujours une inversion de la matrice de corrélation (donc de la matrice de filtrage). Le calcul, à partir de cette réponse identifiée, d'un égaliseur MMSE (resp. ZF) s'obtient par une inversion de la matrice de corrélation (resp. de la matrice de filtrage). Le chemin direct (calculer les égaliseurs directement à partir des SOS) ne se traduirait pas par une inversion et serait alors insensible à la (mauvaise) diversité du canal. Le peu de travaux proposés dans cette direction [Touz 98, Liu 97, GESB 97, Papa 98, GIAN 97] laisse espérer que de meilleures solutions en termes de performances peuvent être développées. L'égalisation directe pose cependant le problème du choix du délai, d'autant plus critique qu'il s'agit d'égaliseurs ZF.

Réception fractionnée et réception multi-capteurs sont toutes les deux traitées dans le cadre unique des canaux SIMO. Cette formulation théorique commune cache des performances très différentes en pratique. Il a souvent été constaté que la réception fractionnée échoue à créer suffisamment de diversité pour identifier les canaux en aveugle et dans des conditions pratiques d'observation. De l'autre côté, la réception multi-capteurs a souvent été citée [Redd 95] comme capable de fournir des conditions favorables à l'identification aveugle des canaux. La borne proposée dans Th. 6.10 fournit un outil

physique y est modélisé comme étant un canal multi-trajets paramétré par les différents délais et atténuations qui deviennent l'inconnue du problème. Dans [Veen 98, Perr 99], cette modélisation est exploitée pour réécrire l'algorithme sous-espace. Plus fine, elle montre de meilleures performances [Veen 98], et éventuellement une robustesse à la surestimation de l'ordre [Perr 99]. Cependant, le problème ainsi formulé n'est plus vraiment *du sous-espace*, puisque l'inconnue du problème n'est plus identifiée comme étant dans un sous-espace de dimension 1. On se ramène à la minimisation d'une fonction coût par le moyen de techniques multi-dimensionnelles, complexes et éventuellement sous-optimales.

³La robustesse à la surestimation de l'ordre est vérifiée aussi par les approches *spectrales* d'identification, tel dans [LI 94, POZI 97]. Ces méthodes déduisent la réponse fréquentielle du canal à partir du cyclo-spectre de sa sortie, et ne nécessitent donc pas la pré-détermination de l'ordre. Elles sont ainsi robustes aux erreurs d'estimation de l'ordre. Par contre, leurs performances sont faibles, même en fort SNR [DING 97].

simple pour l'évaluation de l'une et l'autre des techniques de diversité; chose faite dans ce rapport pour la réception fractionnée. Il est intéressant d'étendre cet emploi au cas de la réception multi-capteurs; et examiner en particulier la borne de Th. 6.10 en fonction des paramètres d'un réseau de capteurs.

Annexe A

Démonstrations de §7

A.1 Démonstration de (7.5)

Montrons que

$$\text{Vec} \left(\delta \mathbf{R}^d \right) = \mathbf{C} \mathbf{d}_{m_1, m_2}^* + \mathbf{K}_{c(m+1)} \mathbf{C} \mathbf{d}_{m_1, m_2} \quad (\text{A.1})$$

$$\text{avec } \mathbf{C}_{SPC} = \left(\mathbf{I}_{c(m+1)} \otimes \mathcal{T}_{m+1, SPC}(\mathbf{h}_m^z) \right) \overline{\mathcal{T}}_{m+1, SPC, M} \mathbf{I}_{SPC}^z \quad (\text{A.2})$$

$$\begin{aligned} \delta \mathbf{R}_{SPC}^d &= \mathcal{T}_{m+1, SPC}(\mathbf{h}_m^z) \mathcal{T}_{m+1, SPC}^H(\mathbf{d}_{m_1, m_2}^z) \\ &\quad + \left[\mathcal{T}_{m+1, SPC}(\mathbf{h}_m^z) \mathcal{T}_{m+1, SPC}^H(\mathbf{d}_{m_1, m_2}^z) \right]^H \\ \text{Vec} \left(\delta \mathbf{R}_{SPC}^d \right) &= \text{Vec} \left(\mathcal{T}_{m+1, SPC}(\mathbf{h}_m^z) \mathcal{T}_{m+1, SPC}^H(\mathbf{d}_{m_1, m_2}^z) \right) \\ &\quad + \mathbf{K}_{c(m+1)} \text{Vec} \left(\mathcal{T}_{m+1, SPC}(\mathbf{h}_m^z) \mathcal{T}_{m+1, SPC}^H(\mathbf{d}_{m_1, m_2}^z) \right)^* \\ \text{Vec} \left(\mathcal{T}_{m+1, SPC}(\mathbf{h}_m^z) \mathcal{T}_{m+1, SPC}^H(\mathbf{d}_{m_1, m_2}^z) \right) &= \left(\mathbf{I}_{c(m+1)} \otimes \mathcal{T}_{m+1, SPC}(\mathbf{h}_m^z) \right) \text{Vec} \left(\mathcal{T}_{m+1, SPC}^T(\mathbf{d}_{m_1, m_2}^{z*}) \right) \\ &= \left(\mathbf{I}_{c(m+1)} \otimes \mathcal{T}_{m+1, SPC}(\mathbf{h}_m^z) \right) \overline{\mathcal{T}}_{m+1, SPC, M} \mathbf{d}_{m_1, m_2, SPC}^{z*} \\ &= \left(\mathbf{I}_{c(m+1)} \otimes \mathcal{T}_{m+1, SPC}(\mathbf{h}_m^z) \right) \overline{\mathcal{T}}_{m+1, SPC, M} \mathbf{I}_{SPC}^z \mathbf{d}_{m_1, m_2, SPC}^* \end{aligned}$$

A.2 Démonstration de (7.4)

Montrons que

$$\text{Vec} \left(\delta \mathbf{R}^d \right) = \mathbf{C}^r \mathbf{d}_{m_1, m_2}$$

avec

$$\mathbf{C}_{TS}^r = \left(\mathbf{K}_{c(m+1)} + \mathbf{I}_{c^2(m+1)^2} \right) \left(\mathbf{I}_{c(m+1)} \otimes \mathcal{T}_{m+1, TS}(\mathbf{h}_m^z) \right) \left(\mathbf{I}_c \otimes \mathbf{C}^\dagger \right) \quad (\text{A.3})$$

$$\mathbf{C}_{ST}^r = \left(\mathbf{K}_{c(m+1)} + \mathbf{I}_{c^2(m+1)^2} \right) \left(\mathbf{K}_{m+1, c} \otimes \mathcal{T}_{m+1, ST}(\mathbf{h}_m^z) \right) \left(\mathbf{I}_c \otimes \mathbf{C}^\dagger \right) \mathbf{K}_{c, M-m} \quad (\text{A.4})$$

$$\text{où } \mathbf{C}^\dagger = \left[\left(\mathbf{1}_m \otimes \begin{bmatrix} \mathbf{I}_{m_1, m_2}^z \\ \mathbf{0}_{m+1, M-m} \end{bmatrix} \right) \right]$$

$$\mathbf{C}_{SPC}^r = \left(\mathbf{K}_{c(m+1)} + \mathbf{I}_{c^2(m+1)^2} \right) \left(\mathbf{I}_{c(m+1)} \otimes \mathcal{T}_{m+1,SPC}(\mathbf{h}_m^z) \right) \overline{\mathcal{T}}_{m+1,SPC,M} \mathbf{I}_{SPC}^z$$

$$\begin{aligned} \mathbf{C}_{TS}^r &= \left(\mathbf{K}_{c(m+1)} + \mathbf{I}_{c^2(m+1)^2} \right) \left(\mathbf{I}_{c(m+1)} \otimes \mathcal{T}_{m+1,TS}(\mathbf{h}_m^z) \right) \overline{\mathcal{T}}_{m+1,TS,M} \mathbf{I}_{TS}^z \\ &= \left(\mathbf{K}_{c(m+1)} + \mathbf{I}_{c^2(m+1)^2} \right) \left(\mathbf{I}_{c(m+1)} \otimes \mathcal{T}_{m+1,TS}(\mathbf{h}_m^z) \right) \left(\mathbf{I}_c \otimes \overline{\mathcal{T}}_{m+1,M} \right) \left(\mathbf{I}_c \otimes \mathbf{I}_{m_1,m_2}^z \right) \\ &= \left(\mathbf{K}_{c(m+1)} + \mathbf{I}_{c^2(m+1)^2} \right) \left(\mathbf{I}_{c(m+1)} \otimes \mathcal{T}_{m+1,TS}(\mathbf{h}_m^z) \right) \left[\left(\mathbf{I}_c \right) \otimes \left(\overline{\mathcal{T}}_{m+1,M} \mathbf{I}_{m_1,m_2}^z \right) \right] \end{aligned}$$

On peut facilement vérifier que $\overline{\mathcal{T}}_{m+1,M} \mathbf{I}_{m_1,m_2}^z$ n'est autre que $\mathbf{C}^\dagger$. D'où (A.3).

$$\begin{aligned} \mathbf{C}_{ST}^r &= \left(\mathbf{K}_{c(m+1)} + \mathbf{I}_{c^2(m+1)^2} \right) \left(\mathbf{I}_{c(m+1)} \otimes \mathcal{T}_{m+1,ST}(\mathbf{h}_m^z) \right) \overline{\mathcal{T}}_{m+1,ST,M} \mathbf{I}_{ST}^z \\ &= \left(\mathbf{K}_{c(m+1)} + \mathbf{I}_{c^2(m+1)^2} \right) \left(\mathbf{I}_{c(m+1)} \otimes \mathcal{T}_{m+1,ST}(\mathbf{h}_m^z) \right) \left(\mathbf{K}_{m+1,c} \otimes \mathbf{I}_{M+m+1} \right) \\ &\quad \left(\mathbf{I}_c \otimes \overline{\mathcal{T}}_{m+1,M} \right) \mathbf{K}_{c,M+1} \left(\mathbf{I}_{m_1,m_2}^z \otimes \mathbf{I}_c \right) \\ &= \left(\mathbf{K}_{c(m+1)} + \mathbf{I}_{c^2(m+1)^2} \right) \left(\mathbf{K}_{m+1,c} \otimes \mathcal{T}_{m+1,ST}(\mathbf{h}_m^z) \right) \left(\mathbf{I}_c \otimes \overline{\mathcal{T}}_{m+1,M} \right) \left(\mathbf{I}_c \otimes \mathbf{I}_{m_1,m_2}^z \right) \mathbf{K}_{c,m_1+m_2} \\ &= \left(\mathbf{K}_{c(m+1)} + \mathbf{I}_{c^2(m+1)^2} \right) \left(\mathbf{K}_{m+1,c} \otimes \mathcal{T}_{m+1,ST}(\mathbf{h}_m^z) \right) \left[\left(\mathbf{I}_c \right) \otimes \left(\overline{\mathcal{T}}_{m+1,M} \mathbf{I}_{m_1,m_2}^z \right) \right] \mathbf{K}_{c,M-m} \end{aligned}$$

d'où (A.4).

A.3 Démonstration de (7.8)

$$\begin{aligned} \hat{\mathbf{R}} &= \frac{1}{N-m} \sum_{n=m}^{N-1} \mathbf{y}_{m+1}(n) \mathbf{y}_{m+1}^H(n) \\ &= \frac{1}{N-m} \sum_{n=m}^{N-1} \left[\mathcal{T}(\mathbf{h}_m^z) \mathbf{s}_{2m+1}(n) + \mathcal{T}(\mathbf{d}_{m_1,m_2}^z) \mathbf{s}_{2m+1}(n) + \mathbf{b}_{m+1}(n) \right] \\ &\quad \left[\mathcal{T}(\mathbf{h}_m^z) \mathbf{s}_{2m+1}(n) + \mathcal{T}(\mathbf{d}_{m_1,m_2}^z) \mathbf{s}_{2m+1}(n) + \mathbf{b}_{m+1}(n) \right]^H \\ &= \frac{1}{N-m} \sum_{n=m}^{N-1} \left[\mathcal{T}(\mathbf{h}_m^z) \mathbf{s}_{2m+1}(n) + \mathbf{b}_{m+1}(n) \right] \left[\mathcal{T}(\mathbf{h}_m^z) \mathbf{s}_{2m+1}(n) + \mathbf{b}_{m+1}(n) \right]^H \quad (\text{A.5}) \end{aligned}$$

$$+ \frac{1}{N-m} \sum_{n=m}^{N-1} \left[\mathcal{T}(\mathbf{d}_{m_1,m_2}^z) \mathbf{s}_{2m+1}(n) \right] \left[\mathcal{T}(\mathbf{d}_{m_1,m_2}^z) \mathbf{s}_{2m+1}(n) \right]^H \quad (\text{A.6})$$

$$+ \frac{1}{N-m} \sum_{n=m}^{N-1} \left[\mathcal{T}(\mathbf{h}_m^z) \mathbf{s}_{2m+1}(n) + \mathbf{b}_{m+1}(n) \right] \left[\mathcal{T}(\mathbf{d}_{m_1,m_2}^z) \mathbf{s}_{2m+1}(n) \right]^H \quad (\text{A.7})$$

$$+ \frac{1}{N-m} \sum_{n=m}^{N-1} \left[\mathcal{T}(\mathbf{d}_{m_1,m_2}^z) \mathbf{s}_{2m+1}(n) \right] \left[\mathcal{T}(\mathbf{h}_m^z) \mathbf{s}_{2m+1}(n) + \mathbf{b}_{m+1}(n) \right]^H \quad (\text{A.8})$$

- Terme (A.5)

$$\frac{1}{N-m} \sum_{n=m}^{N-1} \left[\mathcal{T}(\mathbf{h}_m^z) \mathbf{s}_{2m+1}(n) + \mathbf{b}_{m+1}(n) \right] \left[\mathcal{T}(\mathbf{h}_m^z) \mathbf{s}_{2m+1}(n) + \mathbf{b}_{m+1}(n) \right]^H = \mathbf{R}_{h_m} + \delta \mathbf{R}_{h_m}^N$$

tel que d'après §4.1, $\lim_{N \rightarrow \infty} \left(\left[\text{Vec}(\delta \mathbf{R}_{h_m}^N) \right] \left[\text{Vec}(\delta \mathbf{R}_{h_m}^N) \right]^H \right) = \frac{1}{N} \mathcal{C}_{\mathbf{R}_m, h_m}$

- Terme A.6

$$\frac{1}{N-m} \sum_{n=m}^{N-1} [\mathcal{T}(\mathbf{d}_{m_1, m_2}^z) \mathbf{s}_{2m+1}(n)] [\mathcal{T}(\mathbf{d}_{m_1, m_2}^z) \mathbf{s}_{2m+1}(n)]^H = \mathcal{O}\left(\|\mathbf{d}_{m_1, m_2}\|^2\right).$$

- Terme (A.7)

En généralisant Prop. 4.1 aux intercorrélations, on a

$$\begin{aligned} & \frac{1}{N-m} \sum_{n=m}^{N-1} [\mathcal{T}(\mathbf{h}_m^z) \mathbf{s}_{2m+1}(n) + \mathbf{b}_{m+1}(n)] [\mathcal{T}(\mathbf{d}_{m_1, m_2}^z) \mathbf{s}_{2m+1}(n)]^H \\ &= \mathbb{E} \left([\mathcal{T}(\mathbf{h}_m^z) \mathbf{s}_{2m+1}(n) + \mathbf{b}_{m+1}(n)] [\mathcal{T}(\mathbf{d}_{m_1, m_2}^z) \mathbf{s}_{2m+1}(n)]^H \right) + \frac{1}{N} \mathcal{C}_m + \mathcal{O}\left(\frac{1}{N}\right) \end{aligned}$$

tel que $\mathcal{C}_m = \mathcal{O}\left(\|\mathbf{d}_{m_1, m_2}\|^2\right)$. Donc

$$\begin{aligned} & \frac{1}{N-m} \sum_{n=m}^{N-1} [\mathcal{T}(\mathbf{h}_m^z) \mathbf{s}_{2m+1}(n) + \mathbf{b}_{m+1}(n)] [\mathcal{T}(\mathbf{d}_{m_1, m_2}^z) \mathbf{s}_{2m+1}(n)]^H \\ &= \mathcal{T}(\mathbf{h}_m^z) \mathcal{T}^H(\mathbf{d}_{m_1, m_2}^z) + \mathcal{O}\left(\frac{1}{N} \|\mathbf{d}_{m_1, m_2}\|^2\right) + \mathcal{O}\left(\frac{1}{N}\right) \end{aligned}$$

- Terme (A.8) (pareil que terme (A.7))

On reconnaît $\mathcal{T}(\mathbf{h}_m^z) \mathcal{T}^H(\mathbf{d}_{m_1, m_2}^z) + \mathcal{T}(\mathbf{d}_{m_1, m_2}^z) \mathcal{T}^H(\mathbf{h}_m^z) = \delta \mathbf{R}^d$. D'où le résultat.

A.4 Démonstration de (7.9)

Les grandeurs sont exprimées dans le format ST

$$\begin{aligned} \mathbf{r}_{M+m-1, i} &= \mathcal{T}_m^T(\mathbf{h}_M) \hat{\mathbf{g}}_{m-1, i} \\ &= \mathcal{T}_m^T(\mathbf{h}_m^z) \hat{\mathbf{g}}_{m-1, i} + \mathcal{T}_m^T(\mathbf{d}_{m_1, m_2}^z) \hat{\mathbf{g}}_{m-1, i} \\ &= \begin{bmatrix} \mathbf{0}_{2m_1, 1} \\ \mathcal{T}_m^T(\mathbf{h}_m) \hat{\mathbf{g}}_{m-1, i} \\ \mathbf{0}_{2m_2, 1} \end{bmatrix} + \mathcal{T}_m^T(\mathbf{d}_{m_1, m_2}^z) \hat{\mathbf{g}}_{m-1, i} \end{aligned}$$

$$\begin{aligned} \mathcal{T}_m^T(\mathbf{h}_m) \hat{\mathbf{g}}_{m-1, i} &= \mathcal{T}_m^T(\hat{\mathbf{h}}_m) \hat{\mathbf{g}}_{m-1, i} - \mathcal{T}_m^T(\hat{\mathbf{h}}_m - \mathbf{h}_m) \hat{\mathbf{g}}_{m-1, i} \\ &= \mathbf{e}_{2m, i} - \mathcal{T}_m^T(\hat{\mathbf{h}}_m - \mathbf{h}_m) \mathbf{g}_{m-1, i} - \mathcal{T}_m^T(\hat{\mathbf{h}}_m - \mathbf{h}_m) (\hat{\mathbf{g}}_{m-1, i} - \mathbf{g}_{m-1, i}) \\ &= \mathbf{e}_{2m, i} - \mathcal{T}^T(\mathbf{g}_{m-1, i}) (\hat{\mathbf{h}}_m - \mathbf{h}_m) + \mathcal{O}\left(\|\hat{\mathbf{h}}_m - \mathbf{h}_m\|^2\right) \end{aligned}$$

$$\begin{aligned} \text{Aussi, on a : } \mathcal{T}_m^T(\mathbf{d}_{m_1, m_2}^z) \hat{\mathbf{g}}_{m-1, i} &= \mathcal{T}_m^T(\mathbf{d}_{m_1, m_2}^z) \mathbf{g}_{m-1, i} + \mathcal{T}_m^T(\mathbf{d}_{m_1, m_2}^z) (\hat{\mathbf{g}}_{m-1, i} - \mathbf{g}_{m-1, i}) \\ &= \mathcal{T}_{M+1}^T(\mathbf{g}_{m-1, i}) \mathbf{d}_{m_1, m_2}^z + \mathcal{O}\left((\hat{\mathbf{h}}_m - \mathbf{h}_m) \mathbf{d}_{m_1, m_2}\right) \end{aligned}$$

Ainsi, on obtient

$$\begin{aligned} \mathbf{r}_{M+m-1,i} - \mathbf{e}_{M+m,m_1+i} &= \mathcal{T}_{M+1}^T \left(\mathbf{g}_{m-1,i} \right) \mathbf{d}_{m_1,m_2}^z - \begin{bmatrix} \mathbf{0}_{m_1,c(m+1)} \\ \mathcal{T}^T \left(\mathbf{g}_{m-1,i} \right) \\ \mathbf{0}_{m_2,c(m+1)} \end{bmatrix} \\ &+ \mathcal{O} \left(\left\| \hat{\mathbf{h}}_m - \mathbf{h}_m \right\|^2 \right) + \mathcal{O} \left((\hat{\mathbf{h}}_m - \mathbf{h}_m) \mathbf{d}_{m_1,m_2} \right) \end{aligned} \quad (\text{A.9})$$

d'où le résultat.

A.5 Démonstration de (7.10)

$$\begin{aligned} \text{OEM}_{m_1+i}(\mathbf{r}_{M+m-1,i}) &= \frac{\sum_{k \neq m_1+i} r(k)^2}{r(m_1+i)^2} \\ \text{OEM}'_{m_1+i}(\mathbf{r}_{M+m-1,i}) &= \frac{2}{r(m_1+i)^2} \begin{bmatrix} r(1) \\ \vdots \\ r(m_1+i-1) \\ \frac{-\sum_{k \neq m_1+i} r(k)^2}{r(m_1+i)} \\ r(m_1+i+1) \\ \vdots \\ r(M+m) \end{bmatrix} \\ \text{OEM}''_{m_1+i}(\mathbf{r}_{M+m-1,i}) &= \frac{2}{r(m_1+i)^2} \times \begin{bmatrix} 1 & 0 & \cdots & 0 & \cdots & 0 \\ 0 & \ddots & \ddots & & \ddots & \vdots \\ \vdots & & 1 & & & 0 \\ \frac{-2r(1)}{r(m_1+i)} & \cdots & \frac{-2r(m_1+i-1)}{r(m_1+i)} & \frac{3 \sum_{k \neq m_1+i} r(k)^2}{r(m_1+i)^2} & \frac{-2r(m_1+i+1)}{r(m_1+i)} & \cdots & \frac{-2r(M+m)}{r(m_1+i)} \\ 0 & & & 1 & & \vdots \\ \vdots & \ddots & & & \ddots & 0 \\ 0 & \cdots & 0 & \cdots & 0 & 1 \end{bmatrix}. \end{aligned}$$

Ainsi $\text{OEM}_{m_1+i}(\mathbf{e}_{M+m,m_1+i}) = 0$, $\text{OEM}'_{m_1+i}(\mathbf{e}_{M+m,m_1+i}) = \mathbf{0}_{M+m,1}$ et $\text{OEM}''_{m_1+i}(\mathbf{e}_{M+m,m_1+i}) = 2 \ddot{\mathbf{I}}_i$, avec $\ddot{\mathbf{I}}_i$ est la matrice diagonale donnée par $(\ddot{\mathbf{I}}_i)_{ii} = 1$ pour $i = 1, \dots, m_1+i-1, m_1+i+1, \dots, M+m$ et $(\ddot{\mathbf{I}}_i)_{m_1+i,m_1+i} = 0$.

$$\begin{aligned} \text{OEM}_{m_1+i}(\mathbf{r}_{M+m-1,i}) &= \text{OEM}_{m_1+i}(\mathbf{e}_{M+m,m_1+i} + (\mathbf{r}_{M+m-1,i} - \mathbf{e}_{M+m,m_1+i})) \\ &= \text{OEM}_{m_1+i}(\mathbf{e}_{M+m,m_1+i}) \\ &\quad + (\mathbf{r}_{M+m-1,i} - \mathbf{e}_{M+m,m_1+i})^T \text{OEM}'_{m_1+i}(\mathbf{e}_{M+m,m_1+i}) \\ &\quad + \frac{1}{2}(\mathbf{r}_{M+m-1,i} - \mathbf{e}_{M+m,m_1+i})^T \text{OEM}''_{m_1+i}(\mathbf{e}_{M+m,m_1+i})(\mathbf{r}_{M+m-1,i} - \mathbf{e}_{M+m,m_1+i}) \\ &\quad + \mathcal{O} \left(\left\| \mathbf{r}_{M+m-1,i} - \mathbf{e}_{M+m,m_1+i} \right\|^3 \right) \\ &= (\mathbf{r}_{M+m-1,i} - \mathbf{e}_{M+m,m_1+i})^T \ddot{\mathbf{I}}_i (\mathbf{r}_{M+m-1,i} - \mathbf{e}_{M+m,m_1+i}) \\ &\quad + \mathcal{O} \left(\left\| \mathbf{r}_{M+m-1,i} - \mathbf{e}_{M+m,m_1+i} \right\|^3 \right) \end{aligned}$$

A.6 Démonstration de (7.11)

$$\begin{aligned}
\mathbf{A}_{\mathbf{R}} \text{Vec}(\mathbf{R}) &= -\mathbf{T}_m \left(\left(\mathbf{R}^T - \sigma(\mathbf{R}) \mathbf{I}_{2(m+1)} \right)^{\natural} \otimes \text{vecp}(\mathbf{R})^H \right) \text{Vec}(\mathbf{R}) \\
&= -\mathbf{T}_m \text{Vec} \left(\text{vecp}(\mathbf{R})^H \mathbf{R} \left(\mathbf{R} - \sigma(\mathbf{R}) \mathbf{I}_{2(m+1)} \right)^{\natural} \right) \\
&= -\mathbf{T}_m \text{Vec} \left(\sigma(\mathbf{R}) \text{vecp}(\mathbf{R})^H \left(\mathbf{R} - \sigma(\mathbf{R}) \mathbf{I}_{2(m+1)} \right)^{\natural} \right) \tag{A.10} \\
&= \mathbf{0}_{2(m+1),1} \tag{A.11}
\end{aligned}$$

(A.10) est donnée par $\text{vecp}(\mathbf{R})$ vecteur propre à $\mathbf{R}$. (A.11) est vraie car $\text{vecp}(\mathbf{R})$ est orthogonal à la matrice Hermitienne $\left(\mathbf{R} - \sigma(\mathbf{R}) \mathbf{I}_{2(m+1)} \right)^{\natural}$.

