



HAL
open science

Séparation de sources non linéaires dans le cas des mélanges convolutifs

Marc Castella

► **To cite this version:**

Marc Castella. Séparation de sources non linéaires dans le cas des mélanges convolutifs. Traitement du signal et de l'image [eess.SP]. Université de Marne la Vallée, 2004. Français. NNT : . tel-00714473

HAL Id: tel-00714473

<https://theses.hal.science/tel-00714473>

Submitted on 4 Jul 2012

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



THÈSE

soutenue le 20/12/2004 pour obtenir

le grade de **Docteur en Sciences** de l'Université de Marne la Vallée

Spécialité : Traitement du Signal

par

Marc CASTELLA

Séparation de sources non linéaires dans le cas des mélanges convolutifs

Composition de la commission d'examen :

Président :	Phillip REGALIA
Rapporteurs :	Pierre COMON Christian JUTTEN
Examineurs :	Eric MOREAU Antoine CHEVREUIL (co-encadrant)
Directeur de thèse :	Jean-Christophe PESQUET

Remerciements

D'abord, je tiens tout particulièrement à exprimer ma reconnaissance envers mon directeur de thèse, Jean-Christophe Pesquet. Au cours des trois années de travail commun, il a assuré un encadrement sans faille : sa constante disponibilité, ses conseils scientifiques et ses encouragements ont sans aucun doute été déterminants dans l'aboutissement de cette thèse.

Je souhaite ensuite remercier Antoine Chevreuil qui m'a co-encadré. Ses nombreuses remarques et suggestions m'ont permis de mûrir de nombreuses idées. La clarté de ce document doit beaucoup à sa ténacité, sa patience et sa relecture attentive.

La collaboration avec Eric Moreau a été particulièrement agréable. Je lui sais gré de l'enthousiasme et la bonne humeur dont il a témoigné à cette occasion. Je remercie également l'ensemble des membres du jury et notamment mes rapporteurs, Christian Jutten et Pierre Comon, pour le temps passé à la lecture de ce mémoire et leurs commentaires pertinents.

Naturellement, je ne peux oublier les autres membres du laboratoire, en particulier Philippe Loubaton, dont l'expérience m'a été utile à plusieurs occasions, et les doctorants —ou docteurs maintenant— Sébastien, Jean-Marie, Pascal, Pierre, Julien et Caroline. Ils ont tous contribué à une ambiance de travail agréable qui m'a permis de mener cette thèse dans des conditions idéales.

Enfin, mes parents et ma famille m'ont toujours soutenu et ont supporté mes sautes d'humeur dans les difficiles moments de doute qui ont accompagné ce travail : je leur en suis infiniment reconnaissant.

Table des matières

1	Introduction	11
1.1	Présentation générale	11
1.2	Plan du document	12
1.3	Publications	13
1.4	Problématique et état de l'art	14
1.4.1	Le problème de la séparation de sources	14
1.4.2	Travaux antérieurs	17
2	Approches conjointes	21
2.1	Préblanchiment	21
2.2	Généralisation des contrastes	23
2.2.1	Une approche fréquentielle	24
2.2.2	Retour sur les contrastes temporels	33
2.3	Exemples de signaux non linéaires	37
2.3.1	Signaux de volatilité stochastique	37
2.3.2	Cas particuliers où $a(n)$ est MA	38
2.3.3	Signaux ARCH	39
2.4	Mise en œuvre de la séparation	41
2.4.1	Paramétrisation par une structure en treillis sans perte	41
2.4.2	Optimisation des contrastes	43
2.5	Simulations	45
2.5.1	Propriétés locales des contrastes	46
2.5.2	Propriétés globales des contrastes	47
2.6	Conclusion	51
3	Approches séquentielles	55
3.1	Extension des contrastes MISO	55
3.1.1	Présentation des approches séquentielles	55
3.1.2	Généralisation des contrastes existants	56
3.2	De l'extraction d'une source à l'extraction MIMO	60
3.2.1	Procédé de déflation	60
3.2.2	Maximisation stricte des contrastes sur un sous-ensemble	61
3.3	Etude locale des contrastes	62
3.3.1	Utilisation des maxima locaux pour la séparation	63
3.3.2	Analyse au voisinage d'un filtre MISO séparant	65
3.4	Vers une méthode effective de séparation MIMO	67
3.4.1	Caractérisation des ensembles \mathcal{F}_i	67
3.4.2	Optimisation pour des filtres RIF	68