A.7 Démonstration de (7.12)

$$\begin{aligned}
\mathbf{A}_{\mathbf{R}_{hm}} \text{Vec}(\delta \mathbf{R}^d) &= -\mathbf{T}_{m+1} \left(\left[\mathcal{T}^*(\mathbf{h}_m) \mathcal{T}^T(\mathbf{h}_m) \right]^{\natural} \otimes \left[\mathbf{h}_m^T \mathbf{T}_{m+1} \right] \right) \text{Vec}(\delta \mathbf{R}^d) \\
&= -\mathbf{T}_{m+1} \text{Vec} \left(\mathbf{h}_m^T \mathbf{T}_{m+1} \delta \mathbf{R}^d \left[\mathcal{T}_m(\mathbf{h}_m) \mathcal{T}_m^H(\mathbf{h}_m) \right]^{\natural} \right) \\
&= -\mathbf{T}_{m+1} \left[\mathbf{h}_m^T \mathbf{T}_{m+1} \delta \mathbf{R}^d \left[\mathcal{T}_m(\mathbf{h}_m) \mathcal{T}_m^H(\mathbf{h}_m) \right]^{\natural} \right]^T \\
&= \mathbf{T}_{m+1} \left[\mathcal{T}^*(\mathbf{h}_m) \mathcal{T}^T(\mathbf{h}_m) \right]^{\natural} \delta \mathbf{R}^{d*} \mathbf{T}_{m+1} \mathbf{h}_m \\
&= \mathbf{T}_{m+1} \left[\mathcal{T}^*(\mathbf{h}_m) \mathcal{T}^T(\mathbf{h}_m) \right]^{\natural} \left[\mathcal{T}(\mathbf{h}_m^{z*}) \mathcal{T}^T(\mathbf{d}_{m_1, m_2}^z) \mathbf{T}_{m+1} \mathbf{h}_m + \mathcal{T}(\mathbf{d}_{m_1, m_2}^{z*}) \mathcal{T}^T(\mathbf{h}_m^z) \mathbf{T}_{m+1} \mathbf{h}_m \right] \\
&= \mathbf{T}_{m+1} \left[\mathcal{T}^*(\mathbf{h}_m) \mathcal{T}^T(\mathbf{h}_m) \right]^{\natural} \mathcal{T}(\mathbf{h}_m^{z*}) \mathcal{T}^T(\mathbf{d}_{m_1, m_2}^z) \mathbf{T}_{m+1} \mathbf{h}_m \quad \text{d'après Prop. 4.10} \\
&= \mathbf{T}_{m+1} \left[\mathcal{T}^*(\mathbf{h}_m) \mathcal{T}^T(\mathbf{h}_m) \right]^{\natural} \mathcal{T}(\mathbf{h}_m^{z*}) \mathcal{T}^T(\mathbf{T}_{m+1} \mathbf{h}_m) \mathbf{d}_{m_1, m_2}^z \quad \text{d'après (4.9)} \\
&= \mathbf{T}_{m+1} \left[\mathcal{T}^*(\mathbf{h}_m) \mathcal{T}^T(\mathbf{h}_m) \right]^{\natural} \mathcal{T}(\mathbf{h}_m^{z*}) \mathcal{T}_{M+1}^T(\mathbf{T}_{m+1} \mathbf{h}_m) \mathbf{I}^z \mathbf{d}_{m_1, m_2} \quad \text{d'après (7.3)}
\end{aligned}$$

Ainsi

$$\begin{aligned}
\mathbf{B} &= \mathbf{T}_{m+1} \left[\mathcal{T}^*(\mathbf{h}_m) \mathcal{T}^T(\mathbf{h}_m) \right]^{\natural} \mathcal{T}(\mathbf{h}_m^{z*}) \mathcal{T}_{M+1}^T(\mathbf{T}_{m+1} \mathbf{h}_m) \mathbf{I}^z \\
&= -\mathbf{T}_{m+1} \left[\mathcal{T}^*(\mathbf{h}_m) \mathcal{T}^T(\mathbf{h}_m) \right]^{\natural} \mathcal{T}(\mathbf{h}_m^{z*}) \mathcal{T}_{M+1}^H(\mathbf{v}_m) \mathbf{I}^z
\end{aligned}$$

A.8 Démonstration de Prop. 7.10

Rappelons que $\mathbf{B} = \mathbf{T}_{m+1} \left[\mathcal{T}^*(\mathbf{h}_m) \mathcal{T}^T(\mathbf{h}_m) \right]^{\natural} \mathcal{T}(\mathbf{h}_m^{z*}) \mathcal{T}_{M+1}^T(\mathbf{T}_{m+1} \mathbf{h}_m) \mathbf{I}^z$. Notons d'abord que $\left[\mathcal{T}^*(\mathbf{h}_m) \mathcal{T}^T(\mathbf{h}_m) \right]^{\natural}$ n'est fonction ni de m_1 ni de m_2 . Dans ce qui suit, on exprime tout dans le format TS . On s'intéresse au produit :

$$\mathcal{T}(\mathbf{h}_m^{z*}) \mathcal{T}_{M+1}^H(\mathbf{v}_m) \mathbf{I}^z (\mathbf{I}^z)^T \mathcal{T}_{M+1}(\mathbf{v}_m) \mathcal{T}^T(\mathbf{h}_m^z)$$

$$\mathcal{T}(\mathbf{h}_m^{z*}) \mathcal{T}_{M+1}^H(\mathbf{v}_m) = \begin{bmatrix} \mathcal{T}(\mathbf{h}_m^{z*1}) \\ \mathcal{T}(\mathbf{h}_m^{z*2}) \end{bmatrix} \begin{bmatrix} \mathcal{T}_{M+1}^H(\mathbf{v}_m^1) & \mathcal{T}_{M+1}^H(\mathbf{v}_m^2) \end{bmatrix} = \begin{bmatrix} \mathbf{D}_{11} & \mathbf{D}_{12} \\ \mathbf{D}_{21} & \mathbf{D}_{22} \end{bmatrix}$$

$$\text{où } \mathbf{D}_{ij} = \mathcal{T}(\mathbf{h}_m^{z*i}) \mathcal{T}_{M+1}^H(\mathbf{v}_m^j), \text{ } i \text{ et } j \in \{1, 2\}$$

En regardant une colonne de $\mathcal{T}_{M+1}^H(\mathbf{v}_m^j)$ comme une séquences de données, son produit par $\mathcal{T}(\mathbf{h}_m^{z*i})$ est le vecteur sortie du filtre $\mathbf{h}_m^{z*1}$ excité par cette séquence. Les colonnes de $\mathcal{T}_{M+1}^H(\mathbf{v}_m^j)$ sont des versions décalées les unes des autres, ainsi on peut dire la même chose de $\mathbf{D}_{ij}$. D'où la structure de Toeplitz de $\mathbf{D}_{ij}$, dont chaque élément $(\mathbf{D}_{ij})_{u,v}$ est le produit scalaire de la ligne u de $\mathcal{T}(\mathbf{h}_m^{z*i})$ et la ligne v de $\mathcal{T}_{M+1}^H(\mathbf{v}_m^j)$, i.e., :

$$\left[\underbrace{0 \dots 0}_{u-1+m_1} \quad h_m^i(0) \dots h_m^i(m) \quad \underbrace{0 \dots 0}_{m+1-u+m_2} \right]^H, u \in \{1, \dots, m+1\}$$

et

$$\left[\underbrace{0 \dots 0}_{v-1} \quad v_m^j(0) \dots v_m^j(m) \quad \underbrace{0 \dots 0}_{M+1-v} \right]^H, v \in \{1, \dots, M+1\}$$

Ce produit est nul dans les cas $v \geq u + m_1 + (m+1)$ ou $v \leq u + m_1 - (m+1)$ et éventuellement non nul pour $m_1 - m \leq v - u \leq m_1 + m$ i.e., sur $2m+1$ diagonales :

$$\mathbf{D}_{ij} = \begin{bmatrix} \underbrace{\begin{matrix} \cdot & \cdot & \cdot & d_{i,j}^{2m+1} \\ \cdot & \cdot & \cdot & d_{i,j}^{2m+1} \\ d_{i,j}^2 & \cdot & \cdot & \cdot \\ d_{i,j}^1 & d_{i,j}^2 & \cdot & \cdot \\ 0 & d_{i,j}^1 & \cdot & \cdot \\ 0 & 0 & \cdot & \cdot \end{matrix}}_{m_1} & \overbrace{\begin{matrix} 0 & \cdot & \cdot & \cdot & 0 \\ d_{i,j}^{2m+1} & 0 & \cdot & \cdot & 0 \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & d_{i,j}^{2m+1} & 0 & \cdot \\ \cdot & \cdot & \cdot & \cdot & d_{i,j}^{2m+1} \end{matrix}}^{m_2} \\ \underbrace{\begin{matrix} 0 & \cdot & \cdot & 0 \\ \cdot & \cdot & \cdot & \cdot \\ 0 & \cdot & \cdot & 0 \end{matrix}}_{m_1} & \underbrace{\begin{matrix} d_{i,j}^1 & d_{i,j}^2 & \cdot & \cdot & \cdot \\ d_{i,j}^1 & d_{i,j}^2 & \cdot & \cdot & \cdot \end{matrix}}_{m_1} \end{bmatrix}$$

En particulier, pour $m_1 = m + m'_1 \geq m$, on a :

$$\mathbf{D}_{ij} = \underbrace{\begin{bmatrix} 0 & \cdot & \cdot & 0 & d_{i,j}^1 & d_{i,j}^2 & \cdot & \cdot & d_{i,j}^{2m+1} & 0 & \cdot & 0 \\ \cdot & \cdot & \cdot & 0 & d_{i,j}^1 & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ 0 & \cdot & \cdot & 0 & 0 & \cdot & 0 & d_{i,j}^1 & \cdot & \cdot & \cdot & \cdot \end{bmatrix}}_{\mathbf{0}_{m+1, m'_1}} \underbrace{\begin{bmatrix} \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \end{bmatrix}}_{\mathbf{D}'_{ij}}$$

où la matrice $\mathbf{D}'_{ij}$ n'est pas fonction de m_1 . Par la suite,

$$\begin{aligned} \mathcal{T}(\mathbf{h}_m^{z*}) \mathcal{T}_{M+1}^H(\mathbf{v}_m) \mathbf{I}^z \mathbf{I}_T^z \mathcal{T}_{M+1}(\mathbf{v}_m) \mathcal{T}^T(\mathbf{h}_m^z) &= \begin{bmatrix} \mathbf{D}_{11} & \mathbf{D}_{12} \\ \mathbf{D}_{21} & \mathbf{D}_{22} \end{bmatrix} \mathbf{I}^z (\mathbf{I}^z)^T \begin{bmatrix} \mathbf{D}_{11}^H & \mathbf{D}_{21}^H \\ \mathbf{D}_{12}^H & \mathbf{D}_{22}^H \end{bmatrix} \\ &= \begin{bmatrix} \tilde{\mathbf{D}}_1 & \tilde{\mathbf{D}}_2 \\ \tilde{\mathbf{D}}_3 & \tilde{\mathbf{D}}_4 \end{bmatrix} \end{aligned}$$

Sachant que

$$\begin{aligned}\mathbf{I}^z(\mathbf{I}^z)^T &= (\mathbf{I}_2 \otimes \mathbf{I}^{zz}) \\ \mathbf{I}^{zz} &= \begin{bmatrix} \mathbf{I}_{m_1} & \mathbf{0}_{m_1, m+1} & \mathbf{0}_{m_1, m_2} \\ \mathbf{0}_{m+1, m_1} & \mathbf{0}_{m+1} & \mathbf{0}_{m+1, m_2} \\ \mathbf{0}_{m_2, m_1} & \mathbf{I}_{m_2, m+1} & \mathbf{I}_{m_2} \end{bmatrix} \\ &= \begin{bmatrix} \mathbf{I}_{m'_1} & & & \\ & \mathbf{I}_m & & \\ & & \mathbf{0}_{m+1} & \\ & & & \mathbf{I}_{m_2} \end{bmatrix}\end{aligned}$$

chacune des matrices $\tilde{\mathbf{D}}_j$ est de la forme :

$$\tilde{\mathbf{D}}_w = \mathbf{D}_{i_1 i_2} \mathbf{I}^{zz} \mathbf{D}_{j_1 j_2}^H + \mathbf{D}_{k_1 k_2} \mathbf{I}^{zz} \mathbf{D}_{l_1 l_2}^H, \quad i_1, i_2, j_1, j_2, k_1, k_2, l_1, l_2 \in \{1, 2\}$$

Dans le cas $m_1 = m + m'_1 \geq m, m'_1 \geq 0$, on a

$$\begin{aligned}\mathbf{D}_{i_1 i_2} \mathbf{I}^{zz} \mathbf{D}_{j_1 j_2}^H &= [\mathbf{0}_{m+1, m'_1} \mathbf{D}'_{i_1 i_2}] \begin{bmatrix} \mathbf{I}_{m'_1} & & & \\ & \mathbf{I}_m & & \\ & & \mathbf{0}_{m+1} & \\ & & & \mathbf{I}_{m_2} \end{bmatrix} \begin{bmatrix} \mathbf{0}_{m'_1, m+1} \\ \mathbf{D}_{j_1 j_2}^H \end{bmatrix} \\ &= \mathbf{D}'_{i_1 i_2} \begin{bmatrix} \mathbf{I}_m & & & \\ & \mathbf{0}_{m+1} & & \\ & & & \mathbf{I}_{m_2} \end{bmatrix} \mathbf{D}_{j_1 j_2}^H\end{aligned}$$

ce qui est n'est pas fonction de m_1 . D'où (7.14). En raisonnant par rapport à m_2 , (7.13) se démontre d'une manière similaire. (7.14) est alors évidente.

A.9 Démonstration de Prop. 7.11

On reprend les notations de §A.8. On prend $m_1 = m + m'_1$ et $m_2 = m + m'_2$ avec $m'_1, m'_2 \geq 0$.

$$\mathbf{B} = -\mathbf{T}_m \mathcal{T}_m(\mathbf{h}_m) \mathcal{T}_m^H(\mathbf{h}_m)^{\mathfrak{H}^T} \mathcal{T}(\mathbf{h}_m^{z*}) \mathcal{T}_{M+1}^H(\mathbf{v}_m) \mathbf{I}^z$$

$$\text{Rappelons que } \mathbf{I}^z = (\mathbf{I}_2 \otimes \mathbf{I}_{m_1, m_2}^z) \text{ où } \mathbf{I}_{m_1, m_2}^z \triangleq \begin{bmatrix} \mathbf{I}_{m_1} & \mathbf{0}_{m_1, m_2} \\ \mathbf{0}_{m+1, m_1} & \mathbf{0}_{m+1, m_2} \\ \mathbf{0}_{m_2, m_1} & \mathbf{I}_{m_2} \end{bmatrix}.$$

$$\begin{aligned}\mathcal{T}(\mathbf{h}_m^{z*}) \mathcal{T}_{M+1}^H(\mathbf{v}_m) \mathbf{I}^z &= \begin{bmatrix} \mathbf{D}_{11} & \mathbf{D}_{12} \\ \mathbf{D}_{21} & \mathbf{D}_{22} \end{bmatrix} \mathbf{I}^z \\ &= \begin{bmatrix} \mathbf{D}_{11} \mathbf{I}_{m_1, m_2}^z & \mathbf{D}_{12} \mathbf{I}_{m_1, m_2}^z \\ \mathbf{D}_{21} \mathbf{I}_{m_1, m_2}^z & \mathbf{D}_{22} \mathbf{I}_{m_1, m_2}^z \end{bmatrix}\end{aligned}$$

Puisque $m_1 = m + m'_1 \geq m$ et $m_2 = m + m'_2 \geq m$ alors $\mathbf{D}_{ij} = [\mathbf{0}_{m+1, m'_1} \mathbf{D}'_{ij} \mathbf{0}_{m+1, m'_2}]$ et $\mathbf{D}_{ij} \mathbf{I}_{m_1, m_2}^z =$

$$\mathbf{D}'_{ij} \mathbf{I}_1^z \text{ où } \mathbf{I}_1^z = \begin{bmatrix} \mathbf{0}_{m, m'_1} & \mathbf{I}_m & \mathbf{0}_m & \mathbf{0}_{m, m'_2} \\ \mathbf{0}_{m+1, m'_1} & \mathbf{0}_{m+1, m} & \mathbf{0}_{m+1, m} & \mathbf{0}_{m+1, m'_2} \\ \mathbf{0}_{m, m'_1} & \mathbf{0}_m & \mathbf{I}_m & \mathbf{0}_{m, m'_2} \end{bmatrix}. \text{ Ainsi,}$$

$$\mathcal{T}(\mathbf{h}_m^{z*}) \mathcal{T}_{M+1}^H(\mathbf{v}_m) \mathbf{I}^z = \begin{bmatrix} \mathbf{D}'_{11} \mathbf{I}_1^z & \mathbf{D}'_{12} \mathbf{I}_1^z \\ \mathbf{D}'_{21} \mathbf{I}_1^z & \mathbf{D}'_{22} \mathbf{I}_1^z \end{bmatrix}$$

$$= \begin{bmatrix} \mathbf{D}'_{11} & \mathbf{D}'_{12} \\ \mathbf{D}'_{21} & \mathbf{D}'_{22} \end{bmatrix} (\mathbf{I}_2 \otimes \mathbf{I}_1^z)$$

Ainsi, on peut écrire $\mathbf{B} = \mathbf{B}^\dagger (\mathbf{I}_2 \otimes \mathbf{I}_1^z)$ où la matrice $\mathbf{B}^\dagger \triangleq -\mathbf{T}_m \mathcal{T}_m(\mathbf{h}_m) \mathcal{T}_m^H(\mathbf{h}_m)^\dagger$ n'est fonction ni de m'_1 ni de m'_2 .

Soit le vecteur normé $\mathbf{x}_{m_1+m_2} \in \mathbf{C}^{2(m_1+m_2)}$. On l'écrit comme la somme de deux vecteurs orthogonaux $\mathbf{x}_{m_1+m_2} = \mathbf{x}_a + \mathbf{x}_b$ tel que

$$\mathbf{x}_a = \begin{bmatrix} \mathbf{0}_{1,m'_1} \mathbf{x}_{m_1+m_2}^1(m'_1+1) \cdots \mathbf{x}_{m_1+m_2}^1(m'_1+2m) \mathbf{0}_{1,m'_2} \\ \mathbf{0}_{1,m'_1} \mathbf{x}_{m_1+m_2}^2(m'_1+1) \cdots \mathbf{x}_{m_1+m_2}^2(m'_1+2m) \mathbf{0}_{1,m'_2} \end{bmatrix}^T$$

On a alors $\mathbf{I}_1^z \mathbf{x}_b^i = \mathbf{0}_{3m+1,1}$ donc $(\mathbf{I}_2 \otimes \mathbf{I}_1^z) \mathbf{x}_b = \mathbf{0}_{2(3m+1),1}$. Ainsi, $\mathbf{B} \mathbf{x}_{m_1+m_2} = \mathbf{B} \mathbf{x}_a = \mathbf{B}^\dagger (\mathbf{I}_2 \otimes \mathbf{I}_1^z) \mathbf{x}_a$.

Par conséquent, le vecteur $\mathbf{x}_{m_1+m_2}$ qui maximise $\|\mathbf{B} \mathbf{x}_{m_1+m_2}\|$ est de la forme

$$\begin{bmatrix} \mathbf{0}_{1,m'_1} \left(\mathbf{u}_{m-1}^1\right)^T \left(\mathbf{v}_{m-1}^1\right)^T \mathbf{0}_{1,m'_2} \mathbf{0}_{1,m'_1} \left(\mathbf{u}_{m-1}^2\right)^T \left(\mathbf{v}_{m-1}^2\right)^T \mathbf{0}_{1,m'_2} \end{bmatrix}^T \quad (\text{A.12})$$

$$\begin{aligned} \mathbf{B} \mathbf{x}_{m_1+m_2} &= \mathbf{B}^\dagger (\mathbf{I}_2 \otimes \mathbf{I}_1^z) \mathbf{x}_a \\ &= \mathbf{B}^\dagger \begin{bmatrix} \left(\mathbf{u}_{m-1}^1\right)^T \mathbf{0}_{1,m+1} \left(\mathbf{v}_{m-1}^1\right)^T \left(\mathbf{u}_{m-1}^2\right)^T \mathbf{0}_{1,m+1} \left(\mathbf{v}_{m-1}^2\right)^T \end{bmatrix}^T \\ &= \mathbf{B}^\dagger \left(\mathbf{I}_2 \otimes \begin{bmatrix} \mathbf{I}_m & \mathbf{0}_m \\ \mathbf{0}_{m+1,m} & \mathbf{0}_{m+1,m} \\ \mathbf{0}_m & \mathbf{I}_m \end{bmatrix} \right) \begin{bmatrix} \mathbf{u}_{m-1}^1 \\ \mathbf{v}_{m-1}^1 \\ \mathbf{u}_{m-1}^2 \\ \mathbf{v}_{m-1}^2 \end{bmatrix} \\ &= \mathbf{B}^\dagger \begin{bmatrix} \mathbf{u}_{m-1}^1 \\ \mathbf{v}_{m-1}^1 \\ \mathbf{u}_{m-1}^2 \\ \mathbf{v}_{m-1}^2 \end{bmatrix} \end{aligned}$$

où $\mathbf{B}^\dagger \triangleq \mathbf{B}^\dagger \left(\mathbf{I}_2 \otimes \begin{bmatrix} \mathbf{I}_m & \mathbf{0}_m \\ \mathbf{0}_{m+1,m} & \mathbf{0}_{m+1,m} \\ \mathbf{0}_m & \mathbf{I}_m \end{bmatrix} \right)$ n'est pas fonction de m'_1 ni de m'_2 . Donc elle n'est pas fonction ni de m_1 ni de m_2 .

$\mathbf{x}_{m_1+m_2}$ est un vecteur singulier de $\mathbf{B}$ ssi il vérifie (A.12) et $\begin{bmatrix} \mathbf{u}_{m-1}^1 \\ \mathbf{v}_{m-1}^1 \\ \mathbf{u}_{m-1}^2 \\ \mathbf{v}_{m-1}^2 \end{bmatrix}$ est le vecteur normé singulier à droite de $\mathbf{B}^\dagger$. Il n'est donc pas fonction ni de m_1 ni de m_2 .

Reste à montrer que $\sigma_1(\mathbf{B})$ elle aussi n'est pas fonction ni de m_1 ni de m_2 . Il est facile de voir que

$$\sigma_1(\mathbf{B}) = \|\mathbf{B} \mathbf{x}_{m_1+m_2}\| = \left\| \mathbf{B}^\dagger \begin{bmatrix} \mathbf{u}_{m-1}^1 \\ \mathbf{v}_{m-1}^1 \\ \mathbf{u}_{m-1}^2 \\ \mathbf{v}_{m-1}^2 \end{bmatrix} \right\| = \sigma_1(\mathbf{B}^\dagger) \text{ laquelle n'est pas fonction ni de } m_1 \text{ ni de } m_2.$$

Annexe B

Différentiation des algorithmes d'identification

Dans ce chapitre, on calcule les différentielles, au sens défini dans §7.3, des algorithmes sous-espace (SS), prédiction linéaire (LP) et décomposition en produit extérieur (OPD), rappelés dans §4.4.1, §4.4.2 et §4.4.3, respectivement. Les démonstrations sont faites dans le cas $c = 2$ et font appels à la théorie de perturbation des matrices (§E).

On commence par donner le résultat suivant, utile pour la suite des calculs.