3.4.3	Dimension des ensembles \mathcal{F}_i	69
3.4.4	Implantation de la méthode de séparation	70
3.4.5	Amélioration des performances par une post-optimisation locale	71
3.5	Simulations	71
3.5.1	Sources i.i.d.	71
3.5.2	Sources non i.i.d.	73
3.6	Conclusion	73
4	Applications des approches séquentielles	75
4.1	Séparation et détection d'un mélange de signaux modulés CPM	75
4.1.1	Introduction aux modulations CPM	75
4.1.2	Séparation de CPM dans le cas où la période symbole est égale à la période d'échantillonnage	77
4.1.3	Simulations	80
4.2	Séparation de mélanges convolutifs d'images	87
4.2.1	Bref aperçu des méthodes de séparation/restauration en image	87
4.2.2	Inversibilité des mélanges	87
4.2.3	Application des méthodes de séparation séquentielles dans le cas des images	88
4.2.4	Problèmes liés à l'occupation spectrale des sources	89
4.2.5	Simulations	91
5	Fonctions de contraste avec référence	97
5.1	Présentation	97
5.1.1	Signal de référence	97
5.1.2	Approches liées dans les cas instantané et SISO	98
5.1.3	Retour sur la définition d'un contraste	99
5.2	Contrastes MISO dans le cas convolutif	100
5.2.1	Cas de sources i.i.d.	100
5.2.2	Cas de sources non i.i.d.	104
5.3	Éléments pour la mise en œuvre	109
5.3.1	Expression des formes quadratiques : critère et contrainte	109
5.3.2	Restriction du problème	110
5.3.3	Maximisation exacte par décomposition en valeurs singulières	111
5.3.4	Procédé séquentiel de déflation	112
5.4	Simulations	112
5.4.1	Extraction d'une source à l'aide d'une référence	112
5.4.2	Tirage aléatoire du système de référence	115
5.4.3	Séparation MIMO par déflation	117
5.5	Conclusion	118
6	Conclusion	119
6.1	Résumé des travaux réalisés	119
6.2	Perspectives	120

A	Inversibilité des systèmes MIMO et matrices polynomiales	121
A.1	Introduction et notations	121
A.1.1	Systèmes MIMO et matrices polynomiales	121
A.1.2	Notations	122
A.2	Systèmes MIMO monodimensionnels	122
A.3	Matrices à plusieurs indéterminées	123
A.3.1	Notions de primalité	123
A.3.2	Quelques propriétés	124
A.3.3	Structures algébriques et variétés	124

Notations et acronymes

$:=$ (resp. $=$) Id / Id_N \cdot^T \cdot^* \cdot^H \cdot^\dagger $E\{\cdot\}$ $Cum\{\cdot\}$ Diag Rang $\kappa_x(k_1, k_2, k_3)$ ou $\kappa_{x(n)}(k_1, k_2, k_3)$ $\kappa_4\{x(n), z(n)\}$ δ_n $\gamma_i(k)$ $\Gamma_i(\omega)$ $\Gamma_i^4(\omega_1, \omega_2, \omega_3)$	est égal à, par définition du membre de gauche (resp. droite) matrice identité / matrice identité de taille N transposée complexe conjugué transposée conjuguée (\cdot^{*T}) inverse généralisé de Moore-Penrose espérance mathématique cumulant matrice diagonale carrée formée des éléments donnés rang (d'une matrice) $Cum\{x(n), x^*(n+k_1), x(n+k_2), x^*(n+k_3)\}$ $Cum\{x(n), x^*(n), z(n), z^*(n)\}$ impulsion unité $\delta_0 = 1$ et $\delta_n = 0$ si $n \in \mathbb{Z}^*$ autocorrélation de la $i^{\text{ème}}$ source $s_i(n)$ $\gamma_i(k) := E\{s_i(n)s_i^*(n-k)\}$ densité spectrale de la $i^{\text{ème}}$ source $s_i(n)$ trispectre de la $i^{\text{ème}}$ source $s_i(n)$
MIMO MISO SISO AR ARCH MA TFD RIF RII EQM i.i.d.	multi-entrées / multi-sorties (Multiple Input - Multiple Output) multi-entrées / sortie unique (Multiple Input - Single Output) entrée unique / sortie unique (Single Input - Single Output) auto-régressif auto-régressif à hétéroscédasticité conditionnelle (AutoRegressive Conditional Heteroskedasticity) moyenne mobile (moving-average) transformée de Fourier discrète réponse impulsionnelle finie réponse impulsionnelle infinie erreur quadratique moyenne indépendant et identiquement distribué

Chapitre 1

Introduction

1.1 Présentation générale

Cette thèse a été financée par une allocation de recherche couplée du Ministère de l'Education Nationale et de la Recherche et s'est déroulée au sein de l'Equipe Signal et Communications de l'UMR-CNRS 8049. Cette équipe, composée d'enseignants-chercheurs de l'Université de Marne-la-Vallée, fait partie de l'Institut Gaspard Monge. Une partie du travail mené a également bénéficié d'un soutien du Groupement de Recherche Information, Signal, Images et Vision (GDR ISIS) dans le cadre d'échanges avec l'ISITV de Toulon.

Le travail exposé concerne le problème de la séparation aveugle de sources, un thème de recherche qui a suscité un grand intérêt au cours des dix dernières années et a engendré de nombreuses contributions. La problématique peut se formuler sous une forme générale et permet ainsi d'envisager un grand nombre d'applications telles que les télécommunications, le traitement du son, l'analyse de signaux biomédicaux (magnéto- et électroencéphalographie), ... Ceci explique en partie la grande popularité de cette thématique.

L'objet de la séparation de sources peut se formuler de la façon suivante : considérons N signaux sources $(s_1(n))_{n \in \mathbb{Z}}, \dots, (s_N(n))_{n \in \mathbb{Z}}$ que l'on ne peut pas observer directement, et qui subissent une opération de mélange au travers d'un système \mathfrak{M} fournissant Q observations notées $(x_1(n))_{n \in \mathbb{Z}}, \dots, (x_Q(n))_{n \in \mathbb{Z}}$:

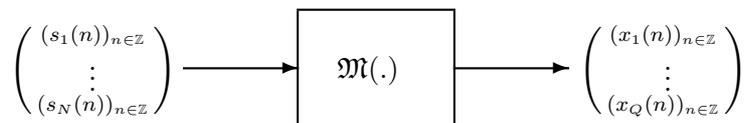


FIG. 1.1 – Modèle général de la séparation de sources.

L'objectif à atteindre est de trouver le moyen d'annuler l'effet du système \mathfrak{M} , c'est-à-dire trouver un inverse \mathfrak{M}^{-1} afin de reconstituer les signaux sources. Pour être en mesure de proposer des solutions, nous sommes en général amenés à poser des hypothèses supplémentaires sur la structure du mélange. En dehors de l'invariance de \mathfrak{M} au cours du temps qui est toujours supposée dans ce document, nous pouvons distinguer les contextes suivants :

- Une hypothèse souvent faite est celle de *linéarité* du système \mathfrak{M} . Le cadre des systèmes non linéaires est bien plus vaste et plus délicat à traiter ; dans ce cas, des hypothèses sur la nature de la non linéarité (mélange post-non linéaire [89], ...)

se révèlent souvent utiles, ne serait-ce que pour assurer l'identifiabilité du système et/ou l'existence d'une solution au problème.