Proposition B.1 *Soient $\mathbf{A}_k$, $k \in \{0, \dots, D\}$ des matrices $c \times c$*

$$\text{Vec} \left(\begin{bmatrix} \mathbf{A}_0 & \mathbf{A}_1 & \cdots & \mathbf{A}_D \\ \mathbf{A}_1 & \mathbf{A}_2 & & \mathbf{0}_c \\ \vdots & & & \vdots \\ \mathbf{A}_D & \mathbf{0}_c & \cdots & \mathbf{0}_c \end{bmatrix} \right) = \overline{\mathcal{S}}_{c,D} \text{Vec} \left(\begin{bmatrix} \mathbf{A}_0 & \mathbf{A}_1 & \cdots & \mathbf{A}_D \\ \mathbf{A}_1^H & \mathbf{A}_0 & \ddots & \vdots \\ \vdots & \ddots & \ddots & \mathbf{A}_1 \\ \mathbf{A}_D^H & \cdots & \mathbf{A}_1^H & \mathbf{A}_0 \end{bmatrix} \right)$$

avec

$$\begin{aligned} \overline{\mathcal{S}}_{c,D} &= \begin{bmatrix} \mathbf{0}_{c^2(D+1)^2, c^2(D+1)D} & \overline{\mathcal{S}}_{c,D}^\dagger \end{bmatrix} \\ \overline{\mathcal{S}}_{c,D}^\dagger &= \begin{bmatrix} (\mathbf{I}_c \otimes \mathbf{I}_{D+1}^\dagger) \\ (\mathbf{I}_c \otimes \mathbf{J}_{D+1}^T \mathbf{I}_{D+1}^\dagger) \\ \vdots \\ (\mathbf{I}_c \otimes \mathbf{J}_{D+1}^{-D} \mathbf{I}_{D+1}^\dagger) \end{bmatrix} \otimes \mathbf{I}_c \\ &= \begin{bmatrix} \begin{bmatrix} \mathbf{I}_{c(D+1)} \\ (\mathbf{I}_c \otimes \mathbf{J}_{D+1}^T) \\ \vdots \\ (\mathbf{I}_c \otimes \mathbf{J}_{D+1}^{-D}) \end{bmatrix} (\mathbf{I}_c \otimes \mathbf{I}_{D+1}^\dagger) \end{bmatrix} \otimes \mathbf{I}_c \\ &= \begin{bmatrix} \begin{bmatrix} \mathbf{I}_{c(D+1)} \\ (\mathbf{I}_c \otimes \mathbf{J}_{D+1}^T) \\ \vdots \\ (\mathbf{I}_c \otimes \mathbf{J}_{D+1}^T)^D \end{bmatrix} (\mathbf{I}_c \otimes \mathbf{I}_{D+1}^\dagger) \end{bmatrix} \otimes \mathbf{I}_c \end{aligned}$$

B.1 Estimation du bruit

La puissance du bruit est estimée par (cf §4.3) la plus petite valeur propre de la matrice de corrélation. Si $\mathbf{X}$ est $2(M+1) \times 2(M+1)$, $\sigma_b^2(\mathbf{X}) = \lambda_{2(M+1)}(\mathbf{X})$ et

$$\begin{aligned}\sigma_b^2(\mathbf{X} + \mathbf{E}) &= \lambda_{2(M+1)}(\mathbf{X}) + \text{vecp}(\mathbf{X})^H \mathbf{E} \text{vecp}(\mathbf{X}) + O(\|\mathbf{E}\|^2) \\ &= \sigma_b^2(\mathbf{X}) + \left(\text{vecp}(\mathbf{X})^T \otimes \text{vecp}(\mathbf{X})^H \right) \text{Vec}(\mathbf{E}) + O(\|\mathbf{E}\|^2)\end{aligned}$$

Proposition B.2 *Soient $\mathbf{X}$ une matrice Hermitienne $2(M+1) \times 2(M+1)$ définie positive et $\mathbf{E}$ une matrice perturbation 'faible'. Alors*

$$\text{Vec} \left(\delta \left(\mathbf{X} - \sigma_b^2(\mathbf{X}) \mathbf{I}_{2(M+1)} \right) \right) = \bar{\mathcal{G}}(\mathbf{X}) \text{Vec}(\mathbf{E}) + O(\|\mathbf{E}\|^2)$$

$$\text{avec}^1 \quad \bar{\mathcal{G}}(\mathbf{X}) = \mathbf{I}_{4(M+1)^2} - \text{Vec} \left(\mathbf{I}_{2(M+1)} \right) \left(\text{vecp}(\mathbf{X})^T \otimes \text{vecp}(\mathbf{X})^H \right).$$

Démonstration:

$$\begin{aligned}\text{Vec} \left(\mathbf{E} - \delta \sigma_b^2(\mathbf{X}) \mathbf{I}_{2(M+1)} \right) &= \text{Vec}(\mathbf{E}) - \left[\left(\text{vecp}(\mathbf{X})^T \otimes \text{vecp}(\mathbf{X})^H \right) \text{Vec}(\mathbf{E}) \right] \text{Vec} \left(\mathbf{I}_{2(M+1)} \right) + O(\|\mathbf{E}\|^2) \\ &= \text{Vec}(\mathbf{E}) - \text{Vec} \left(\mathbf{I}_{2(M+1)} \right) \left[\left(\text{vecp}(\mathbf{X})^T \otimes \text{vecp}(\mathbf{X})^H \right) \text{Vec}(\mathbf{E}) \right] + O(\|\mathbf{E}\|^2) \\ &= \text{Vec}(\mathbf{E}) - \text{Vec} \left(\mathbf{I}_{2(M+1)} \right) \left(\text{vecp}(\mathbf{X})^T \otimes \text{vecp}(\mathbf{X})^H \right) \text{Vec}(\mathbf{E}) + O(\|\mathbf{E}\|^2)\end{aligned}$$

■

Proposition B.3

$$\bar{\mathcal{G}}(\mathbf{X}) \mathbf{K}_{2(M+1)} = \mathbf{K}_{2(M+1)} \bar{\mathcal{G}}^*(\mathbf{X})$$

Démonstration:

$$\begin{aligned}\bar{\mathcal{G}}(X) \mathbf{K}_{2(M+1)} &= \mathbf{K}_{2(M+1)} - \text{Vec} \left(\mathbf{I}_{2(M+1)} \right) \left(\text{vecp}(\mathbf{X})^T \otimes \text{vecp}(\mathbf{X})^H \right) \mathbf{K}_{2(M+1)} \\ &= \mathbf{K}_{2(M+1)} - \text{Vec} \left(\mathbf{I}_{2(M+1)} \right) \left(\text{vecp}(\mathbf{X})^H \otimes \text{vecp}(\mathbf{X})^T \right) \\ &= \mathbf{K}_{2(M+1)} - \text{Vec} \left(\mathbf{I}_{2(M+1)}^T \right) \left(\text{vecp}(\mathbf{X})^H \otimes \text{vecp}(\mathbf{X})^T \right) \\ &= \mathbf{K}_{2(M+1)} \left[\mathbf{I}_{4(M+1)^2} - \text{Vec} \left(\mathbf{I}_{2(M+1)} \right) \left(\text{vecp}(\mathbf{X})^H \otimes \text{vecp}(\mathbf{X})^T \right) \right] \\ &= \mathbf{K}_{2(M+1)} \bar{\mathcal{G}}^*(X)\end{aligned}$$

■

¹ $\bar{\mathcal{G}}(\mathbf{X})$ est indépendante du choix (à une phase près) de $\text{vecp}(\mathbf{X})$.

B.2 L'algorithme sous-espace

L'application de l'algorithme SS à une matrice $\mathbf{X}$ Hermitienne donne $\text{alg}(\mathbf{X}) = \mathbf{T}_{M+1} \text{vecp}(\mathbf{X})^*$.

Proposition B.4 *Soit $\mathbf{X} \in \mathbf{C}^{2(M+1)} \times \mathbf{C}^{2(M+1)}$ Hermitienne pour laquelle la valeur propre la plus petite est simple (de multiplicité 1). Alors*

$$\begin{aligned} \mathbf{A}_{\mathbf{X}}^1 &= \mathbf{0}_{2(M+1), 4(M+1)^2} \\ \mathbf{A}_{\mathbf{X}}^2 &= -\mathbf{T}_{M+1} \left(\text{vecp}(\mathbf{X})^H \otimes \left(\mathbf{X}^T - \lambda_{2(M+1)}(\mathbf{X})^* \mathbf{I}_{2(M+1)} \right)^{\natural} \right) \\ \mathbf{A}_{\mathbf{X}} &= -\mathbf{T}_{M+1} \left(\left(\mathbf{X}^T - \lambda_{2(M+1)}^*(\mathbf{X}) \mathbf{I}_{2(M+1)} \right)^{\natural} \otimes \text{vecp}(\mathbf{X})^H \right) \end{aligned}$$

Démonstration:

Soit $\mathbf{E} \in \mathbf{C}^{2(M+1)} \times \mathbf{C}^{2(M+1)}$ une matrice perturbation 'faible'. On a la relation suivante :

$$\begin{aligned} \text{alg}(\mathbf{X} + \mathbf{E}) &= \text{alg}(\mathbf{X}) - \mathbf{T}_{M+1} \left[\left(\mathbf{X} - \lambda_{2(M+1)}(\mathbf{X}) \mathbf{I}_{2(M+1)} \right)^{\natural} \mathbf{E} \text{vecp}(\mathbf{X}) \right]^* + \mathcal{O}(\|\mathbf{E}\|^2) \\ &= \text{alg}(\mathbf{X}) - \mathbf{T}_{M+1} \text{Vec} \left(\left(\mathbf{X} - \lambda_{2(M+1)}(\mathbf{X}) \mathbf{I}_{2(M+1)} \right)^{\natural} \mathbf{E} \text{vecp}(\mathbf{X}) \right)^* + \mathcal{O}(\|\mathbf{E}\|^2) \\ &= \text{alg}(\mathbf{X}) - \mathbf{T}_{M+1} \left(\text{vecp}(\mathbf{X})^H \otimes \left(\mathbf{X} - \lambda_{2(M+1)}(\mathbf{X}) \mathbf{I}_{2(M+1)} \right)^{\natural*} \right) \text{Vec}(\mathbf{E}^*) + \mathcal{O}(\|\mathbf{E}\|^2) \\ &= \text{alg}(\mathbf{X}) - \mathbf{T}_{M+1} \left(\text{vecp}(\mathbf{X})^H \otimes \left(\mathbf{X} - \lambda_{2(M+1)}(\mathbf{X}) \mathbf{I}_{2(M+1)} \right)^{* \natural} \right) \text{Vec}(\mathbf{E}^*) + \mathcal{O}(\|\mathbf{E}\|^2) \\ &= \text{alg}(\mathbf{X}) - \mathbf{T}_{M+1} \left(\text{vecp}(\mathbf{X})^H \otimes \left(\mathbf{X}^T - \lambda_{2(M+1)}(\mathbf{X})^* \mathbf{I}_{2(M+1)} \right)^{\natural} \right) \text{Vec}(\mathbf{E}^*) + \mathcal{O}(\|\mathbf{E}\|^2) \end{aligned}$$

■

B.3 L'algorithme prédiction linéaire

On rappelle ici les étapes et les paramètres intermédiaires au calcul de $\text{alg}(\mathbf{X}) = LP(\mathbf{X})$ où $\mathbf{X}$ est une matrice Hermitienne définie positive $c(M+1) \times c(M+1)$.

$$\begin{aligned} \mathbf{Y} &\triangleq \mathbf{X} - \sigma_b^2(\mathbf{X}) \mathbf{I}_{c(M+1)} \quad \text{Hermitienne définie positive} \\ \mathbf{Z} &\triangleq [\mathbf{0}_{cM, c} \quad \mathbf{I}_{cM}] \mathbf{Y} [\mathbf{0}_{cM, c} \quad \mathbf{I}_{cM}]^T \\ &= [\mathbf{0}_{cM, c} \quad \mathbf{I}_{cM}] \mathbf{X} [\mathbf{0}_{cM, c} \quad \mathbf{I}_{cM}]^T - \sigma_b^2(\mathbf{X}) \mathbf{I}_{cM} \quad \text{Hermitienne définie positive} \\ &= \mathbf{X}(c+1 : c(M+1), c+1 : c(M+1)) \quad (\text{notations MATLAB}) \\ [\gamma(0)\gamma(1)\cdots\gamma(M)] &\triangleq [\mathbf{I}_c \quad \mathbf{0}_{c, cM}] \mathbf{Y} \quad , \text{ où } \gamma(i) \text{ est } c \times c \\ [\mathbf{A}_1 \cdots \mathbf{A}_M] &\triangleq -[\gamma(1) \cdots \gamma(M)] \mathbf{Z}^{\natural} \\ \mathbf{D} &\triangleq [\mathbf{I}_c \mathbf{A}_1 \cdots \mathbf{A}_M] [\gamma(0)\gamma(1)\cdots\gamma(M)]^H \\ &= \gamma(0) - [\gamma(1) \cdots \gamma(M)] \mathbf{Z}^{\natural} [\gamma(1) \cdots \gamma(M)]^H \end{aligned}$$

différence entre deux matrices Hermitiennes définies positives $c \times c$

On émet l'hypothèse non restrictive qu'elle n'est pas strictement définie négative

$$\begin{aligned}
(\lambda, \mathbf{l}) &\triangleq \text{la plus grande valeur propre de } \mathbf{D} \text{ et le vecteur propre normé associé} \\
\mathbf{f} &\triangleq \frac{1}{\sqrt{\lambda}} \mathbf{l} \\
\mathbf{g} &\triangleq [\mathbf{I}_c \mathbf{A}_1 \cdots \mathbf{A}_M]^T \mathbf{f}^* \\
\mathbf{Y}^\circ &\triangleq \mathbf{Han}(\gamma(0), \dots, \gamma(M)) = \begin{bmatrix} \gamma(0) & \gamma(1) & \cdots & \gamma(M) \\ \gamma(1) & \gamma(2) & \cdots & \mathbf{0}_c \\ \vdots & & & \vdots \\ \gamma(M) & \mathbf{0}_c & \cdots & \mathbf{0}_c \end{bmatrix} \\
\text{alg}(\mathbf{X}) &\triangleq \mathbf{Y}^\circ \mathbf{g}^*
\end{aligned} \tag{B.1}$$

On peut montrer que

$$\begin{aligned}
\mathbf{Y} &= \begin{bmatrix} \gamma(0) & [\gamma(1) \cdots \gamma(M)] \\ [\gamma(1) \cdots \gamma(M)]^H & \mathbf{Z} \end{bmatrix} \\
[\gamma(1) \cdots \gamma(M)] &= [\mathbf{I}_c \mathbf{0}_{c,cM}] \mathbf{Y} [\mathbf{0}_{cM,c} \mathbf{I}_{cM}]^T \\
[\mathbf{A}_1 \cdots \mathbf{A}_M] &= [\mathbf{I}_c \mathbf{A}_1 \cdots \mathbf{A}_M] [\mathbf{0}_{cM,c} \mathbf{I}_{cM}]^T \\
[\mathbf{I}_c \mathbf{A}_1 \cdots \mathbf{A}_M] \mathbf{Y} &= \left[[\mathbf{I}_c \mathbf{A}_1 \cdots \mathbf{A}_M] [\gamma(0) \gamma(1) \cdots \gamma(M)]^H \quad [\gamma(1) \cdots \gamma(M)] - [\mathbf{A}_1 \cdots \mathbf{A}_M] \mathbf{Z} \right] \\
&= [\mathbf{D} \quad \mathbf{0}_{c,cM}]
\end{aligned} \tag{B.2}$$

Proposition B.5 Soit $\mathbf{X}$ une matrice Hermitienne définie positive $2(M+1) \times 2(M+1)$.

$$\begin{aligned}
\mathbf{A}_\mathbf{X}^1 &= \Delta_6 \overline{\mathcal{G}}(\mathbf{X}) \\
\mathbf{A}_\mathbf{X}^2 &= - \left[\left(\mathbf{Y}^\circ \begin{bmatrix} \mathbf{0}_c & \mathbf{0}_{c,cM} \\ \mathbf{0}_{cM,c} & \mathbf{Z}^{-1} \end{bmatrix} \right) \otimes (\mathbf{g}^H) \right] \overline{\mathcal{G}}^*(\mathbf{X}) \\
\mathbf{A}_\mathbf{X} &= \left(\Delta_6 - \left[(\mathbf{g}^H) \otimes \left(\mathbf{Y}^\circ \begin{bmatrix} \mathbf{0}_c & \mathbf{0}_{c,cM} \\ \mathbf{0}_{cM,c} & \mathbf{Z}^{-1} \end{bmatrix} \right) \right] \right) \overline{\mathcal{G}}(\mathbf{X})
\end{aligned}$$

avec

$$\begin{aligned}
\Delta_6 &\triangleq (\mathbf{g}^H \otimes \mathbf{I}_{c(M+1)}) \overline{\mathcal{S}}_{2,M} + \mathbf{Y}^\circ [\mathbf{I}_2 \mathbf{A}_1 \cdots \mathbf{A}_M]^H \Delta_3 \\
\Delta_3 &\triangleq -\frac{1}{\sqrt{\lambda}} \left[(\mathbf{I}^T \otimes (\mathbf{D} - \lambda \mathbf{I}_2)^\natural) + \frac{1}{2\lambda} [(\mathbf{I}^T) \otimes (\mathbf{I}^H)] \right] ([\mathbf{I}_2 \mathbf{A}_1 \cdots \mathbf{A}_M]^* \otimes [\mathbf{I}_2 \mathbf{A}_1 \cdots \mathbf{A}_M])
\end{aligned}$$

Démonstration:

²Soit $\mathbf{E}$ une matrice 'perturbation' faible.

$$\begin{aligned}
\delta(\mathbf{Z}^\natural) &= -\mathbf{Z}^\natural \delta \mathbf{Z} \mathbf{Z}^\natural + \mathbf{Z}^{\natural^2} \delta \mathbf{Z}^H \mathbf{P} - \mathbf{P} \delta \mathbf{Z}^H \mathbf{Z}^{\natural^2} + \mathcal{O}(\|\delta \mathbf{Y}\|^2) \\
\delta[\mathbf{A}_1 \cdots \mathbf{A}_m] &= -\delta[\gamma(1) \cdots \gamma(M)] \mathbf{Z}^\natural - [\gamma(1) \cdots \gamma(M)] \delta(\mathbf{Z}^\natural)
\end{aligned}$$

²Démonstration inspirée du calcul de différentiation dans [ABED 97] fait en réel et autour d'une matrice $\mathbf{X}$ dans $\mathcal{M}_M$.

$$\begin{aligned}
& + O\left(\delta[\gamma(1) \cdots \gamma(M)] \delta(\mathbf{Z}^\natural)\right) \\
& = -[\mathbf{I}_c \mathbf{0}_{c,cM}] \delta \mathbf{Y} [\mathbf{0}_{cM,c} \mathbf{I}_{cM}]^T \mathbf{Z}^\natural - [\gamma(1) \cdots \gamma(M)] \delta(\mathbf{Z}^\natural) + O\left(\|\delta \mathbf{Y}\|^2\right) \\
[\gamma(1) \cdots \gamma(M)] \delta(\mathbf{Z}^\natural) & = -[\gamma(1) \cdots \gamma(M)] \mathbf{Z}^\natural \delta \mathbf{Z} \mathbf{Z}^\natural + [\gamma(1) \cdots \gamma(M)] \mathbf{Z}^{\natural^2} \delta \mathbf{Z}^H \mathbf{P} \\
& + [\mathbf{A}_1 \cdots \mathbf{A}_M] \mathbf{Z}^\natural \mathbf{P} \delta \mathbf{Z}^H \mathbf{Z}^{\natural^2} + O\left(\|\delta \mathbf{Y}\|^2\right) \\
& = [\mathbf{A}_1 \cdots \mathbf{A}_M] \delta \mathbf{Z} \mathbf{Z}^\natural - [\mathbf{A}_1 \cdots \mathbf{A}_M] \mathbf{Z}^\natural \delta \mathbf{Z}^H \mathbf{P} + O\left(\|\delta \mathbf{Y}\|^2\right)
\end{aligned}$$

On fait l'hypothèse non restrictive que $\mathbf{Z}$ est inversible, réduisant $\mathbf{P}$ à une matrice nulle et $\mathbf{Z}^\natural$ à $\mathbf{Z}^{-1}$.

$$\begin{aligned}
[\gamma(1) \cdots \gamma(M)] \delta(\mathbf{Z}^{-1}) & = [\mathbf{A}_1 \cdots \mathbf{A}_M] \delta \mathbf{Z} \mathbf{Z}^{-1} + O\left(\|\delta \mathbf{Y}\|^2\right) \\
\delta[\mathbf{A}_1 \cdots \mathbf{A}_m] & = -[\mathbf{I}_c \mathbf{0}_{c,cM}] \delta \mathbf{Y} [\mathbf{0}_{cM,c} \mathbf{I}_{cM}]^T \mathbf{Z}^{-1} - [\mathbf{A}_1 \cdots \mathbf{A}_M] \delta \mathbf{Z} \mathbf{Z}^{-1} + O\left(\|\delta \mathbf{Y}\|^2\right) \\
& = -[\mathbf{I}_c \mathbf{0}_{c,cM}] \delta \mathbf{Y} [\mathbf{0}_{cM,c} \mathbf{I}_{cM}]^T \mathbf{Z}^{-1} \\
& \quad - [\mathbf{A}_1 \cdots \mathbf{A}_M] [\mathbf{0}_{cM,c} \mathbf{I}_{cM}] \delta \mathbf{Y} [\mathbf{0}_{cM,c} \mathbf{I}_{cM}]^T \mathbf{Z}^{-1} + O\left(\|\delta \mathbf{Y}\|^2\right) \\
& = -([\mathbf{I}_c \mathbf{0}_{c,cM}] + [\mathbf{0}_c \mathbf{A}_1 \cdots \mathbf{A}_M]) \delta \mathbf{Y} [\mathbf{0}_{cM,c} \mathbf{I}_{cM}]^T \mathbf{Z}^{-1} + O\left(\|\delta \mathbf{Y}\|^2\right) \\
& = -[\mathbf{I}_c \mathbf{A}_1 \cdots \mathbf{A}_M] \delta \mathbf{Y} [\mathbf{0}_{cM,c} \mathbf{I}_{cM}]^T \mathbf{Z}^{-1} + O\left(\|\delta \mathbf{Y}\|^2\right) \\
\text{Vec}(\delta[\mathbf{A}_1 \cdots \mathbf{A}_m]) & = \Delta_1 \text{Vec}(\delta \mathbf{Y}) + O\left(\|\delta \mathbf{Y}\|^2\right) \\
\Delta_1 & = -\left[\left(\mathbf{Z}^{-T} [\mathbf{0}_{cM,c} \mathbf{I}_{cM}]\right) \otimes ([\mathbf{I}_c \mathbf{A}_1 \cdots \mathbf{A}_M])\right] \\
\delta \mathbf{D} & = [\mathbf{I}_c \mathbf{A}_1 \cdots \mathbf{A}_M] \delta[\gamma(0) \gamma(1) \cdots \gamma(M)]^H + \\
& \quad \delta[\mathbf{A}_1 \cdots \mathbf{A}_m] [\gamma(1) \cdots \gamma(M)]^H + O\left(\|\delta \mathbf{Y}\|^2\right)
\end{aligned}$$

$$\begin{aligned}
\text{En utilisant } [\gamma(1) \cdots \gamma(M)] \mathbf{Z}^{-1} & = -[\mathbf{A}_1 \cdots \mathbf{A}_M] \\
\text{et } [\gamma(1) \cdots \gamma(M)] \mathbf{P} & = -[\mathbf{A}_1 \cdots \mathbf{A}_M] \mathbf{Z} \mathbf{P} = \mathbf{0}
\end{aligned}$$

On montre que :

$$\begin{aligned}
& \delta[\mathbf{A}_1 \cdots \mathbf{A}_m] [\gamma(1) \cdots \gamma(M)]^H \\
& = -\delta[\gamma(1) \cdots \gamma(M)] \mathbf{Z}^{-1} [\gamma(1) \cdots \gamma(M)]^H - [\gamma(1) \cdots \gamma(M)] \delta \mathbf{Z}^{-1} [\gamma(1) \cdots \gamma(M)]^H + O\left(\|\delta \mathbf{Y}\|^2\right) \\
& = \delta[\gamma(1) \cdots \gamma(M)] [\mathbf{A}_1 \cdots \mathbf{A}_M]^H + [\gamma(1) \cdots \gamma(M)] \mathbf{Z}^{-1} \delta \mathbf{Z} \mathbf{Z}^{-1} [\gamma(1) \cdots \gamma(M)]^H \\
& \quad + [\gamma(1) \cdots \gamma(M)] \mathbf{P} \delta \mathbf{Z} \mathbf{Z}^{\natural^2} [\gamma(1) \cdots \gamma(M)]^H \\
& \quad - [\gamma(1) \cdots \gamma(M)] \mathbf{Z}^{\natural^2} \delta \mathbf{Z} \mathbf{P} [\gamma(1) \cdots \gamma(M)]^H + O\left(\|\delta \mathbf{Y}\|^2\right) \\
& = \delta[\gamma(1) \cdots \gamma(M)] [\mathbf{A}_1 \cdots \mathbf{A}_M]^H + [\mathbf{A}_1 \cdots \mathbf{A}_M] \delta \mathbf{Z} [\mathbf{A}_1 \cdots \mathbf{A}_M]^H + O\left(\|\delta \mathbf{Y}\|^2\right)
\end{aligned}$$

$$\begin{aligned}
\delta \mathbf{D} & = [\mathbf{I}_c \mathbf{A}_1 \cdots \mathbf{A}_M] \delta \mathbf{Y} [\mathbf{I}_c \mathbf{0}_{c,cM}]^T + [\mathbf{I}_c \mathbf{0}_{c,cM}] \delta \mathbf{Y} [\mathbf{0}_{cM,c} \mathbf{I}_{cM}]^T [\mathbf{A}_1 \cdots \mathbf{A}_M]^H \\
& \quad + [\mathbf{A}_1 \cdots \mathbf{A}_M] [\mathbf{0}_{cM,c} \mathbf{I}_{cM}] \delta \mathbf{Y} [\mathbf{0}_{cM,c} \mathbf{I}_{cM}]^T [\mathbf{A}_1 \cdots \mathbf{A}_M]^H + O\left(\|\delta \mathbf{Y}\|^2\right) \\
& = [\mathbf{I}_c \mathbf{A}_1 \cdots \mathbf{A}_M] \delta \mathbf{Y} [\mathbf{I}_c \mathbf{0}_{c,cM}]^T + [\mathbf{I}_c \mathbf{0}_{c,cM}] \delta \mathbf{Y} [\mathbf{0}_c \mathbf{A}_1 \cdots \mathbf{A}_M]^H \\
& \quad + [\mathbf{0}_c \mathbf{A}_1 \cdots \mathbf{A}_M] \delta \mathbf{Y} [\mathbf{0}_c \mathbf{A}_1 \cdots \mathbf{A}_M]^H + O\left(\|\delta \mathbf{Y}\|^2\right) \\
& = [\mathbf{I}_c \mathbf{A}_1 \cdots \mathbf{A}_M] \delta \mathbf{Y} [\mathbf{I}_c \mathbf{0}_{c,cM}]^T + [\mathbf{I}_c \mathbf{A}_1 \cdots \mathbf{A}_M] \delta \mathbf{Y} [\mathbf{0}_c \mathbf{A}_1 \cdots \mathbf{A}_M]^H + O\left(\|\delta \mathbf{Y}\|^2\right)
\end{aligned}$$

$$\begin{aligned}
&= [\mathbf{I}_c \mathbf{A}_1 \cdots \mathbf{A}_M] \delta \mathbf{Y} [\mathbf{I}_c \mathbf{A}_1 \cdots \mathbf{A}_M]^H + \mathcal{O}(\|\delta \mathbf{Y}\|^2) \\
\text{Vec}(\delta \mathbf{D}) &= \Delta_2 \text{Vec}(\delta \mathbf{Y}) + \mathcal{O}(\|\delta \mathbf{Y}\|^2) \\
\Delta_2 &= ([\mathbf{I}_c \mathbf{A}_1 \cdots \mathbf{A}_M]^* \otimes [\mathbf{I}_c \mathbf{A}_1 \cdots \mathbf{A}_M])
\end{aligned}$$

On fait l'hypothèse non restrictive que la valeur propre la plus grande de $\mathbf{D}$ est de multiplicité 1 et est réelle positive.

$$\begin{aligned}
\delta \mathbf{l} &= -(\mathbf{D} - \lambda \mathbf{I}_c)^\sharp \delta \mathbf{D} \mathbf{l} + \mathcal{O}(\|\delta \mathbf{Y}\|^2) \\
\delta \lambda &= \mathbf{l}^H \delta \mathbf{D} \mathbf{l} + \mathcal{O}(\|\delta \mathbf{Y}\|^2) \\
\delta \mathbf{f} &= \frac{1}{\sqrt{\lambda}} \delta \mathbf{l} - \frac{1}{2\lambda^{3/2}} \mathbf{l} \delta \lambda \\
&= -\frac{1}{\sqrt{\lambda}} (\mathbf{D} - \lambda \mathbf{I}_c)^\sharp \delta \mathbf{D} \mathbf{l} - \frac{1}{2\lambda^{3/2}} \mathbf{l}^H \delta \mathbf{D} \mathbf{l} + \mathcal{O}(\|\delta \mathbf{Y}\|^2) \\
&= \Delta_3 \text{Vec}(\delta \mathbf{Y}) + \mathcal{O}(\|\delta \mathbf{Y}\|^2) \\
\Delta_3 &= -\frac{1}{\sqrt{\lambda}} \left[(\mathbf{l}^T \otimes (\mathbf{D} - \lambda \mathbf{I}_c)^\sharp) + \frac{1}{2\lambda} [(\mathbf{l}^T) \otimes (\mathbf{l}^H)] \right] \Delta_2 \\
\delta \mathbf{g} &= [\mathbf{0}_c \ \delta [\mathbf{A}_1 \cdots \mathbf{A}_m]]^T \mathbf{f}^* + [\mathbf{I}_c \mathbf{A}_1 \cdots \mathbf{A}_M]^T \delta \mathbf{f}^* \\
&= [\mathbf{0}_{cM,c} \mathbf{I}_{cM}]^T \delta [\mathbf{A}_1 \cdots \mathbf{A}_m]^T \mathbf{f}^* + [\mathbf{I}_c \mathbf{A}_1 \cdots \mathbf{A}_M]^T \delta \mathbf{f}^* \\
&= \Delta_4 \text{Vec}(\delta \mathbf{Y}) + \Delta_5 \text{Vec}(\delta \mathbf{Y})^* + \mathcal{O}(\|\delta \mathbf{Y}\|^2) \\
\Delta_4 &= (\mathbf{f}^H \otimes [\mathbf{0}_{cM,c} \mathbf{I}_{cM}]^T) \mathbf{K}_{c,cM} \Delta_1 \\
&= ([\mathbf{0}_{cM,c} \mathbf{I}_{cM}]^T \otimes \mathbf{f}^H) \Delta_1 \\
&= -([\mathbf{0}_{cM,c} \mathbf{I}_{cM}]^T \otimes \mathbf{f}^H) \left[(\mathbf{Z}^{-T} [\mathbf{0}_{cM,c} \mathbf{I}_{cM}]) \otimes ([\mathbf{I}_c \mathbf{A}_1 \cdots \mathbf{A}_M]) \right] \\
&= -\left[([\mathbf{0}_{cM,c} \mathbf{I}_{cM}]^T \mathbf{Z}^{-T} [\mathbf{0}_{cM,c} \mathbf{I}_{cM}]) \otimes (\mathbf{g}^T) \right] \\
&= -\left(\begin{bmatrix} \mathbf{0}_c & \mathbf{0}_{c,cM} \\ \mathbf{0}_{cM,c} & \mathbf{Z}^{-T} \end{bmatrix} \otimes \mathbf{g}^T \right) \\
\Delta_5 &= [\mathbf{I}_c \mathbf{A}_1 \cdots \mathbf{A}_M]^T \Delta_3^* \\
\delta \mathbf{h} &= \delta \mathbf{Y}^\circ \mathbf{g}^* + \mathbf{Y}^\circ \delta \mathbf{g}^* \\
&= (\mathbf{g}^H \otimes \mathbf{I}_{c(M+1)}) \overline{\mathcal{S}}_{c,M} \text{Vec}(\delta \mathbf{Y}) + \mathbf{Y}^\circ (\Delta_4 \text{Vec}(\delta \mathbf{Y}) + \Delta_5 \text{Vec}(\delta \mathbf{Y})^*)^* + \mathcal{O}(\|\delta \mathbf{Y}\|^2) \\
&= \Delta_6 \text{Vec}(\delta \mathbf{Y}) + \Delta_7 \text{Vec}(\delta \mathbf{Y})^* + \mathcal{O}(\|\delta \mathbf{Y}\|^2) \\
\Delta_6 &= (\mathbf{g}^H \otimes \mathbf{I}_{c(M+1)}) \overline{\mathcal{S}}_{c,M} + \mathbf{Y}^\circ \Delta_5^* \\
\Delta_7 &= \mathbf{Y}^\circ \Delta_4^* \\
&= -\mathbf{Y}^\circ \left(\begin{bmatrix} \mathbf{0}_c & \mathbf{0}_{c,cM} \\ \mathbf{0}_{cM,c} & \mathbf{Z}^{-H} \end{bmatrix} \otimes \mathbf{g}^H \right) \\
&= -(1 \otimes \mathbf{Y}^\circ) \left(\begin{bmatrix} \mathbf{0}_c & \mathbf{0}_{c,cM} \\ \mathbf{0}_{cM,c} & \mathbf{Z}^{-1} \end{bmatrix} \otimes \mathbf{g}^H \right) \\
&= -\left[\left(\mathbf{Y}^\circ \begin{bmatrix} \mathbf{0}_c & \mathbf{0}_{c,cM} \\ \mathbf{0}_{cM,c} & \mathbf{Z}^{-1} \end{bmatrix} \right) \otimes (\mathbf{g}^H) \right]
\end{aligned}$$

Si $c = 2$

$$\text{Vec}(\delta \mathbf{Y}) = \overline{\mathcal{G}}(\mathbf{X}) \text{Vec}(\mathbf{E}) + \mathcal{O}(\|\mathbf{E}\|^2)$$

$$\begin{aligned}
O(\|\delta \mathbf{Y}\|) &= O(\|\mathbf{E}\|) \\
\delta \mathbf{h} &= \Delta_6 \text{Vec}(\delta \mathbf{Y}) + \Delta_7 \text{Vec}(\delta \mathbf{Y})^* + O(\|\delta \mathbf{Y}\|^2) \\
&= \Delta_6 \overline{\mathcal{G}}(\mathbf{X}) \text{Vec}(\mathbf{E}) + \Delta_7 \overline{\mathcal{G}}^*(\mathbf{X}) \text{Vec}(\mathbf{E})^* + O(\|\delta \mathbf{Y}\|^2)
\end{aligned}$$

d'où les expressions de $\mathbf{A}_{\mathbf{X}}^1$ et $\mathbf{A}_{\mathbf{X}}^2$.

$$\begin{aligned}
\mathbf{A}_{\mathbf{X}} &= \mathbf{A}_{\mathbf{X}}^1 + \mathbf{A}_{\mathbf{X}}^2 \mathbf{K}_{2(M+1)} \\
\mathbf{A}_{\mathbf{X}}^2 \mathbf{K}_{2(M+1)} &= \Delta_7 \overline{\mathcal{G}}^*(\mathbf{X}) \mathbf{K}_{2(M+1)} \\
&= \Delta_7 \mathbf{K}_{2(M+1)} \overline{\mathcal{G}}(\mathbf{X}) \\
\Delta_7 \mathbf{K}_{2(M+1)} &= - \left[\left(\mathbf{Y}^\circ \begin{bmatrix} \mathbf{0}_c & \mathbf{0}_{c,cM} \\ \mathbf{0}_{cM,c} & \mathbf{Z}^{-1} \end{bmatrix} \right) \otimes (\mathbf{g}^H) \right] \mathbf{K}_{2(M+1)} \\
&= - \left[(\mathbf{g}^H) \otimes \left(\mathbf{Y}^\circ \begin{bmatrix} \mathbf{0}_c & \mathbf{0}_{c,cM} \\ \mathbf{0}_{cM,c} & \mathbf{Z}^{-1} \end{bmatrix} \right) \right] \\
\mathbf{A}_{\mathbf{X}} &= \Delta_6 \overline{\mathcal{G}}(\mathbf{X}) - \left[(\mathbf{g}^H) \otimes \left(\mathbf{Y}^\circ \begin{bmatrix} \mathbf{0}_c & \mathbf{0}_{c,cM} \\ \mathbf{0}_{cM,c} & \mathbf{Z}^{-1} \end{bmatrix} \right) \right] \overline{\mathcal{G}}(\mathbf{X})
\end{aligned}$$

■

Proposition B.6 Soient $\mathbf{x}_M \in \mathbf{C}^{c(M+1)}$ et $x \in \mathbf{R}^+$.

$$\begin{aligned}
\text{Soit } \mathbf{X} &= \mathcal{T}_M(\mathbf{x}_M) \mathcal{T}_M^H(\mathbf{x}_M) + x \mathbf{I}_{c(M+1)} \in \mathcal{M}_M \\
\text{Alors } \Delta_6 &= (\mathbf{g}^H \otimes \mathbf{I}_{c(M+1)}) \overline{\mathcal{S}}_{c,M} - \frac{1}{2} [(\mathbf{g}^H) \otimes (\mathbf{x}_M \mathbf{g}^T)]
\end{aligned}$$

Démonstration:

$$\begin{aligned}
\mathbf{Y} &= \sigma_s^2 \mathcal{T}_{M+1}(\mathbf{x}_M) \mathcal{T}_{M+1}^H(\mathbf{x}_M) \\
\mathbf{Y}^\circ &= \sigma_s^2 \mathcal{T}^\circ(\mathbf{x}_M) \mathcal{T}_{M+1}^H(\mathbf{x}_M) \\
\mathbf{D} &= \lambda \mathbb{I}^H \\
\mathbf{D} - \lambda \mathbf{I}_c &= \lambda (\mathbb{I}^H - \mathbf{I}_c)
\end{aligned}$$

$(\mathbb{I}^H - \mathbf{I}_c)$ est un projecteur. Il est égale à sa propre pseudo-inverse. $(\mathbf{D} - \lambda \mathbf{I}_c)^\sharp = \frac{1}{\lambda} (\mathbb{I}^H - \mathbf{I}_c)$.

$$\begin{aligned}
\Delta_6 &= (\mathbf{g}^H \otimes \mathbf{I}_{c(M+1)}) \overline{\mathcal{S}}_{c,M} + \mathbf{Y}^\circ \Delta_5^* \\
\Delta_5 &= [\mathbf{I}_c \mathbf{A}_1 \cdots \mathbf{A}_M]^T \Delta_3^* \\
\Delta_5^* &= [\mathbf{I}_c \mathbf{A}_1 \cdots \mathbf{A}_M]^H \Delta_3 \\
\Delta_3 &= -\frac{1}{\sqrt{\lambda}} \left[(\mathbf{l}^T \otimes (\mathbf{D} - \lambda \mathbf{I}_c)^\sharp) - \frac{1}{2\lambda} [(\mathbf{l}^T) \otimes (\mathbb{I}^H)] \right] \Delta_2
\end{aligned}$$

$$\begin{aligned}
& [\mathbf{I}_c \mathbf{A}_1 \cdots \mathbf{A}_M]^H \left[(\mathbf{l}^T) \otimes \left((\mathbf{D} - \lambda \mathbf{I}_c)^\natural \right) \right] \\
&= \left(\mathbf{1} \otimes [\mathbf{I}_c \mathbf{A}_1 \cdots \mathbf{A}_M]^H \right) \left[(\mathbf{l}^T) \otimes \left((\mathbf{D} - \lambda \mathbf{I}_c)^\natural \right) \right] \\
&= \left[(\mathbf{l}^T) \otimes \left([\mathbf{I}_c \mathbf{A}_1 \cdots \mathbf{A}_M]^H (\mathbf{D} - \lambda \mathbf{I}_c)^\natural \right) \right] \\
&= \frac{1}{\lambda} \left[(\mathbf{l}^T) \otimes \left([\mathbf{I}_c \mathbf{A}_1 \cdots \mathbf{A}_M]^H (\mathbb{I}^H - \mathbf{I}_c) \right) \right]
\end{aligned}$$

$$\begin{aligned}
(B.2) \quad & \Rightarrow \quad (\mathbf{I}_c - \mathbb{I}^H) [\mathbf{I}_c \mathbf{A}_1 \cdots \mathbf{A}_M] \mathbf{Y} = \mathbf{0}_{c,c(M+1)} \\
& \Rightarrow \quad \mathcal{T}_{M+1}(\mathbf{x}_M) \mathcal{T}_{M+1}^H(\mathbf{x}_M) [\mathbf{I}_c \mathbf{A}_1 \cdots \mathbf{A}_M]^H (\mathbf{I}_c - \mathbb{I}^H) = \mathbf{0}_{c(M+1),c} \\
\mathcal{T}_{M+1}(\mathbf{x}_M) \quad & \text{est de rang colonne plein} \\
& \Rightarrow \quad \mathcal{T}_{M+1}^H(\mathbf{x}_M) [\mathbf{I}_c \mathbf{A}_1 \cdots \mathbf{A}_M]^H (\mathbf{I}_c - \mathbb{I}^H) = \mathbf{0}_{2M+1,c} \\
\text{Ou encore} \quad & \mathbf{Y}^\circ [\mathbf{I}_c \mathbf{A}_1 \cdots \mathbf{A}_M]^H (\mathbf{I}_c - \mathbb{I}^H) = \mathbf{0}_{2(M+1),c}
\end{aligned}$$

$$\begin{aligned}
(B.1) \text{ implique} \quad & [\mathbf{I}_c \mathbf{A}_1 \cdots \mathbf{A}_M]^H \mathbf{l} = \sqrt{\lambda} \mathbf{g}^* \\
\text{ou bien} \quad & \mathbf{l}^H [\mathbf{I}_c \mathbf{A}_1 \cdots \mathbf{A}_M] = \sqrt{\lambda} \mathbf{g}^T \\
\text{Ainsi} \quad & \mathbf{l}^H [\mathbf{A}_1 \cdots \mathbf{A}_M] = \mathbf{l}^H [\mathbf{I}_c \mathbf{A}_1 \cdots \mathbf{A}_M] [\mathbf{0}_{cM,c} \quad \mathbf{I}_{cM}]^T \\
& = \sqrt{\lambda} \mathbf{g}^T [\mathbf{0}_{cM,c} \quad \mathbf{I}_{cM}]^T
\end{aligned}$$

$$\begin{aligned}
\mathbf{Y}^\circ \Delta_5^* &= -\frac{1}{2\lambda^{3/2}} \mathbf{Y}^\circ [\mathbf{I}_c \mathbf{A}_1 \cdots \mathbf{A}_M]^H \left[(\mathbf{l}^T) \otimes (\mathbb{I}^H) \right] \Delta_2 \\
&= -\frac{1}{2\lambda^{3/2}} \mathbf{Y}^\circ \left(\mathbf{1} \otimes [\mathbf{I}_c \mathbf{A}_1 \cdots \mathbf{A}_M]^H \right) \left[(\mathbf{l}^T) \otimes (\mathbb{I}^H) \right] \Delta_2 \\
&= -\frac{1}{2\lambda^{3/2}} \mathbf{Y}^\circ \left[(\mathbf{l}^T) \otimes \left([\mathbf{I}_c \mathbf{A}_1 \cdots \mathbf{A}_M]^H \mathbb{I}^H \right) \right] \Delta_2 \\
&= -\frac{1}{2\lambda} \mathbf{Y}^\circ \left[(\mathbf{l}^T) \otimes (\mathbf{g}^* \mathbf{l}^H) \right] \Delta_2 \\
&= -\frac{1}{2\lambda} (\mathbf{1} \otimes \mathbf{Y}^\circ) \left[(\mathbf{l}^T) \otimes (\mathbf{g}^* \mathbf{l}^H) \right] \Delta_2 \\
&= -\frac{1}{2\lambda} \left[(\mathbf{l}^T) \otimes (\mathbf{Y}^\circ \mathbf{g}^* \mathbf{l}^H) \right] \Delta_2 \\
&= -\frac{1}{2\lambda} \left[(\mathbf{l}^T) \otimes (\mathbf{x}_M \mathbf{l}^H) \right] \Delta_2 \\
&= -\frac{1}{2\lambda} \left[(\mathbf{l}^T) \otimes (\mathbf{x}_M \mathbf{l}^H) \right] ([\mathbf{I}_c \mathbf{A}_1 \cdots \mathbf{A}_M]^* \otimes [\mathbf{I}_c \mathbf{A}_1 \cdots \mathbf{A}_M]) \\
&= -\frac{1}{2\lambda} \left[(\mathbf{l}^T [\mathbf{I}_c \mathbf{A}_1 \cdots \mathbf{A}_M]^*) \otimes (\mathbf{x}_M \mathbf{l}^H [\mathbf{I}_c \mathbf{A}_1 \cdots \mathbf{A}_M]) \right] \\
&= -\frac{1}{2} \left[(\mathbf{g}^H) \otimes (\mathbf{x}_M \mathbf{g}^T) \right]
\end{aligned}$$

$$\Delta_6 = (\mathbf{g}^H \otimes \mathbf{I}_{c(M+1)}) \overline{\mathcal{S}}_{c,M} - \frac{1}{2} \left[(\mathbf{g}^H) \otimes (\mathbf{x}_M \mathbf{g}^T) \right]$$

■

B.4 L'algorithme décomposition en produit extérieur

Soit $\mathbf{X}$ une matrice Hermitienne définie positive $c(M+1) \times c(M+1)$. L'application de l'algorithme OPD à $\mathbf{X}$ suit les étapes suivantes (cf §4.4.3)

$$\begin{aligned}\mathbf{Y} &\triangleq \mathbf{X} - \sigma_b^2(\mathbf{X})\mathbf{I}_{c(M+1)} \quad \text{Hermitienne définie positive} \\ [\gamma(0)\gamma(1)\cdots\gamma(M)] &\triangleq [\mathbf{I}_c \mathbf{0}_{c,cM}]\mathbf{Y} \quad , \text{ où } \gamma(i) \text{ est } c \times c \\ \mathbf{Y}^\circ &\triangleq \mathbf{Han}(\gamma(0), \dots, \gamma(M)) \\ \mathbf{D}_1 &\triangleq \frac{1}{\sigma_s^2} \mathbf{Y}^\circ \mathbf{Y}^{\dagger} \mathbf{Y}^{\circ H}\end{aligned}\tag{B.3}$$

$$\begin{aligned}\mathbf{D}_2 &\triangleq (\mathbf{J}_{M+1} \otimes \mathbf{I}_c) \mathbf{D}_1 (\mathbf{J}_{M+1} \otimes \mathbf{I}_c)^T \\ \lambda &\triangleq \lambda_1(\mathbf{D}_1 - \mathbf{D}_2) \\ \text{alg}(\mathbf{X}) &\triangleq \text{VECP}(\mathbf{D}_1 - \mathbf{D}_2)\end{aligned}\tag{B.4}$$

Si $\mathbf{X} = \sigma_s^2 \mathcal{T}_M(\mathbf{x}_M) \mathcal{T}_M^H(\mathbf{x}_M) + x \mathbf{I}_{c(M+1)} \in \mathcal{M}_M$ et $c = 2$,

$$\begin{aligned}\mathbf{Y} &= \sigma_s^2 \mathcal{T}_M(\mathbf{x}_M) \mathcal{T}_M^H(\mathbf{x}_M) \\ \mathbf{Y}^\circ &= \sigma_s^2 \mathcal{T}^\circ(\mathbf{x}_M) \mathcal{T}_M^H(\mathbf{x}_M) \\ \mathbf{D}_1 &= \mathcal{T}^\circ(\mathbf{x}_M) \mathcal{T}^{\circ H}(\mathbf{x}_M) \\ (\mathbf{D}_1 - \mathbf{D}_2) &= \mathbf{x}_M \mathbf{x}_M^H \\ \text{VECP}(\mathbf{D}_1 - \mathbf{D}_2) &= \mathbf{x}_M \\ \lambda &= \|\mathbf{x}_M\|^2\end{aligned}$$

Proposition B.7 Soit $\mathbf{X}$ une matrice Hermitienne définie positive $2(M+1) \times 2(M+1)$. Si $\mathbf{X}$ est inversible

$$\begin{aligned}\mathbf{A}_X^1 &= -\frac{1}{\sigma_s^2} \left([\text{VECP}(\mathbf{D}_1 - \mathbf{D}_2)]^T \otimes (\mathbf{D}_1 - \mathbf{D}_2 - \lambda \mathbf{I}_{2(M+1)})^{\natural} \right) \\ &\quad \left(\mathbf{I}_{4(M+1)^2} - ((\mathbf{J}_{M+1} \otimes \mathbf{I}_2) \otimes (\mathbf{J}_{M+1} \otimes \mathbf{I}_2)) \right) \\ &\quad \Delta_2 \overline{\mathcal{G}}(\mathbf{X}) \\ \mathbf{A}_X^2 &= -\frac{1}{\sigma_s^2} \left([\text{VECP}(\mathbf{D}_1 - \mathbf{D}_2)]^T \otimes (\mathbf{D}_1 - \mathbf{D}_2 - \lambda \mathbf{I}_{2(M+1)})^{\natural} \right) \\ &\quad \left(\mathbf{I}_{4(M+1)^2} - ((\mathbf{J}_{M+1} \otimes \mathbf{I}_2) \otimes (\mathbf{J}_{M+1} \otimes \mathbf{I}_2)) \right) \\ &\quad \Delta_3 \overline{\mathcal{G}}^*(\mathbf{X}) \\ \mathbf{A}_X &= -\frac{1}{\sigma_s^2} \left([\text{VECP}(\mathbf{D}_1 - \mathbf{D}_2)]^T \otimes (\mathbf{D}_1 - \mathbf{D}_2 - \lambda \mathbf{I}_{2(M+1)})^{\natural} \right) \\ &\quad \left(\mathbf{I}_{4(M+1)^2} - ((\mathbf{J}_{M+1} \otimes \mathbf{I}_2) \otimes (\mathbf{J}_{M+1} \otimes \mathbf{I}_2)) \right) \\ &\quad (\Delta_2 + \Delta_3 \mathbf{K}_{2(M+1)}) \overline{\mathcal{G}}(\mathbf{X})\end{aligned}$$

avec

$$\begin{aligned}\Delta_2 &\triangleq \left[(\mathbf{Y}^{\circ*} \mathbf{Y}^{\dagger*}) \otimes (\mathbf{I}_{2(M+1)}) \right] \overline{\mathcal{S}}_{2,M} - \left[(\mathbf{Y}^{\circ*} \mathbf{Y}^{\dagger*}) \otimes (\mathbf{Y}^\circ \mathbf{Y}^{\dagger}) \right] \\ \Delta_3 &\triangleq \left[(\mathbf{I}_{2(M+1)}) \otimes (\mathbf{Y}^\circ \mathbf{Y}^{\dagger}) \right] \mathbf{K}_{2(M+1)} \overline{\mathcal{S}}_{2,M}^* + (\mathbf{Y}^{\circ*} \otimes \mathbf{Y}^\circ) \Delta_1 \mathbf{K}_{2(M+1)} \\ \Delta_1 &\triangleq (\mathbf{P}^T \otimes \mathbf{Y}^{\dagger^2}) - (\mathbf{Y}^{\dagger^2^T} \otimes \mathbf{P})\end{aligned}$$

Démonstration:

Soit $\mathbf{E}$ une matrice de 'perturbation' faible. On suppose que λ soit de multiplicité 1.

$$\begin{aligned}
& ALG(\mathbf{X} + \mathbf{E}) \\
&= VEC P(\mathbf{D}_1 - \mathbf{D}_2) - \left(\mathbf{D}_1 - \mathbf{D}_2 - \lambda \mathbf{I}_{c(M+1)} \right)^{\natural} \delta(\mathbf{D}_1 - \mathbf{D}_2) VEC P(\mathbf{D}_1 - \mathbf{D}_2) + O(\|\mathbf{E}\|^2) \\
&= ALG(\mathbf{X}) - \left(\mathbf{D}_1 - \mathbf{D}_2 - \lambda \mathbf{I}_{c(M+1)} \right)^{\natural} \delta(\mathbf{D}_1 - \mathbf{D}_2) VEC P(\mathbf{D}_1 - \mathbf{D}_2) + O(\|\mathbf{E}\|^2) \\
&= ALG(\mathbf{X}) - \text{Vec} \left(\left(\mathbf{D}_1 - \mathbf{D}_2 - \lambda \mathbf{I}_{c(M+1)} \right)^{\natural} \delta(\mathbf{D}_1 - \mathbf{D}_2) VEC P(\mathbf{D}_1 - \mathbf{D}_2) \right) + O(\|\mathbf{E}\|^2) \\
&= ALG(\mathbf{X}) - \left([VEC P(\mathbf{D}_1 - \mathbf{D}_2)]^T \otimes \left(\mathbf{D}_1 - \mathbf{D}_2 - \lambda \mathbf{I}_{c(M+1)} \right)^{\natural} \right) \text{Vec}(\delta(\mathbf{D}_1 - \mathbf{D}_2)) + O(\|\mathbf{E}\|^2)
\end{aligned}$$

$$\begin{aligned}
\text{Vec}(\delta(\mathbf{D}_1 - \mathbf{D}_2)) &= \left(\mathbf{I}_{c^2(M+1)^2} - \left((\mathbf{J}_{M+1}^T \otimes \mathbf{I}_c) \otimes (\mathbf{J}_{M+1}^T \otimes \mathbf{I}_c) \right) \right) \text{Vec}(\delta \mathbf{D}_1) \\
\sigma_s^2 \delta \mathbf{D}_1 &= \delta \mathbf{Y}^\circ \mathbf{Y}^{\natural} \mathbf{Y}^{\circ H} + \mathbf{Y}^\circ \delta(\mathbf{Y}^{\natural}) \mathbf{Y}^{\circ H} + \mathbf{Y}^\circ \mathbf{Y}^{\natural} \delta(\mathbf{Y}^{\circ H}) + O(\|\mathbf{Y}\|^2) \\
\text{Vec}(\delta \mathbf{Y}^\circ) &= \overline{\mathcal{S}}_{c,M} \text{Vec}(\delta \mathbf{Y}) \\
\text{Vec}(\delta(\mathbf{Y}^{\circ H})) &= \mathbf{K}_{c(M+1)} \text{Vec}(\delta \mathbf{Y}^{\circ*}) \\
&= \mathbf{K}_{c(M+1)} \overline{\mathcal{S}}_{c,M}^* \text{Vec}(\delta \mathbf{Y})^* \\
\delta(\mathbf{Y}^{\natural}) &= -\mathbf{Y}^{\natural} \delta \mathbf{Y} \mathbf{Y}^{\natural} + \mathbf{Y}^{\natural^2} \delta \mathbf{Y}^H \mathbf{P} - \mathbf{P} \delta \mathbf{Y}^H \mathbf{Y}^{\natural^2} + O(\|\mathbf{Y}\|^2) \\
\text{Vec}(\delta(\mathbf{Y}^{\natural})) &= -(\mathbf{Y}^{\natural^T} \otimes \mathbf{Y}^{\natural}) \text{Vec}(\delta \mathbf{Y}) + (\mathbf{P}^T \otimes \mathbf{Y}^{\natural^2}) \mathbf{K}_{c(M+1)} \text{Vec}(\delta \mathbf{Y})^* \\
&\quad - (\mathbf{Y}^{\natural^{2T}} \otimes \mathbf{P}) \mathbf{K}_{c(M+1)} \text{Vec}(\delta \mathbf{Y})^* + O(\|\mathbf{Y}\|^2) \\
&= -(\mathbf{Y}^{\natural^T} \otimes \mathbf{Y}^{\natural}) \text{Vec}(\delta \mathbf{Y}) + \Delta_1 \mathbf{K}_{c(M+1)} \text{Vec}(\delta \mathbf{Y})^* + O(\|\mathbf{Y}\|^2) \\
\Delta_1 &\triangleq (\mathbf{P}^T \otimes \mathbf{Y}^{\natural^2}) - (\mathbf{Y}^{\natural^{2T}} \otimes \mathbf{P})
\end{aligned}$$

$\mathbf{P}$ étant le projecteur sur le noyau de $\mathbf{Y}$.

$$\begin{aligned}
\sigma_s^2 \text{Vec}(\delta \mathbf{D}_1) &= \left[(\mathbf{Y}^{\circ*} \mathbf{Y}^{\natural^T}) \otimes (\mathbf{I}_{c(M+1)}) \right] \overline{\mathcal{S}}_{c,M} \text{Vec}(\delta \mathbf{Y}) \\
&\quad + (\mathbf{Y}^{\circ*} \otimes \mathbf{Y}^\circ) \left(-(\mathbf{Y}^{\natural^T} \otimes \mathbf{Y}^{\natural}) \text{Vec}(\delta \mathbf{Y}) + \Delta_1 \mathbf{K}_{c(M+1)} \text{Vec}(\delta \mathbf{Y})^* \right) \\
&\quad + \left[(\mathbf{I}_{c(M+1)}) \otimes (\mathbf{Y}^\circ \mathbf{Y}^{\natural}) \right] \mathbf{K}_{c(M+1)} \overline{\mathcal{S}}_{c,M}^* \text{Vec}(\delta \mathbf{Y})^* + O(\|\mathbf{Y}\|^2) \\
\text{Vec}(\delta \mathbf{D}_1) &= \frac{1}{\sigma_s^2} [\Delta_2 \text{Vec}(\delta \mathbf{Y}) + \Delta_3 \text{Vec}(\delta \mathbf{Y})^*] + O(\|\mathbf{Y}\|^2) \\
\Delta_2 &\triangleq \left[(\mathbf{Y}^{\circ*} \mathbf{Y}^{\natural^T}) \otimes (\mathbf{I}_{c(M+1)}) \right] \overline{\mathcal{S}}_{c,M} - \left[(\mathbf{Y}^{\circ*} \mathbf{Y}^{\natural^T}) \otimes (\mathbf{Y}^\circ \mathbf{Y}^{\natural}) \right] \\
&= \left[(\mathbf{Y}^{\circ*} \mathbf{Y}^{\natural^*}) \otimes (\mathbf{I}_{c(M+1)}) \right] \overline{\mathcal{S}}_{c,M} - \left[(\mathbf{Y}^{\circ*} \mathbf{Y}^{\natural^*}) \otimes (\mathbf{Y}^\circ \mathbf{Y}^{\natural}) \right] \\
\Delta_3 &\triangleq \left[(\mathbf{I}_{c(M+1)}) \otimes (\mathbf{Y}^\circ \mathbf{Y}^{\natural}) \right] \mathbf{K}_{c(M+1)} \overline{\mathcal{S}}_{c,M}^* + (\mathbf{Y}^{\circ*} \otimes \mathbf{Y}^\circ) \Delta_1 \mathbf{K}_{c(M+1)}
\end{aligned}$$

Si $c = 2$

$$\begin{aligned}
\text{Vec}(\delta \mathbf{Y}) &= \overline{\mathcal{G}}(\mathbf{X}) \text{Vec}(\mathbf{E}) + O(\|\mathbf{E}\|^2) \\
\sigma_s^2 \text{Vec}(\delta \mathbf{D}_1) &= \Delta_2 \overline{\mathcal{G}}(\mathbf{X}) \text{Vec}(\mathbf{E}) + \Delta_3 \overline{\mathcal{G}}^*(\mathbf{X}) \text{Vec}(\mathbf{E})^* + O(\|\mathbf{E}\|^2)
\end{aligned}$$

d'où les expressions de $\mathbf{A}_{\mathbf{X}}^1$ et $\mathbf{A}_{\mathbf{X}}^2$. En remarquant que $\overline{\mathcal{G}}^*(\mathbf{X})\mathbf{K}_{2(M+1)} = \mathbf{K}_{2(M+1)}\overline{\mathcal{G}}(\mathbf{X})$, on déduit $\mathbf{A}_{\mathbf{X}}$. ■

Proposition B.8

$$Si \quad \mathbf{X} \in \mathcal{M}_M, \quad alors \quad \Delta_3 = \left[\left(\mathbf{I}_{2(M+1)} \right) \otimes \left(\mathbf{Y}^\circ \mathbf{Y}^\natural \right) \right] \mathbf{K}_{2(M+1)} \overline{\mathcal{S}}_{2,M}^*$$

Démonstration:

Si $\mathbf{X} = \sigma_s^2 \mathcal{T}_{M+1}(\mathbf{x}_M) \mathcal{T}_{M+1}^H(\mathbf{x}_M) + x \mathbf{I}_{c(M+1)} \in \mathcal{R}_M$, $\mathbf{Y} = \sigma_s^2 \mathcal{T}_{M+1}(\mathbf{x}_M) \mathcal{T}_{M+1}^H(\mathbf{x}_M)$ et $\mathbf{P} = -\mathbf{T}_{M+1} \mathbf{x}_M^* \mathbf{x}_M^H \mathbf{T}_{M+1}$. On a $\mathbf{Y}\mathbf{P} = \mathbf{0}_{c(M+1)}$ et $\mathcal{T}_{M+1}^H(\mathbf{x}_M) \mathbf{P} = \mathbf{0}_{c(M+1)}$.

$$\begin{aligned} \mathbf{Y}^\circ \mathbf{P} &= \sigma_s^2 \mathcal{T}^\circ(\mathbf{x}_M) \mathcal{T}_{M+1}^H(\mathbf{x}_M) \mathbf{P} = \mathbf{0}_{c(M+1)} \\ (\mathbf{Y}^{\circ*} \otimes \mathbf{Y}^\circ) \Delta_2 &= (\mathbf{Y}^{\circ*} \otimes \mathbf{Y}^\circ) \left[\left(\mathbf{P}^T \otimes \mathbf{Y}^{\natural^2} \right) - \left(\mathbf{Y}^{\natural^{2^T}} \otimes \mathbf{P} \right) \right] \\ &= (\mathbf{Y}^{\circ*} \otimes \mathbf{Y}^\circ) \left[\left(\mathbf{P}^* \otimes \mathbf{Y}^{\natural^2} \right) - \left(\mathbf{Y}^{\natural^{2^T}} \otimes \mathbf{P} \right) \right] \\ &= \mathbf{0}_{c(M+1), c(M+1)} \end{aligned}$$

■

Annexe C

Démonstration de Prop. 4.1

Résultat C.1

$$\sum_{n,m=A}^B f(m-n) = \sum_{h=-(B-A)}^{B-A} (B-A+1-|h|) f(h)$$

On note r_b^a le reste de la division euclidienne de a par b . Ainsi, $r_b^a = a - \lfloor \frac{a}{b} \rfloor b$.

Résultat C.2

$$\forall k \in \{1, \dots, cD\}, \left[\mathbf{x}_{D,ST}(n) \right]_k = x^{c(k)} (n - T(k)) \quad \text{avec} \quad c(k) \hat{=} r_c^{k-1} + 1 \quad \text{et} \quad T(k) \hat{=} \lfloor \frac{k-1}{c} \rfloor$$

Il s'en suit que

$$\begin{aligned} (\mathbf{R}_{D,ST})_{kl} &= \mathbb{E} \left[x^{c(k)} (n - T(k)) x^{c(l)} (n - T(l)) \right] = \gamma_{c(k)c(l)} [T(l) - T(k)] \\ (\hat{\mathbf{R}}_{D,ST})_{kl} &= \frac{1}{N-D+1} \sum_{n=D-1}^{N-1} \left[\mathbf{x}_{D,ST}(n) \right]_i \left[\mathbf{x}_{D,ST}(n) \right]_j \\ &= \frac{1}{N-D+1} \sum_{n=D-1}^{N-1} x^{c(k)} (n - T(k)) x^{c(l)*} (n - T(l)) \end{aligned}$$

Résultat C.3

$$\text{Vec}(\mathbf{R}_{D,ST})_k = (\mathbf{R}_{D,ST})_{a(k)b(k)} \quad \text{avec} \quad a(k) \hat{=} r_{cD}^{k-1} + 1 \quad \text{et} \quad b(k) \hat{=} \lfloor \frac{k-1}{cD} \rfloor + 1$$

Par définition du reste, on a

Résultat C.4

$$\begin{aligned} k &= c(k) + T(k)c \\ &= a(k) + (b(k) - 1)cD \end{aligned}$$

Résultat C.5

$$\mathbb{E} \left[\text{Vec} \left(\hat{\mathbf{R}}_{D,ST} - \mathbf{R}_{D,ST} \right) \text{Vec} \left(\hat{\mathbf{R}}_{D,ST} - \mathbf{R}_{D,ST} \right)^H \right] = \mathbf{A}_N + \mathbf{B}_N + \mathbf{C}_N$$

avec

$$\begin{aligned} \mathbf{A}_{N_{k,l}} &= \frac{1}{(N-D+1)^2} \sum_{n,m=D-1}^{N-1} \gamma_{c(a(k))c(a(l))} (n-m-T(a(k))+T(a(l))) \\ &\quad \gamma_{c(b(l))c(b(k))} (m-n-T(b(l))+T(b(k))) \\ \mathbf{B}_{N_{k,l}} &= \frac{1}{(N-D+1)^2} \sum_{n,m=D-1}^{N-1} \tau_{c(a(k))c(b(l))} (n-m-T(a(k))+T(b(l))) \\ &\quad \tau_{c(b(k))c(a(l))} (n-m-T(b(k))+T(a(l)))^* \\ &= \frac{1}{(N-D+1)^2} \sum_{n,m=D-1}^{N-1} \tau_{c(b(l))c(a(k))} (m-n-T(b(l))+T(a(k))) \\ &\quad \tau_{c(b(k))c(a(l))} (n-m-T(b(k))+T(a(l)))^* \\ \mathbf{C}_{N_{k,l}} &= \frac{1}{(N-D+1)^2} \sum_{n,m=D-1}^{N-1} \text{cum} \left[x^{c(a(k))} (n-T(a(k))), x^{c(b(l))} (m-T(b(l))), \right. \\ &\quad \left. x^{c(b(k))} (n-T(b(k)))^*, x^{c(a(l))} (m-T(a(l)))^* \right] \end{aligned}$$

Démonstration:

$$\begin{aligned} &\mathbb{E} \left[\text{Vec} \left(\hat{\mathbf{R}}_{D,ST} - \mathbf{R}_{D,ST} \right) \text{Vec} \left(\hat{\mathbf{R}}_{D,ST} - \mathbf{R}_{D,ST} \right)^H \right] \\ &= \mathbb{E} \left[\text{Vec} \left(\hat{\mathbf{R}}_{D,ST} \right) \text{Vec} \left(\hat{\mathbf{R}}_{D,ST} \right)^H \right] - \text{Vec} \left(\mathbf{R}_{D,ST} \right) \text{Vec} \left(\mathbf{R}_{D,ST} \right)^H. \end{aligned}$$

Calculons les éléments de la matrice $\mathbb{E} \left[\text{Vec} \left(\hat{\mathbf{R}}_{D,ST} \right) \text{Vec} \left(\hat{\mathbf{R}}_{D,ST} \right)^H \right]$, soit

$$\begin{aligned} \left(\mathbb{E} \left[\text{Vec} \left(\hat{\mathbf{R}}_{D,ST} \right) \text{Vec} \left(\hat{\mathbf{R}}_{D,ST} \right)^H \right] \right)_{kl} &= \mathbb{E} \left[\text{Vec} \left(\hat{\mathbf{R}}_{D,ST} \right)_k \text{Vec} \left(\hat{\mathbf{R}}_{D,ST} \right)_l^* \right] \\ &= \mathbb{E} \left[\left(\hat{\mathbf{R}}_{D,ST} \right)_{a(k)b(k)} \left(\hat{\mathbf{R}}_{D,ST} \right)_{a(l)b(l)}^* \right] \end{aligned}$$

$$\begin{aligned} \text{Ainsi } (N-D+1)^2 \left(\mathbb{E} \left[\text{Vec} \left(\hat{\mathbf{R}}_{D,ST} \right) \text{Vec} \left(\hat{\mathbf{R}}_{D,ST} \right)^H \right] \right)_{kl} \\ &= \mathbb{E} \left[\sum_{n,m=D-1}^{N-1} x^{c(a(k))} (n-T(a(k))) x^{c(b(k))} (n-T(b(k)))^* x^{c(a(l))} (m-T(a(l)))^* x^{c(b(l))} (m-T(b(l))) \right] \\ &= \sum_{n,m=D-1}^{N-1} \mathbb{E} \left[x^{c(a(k))} (n-T(a(k))) x^{c(b(l))} (m-T(b(l))) x^{c(b(k))} (n-T(b(k)))^* x^{c(a(l))} (m-T(a(l)))^* \right] \end{aligned}$$

$\mathbf{x}(n)$ étant supposé centré, ainsi

$$\begin{aligned} &\mathbb{E} \left[x^{c(a(k))} (n-T(a(k))) x^{c(b(l))} (m-T(b(l))) x^{c(b(k))} (n-T(b(k)))^* x^{c(a(l))} (m-T(a(l)))^* \right] \\ &= \text{cum} \left(x^{c(a(k))} (n-T(a(k))), x^{c(b(l))} (m-T(b(l))), x^{c(b(k))} (n-T(b(k)))^*, x^{c(a(l))} (m-T(a(l)))^* \right) \end{aligned}$$

$$\begin{aligned}
& + \mathbb{E} \left[x^{c(a(k))} (n - T(a(k))) x^{c(b(l))} (m - T(b(l))) \right] \mathbb{E} \left[x^{c(b(k))} (n - T(b(k)))^* x^{c(a(l))} (m - T(a(l)))^* \right] \\
& + \mathbb{E} \left[x^{c(a(k))} (n - T(a(k))) x^{c(b(k))} (n - T(b(k)))^* \right] \mathbb{E} \left[x^{c(b(l))} (m - T(b(l))) x^{c(a(l))} (m - T(a(l)))^* \right] \\
& + \mathbb{E} \left[x^{c(a(k))} (n - T(a(k))) x^{c(a(l))} (m - T(a(l)))^* \right] \mathbb{E} \left[x^{c(b(l))} (m - T(b(l))) x^{c(b(k))} (n - T(b(k)))^* \right]
\end{aligned}$$

$$\begin{aligned}
\text{Donc } & (N - D + 1)^2 \left(\mathbb{E} \left[\text{Vec} \left(\hat{\mathbf{R}}_{D,ST} \right) \text{Vec} \left(\hat{\mathbf{R}}_{D,ST} \right)^H \right] \right)_{kl} \\
& = \sum_{n,m=D-1}^{N-1} \text{cum} \left(x^{c(a(k))} (n - T(a(k))), x^{c(b(l))} (m - T(b(l))), x^{c(b(k))} (n - T(b(k)))^*, x^{c(a(l))} (m - T(a(l)))^* \right) \\
& + \sum_{n,m=D-1}^{N-1} \tau_{c(a(k))c(b(l))} (n - m - T(a(k)) + T(b(l))) \tau_{c(b(k))c(a(l))} (n - m - T(b(k)) + T(a(l)))^* \\
& + \sum_{n,m=D-1}^{N-1} \gamma_{c(a(k))c(b(k))} (T(b(k)) - T(a(k))) \gamma_{c(b(l))c(a(l))} (T(a(l)) - T(b(l))) \\
& + \sum_{n,m=D-1}^{N-1} \gamma_{c(a(k))c(a(l))} (n - m - T(a(k)) + T(a(l))) \gamma_{c(b(l))c(b(k))} (m - n - T(b(l)) + T(b(k)))
\end{aligned}$$

$$\begin{aligned}
\text{On a } & \frac{1}{(N-D+1)^2} \sum_{n,m=D-1}^{N-1} \gamma_{c(a(k))c(b(k))} (T(b(k)) - T(a(k))) \gamma_{c(b(l))c(a(l))} (T(a(l)) - T(b(l))) \\
& = \gamma_{c(a(k))c(b(k))} (T(b(k)) - T(a(k))) \gamma_{c(b(l))c(a(l))} (T(a(l)) - T(b(l))) \\
& = (\mathbf{R}_{D,ST})_{a(k)b(k)} (\mathbf{R}_{D,ST})_{b(l)a(l)} \\
& = (\mathbf{R}_{D,ST})_{a(k)b(k)} (\mathbf{R}_{D,ST})_{a(l)b(l)}^* \\
& = \text{Vec} (\mathbf{R}_{D,ST})_k \text{Vec} (\mathbf{R}_{D,ST})_l^* \\
& = \left[\text{Vec} (\mathbf{R}_{D,ST}) \text{Vec} (\mathbf{R}_{D,ST})^H \right]_{kl}
\end{aligned}$$

$$\begin{aligned}
\text{Finalement } & \mathbb{E} \left[\text{Vec} \left(\hat{\mathbf{R}}_{D,ST} - \mathbf{R}_{D,ST} \right) - \text{Vec} \left(\hat{\mathbf{R}}_{D,ST} - \mathbf{R}_{D,ST} \right)^H \right] \\
& = \mathbb{E} \left[\text{Vec} \left(\hat{\mathbf{R}}_{D,ST} \right) \text{Vec} \left(\hat{\mathbf{R}}_{D,ST} \right)^H \right] - \left[\text{Vec} (\mathbf{R}_{D,ST}) \text{Vec} (\mathbf{R}_{D,ST})^H \right] \\
& = \mathbf{A}_N + \mathbf{B}_N + \mathbf{C}_N
\end{aligned}$$

avec

$$\begin{aligned}
\mathbf{A}_{N_{k,l}} & = \frac{1}{(N - D + 1)^2} \sum_{n,m=D-1}^{N-1} \gamma_{c(a(k))c(a(l))} (n - m - T(a(k)) + T(a(l))) \\
& \quad \gamma_{c(b(l))c(b(k))} (m - n - T(b(l)) + T(b(k))) \\
\mathbf{B}_{N_{k,l}} & = \frac{1}{(N - D + 1)^2} \sum_{n,m=D-1}^{N-1} \tau_{c(a(k))c(b(l))} (n - m - T(a(k)) + T(b(l))) \\
& \quad \tau_{c(b(k))c(a(l))} (n - m - T(b(k)) + T(a(l)))^* \\
& = \frac{1}{(N - D + 1)^2} \sum_{n,m=D-1}^{N-1} \tau_{c(b(l))c(a(k))} (m - n - T(b(l)) + T(a(k))) \\
& \quad \tau_{c(b(k))c(a(l))} (n - m - T(b(k)) + T(a(l)))^* \\
\mathbf{C}_{N_{k,l}} & = \frac{1}{(N - D + 1)^2} \sum_{n,m=D-1}^{N-1} \text{cum} \left[x^{c(a(k))} (n - T(a(k))), x^{c(b(l))} (m - T(b(l))), \right. \\
& \quad \left. x^{c(b(k))} (n - T(b(k)))^*, x^{c(a(l))} (m - T(a(l)))^* \right]
\end{aligned}$$

■

On s'intéresse au calcul de $\mathbf{A} \hat{=} \lim_{N \rightarrow +\infty} N \mathbf{A}_N$, $\mathbf{B} \hat{=} \lim_{N \rightarrow +\infty} N \mathbf{B}_N$, $\mathbf{C} \hat{=} \lim_{N \rightarrow +\infty} N \mathbf{C}_N$.

Résultat C.6

$$\mathbf{A}_{kl} = \frac{1}{2\pi} \int_{w=-\pi}^{\pi} \mathbf{S}_{c(a(k))c(a(l))}(w) \mathbf{S}_{c(b(l))c(b(k))}(w) e^{jw(T(a(l))-T(a(k))+T(b(k))-T(b(l)))} dw$$

Démonstration:

$$\begin{aligned} \mathbf{A}_{N_{k,l}} &= \frac{1}{(N-D+1)^2} \sum_{n,m=D-1}^{N-1} \gamma_{c(a(k))c(a(l))}(n-m-T(a(k))+T(a(l))) \\ &\quad \gamma_{c(b(l))c(b(k))}(m-n-T(b(l))+T(b(k))) \\ &= \frac{1}{(N-D+1)^2} \sum_{n=D-1}^{N-1} \sum_{m'=n+T_1-N+1}^{n+T_1-D+1} \gamma_{c(a(k))c(a(l))}(m') \gamma_{c(b(l))c(b(k))}(T_2-m') \\ &\quad \text{où } T_1 \hat{=} T(a(l)) - T(a(k)) \\ &\quad T_2 \hat{=} T_1 + T(b(k)) - T(b(l)) \end{aligned}$$

On définit $f_n^N(m') = 1$ si $m' \in \{n+T_1-N+1 \cdots n+T_1-D+1\}$, nulle si non.

$$\begin{aligned} \mathbf{A}_{N_{k,l}} &= \frac{1}{(N-D+1)^2} \sum_{n=D-1}^{N-1} \sum_{m' \in \mathbf{Z}} \gamma_{c(a(k))c(a(l))}(m') \gamma_{c(b(l))c(b(k))}(T_2-m') f_n^N(m') \\ &= \sum_{m' \in \mathbf{Z}} \gamma_{c(a(k))c(a(l))}(m') \gamma_{c(b(l))c(b(k))}(T_2-m') \left(\frac{1}{(N-D+1)^2} \sum_{n=D-1}^{N-1} f_n^N(m') \right) \\ \mathbf{A}_{k,l} &= \sum_{m' \in \mathbf{Z}} \gamma_{c(a(k))c(a(l))}(m') \gamma_{c(b(l))c(b(k))}(T_2-m') f(m') \\ \text{avec } f(m') &\hat{=} \lim_{N \rightarrow \infty} \frac{N}{(N-D+1)^2} \sum_{n=D-1}^{N-1} f_n^N(m') \end{aligned}$$

Soit N 'très grand' et $m' \in \mathbf{Z}$, alors $m' > n+T_1-N+1$.

$$\begin{aligned} f_n^N(m') &= 1 \quad \text{si } m' \leq n+T_1-D+1 \text{ i.e., si } n \geq m'+D-T_1 \\ &= 0 \quad \text{si } m' > n+T_1-D+1 \text{ i.e., si non} \\ \sum_{n=D-1}^{N-1} f_n^N(m') &= \sum_{n=\max(D, m'+D-T_1)-1}^{N-1} 1 \\ f(m') &= \lim_{N \rightarrow \infty} \frac{N}{(N-D+1)^2} (N - \max(D, m'+D-T_1) + 1) \\ &= 1 \end{aligned}$$

Finalement

$$\mathbf{A}_{k,l} = \sum_{m' \in \mathbf{Z}} \gamma_{c(a(k))c(a(l))}(m') \gamma_{c(b(l))c(b(k))}(T_2-m')$$

$$\begin{aligned}
&= \left(\gamma_{c(a(k))c(a(l))} * \gamma_{c(b(l))c(b(k))} \right) (T_2) \\
&= \left(\gamma_{c(a(k))c(a(l))} * \gamma_{c(b(l))c(b(k))} \right) (T(a(l)) - T(a(k)) + T(b(k)) - T(b(l))) \\
&= \frac{1}{2\pi} \int_{w=-\pi}^{\pi} TF \left(\gamma_{c(a(k))c(a(l))} * \gamma_{c(b(l))c(b(k))} \right) (w) e^{jwT_2} dw \\
&= \frac{1}{2\pi} \int_{w=-\pi}^{\pi} \mathbf{S}_{c(a(k))c(a(l))}(w) \mathbf{S}_{c(b(l))c(b(k))}(w) e^{jw(T(a(l))-T(a(k))+T(b(k))-T(b(l)))} dw
\end{aligned}$$

■

Résultat C.7

$$\mathbf{B}_{kl} = \frac{1}{2\pi} \int_{w=-\pi}^{\pi} \mathbf{T}_{c(b(l))c(a(k))}(w) \mathbf{T}_{c(a(l))c(b(k))}^*(w) e^{jw(T(a(l))-T(a(k))+T(b(l))-T(b(k)))} dw$$

Démonstration:

$$\begin{aligned}
\mathbf{B}_{N_{k,l}} &= \frac{1}{(N-D+1)^2} \sum_{n,m=D-1}^{N-1} \tau_{c(b(l))c(a(k))}(m-n-T(b(l))+T(a(k))) \\
&\quad \tau_{c(b(k))c(a(l))}(n-m-T(b(k))+T(a(l)))^* \\
&= \frac{1}{(N-D+1)^2} \sum_{n=D-1}^{N-1} \sum_{m'=-n+T'_1+D}^{-n+T'_1+N-1} \tau_{c(b(l))c(a(k))}(m') \tau_{c(b(k))c(a(l))}(T'_2-m')^*
\end{aligned}$$

$$\begin{aligned}
\text{où } T'_1 &\triangleq T(a(k)) - T(b(l)) \\
T'_2 &\triangleq T'_1 + T(a(l)) - T(b(k))
\end{aligned}$$

On définit $g_n^N(m') = 1$ si $m' \in \{-n+T'_1+D-1 \dots -n+T'_1+N-1\}$, nulle si non.

$$\begin{aligned}
\mathbf{B}_{N_{k,l}} &= \frac{1}{(N-D+1)^2} \sum_{n=D-1}^{N-1} \sum_{m' \in \mathbf{Z}} \tau_{c(b(l))c(a(k))}(m') \tau_{c(b(k))c(a(l))}(T'_2-m')^* g_n^N(m') \\
&= \sum_{m' \in \mathbf{Z}} \tau_{c(b(l))c(a(k))}(m') \tau_{c(b(k))c(a(l))}(T'_2-m')^* \left(\frac{1}{(N-D+1)^2} \sum_{n=D-1}^{N-1} g_n^N(m') \right) \\
&= \sum_{m' \in \mathbf{Z}} \tau_{c(b(l))c(a(k))}(m') \tau_{c(b(k))c(a(l))}(T'_2-m')^* g(m') \\
\text{avec } g(m') &\triangleq \lim_{N \rightarrow \infty} \frac{N}{(N-D+1)^2} \sum_{n=D-1}^{N-1} g_n^N(m')
\end{aligned}$$

Soit N 'très grand' et $m' \in \mathbf{Z}$, alors $m' \leq -n+T'_1+N-1$.

$$\begin{aligned}
g_n^N(m') &= 1 \quad \text{si } m' \geq -n+T'_1+D-1 \text{ i.e., si } n \geq -m'+T'_1+D-1 \\
&= 0 \quad \text{si } m' < -n+T'_1+D-1 \text{ i.e., si non}
\end{aligned}$$

$$\begin{aligned}
\sum_{n=D-1}^{N-1} g_n^N(m') &= \sum_{n=\max(D, -m' + T'_1 + D) - 1}^{N-1} 1 \\
g(m') &= \lim_{N \rightarrow \infty} \frac{N}{(N - D + 1)^2} (N - \max(D, -m' + T'_1 + D) + 1) \\
&= 1
\end{aligned}$$

Finalement

$$\begin{aligned}
\mathbf{B}_{k,l} &= \sum_{m' \in \mathbf{Z}} \tau_{c(b(l))c(a(k))}(m') \tau_{c(b(k))c(a(l))} (T'_2 - m')^* \\
&= \left(\tau_{c(b(l))c(a(k))} * \tau_{c(b(k))c(a(l))}^* \right) (T'_2) \\
&= \left(\tau_{c(b(l))c(a(k))} * \tau_{c(b(k))c(a(l))}^* \right) (T(a(l)) + T(a(k)) - T(b(l)) - T(b(k))) \\
&= \frac{1}{2\pi} \int_{w=-\pi}^{\pi} TF \left(\tau_{c(b(l))c(a(k))} * \tau_{c(b(k))c(a(l))}^* \right) (w) e^{jwT'_2} dw \\
&= \frac{1}{2\pi} \int_{w=-\pi}^{\pi} \mathbf{T}_{c(b(l))c(a(k))}(w) TF \left(\tau_{c(b(k))c(a(l))}^* \right) e^{jw(T(a(l)) + T(a(k)) - T(b(l)) - T(b(k)))} dw
\end{aligned}$$

$$\begin{aligned}
TF \left(\tau_{c(b(k))c(a(l))}^* \right) &= \sum_k \tau_{c(b(k))c(a(l))}^*(k) e^{-jwk} \\
&= \sum_k \tau_{c(a(l))c(b(k))}^*(-k) e^{-jwk} \\
&= \sum_k \left(\tau_{c(a(l))c(b(k))}(-k) e^{jwk} \right)^* \\
&= \sum_k \left(\tau_{c(a(l))c(b(k))}(k) e^{-jwk} \right)^* \\
&= \mathbf{T}_{c(a(l))c(b(k))}^*(w)
\end{aligned}$$

■

Dans ce qui suit, on cherche à factoriser les matrices $\mathbf{A}$ et $\mathbf{B}$.

Résultat C.8

$$\mathbf{A} = \frac{1}{2\pi} \int_{w=-\pi}^{\pi} (\mathbf{M}_{ST}^*(w) \otimes \mathbf{M}_{ST}(w)) dw$$

Démonstration:

On note $\mathbf{I}_z^c \triangleq \{cz + 1, \dots, c(z + 1)\}$. Soit $x \in \mathbf{I}_z^c$. On peut écrire $x - 1 = cz + r_c^{x-1}$, z et r_c^{x-1} sont respectivement le quotient et le reste de la division euclidienne de $x - 1$ par c . Pour un tel x , déterminons ce que vaut $a(x)$, $b(x)$, $c(a(x))$, $c(b(x))$, $T(a(x))$ et $T(b(x))$.

$$\begin{aligned}
b(x) &= \lfloor \frac{x-1}{cD} \rfloor + 1 = \lfloor \frac{cz}{cD} + \frac{r_c^{x-1}}{cD} \rfloor + 1 = \lfloor \frac{z}{D} + \frac{r_c^{x-1}}{cD} \rfloor + 1 = \lfloor \frac{z}{D} + \frac{r_c^{x-1}}{cD} \rfloor + 1. \text{ Puisque } z = \lfloor \frac{z}{D} \rfloor D + r_D^z, \text{ on} \\
a(b(x)) &= \lfloor \lfloor \frac{z}{D} \rfloor + \frac{r_D^z}{D} + \frac{r_c^{x-1}}{cD} \rfloor + 1 = \lfloor \lfloor \frac{z}{D} \rfloor + \frac{cr_D^z + r_c^{x-1}}{cD} \rfloor + 1.
\end{aligned}$$

$r_c^{x-1} \leq c-1$ et $r_D^z \leq D-1$ donc $cr_D^z + r_c^{x-1} \leq cD-1 < cD$, soit $\frac{cr_D^z + r_c^{x-1}}{cD} < 1$. Et puisque $\lfloor \frac{z}{D} \rfloor$ est un entier, $\lfloor \lfloor \frac{z}{D} \rfloor + \frac{cr_D^z + r_c^{x-1}}{cD} \rfloor = \lfloor \frac{z}{D} \rfloor$. Ainsi

$$b(x) = \lfloor \frac{z}{D} \rfloor + 1, \forall x \in \mathbf{I}_z^c$$

$$\begin{aligned} a(x) - 1 &= r_{cD}^{x-1} \\ &= (x-1) - \lfloor \frac{x-1}{cD} \rfloor cD \\ &= cz + r_c^{x-1} - (b(x)-1)cD \\ &= cz + r_c^{x-1} - \lfloor \frac{z}{D} \rfloor cD \\ &= c \left[z - \lfloor \frac{z}{D} \rfloor D \right] + r_c^{x-1} \end{aligned}$$

Ainsi

$$a(x) = c \left[z - \lfloor \frac{z}{D} \rfloor D \right] + r_c^{x-1} + 1, \forall x \in \mathbf{I}_z^c$$

$$\begin{aligned} c(a(x)) - 1 &= r_c^{a(x)-1} = r_c^{r_c^{x-1}} = r_c^{x-1} \\ c(b(x)) - 1 &= r_c^{b(x)-1} = r_c^{\lfloor \frac{z}{D} \rfloor} \\ T(a(x)) &= \lfloor \frac{a(x)-1}{c} \rfloor = z - \lfloor \frac{z}{D} \rfloor D \\ T(b(x)) &= \lfloor \frac{b(x)-1}{c} \rfloor = \lfloor \frac{\lfloor \frac{z}{D} \rfloor}{c} \rfloor \end{aligned}$$

Soit $k, l \in \{1, \dots, c^2 D^2\}$. On s'intéresse aux sous-blocs de $\mathbf{A}$ définis par $k \in \mathbf{I}_u^c$ et $l \in \mathbf{I}_v^c$, où $u, v \in \{0, \dots, cD^2 - 1\}$.

$$\begin{aligned} \mathbf{A}_{kl} &= \frac{1}{2\pi} \int_{w=-\pi}^{\pi} \mathbf{S}_{c(a(k))c(a(l))}(w) \mathbf{S}_{c(b(l))c(b(k))}(w) e^{jwT_2} dw \\ c(b(k)) &= r_c^{\lfloor \frac{u}{D} \rfloor} + 1 \\ c(b(l)) &= r_c^{\lfloor \frac{v}{D} \rfloor} + 1 \\ T_2 &= T(a(l)) - T(a(k)) + T(b(k)) - T(b(l)) \\ &= v - \lfloor \frac{v}{D} \rfloor D - u + \lfloor \frac{u}{D} \rfloor D + \lfloor \frac{\lfloor \frac{u}{D} \rfloor}{c} \rfloor - \lfloor \frac{\lfloor \frac{v}{D} \rfloor}{c} \rfloor \end{aligned}$$

Le terme $\mathbf{S}_{c(b(l))c(b(k))}(w) e^{jw(T_2)}$ est fonction de u et v et est donc constant sur le sous-bloc en question. Quand k, l décrivent $\mathbf{I}_u^c \times \mathbf{I}_v^c$, Le terme $\mathbf{S}_{c(a(k))c(a(l))}(w) = \mathbf{S}_{r_c^{k-1}+1, r_c^{l-1}+1}(w)$ décrit la matrice

$$\begin{bmatrix} \mathbf{S}_{11}(w) & \mathbf{S}_{12}(w) & \cdots & \mathbf{S}_{1c}(w) \\ \mathbf{S}_{21}(w) & \ddots & & \mathbf{S}_{2c}(w) \\ \vdots & & & \vdots \\ \mathbf{S}_{c1}(w) & \cdots & \mathbf{S}_{cc}(w) \end{bmatrix} = \mathbf{S}(w)$$

On peut alors écrire $\mathbf{A} = \frac{1}{2\pi} \int_{w=-\pi}^{\pi} \left[\left(\mathbf{A}^{(1)} \right) \otimes (\mathbf{S}(w)) \right] dw$, où $\mathbf{A}^{(1)}$ est une matrice carrée cD^2 -dim définie par ses termes

$$\mathbf{A}_{kl}^{(1)} \triangleq \mathbf{S}_{r_c^{\lfloor \frac{l-1}{D} \rfloor + 1, r_c^{\lfloor \frac{k-1}{D} \rfloor + 1}}(w) e^{jw \left((l-1) - \lfloor \frac{l-1}{D} \rfloor D - (k-1) + \lfloor \frac{k-1}{D} \rfloor D + \lfloor \frac{\lfloor \frac{k-1}{D} \rfloor}{c} \rfloor - \lfloor \frac{\lfloor \frac{l-1}{D} \rfloor}{c} \rfloor \right)}$$

On factorise $\mathbf{A}^{(1)}$ à son tour. On note $\mathbf{I}_z^D \triangleq \{Dz + 1, \dots, D(z+1)\}$. Soit $x \in \mathbf{I}_z^D$. On peut écrire $x - 1 = Dz + r_D^{x-1}$, z et r_D^{x-1} sont respectivement le quotient et le reste de la division euclidienne de $x - 1$ par D .

On regarde les sous-blocs de $\mathbf{A}^{(1)}$ définis par $\mathbf{A}_{kl}^{(1)}$, $k \in \mathbf{I}_u^D$ et $l \in \mathbf{I}_v^D$. $u, v \in \{0, \dots, cD - 1\}$

$$\begin{aligned} \mathbf{A}_{kl}^{(1)} &= \mathbf{S}_{r_c^v + 1, r_c^u + 1}(w) e^{jw \left((l-1) - vD - (k-1) + uD + \lfloor \frac{u}{c} \rfloor - \lfloor \frac{v}{c} \rfloor \right)} \\ &= \mathbf{S}_{r_c^v + 1, r_c^u + 1}(w) e^{jw \left((l-k) + (u-v)D + \lfloor \frac{u}{c} \rfloor - \lfloor \frac{v}{c} \rfloor \right)} \\ &= \mathbf{S}_{r_c^v + 1, r_c^u + 1}(w) e^{jw \left((v-u)D + (r_D^l - r_D^k) + (u-v)D + \lfloor \frac{u}{c} \rfloor - \lfloor \frac{v}{c} \rfloor \right)} \\ &= \mathbf{S}_{r_c^v + 1, r_c^u + 1}(w) e^{jw \left(\lfloor \frac{u}{c} \rfloor - \lfloor \frac{v}{c} \rfloor \right)} e^{jw (r_D^l - r_D^k)} \end{aligned}$$

On peut écrire $\mathbf{A}^{(1)} = \left[\left(\mathbf{A}^{(2)} \right) \otimes \left(\mathbf{A}^{(3)} \right) \right]$, où $\mathbf{A}^{(2)}$ est une matrice carrée cD -dim et $\mathbf{A}^{(3)}$ est une matrice carrée D -dim, définies comme suit

$$\begin{aligned} \mathbf{A}_{kl}^{(3)} &\triangleq e^{jw(l-k)} = e^{-jw(k-1)} \left[e^{-jw(l-1)} \right]^* = [\mathbf{e}_D(w)]_k [\mathbf{e}_D(w)]_l^* \\ \text{Donc } \mathbf{A}^{(3)} &= \mathbf{e}_D(w) \mathbf{e}_D^H(w) \end{aligned}$$

$\mathbf{A}_{kl}^{(2)} \triangleq \mathbf{S}_{r_c^{l-1} + 1, r_c^{k-1} + 1}(w) e^{jw \left(\lfloor \frac{k-1}{c} \rfloor - \lfloor \frac{l-1}{c} \rfloor \right)}$. Factorisons $\mathbf{A}^{(2)}$. Lorsque k et l décrivent le sous-blocs $k \in \mathbf{I}_u^c$ et $l \in \mathbf{I}_v^c$, où $u, v \in \{0, \dots, D - 1\}$, on a $\mathbf{A}_{kl}^{(2)} = \mathbf{S}_{r_c^l + 1, r_c^k + 1}(w) e^{jw(u-v)}$, où $r_c^k + 1$ et $r_c^l + 1$ décrivent $1, \dots, c$. Ainsi

$$\begin{aligned} \mathbf{A}^{(2)} &= \left[\left(\mathbf{A}^{(4)} \right) \otimes \left(\mathbf{S}^T(w) \right) \right] \\ \text{où } \mathbf{A}_{kl}^{(4)} &\triangleq e^{jw(k-l)} = [\mathbf{e}_D(w)]_k^* [\mathbf{e}_D(w)]_l = ([\mathbf{e}_D(w)]_k [\mathbf{e}_D(w)]_l^*)^* \\ \text{soit, } \mathbf{A}^{(4)} &= \left[\mathbf{e}_D(w) \mathbf{e}_D^H(w) \right]^* \end{aligned}$$

Finalement

$$\begin{aligned} \mathbf{A} &= \frac{1}{2\pi} \int_{w=-\pi}^{\pi} \left[\left(\mathbf{A}^{(1)} \right) \otimes (\mathbf{S}(w)) \right] dw \\ &= \frac{1}{2\pi} \int_{w=-\pi}^{\pi} \left[\left(\left[\left(\mathbf{A}^{(2)} \right) \otimes \left(\mathbf{e}_D(w) \mathbf{e}_D^H(w) \right) \right] \right) \otimes (\mathbf{S}(w)) \right] dw \\ &= \frac{1}{2\pi} \int_{w=-\pi}^{\pi} \left[\left(\left[\left(\left[\left(\left[\mathbf{e}_D(w) \mathbf{e}_D^H(w) \right]^* \right) \otimes \left(\mathbf{S}^T(w) \right) \right] \right) \otimes \left(\mathbf{e}_D(w) \mathbf{e}_D^H(w) \right) \right] \otimes (\mathbf{S}(w)) \right] dw \end{aligned}$$

En rappelant que $\mathbf{S}(w)$ est Hermitienne et en définissant la matrice Hermitienne

$$\mathbf{M}_{ST}(w) \triangleq \left[\left(\mathbf{e}_D(w) \mathbf{e}_D^H(w) \right) \otimes (\mathbf{S}(w)) \right]$$

on a

$$\begin{aligned} \mathbf{A} &= \frac{1}{2\pi} \int_{w=-\pi}^{\pi} (\mathbf{M}_{ST}^*(w) \otimes \mathbf{M}_{ST}(w)) dw \\ &= \frac{1}{2\pi} \int_{w=-\pi}^{\pi} (\mathbf{M}_{ST}^T(w) \otimes \mathbf{M}_{ST}(w)) dw \end{aligned}$$

■

Résultat C.9

$$\mathbf{B} = \frac{1}{2\pi} \int_{w=-\pi}^{\pi} (\mathbf{N}_{ST}^*(w) \otimes \mathbf{N}_{ST}(w)) dw \mathbf{K}_{cD, cD}$$

La démonstration est analogue à celle de Res. C.8.

Résultat C.10

$$\mathbf{C}_{k,l} = \sum_{h=-\infty}^{\infty} \text{cum} \left(x^{c(a(k))} (-T(a(k))), x^{c(b(l))} (h - T(b(l))), x^{c(b(k))} (-T(b(k)))^*, x^{c(a(l))} (h - T(a(l)))^* \right)$$

Démonstration:

$$\begin{aligned} &(N - D + 1)^2 \mathbf{C}_{N_{k,l}} = \\ &\sum_{n,m=D-1}^{N-1} \text{cum} \left(x^{c(a(k))} (n - T(a(k))), x^{c(b(l))} (m - T(b(l))), x^{c(b(k))} (n - T(b(k)))^*, x^{c(a(l))} (m - T(a(l)))^* \right). \end{aligned}$$

Par stationnarité de $\mathbf{x}(n)$, on a

$$\begin{aligned} &\text{cum} \left(x^{c(a(k))} (n - T(a(k))), x^{c(b(l))} (m - T(b(l))), x^{c(b(k))} (n - T(b(k)))^*, x^{c(a(l))} (m - T(a(l)))^* \right) \\ &= \text{cum} \left(x^{c(a(k))} (-T(a(k))), x^{c(b(l))} (m - n - T(b(l))), x^{c(b(k))} (-T(b(k)))^*, x^{c(a(l))} (m - n - T(a(l)))^* \right) \end{aligned}$$

$$\begin{aligned} &\mathbf{C}_{N_{k,l}} \\ &= \frac{1}{(N - D + 1)^2} \sum_{n,m=D-1}^{N-1} \text{cum} \left(x^{c(a(k))} (-T(a(k))), x^{c(b(l))} (m - n - T(b(l))), x^{c(b(k))} (-T(b(k)))^*, x^{c(a(l))} (m - n - T(a(l)))^* \right) \\ &= \frac{1}{(N - D + 1)^2} \sum_{h=-(N-D)}^{N-D} (N - D + 1 - |h|) \\ &\quad \text{cum} \left(x^{c(a(k))} (-T(a(k))), x^{c(b(l))} (h - T(b(l))), x^{c(b(k))} (-T(b(k)))^*, x^{c(a(l))} (h - T(a(l)))^* \right) \end{aligned}$$

$$\begin{aligned}
& \mathbf{C}_{k,l} \\
&= \lim_{N \rightarrow \infty} \frac{N}{(N-D+1)^2} \sum_{h=-(N-D)}^{N-D} (N-D+1-|h|) \\
& \quad \text{cum} \left(x^{c(a(k))} (-T(a(k))), x^{c(b(l))} (h-T(b(l))), x^{c(b(k))} (-T(b(k)))^*, x^{c(a(l))} (h-T(a(l)))^* \right) \\
&= \lim_{N \rightarrow \infty} \frac{N}{N-D+1} \sum_{h=-(N-D)}^{N-D} \frac{N-D+1-|h|}{N-D+1} \\
& \quad \text{cum} \left(x^{c(a(k))} (-T(a(k))), x^{c(b(l))} (h-T(b(l))), x^{c(b(k))} (-T(b(k)))^*, x^{c(a(l))} (h-T(a(l)))^* \right)
\end{aligned}$$

On peut faire l'hypothèse qu'il existe N_0 telle que $\forall |h| > N_0 - D$,

$$\text{cum} \left(x^{c(a(k))} (-T(a(k))), x^{c(b(l))} (h-T(b(l))), x^{c(b(k))} (-T(b(k)))^*, x^{c(a(l))} (h-T(a(l)))^* \right) = 0$$

$$\begin{aligned}
& \mathbf{C}_{k,l} \\
&= \lim_{N \rightarrow \infty} \frac{N}{N-D+1} \sum_{h=-(N_0-D)}^{N_0-D} \frac{N-D+1-|h|}{N-D+1} \\
& \quad \text{cum} \left(x^{c(a(k))} (-T(a(k))), x^{c(b(l))} (h-T(b(l))), x^{c(b(k))} (-T(b(k)))^*, x^{c(a(l))} (h-T(a(l)))^* \right) \\
&= \sum_{h=-(N_0-D)}^{N_0-D} \text{cum} \left(x^{c(a(k))} (-T(a(k))), x^{c(b(l))} (h-T(b(l))), x^{c(b(k))} (-T(b(k)))^*, x^{c(a(l))} (h-T(a(l)))^* \right) \\
&= \sum_{h=-\infty}^{\infty} \text{cum} \left(x^{c(a(k))} (-T(a(k))), x^{c(b(l))} (h-T(b(l))), x^{c(b(k))} (-T(b(k)))^*, x^{c(a(l))} (h-T(a(l)))^* \right)
\end{aligned}$$

■

On déduit de ce qui précède le résultat suivant:

Résultat C.11

$$\lim_{N \rightarrow \infty} \mathbb{E} \left[\text{Vec} \left(\hat{\mathbf{R}}_{D,ST} - \mathbf{R}_{D,ST} \right) \text{Vec} \left(\hat{\mathbf{R}}_{D,ST} - \mathbf{R}_{D,ST} \right)^H \right] = \mathbf{A} + \mathbf{B} + \mathbf{C}$$

Résultat C.12

$$\lim_{N \rightarrow \infty} \mathbb{E} \left[\text{Vec} \left(\hat{\mathbf{R}}_{D,ST} - \mathbf{R}_{D,ST} \right) \text{Vec} \left(\hat{\mathbf{R}}_{D,ST} - \mathbf{R}_{D,ST} \right)^T \right] = (\mathbf{A} + \mathbf{B} + \mathbf{C}) \mathbf{K}_{cD}$$

Démonstration:

$$\begin{aligned}
& \mathbb{E} \left[\text{Vec} \left(\hat{\mathbf{R}}_{D,ST} - \mathbf{R}_{D,ST} \right) \text{Vec} \left(\hat{\mathbf{R}}_{D,ST} - \mathbf{R}_{D,ST} \right)^T \right] \\
&= \mathbb{E} \left[\text{Vec} \left(\hat{\mathbf{R}}_{D,ST} - \mathbf{R}_{D,ST} \right) \left[\text{Vec} \left(\hat{\mathbf{R}}_{D,ST} - \mathbf{R}_{D,ST} \right)^* \right]^H \right] \\
&= \mathbb{E} \left[\text{Vec} \left(\hat{\mathbf{R}}_{D,ST} - \mathbf{R}_{D,ST} \right) \left[\mathbf{K}_{cD} \text{Vec} \left(\hat{\mathbf{R}}_{D,ST} - \mathbf{R}_{D,ST} \right) \right]^H \right] \quad \text{cf. §4.1.3}
\end{aligned}$$

$$\begin{aligned}
&= \mathbb{E} \left[\text{Vec} \left(\hat{\mathbf{R}}_{D,ST} - \mathbf{R}_{D,ST} \right) \left[\text{Vec} \left(\hat{\mathbf{R}}_{D,ST} - \mathbf{R}_{D,ST} \right) \right]^H \mathbf{K}_{cD} \right] \\
&= \mathbb{E} \left[\text{Vec} \left(\hat{\mathbf{R}}_{D,ST} - \mathbf{R}_{D,ST} \right) \text{Vec} \left(\hat{\mathbf{R}}_{D,ST} - \mathbf{R}_{D,ST} \right) \right] \mathbf{K}_{cD}
\end{aligned}$$

■

Les matrices $\mathbf{A}$, $\mathbf{B}$ et $\mathbf{C}$ ci-dessus, sont relatives au format ST . On les note $\mathbf{A}_{ST}$, $\mathbf{B}_{ST}$ et $\mathbf{C}_{ST}$. On note $\mathbf{A}_{TS}$, $\mathbf{B}_{TS}$ et $\mathbf{C}_{TS}$ les matrices correspondantes dans le format TS . On montre alors,

Résultat C.13

$$\begin{aligned}
\mathbf{A}_{TS} &= \frac{1}{2\pi} \int_{w=-\pi}^{\pi} (\mathbf{M}_{TS}^*(w) \otimes \mathbf{M}_{TS}(w)) dw \\
\mathbf{B}_{TS} &= \frac{1}{2\pi} \int_{w=-\pi}^{\pi} (\mathbf{N}_{TS}^*(w) \otimes \mathbf{N}_{TS}(w)) dw \mathbf{K}_{cD} \\
\mathbf{M}_{TS}(w) &= \left[(\mathbf{S}(w)) \otimes (\mathbf{e}_D(w) \mathbf{e}_D^H(w)) \right] \\
\mathbf{N}_{TS}(w) &= \left[(\mathbf{T}(w)) \otimes (\mathbf{e}_D(w) \mathbf{e}_D^H(w)) \right] \\
\mathbf{C}_{TS} &= (\mathbf{K}_{c,D} \otimes \mathbf{K}_{c,D}) \mathbf{C}_{ST} (\mathbf{K}_{D,c} \otimes \mathbf{K}_{D,c})
\end{aligned}$$

Démonstration:

$$\begin{aligned}
&\mathbb{E} \left[\text{Vec} \left(\hat{\mathbf{R}}_{D,TS} - \mathbf{R}_{D,TS} \right) \text{Vec} \left(\hat{\mathbf{R}}_{D,TS} - \mathbf{R}_{D,TS} \right)^H \right] \\
&= \mathbb{E} \left[\text{Vec} \left(\mathbf{K}_{c,D} (\hat{\mathbf{R}}_{D,ST} - \mathbf{R}_{D,ST}) \mathbf{K}_{D,c} \right) \text{Vec} \left(\mathbf{K}_{c,D} (\hat{\mathbf{R}}_{D,ST} - \mathbf{R}_{D,ST}) \mathbf{K}_{D,c} \right)^H \right] \\
&= \mathbb{E} \left[\text{Vec} \left(\mathbf{K}_{c,D} (\hat{\mathbf{R}}_{D,ST} - \mathbf{R}_{D,ST}) \mathbf{K}_{D,c} \right) \text{Vec} \left(\mathbf{K}_{c,D} (\hat{\mathbf{R}}_{D,ST} - \mathbf{R}_{D,ST}) \mathbf{K}_{D,c} \right)^H \right] \\
&= \mathbb{E} \left[(\mathbf{K}_{c,D} \otimes \mathbf{K}_{c,D}) \text{Vec} \left((\hat{\mathbf{R}}_{D,ST} - \mathbf{R}_{D,ST}) \right) \left[(\mathbf{K}_{c,D} \otimes \mathbf{K}_{c,D}) \text{Vec} \left((\hat{\mathbf{R}}_{D,ST} - \mathbf{R}_{D,ST}) \right) \right]^H \right] \\
&= \mathbb{E} \left[(\mathbf{K}_{c,D} \otimes \mathbf{K}_{c,D}) \text{Vec} \left((\hat{\mathbf{R}}_{D,ST} - \mathbf{R}_{D,ST}) \right) \text{Vec} \left((\hat{\mathbf{R}}_{D,ST} - \mathbf{R}_{D,ST}) \right)^H (\mathbf{K}_{D,c} \otimes \mathbf{K}_{D,c}) \right] \\
&= (\mathbf{K}_{c,D} \otimes \mathbf{K}_{c,D}) \mathbb{E} \left[\text{Vec} \left((\hat{\mathbf{R}}_{D,ST} - \mathbf{R}_{D,ST}) \right) \text{Vec} \left((\hat{\mathbf{R}}_{D,ST} - \mathbf{R}_{D,ST}) \right)^H \right] (\mathbf{K}_{D,c} \otimes \mathbf{K}_{D,c}) \\
&= (\mathbf{K}_{c,D} \otimes \mathbf{K}_{c,D}) \mathbb{E} \left[\text{Vec} \left((\hat{\mathbf{R}}_{D,ST} - \mathbf{R}_{D,ST}) \right) \text{Vec} \left((\hat{\mathbf{R}}_{D,ST} - \mathbf{R}_{D,ST}) \right)^H \right] (\mathbf{K}_{D,c} \otimes \mathbf{K}_{D,c})
\end{aligned}$$

On peut écrire

$$\begin{aligned}
\mathbf{A}_{TS} &= (\mathbf{K}_{c,D} \otimes \mathbf{K}_{c,D}) \mathbf{A}_{ST} (\mathbf{K}_{D,c} \otimes \mathbf{K}_{D,c}) \\
\mathbf{B}_{TS} &= (\mathbf{K}_{c,D} \otimes \mathbf{K}_{c,D}) \mathbf{B}_{ST} (\mathbf{K}_{D,c} \otimes \mathbf{K}_{D,c}) \\
\mathbf{C}_{TS} &= (\mathbf{K}_{c,D} \otimes \mathbf{K}_{c,D}) \mathbf{C}_{ST} (\mathbf{K}_{D,c} \otimes \mathbf{K}_{D,c})
\end{aligned}$$

$$\mathbf{A}_{TS} = (\mathbf{K}_{c,D} \otimes \mathbf{K}_{c,D}) \mathbf{A}_{ST} (\mathbf{K}_{D,c} \otimes \mathbf{K}_{D,c})$$

$$\begin{aligned}
&= (\mathbf{K}_{c,D} \otimes \mathbf{K}_{c,D}) \frac{1}{2\pi} \int_{w=-\pi}^{\pi} (\mathbf{M}_{ST}^*(w) \otimes \mathbf{M}_{ST}(w)) dw (\mathbf{K}_{D,c} \otimes \mathbf{K}_{D,c}) \\
&= \frac{1}{2\pi} \int_{w=-\pi}^{\pi} (\mathbf{K}_{c,D} \otimes \mathbf{K}_{c,D}) (\mathbf{M}_{ST}^*(w) \otimes \mathbf{M}_{ST}(w)) (\mathbf{K}_{D,c} \otimes \mathbf{K}_{D,c}) dw \\
&= \frac{1}{2\pi} \int_{w=-\pi}^{\pi} [(\mathbf{K}_{c,D} \mathbf{M}_{ST}^*(w) \mathbf{K}_{D,c}) \otimes (\mathbf{K}_{c,D} \mathbf{M}_{ST}(w) \mathbf{K}_{D,c})] dw \\
&= \frac{1}{2\pi} \int_{w=-\pi}^{\pi} (\mathbf{M}_{TS}^*(w) \otimes \mathbf{M}_{TS}(w)) dw
\end{aligned}$$

avec

$$\begin{aligned}
\mathbf{M}_{TS}(w) &= \mathbf{K}_{c,D} \mathbf{M}_{ST}(w) \mathbf{K}_{D,c} \\
&= \mathbf{K}_{c,D} \left[\left(\mathbf{e}_D(w) \mathbf{e}_D^H(w) \right) \otimes (\mathbf{S}(w)) \right] \mathbf{K}_{D,c} \\
&= \left[(\mathbf{S}(w)) \otimes \left(\mathbf{e}_D(w) \mathbf{e}_D^H(w) \right) \right]
\end{aligned}$$

$$\begin{aligned}
\mathbf{B}_{TS} &= (\mathbf{K}_{c,D} \otimes \mathbf{K}_{c,D}) \mathbf{B}_{ST} (\mathbf{K}_{D,c} \otimes \mathbf{K}_{D,c}) \\
&= (\mathbf{K}_{c,D} \otimes \mathbf{K}_{c,D}) \frac{1}{2\pi} \int_{w=-\pi}^{\pi} (\mathbf{N}_{ST}^*(w) \otimes \mathbf{N}_{ST}(w)) dw \mathbf{K}_{cD} (\mathbf{K}_{D,c} \otimes \mathbf{K}_{D,c}) \\
&= \frac{1}{2\pi} \int_{w=-\pi}^{\pi} (\mathbf{K}_{c,D} \otimes \mathbf{K}_{c,D}) (\mathbf{N}_{ST}^*(w) \otimes \mathbf{N}_{ST}(w)) \mathbf{K}_{cD} (\mathbf{K}_{D,c} \otimes \mathbf{K}_{D,c}) dw \\
&= \frac{1}{2\pi} \int_{w=-\pi}^{\pi} (\mathbf{K}_{c,D} \otimes \mathbf{K}_{c,D}) (\mathbf{N}_{ST}^*(w) \otimes \mathbf{N}_{ST}(w)) (\mathbf{K}_{D,c} \otimes \mathbf{K}_{D,c}) \mathbf{K}_{cD} dw \\
&= \frac{1}{2\pi} \int_{w=-\pi}^{\pi} (\mathbf{N}_{TS}^*(w) \otimes \mathbf{N}_{TS}(w)) \mathbf{K}_{cD} dw
\end{aligned}$$

avec

$$\begin{aligned}
\mathbf{N}_{TS}(w) &= \mathbf{K}_{c,D} \mathbf{N}_{ST}(w) \mathbf{K}_{D,c} \\
&= \mathbf{K}_{c,D} \left[\left(\mathbf{e}_D(w) \mathbf{e}_D^H(w) \right) \otimes (\mathbf{T}(w)) \right] \mathbf{K}_{D,c} \\
&= \left[(\mathbf{T}(w)) \otimes \left(\mathbf{e}_D(w) \mathbf{e}_D^H(w) \right) \right]
\end{aligned}$$

■

Annexe D

Démonstration de Prop. 4.11

$$\begin{aligned}
& \mathbf{C}_{k,l} \\
&= \sum_{u=-\infty}^{\infty} \text{cum} \left(x^{c(a(k))} (-T(a(k))), x^{c(b(l))} (u - T(b(l))), x^{c(b(k))} (-T(b(k)))^*, x^{c(a(l))} (u - T(a(l)))^* \right) \\
&= \sum_{u=-\infty}^{\infty} \text{cum} \left(\sum_{t_1=-\infty}^{\infty} s_{t_1} h^{c(a(k))} (-T(a(k)) - t_1) \sum_{t_2=-\infty}^{\infty} s_{t_2} h^{c(b(l))} (u - T(b(l)) - t_2) \right. \\
&\quad \left. \sum_{t_3=-\infty}^{\infty} s_{t_3}^* h^{c(b(k))} (-T(b(k)) - t_3)^* \sum_{t_4=-\infty}^{\infty} s_{t_4}^* h^{c(a(l))} (u - T(a(l)) - t_4)^* \right) \\
&= \sum_{u=-\infty}^{\infty} \sum_{t_1, t_2, t_3, t_4=-\infty}^{\infty} h^{c(a(k))} (-T(a(k)) - t_1) h^{c(b(l))} (u - T(b(l)) - t_2) \\
&\quad h^{c(b(k))} (-T(b(k)) - t_3)^* h^{c(a(l))} (u - T(a(l)) - t_4)^* \text{cum}(s_{t_1}, s_{t_2}, s_{t_3}^*, s_{t_4}^*)
\end{aligned}$$

Si parmi t_1, t_2, t_3 et t_4 , deux sont différents, $\text{cum}(s_{t_1}, s_{t_2}, s_{t_3}^*, s_{t_4}^*)$ sera nul. De plus, par stationnarité, $\text{cum}(s_{t_1}, s_{t_1}, s_{t_1}^*, s_{t_1}^*) = \text{cum}(s_0, s_0, s_0^*, s_0^*)$. On note $\mathcal{K}_{s_4} \triangleq \text{cum}(s_0, s_0, s_0^*, s_0^*)$.

$$\begin{aligned}
\mathbf{C}_{k,l} &= \mathcal{K}_{s_4} \sum_{u=-\infty}^{\infty} \sum_{t=-\infty}^{\infty} \\
&\quad h^{c(a(k))} (-T(a(k)) - t) h^{c(b(l))} (u - T(b(l)) - t) h^{c(b(k))} (-T(b(k)) - t)^* h^{c(a(l))} (u - T(a(l)) - t)^* \\
&= \mathcal{K}_{s_4} \sum_{t=-\infty}^{\infty} \sum_{u=-\infty}^{\infty} \\
&\quad h^{c(a(k))} (-T(a(k)) - t) h^{c(b(l))} (u - T(b(l)) - t) h^{c(b(k))} (-T(b(k)) - t)^* h^{c(a(l))} (u - T(a(l)) - t)^* \\
&= \mathcal{K}_{s_4} \sum_{t=-\infty}^{\infty} h^{c(a(k))} (-T(a(k)) - t) h^{c(b(k))} (-T(b(k)) - t)^* \\
&\quad \sum_{u=-\infty}^{\infty} h^{c(b(l))} (u - T(b(l)) - t) h^{c(a(l))} (u - T(a(l)) - t)^* \\
&= \mathcal{K}_{s_4} \sum_{t=-\infty}^{\infty} h^{c(a(k))} (-T(a(k)) - t) h^{c(b(k))} (-T(b(k)) - t)^* \\
&\quad \sum_{u=-\infty}^{\infty} h^{c(b(l))} (u - T(b(l))) h^{c(a(l))} (u - T(a(l)))^*
\end{aligned}$$

$$\begin{aligned}
&= \mathcal{K}_{s_4} \sum_{t=-\infty}^{\infty} h^{c(a(k))}(t) h^{c(b(k))}(t + T(a(k)) - T(b(k)))^* \\
&\quad \left(\sum_{u=-\infty}^{\infty} h^{c(a(l))}(u) h^{c(b(l))}(u + T(a(l)) - T(b(l)))^* \right)^*
\end{aligned}$$

On défini le vecteur $\mathbf{v} \in \mathbf{C}^{c^2 D^2}$ tel que

$$\begin{aligned}
\mathbf{v}_k &\triangleq \sum_{t=-\infty}^{\infty} h^{c(a(k))}(t) h^{c(b(k))}(t + T(a(k)) - T(b(k)))^* \\
&= \sum_{t=0}^M h^{c(a(k))}(t) h^{c(b(k))}(t + T(a(k)) - T(b(k)))^* \\
&= \gamma_{c(a(k)), c(b(k))}(T(b(k)) - T(a(k))) \\
&= (\mathbf{R}_{D,ST})_{a(k), a(k)} \\
&= \text{Vec}(\mathbf{R}_{D,ST})_k \\
\text{Ainsi } \mathbf{v} &= \text{Vec}(\mathbf{R}_{D,ST})
\end{aligned}$$

$$\begin{aligned}
\text{Par conséquent, } \mathbf{C}_{k,l} &= \mathcal{K}_{s_4} \mathbf{v}_k \mathbf{v}_l^* \\
\text{et donc } \mathbf{C} &= \mathcal{K}_{s_4} \text{Vec}(\mathbf{R}_{D,ST}) \text{Vec}(\mathbf{R}_{D,ST})^H
\end{aligned}$$

$\mathbf{C}$ ci-dessus calculé correspond au format ST , soit

$$\begin{aligned}
\mathbf{C}_{ST} &= \mathcal{K}_{s_4} \text{Vec}(\mathbf{R}_{D,ST}) \text{Vec}(\mathbf{R}_{D,ST})^H \\
\text{On déduit } \mathbf{C}_{TS} &= (\mathbf{K}_{c,D} \otimes \mathbf{K}_{c,D}) \mathbf{C}_{ST} (\mathbf{K}_{D,c} \otimes \mathbf{K}_{D,c}) \\
&= \mathcal{K}_{s_4} (\mathbf{K}_{c,D} \otimes \mathbf{K}_{c,D}) \text{Vec}(\mathbf{R}_{D,ST}) \text{Vec}(\mathbf{R}_{D,ST})^H (\mathbf{K}_{D,c} \otimes \mathbf{K}_{D,c}) \\
&= \mathcal{K}_{s_4} (\mathbf{K}_{c,D} \otimes \mathbf{K}_{c,D}) \text{Vec}(\mathbf{R}_{D,ST}) ((\mathbf{K}_{c,D} \otimes \mathbf{K}_{c,D}) \text{Vec}(\mathbf{R}_{D,ST}))^H
\end{aligned}$$

Or $(\mathbf{K}_{c,D} \otimes \mathbf{K}_{c,D}) \text{Vec}(\mathbf{R}_{D,ST}) = \text{Vec}(\mathbf{K}_{c,D} \mathbf{R}_{D,ST} \mathbf{K}_{D,c}) = \text{Vec}(\mathbf{R}_{D,TS})$, et donc

$$\mathbf{C}_{TS} = \mathcal{K}_{s_4} \text{Vec}(\mathbf{R}_{D,TS}) \text{Vec}(\mathbf{R}_{D,TS})^H$$

Annexe E

Théorie de perturbation des matrices

Soit $\mathbf{X}$ une matrice $m \times n$ avec $m \geq n$. k étant le rang de $\mathbf{X}$. le sous-espace colonne $\mathcal{R}(\mathbf{X}) \subset \mathbf{C}^m$ et le sous espace ligne $\mathcal{R}(\mathbf{X}^H) \subset \mathbf{C}^n$ ont la même dimension k . Les projecteurs respectifs sont notés $\mathbf{P}$ ($m \times m$) et $\mathbf{R}$ ($n \times n$). On a $\mathbf{P} = \mathbf{X}\mathbf{X}^\dagger$ et $\mathbf{R} = \mathbf{X}^\dagger\mathbf{X}$. Dans la suite, on se donne $\mathbf{E}$ matrice de 'perturbation faible'.

Proposition E.1 Perturbation de la Pseudo-inverse¹ [STEW 90, (3.35), p.150]

$$(\mathbf{X} + \mathbf{E})^\dagger = \mathbf{X}^\dagger - \mathbf{X}^\dagger \mathbf{P} \mathbf{E} \mathbf{R} \mathbf{X}^\dagger + (\mathbf{X}^H \mathbf{X})^\dagger \mathbf{R} \mathbf{E}^H (\mathbf{I}_m - \mathbf{P}) - (\mathbf{I}_n - \mathbf{R}) \mathbf{E}^H \mathbf{P} (\mathbf{X} \mathbf{X}^H)^\dagger + O(\|\mathbf{E}\|^2)$$

Si $\mathbf{X}$ est Hermitienne, alors

$$(\mathbf{X} + \mathbf{E})^\dagger = \mathbf{X}^\dagger - \mathbf{X}^\dagger \mathbf{E} \mathbf{X}^\dagger + \mathbf{X}^{\dagger^2} \mathbf{E}^H (\mathbf{I}_m - \mathbf{P}) - (\mathbf{I}_n - \mathbf{P}) \mathbf{E}^H \mathbf{X}^{\dagger^2} + O(\|\mathbf{E}\|^2)$$

Proposition E.2 Perturbation du projecteur image² [STEW 90, p.155]

$$\hat{\mathbf{P}} = \mathbf{P} + (\mathbf{I}_m - \mathbf{P}) \mathbf{E} \mathbf{R} \mathbf{X}^\dagger + \mathbf{X}^{\dagger^H} \mathbf{R} \mathbf{E}^H (\mathbf{I}_m - \mathbf{P}) + O(\|\mathbf{E}\|^2)$$

Si $\mathbf{X}$ est Hermitienne

$$\hat{\mathbf{P}} = \mathbf{P} + (\mathbf{I}_m - \mathbf{P}) \mathbf{E} \mathbf{X}^\dagger + \mathbf{X}^\dagger \mathbf{E}^H (\mathbf{I}_m - \mathbf{P}) + O(\|\mathbf{E}\|^2)$$

Proposition E.3 Perturbation des vecteurs propres³ simples [STEW 90, §2.3, p.240]

Si $\mathbf{X}$ est carrée et si λ est une valeur propre de multiplicité⁴ 1 associée au vecteur propre $\mathbf{x}$, alors

$$\hat{\mathbf{x}} = \mathbf{x} - (\mathbf{X} - \lambda \mathbf{I}_n)^\dagger \mathbf{E} \mathbf{x} + O(\|\mathbf{E}\|^2)$$

où $(.)^\dagger$ est la pseudo-inverse de groupe ou de Drazin [STEW 90, p.133], qui est égale à la pseudo-inverse de Moore $(.)^\dagger$ lorsque la matrice est Hermitienne.

Le projecteur $\mathbf{P}_x \triangleq \mathbf{x} \mathbf{x}^H$ associé est perturbé au sens

$$\hat{\mathbf{P}}_x \triangleq \hat{\mathbf{x}} \hat{\mathbf{x}}^H = \mathbf{P}_x - \mathbf{P}_x \mathbf{E} (\mathbf{X} - \lambda \mathbf{I}_n)^\dagger - (\mathbf{X} - \lambda \mathbf{I}_n)^\dagger \mathbf{E} \mathbf{P}_x + O(\|\mathbf{E}\|^2)$$

¹au sens de Moore. Lorsque $\mathbf{X}$ est inversible, remplacer $\mathbf{P}$ et $\mathbf{R}$ par $\mathbf{I}_m$ [STEW 90, p.118].

²C'est la même perturbation sur le noyau vu que $(\mathbf{I}_m - \hat{\mathbf{P}}) - (\mathbf{I}_m - \mathbf{P}) = (\mathbf{P} - \hat{\mathbf{P}})$.

³à gauche.

⁴multiplicité géométrique [STEW 90, p.15].

Bibliographie

- [ABED 95] Karim Abed-Meraim, “Identification-égalisation autodidacte au second ordre : algorithmes et performances”, Thèse de Doctorat, École Nationale Supérieure des Télécommunications, 1995.
- [ABED 97] Karim Abed-Meraim, Eric Moulines and Philippe Loubaton, “Prediction Error Method for Second-Order Blind Identification”, IEEE Transactions on Signal Processing, vol. 45, no. 3, pp. 694-705, March 1997.
- [ABEDa97] Karim Abed-Meraim, Jean Francois Cardoso, Alexei Gorokhov, Philippe Loubaton and Eric Moulines, “On Subspace Methods for the Blind Identification of Single-Input Multiple-Output FIR systems”, IEEE Transactions on Signal Processing, vol. 45, no. 1, pp. 42-55, January 1997.
- [Abed 97] Karim Abed-Meraim, Pierre Duhamel, David Gesbert, Philippe Loubaton, Sylvie Mayrargue, Eric Moulines and Dirk T.M. Slock, “Prediction Error Method for Time-Domain Blind Identification of Multichannel FIR Filters”, IEEE ICASSP Conference, Atlanta, GA, May 1995, Conference Proceedings, pp. 1968-1971.
- [AfkH 95] Kaywan H. Afkhamie and Zhi-Quan Luo, “Blind Equalization Using Second-Order Statistics”, IEEE ICASSP Conference, Atlanta, GA, May 1995, Conference Proceedings, pp. 1053-1056.
- [BJRK 96] Åke Björck, “Numerical Methods for Least Squares Problems”, SIAM, 1996.
- [BENE 80] Albert Beneviste, Maurice Goursat and Gabriel Ruget, “Robust Identification of a Non-minimum Phase System: Blind Adjustment of a Linear Equalizer in Data Communications”, IEEE Transactions on Automatic Control, vol. AC-25, no. 3, pp. 385-399, June 1980.
- [BOSE 98] N. K. Bose and K. J. Boo, “Asymptotic Eigenvalue Distribution of Block-Toeplitz Matrices”, IEEE Transactions on Information Theory, vol. 44, no. 2, pp. 858-861, March 1998.
- [BROS 97] Jean Marc Brossier, “Signal et communication numérique : égalisation et synchronisation”, Hermes, 1997.
- [Carv 98] Elisabeth de Carvalho and Dirk T.M. Slock, “Identifiability Conditions for Blind and Semi-Blind Multichannel Estimation”, EUSIPCO Conference, Greece, September 1998.

- [Cibl 98] Philippe Ciblat and Philippe Loubaton, "Second Order Blind Equalization: The Bandlimited case", IEEE ICASSP Conference, Seattle, WA, May 1998.
- [DING 91] Zhi Ding, Rodney A. Kennedy, Brian D. O. Anderson and C. Richard Johnson, Jr., "Ill-Convergence of Godard Blind Equalizers in Data Communication Systems", IEEE Transactions on Communications, vol. 39, no. 9, pp. 1313-1327, September 1991.
- [DING 92] Zhi Ding, C. Richard Johnson, Jr. and Rodney A. Kennedy, "On the (Non)Existence of Undesirable Equilibria of Godard Blind Equalizers", IEEE Transactions on Signal Processing, vol. 40, no. 10, pp. 2425-2432, October 1992.
- [Ding 95] Zhi Ding and Z. Mao, "Knowledge Based Identification of Fractionally Sampled Channels", IEEE ICASSP Conference, Atlanta, GA, May 1995, Conference Proceedings, pp. 1996-1999.
- [DING 96] Zhi Ding, "Characteristics of Band-Limited Channels Unidentifiable From Second-Order Cyclostationary Statistics", IEEE Signal Processing Letters, vol. 3, no. 5, pp. 150-152, May 1996.
- [DING 97] Zhi Ding, "Multipath Channel Identification Based on Partial System Information", IEEE Transactions on Signal Processing, vol. 45, no. 1, pp. 235-240, January 1997.
- [DINGa97] Zhi Ding, "Matrix Outer Product Decomposition Method for Blind Multiple Channel Identification", IEEE Transactions on Signal Processing, vol. 45, no. 12, pp. 3053-3061, December 1997.
- [FIJA 97] Inbar Fijalkow, Azzèdine Touzni and John R. Treichler, "Fractionally Spaced Equalization Using CMA: Robustness to Channel Noise and Lack of Disparity", IEEE Transactions on Signal Processing, vol. 45, no. 1, pp. 56-66, January 1997.
- [GESB 97] David Gesbert, Pierre Duhamel and Sylvie Mayrargue, "On-line Blind Multichannel Equalization Based on Mutually Referenced Filters", IEEE Transactions on Signal Processing, vol. 45, no. 9, pp. 2307-2317, September 1997.
- [Gesb 97] David Gesbert and Pierre Duhamel, "Robust Blind Channel Identification and Equalization Based on Multi-Step Predictors", IEEE ICASSP Conference, Munich, Germany, April 1997, Conference Proceedings, pp. 3621-3624.
- [GESB 00] David Gesbert and Pierre Duhamel, "Unbiased Blind Adaptive Channel Identification and Equalization of Multichannel FIR Filters", IEEE Transactions on Signal Processing, vol. 48, no. 1, pp. 148-158, January 2000.
- [GIAN 97] Georgios B. Giannakis and Steven D. Halford, "Blind Fractionally Spaced Equalization of Noisy FIR Channels: Direct and Adaptive Solutions", IEEE Transactions on Signal Processing, vol. 45, no. 9, pp. 2277-2292, September 1997.
- [Gian 97] Georgios B. Giannakis and Steven D. Halford, "Blind Fractionally-Spaced Equalization of Noisy FIR Channels: Adaptive and Optimal Solutions", IEEE ICASSP Conference, Atlanta, GA, May 1995, Conference Proceedings, pp. 1972-1975.

- [GODR 80] Dominique N. Godard, "Self-Recovering Equalization and Carrier Tracking in Two-Dimensional Data Communication Systems", IEEE Transactions on Communications, vol. COM-28, no. 11, pp. 1867-1875, November 1980.
- [GORO 97] Alexei Gorokhov, "Séparation autodidacte de mélanges convolutifs : méthodes de second ordre - Blind separation of convolutive mixtures : second order methods", Thèse de Doctorat, École Nationale Supérieure des Télécommunications, 1997.
- [GORO 99] Alexei Gorokhov and Philippe Loubaton, "Blind Identification of MIMO-FIR systems: A generalized prediction approach", Signal Processing, 73 (1999), pp. 105-124.
- [GRAY 72] Robert Molten Gray, "On the Asymptotic Eigenvalue Distribution of Toeplitz Matrices", IEEE Transactions on Information Theory, vol. IT-18, no. 6, pp. 725-730, November 1972.
- [GRAY 98] Robert Molten Gray, "Toeplitz and Circulant Matrices: A review", Information Theory Laboratory. Stanford University. Technical report, July 1998.
- [GREN 84] Ulf Grenander and Gabor Szegő, "Toeplitz Forms and Their Applications", Chelsea publishing company, N. Y., 1984.
- [Gzh 98] Jean-Pierre Delmas, Houcem Gazzah and A. P. Liavas, "Statistical Analysis of Second-Order Blind Channel Identification Under Practical Conditions, Application to the Least Squares and Subspace Methods", IMA Conference on Mathematics in Communications, Loughborough, UK, 15-17 December 1998.
- [Gzh 00] Houcem Gazzah, Phillip A. Regalia and Jean-Pierre Delmas, "A Blind Identification Algorithm Robust to Order Over Estimation", IEEE ICASSP Conference, Istanbul, Turkey, June 2000.
- [GZH a] Jean-Pierre Delmas, Houcem Gazzah and A. P. Liavas, "Statistical Analysis of Some Second-Order Methods for Blind Channel Identification/Equalization with Respect to Channel Undermodeling", IEEE Transactions On Signal Processing, vol. 48, no. 7, pp. 1984-1998, July 2000.
- [GZH b] Houcem Gazzah, Phillip A. Regalia and Jean-Pierre Delmas, "Asymptotic Eigenvalue Distribution of Block Toeplitz Matrices and Application to Blind SIMO Channels Identification", à paraître dans IEEE Transactions on Information Theory, janvier 2001.
- [GZH c] Houcem Gazzah, Phillip A. Regalia, Jean-Pierre Delmas and Karim Abed-Meraim, "A Blind Multichannel Identification Algorithm Robust to Order Over Estimation", soumis à IEEE Transactions on Signal Processing, octobre 2000.
- [Gzh 01] Houcem Gazzah, Phillip A. Regalia and Jean-Pierre Delmas, "On Blind (non)Identifiability of Dispersive Bandlimited Channels", soumis à IEEE ICASSP, Salt Lake City, Utah, mai 2001.
- [HANN 70] , E. J. Hannan, "Multiple Time Series", Wiley, 1970.

- [HEND 81] Harold V. Henderson and S. R. Searle, "The Vec-Permutation Matrix, The Vec Operator and Kronecker Products: A Review", *Linear and Multilinear Algebra*, 1981, vol. 9, pp. 271-288.
- [HORN 91] Roger A. Horn and Charles R. Johnson, "Matrix Analysis", Cambridge University Press, 1991.
- [HORN91] Roger A. Horn and Charles R. Johnson, "Topics in Matrix Analysis", Cambridge University Press, 1991.
- [KNDY] Robert Kennedy, "Fading Dispersive Communication Channels", J. Wiley and Sons.
- [LAC 97] J. L. Lacoume, P.O. Amblard and P. Comon "Statistiques d'ordre supérieur pour le traitement du signal", Masson, 1997.
- [LEE 95] William C. Y. Lee, "Mobile Cellular Telecommunications, Analog and Digital Systems", McGRAW-HILL, Inc, 1995.
- [LI 94] Ye Li and K. J. Ray Liu, "ARMA System Identification Based on Second Order Cyclostationarity", *IEEE Transactions on Signal Processing*, vol. 42, no. 12, pp. 3483-3494, Decembre 1994.
- [LI 96] Ye Li and K. J. Ray Liu, "Blind Adaptive Equalization of SIMO Channels based on Second-Order Statistics", Technical Report, University of Maryland, December, 1996.
- [LIAV 99] A. P. Liavas, P. A. Regalia and J.P. Delmas, "Robustness of Least Squares and Subspace Methods for Blind Channel Identification/Equalization with Respect to Effective Channel Undermodeling", *IEEE Transactions on Signal Processing*, vol. 47, no. 6, pp. 1636-1645, June 1999.
- [LIAV 99] A. P. Liavas, P. A. Regalia and J.P. Delmas, "Blind Channel Approximation: Effective Channel Order Determination", *IEEE Transactions on Signal Processing*, vol. 47, no. 12, pp. 3336-3344, December 1999.
- [LIAV 00] A. P. Liavas, P. A. Regalia and J.P. Delmas, "On the Robustness of the Linear Prediction Method for Blind Channel Identification with Respect to Effective Channel Undermodeling/Overmodeling", *IEEE Transactions on Signal Processing*, vol. 48, no. 5, pp. 1477-1480, May 2000.
- [LIND 99] Erik Lindskog, "Space-Time Processing and Equalization for Wireless Communications", Thèse de Doctorat, Uppsala University, 1999.
- [LIU 94] Hui Liu and Guanghan Xu, "A Deterministic Approach to Blind Symbol Estimation", *IEEE Signal Processing Letters*, vol. 1, no. 12, pp. 205-207, December 1994.
- [Liu 95] Hui Liu and Guanghan Xu, "Multiuser Blind Channel Estimation and Spatial Channel Pre-Equalization", *IEEE ICASSP Conference, Atlanta, GA, May 1995, Conference Proceedings*, pp. 1756-1759.

- [Liu 97] Ruey-Wen Liu and Hui Luo, "Some Results on Multiple-Channel Blind Equalization", European Conferences on Circuit Theory and Design, Budapest, Hungary, September 1997, Conference Proceedings, pp. 1347-1351.
- [MACC 98] Odile Macchi, "L'égalisation numérique en communications", Ann. Télécommun., 53, n 1-2, pp. 39-58, 1998.
- [MEND 91] Jerry M. Mendel, "Tutorial on Higher-Order Statistics (Spectra) in Signal Processing and System Theory: Theoretical Results and Some Applications", Proceedings of the IEEE, vol. 79, no. 3, pp. 278-305, March 1991.
- [MEYE 88] Carl D. Meyer and G. W. Stewart, "Derivatives and Perturbations of Eigenvectors", SIAM Journal on Numerical Analysis, vol. 25, no. 3, pp. 679-691, June 1988.
- [MIRA 96] Michele Miranda and Paolo Tilli, "Block Toeplitz Matrices and Preconditioning", Calcolo, 33, pp. 79-86, 1996.
- [MORG 98] Dennis R. Morgan, Jacob Benesty and M. Mohan Sondhi, "On the Evaluation of Estimated Impulse Responses", IEEE Signal Processing Letters, vol. 5, no. 7, pp. 174-176, July 1998.
- [MOUL 95] Eric Moulines, Pierre Duhamel, Jean Francois Cardoso and Sylvie Mayrargue, "Subspace Methods for the Blind Identification of Multichannel FIR Filters", IEEE Transactions on Signal Processing, vol. 43, no. 2, pp. 516-525, February 1995.
- [NAGB 96] Ayman F. Naguib, "Adaptive Antennas for CDMA Wireless Networks", Thèse de Doctorat, Stanford University, 1996.
- [PAPA 95] Constantinos B. Papadias, "Méthodes d'égalisation aveugle pour les canaux linéaires", Thèse de Doctorat, École Nationale Supérieure des Télécommunications, Mars 1995.
- [PAPA 98] Constantinos B. Papadias, David Gesbert and Arogyaswami Paulraj, "Robust Second-Order Blind Equalization of Polyphase Channels", EUSIPCO Conference, Greece, 8-11 September 1998, Conference Proceedings, pp. 2045-2048.
- [PAUL 97] Arogyaswami Paulraj and Constantinos B. Papadias, "Array Processing for Mobile Communications", in "Handbook in Signal Processing", CRC Press, 1997.
- [PERR 99] Lisa Perros-Meilhac, Eric Moulines, Pascal Chevalier et Pierre Duhamel, "Identification paramétrique et égalisation autodidacte d'un canal spéculaire", conférence GRETSI, Vannes, France, 13-17 septembre 1999.
- [POZI 97] Haralambos Pozidis and Athina P. Petropulu, "Cross-Spectrum Based Blind Channel Identification", IEEE Transactions on Signal Processing, vol. 45, no. 12, pp. 2977-2993, December 1997.
- [QIU 96] Wanzhi Qiu, Yingbo Hua, "Performance analysis of the Subspace Method for Blind Channel Identification", Signal Processing, 50 (1996), pp. 71-81.

- [QIU 97] Wanzhi Qiu and Yingb Hua, “Performance Analysis of the Matrix Pair Method for Blind Channel Identification”, IEEE Transactions on Information Theory, vol. 43, no. 4, pp. 1245-1253, July 1997.
- [Redd 95] V. U. Reddy, C. B. Papadias and A. Paulraj, “Second-Order Blind Identifiability of Certain Classes of Multipath Channels Using Antenna Arrays”, IEEE ICASSP Conference, Munich, Germany, April 1997, Conference Proceedings, pp. 3465-3468.
- [SATO 75] Y. Sato, “A Method of Self-Recovering Equalization”, IEEE Transactions on Communications, vol. 23(6), pp. 679-682, June 1975.
- [SCHL 94] S. Schell, D. L. Smith and S. Roy, “Blind Channel Identification and Equalization Using Subchannel Response Matching”, 26th Conference on Information Sciences and Systems , Princeton, NJ, March 1994.
- [SCHO 97] James R. Schott, “Matrix Analysis for Statistics”, Wiley, 1997.
- [SERR 98] Stefano Serra, “On the Extreme Eigenvalues of Hermitian (Block) Toeplitz Matrices”, Linear Algebra and its Applications, 270, pp.109-129, 1998.
- [SHAL 90] Ofir Shalvi and Ehud Weinstein, “New Criteria for Blind Deconvolution of Nonminimum Phase Systems (Channels)”, IEEE Transactions on Information Theory, vol. 36, no. 2, pp. 312-321, March 1990.
- [Slock 94] D. Slock, “Blind fractionally-spaced equalization, perfect reconstruction filterbanks, and multilinear prediction”, IEEE ICASSP Conference, Adelaide, Australia, April 1994.
- [Slock 95] Dirk T.M. Slock and Constantinos B. Papadias, “Further Results on Blind Identification and Equalization of Multiple FIR Channels”, IEEE ICASSP Conference, Atlanta, GA, May 1995, Conference Proceedings, pp. 1964-1967.
- [SOUL 93] Antoine Souloumiac, “Utilisation des statistiques d’ordre supérieur pour le filtrage et la séparation des sources en traitement d’antennes”, Thèse de Doctorat, École Nationale Supérieure des Télécommunications, 1993.
- [STEW 90] G. W. Stewart and Ji-guang Sun, “Matrix Perturbation Theory”, Academic Press, 1990.
- [STOI 89] Torsten Söderström and Petre Stoica, “System Identification”, Prentice Hall, 1989.
- [TILL 98] Paolo Tilli, “Singular values and eigenvalues of non-Hermitian block Toeplitz matrices”, Linear Algebra and its Applications, 272, pp. 59-89, 1998.
- [TILL 99] Paolo Tilli, “Asymptotic Spectral Distribution of Toeplitz-Related Matrices”, in “Fast Reliable Algorithms for Matrices with Structure”, pp. 153-187, Edited by Thomas Kailath and A. H. Sayed, SIAM 1999.
- [TILL 97] Paolo Tilli, “On the Asymptotic Spectrum of Hermitian Block Toeplitz Matrices with Toeplitz Blocks”, Mathematics of Computation, vol. 6, no. 219, 1997, pp. 1147-1159.
- [TILL] Paolo Tilli, “Asymptotic Spectrum of Hermitian Block Toeplitz Matrices and Preconditioning Results”, SIAM J. Matrix Analysis Applications, to appear.

- [TONG 94] Lang Tong, Guanghan Xu and Thomas Kailath, “Blind Identification and Equalization Based on Second-Order Statistics : A Time Domain Approach”, IEEE Transactions on Information Theory, vol. 40, no. 2, pp. 340-349, March 1994.
- [TONG 95] Lang Tong, Guanghan Xu, B Hassibi and Thomas Kailath, “Blind Identification and Equalization Based on Second-Order Statistics : A Frequency-Domain Approach”, IEEE Transactions on Information Theory, vol. 41, no. 1, pp. 329-334, January 1995.
- [TONG 98] Lang Tong and Sylvie Perreau, “Multichannel Blind Identification: From Subspace to Maximum Likelihood Methods”, Proceedings of the IEEE , vol. 86, no. 10, October 1998.
- [Touz 98] Azzèdine Touzni and Inbar Fijalkow, “Blind Multichannel Equalization with Controlled Delay”, EUSIPCO Conference, Greece, September 1998, Conference Proceedings, pp. 1661-1664.
- [TOUZ 98] Azzèdine Touzni “Performance et robustesse en galisation aveugle spatio-temporelle”, Thèse de Doctorat, Université de Cergy-Pontoise, 1998.
- [TRCH 83] John .R. Treichler and Brian G. Agee, “A New Approach to Multipath Correction of Constant Modulus Signals”, IEEE Transactions on Acoustics, Speech, and Signal Processing, vol. ASSP-31, no. 2, pp. 459-471, April 1983.
- [TRCH 96] J.R. Treichler, I. Fijalkow and C.R. Johnson, Jr., “Fractionally Spaced Equalizers, How long should they really be”, IEEE Signal Processing Magazine, pp. 65-81, May 1996.
- [TUGN 95] Jitendra K. Tugnait, “On Blind Identifiability of Multipath Channels Using Fractional Sampling and Second-Order Cyclostationary Statistics”, IEEE Transactions on Information Theory, vol. 41, no. 1, pp. 308-311, January 1995.
- [VAID 93] P. P. Vaidyanathan, “Multirate Systems and Filter Banks”, Prentice Hall, 1993.
- [Veen 95] Alle-Jan van der Veen, “Resolution Limits of Blind Multi-user Multi-channel Identification Schemes - The Bandlimited Case”, IEEE ICASSP Conference, Atlanta, GA, May 1995, Conference Proceedings, pp. 2722-2727.
- [Veena95] Alle-Jan van der Veen, Shilpa Talwar and Arogyaswami Paulraj, “Blind Identification of FIR Channels Carrying Multiple Finite Alphabet Signals”, IEEE ICASSP Conference, Atlanta, GA, May 1995, Conference Proceedings, pp. 1213-1216.
- [Veen 98] Michaela C. Vanderveen and Arogyaswami Paulraj, “Improved Blind Channel Identification Using a Parametric Approach”, IEEE Communications Letters, vol. 2, no. 8, pp. 226-228, August 1998.
- [VOOI 96] Paul A. Voois, “A Theorem on the Asymptotic Eigenvalue Distribution of Toeplitz-Block-Toeplitz Matrices”, IEEE Transactions on Signal Processing, vol. 44, no. 7, pp. 1837-1841, July 1996.
- [WAX 85] Mati Wax and Thomas Kailath, “Detection of Signals by Information Theoretic Criteria”, IEEE Transactions on Acoustics, Speech and Signal Processing, vol. ASSP-33, no. 2, pp. 387-392, April 1985.

- [XU 95] Guanghai Xu, Hui Liu, Lang Tong and Thomas Kailath, “A Least Square Approach to Blind Channel Identification”, IEEE Transactions on Signal Processing, vol. 43, no. 12, pp. 2982-2993, December 1995.
- [YANG 97] Howard Hua Yang and Yingbo Hua, “On Performance of Cross-Relation Method for Blind Channel Identification”, Signal Processing, 62 (1997), pp. 187-205.
- [ZENG 96] Hanks H. Zeng and Lang Tong, “Connections Between the Least Squares and the Subspace Approaches to Blind Channel Estimation”, IEEE Transactions on Signal Processing, vol. 44, no. 6, pp. 1593-1596, June 1996.