- L'*instantanéité* du système se traduit par le fait que les observations à un instant $n \in \mathbb{Z}$ donné ne dépendent que des sources au même instant, ce qui peut s'écrire :

$$\forall n \in \mathbb{Z} \quad (x_1(n), \dots, x_Q(n)) = \mathfrak{M}(s_1(n), \dots, s_N(n)) \quad (1.1)$$

Le cas d'un système linéaire instantané se ramène à l'étude d'un mélange constitué d'un simple produit matriciel. Par opposition, on parle de mélange linéaire *convolutif* dans le cas linéaire mais non instantané : c'est le cadre étudié dans ce document.

Diverses informations peuvent être utilisées dans un but d'identification de \mathfrak{M} ou de son inverse. Par exemple, des connaissances *a priori* peuvent être disponibles sur le système de mélange (modèle physique, ...). Dans d'autres cas, il est possible de disposer à la fois des entrées et sorties du système \mathfrak{M} . Cette thèse se situe au contraire dans le cadre dit *aveugle*, ce qui signifie, d'une part qu'aucune connaissance n'est disponible sur \mathfrak{M} en dehors des hypothèses de structure (c'est-à-dire de modèle), et d'autre part que les sources sont inobservables. Ce défaut de connaissance est toutefois compensé par une hypothèse statistique forte qui est celle d'indépendance mutuelle des sources.

1.2 Plan du document

Ce mémoire de thèse se divise en six chapitres et une annexe. Le présent chapitre constitue une introduction au problème traité. Les notations utilisées dans l'ensemble du document y sont données et un état de l'art est rapidement exposé.

Les méthodes de séparation de sources considérées peuvent se classer en deux catégories. Les approches conjointes d'une part —*i.e.* qui extraient toutes les sources simultanément— sont traitées dans le deuxième chapitre, tandis que les méthodes séquentielles —*i.e.* où les sources sont extraites successivement— sont abordées dans les chapitres 3 à 5. Partant d'une approche fréquentielle, le chapitre 2 établit des critères de contraste et en étudie les propriétés. Une alternative aux méthodes dites «par déflation» est présentée dans le chapitre 3 qui constitue également l'occasion de faire un point approfondi sur des résultats connus et de les étendre. Le chapitre 4 propose alors deux applications du chapitre précédent :

- La séparation des signaux de communication dits CPM (Continuous Phase Modulation). Ces derniers constituent un exemple important en pratique de signal non i.i.d. L'utilisation de résultats récents a permis par ailleurs de dépasser la simple séparation et d'envisager la détection des symboles.
- Une extension des méthodes de séparation au cas des images est présentée. Cette dernière est rendue possible par l'utilisation des résultats de l'annexe A sur les matrices polynomiales à plusieurs indéterminées.

Le chapitre 5 propose une nouvelle classe de fonctions de contrastes dont la dépendance est quadratique en les paramètres cherchés. L'efficacité de l'optimisation —possible en un nombre fini d'itérations— justifie à elle seule l'intérêt majeur de ce chapitre. Enfin, le chapitre 6 conclut ce travail.

Nous avons fait le choix d'inclure systématiquement les preuves au cœur du texte et de ne pas les reléguer en annexe. Cependant, la plupart d'entre elles peuvent être laissées de côté au cours d'une première lecture, sans nuire pour autant à la compréhension globale du texte. Pour cette raison, les passages de démonstration, situés entre le mot *Preuve* et un carré plein de fin de démonstration (■), ont été mis dans une police de taille inférieure.

Après avoir pris connaissance des notations et quelques rappels exposés dans le chapitre 1, il est possible d'aborder indépendamment les chapitres 2, 3 ou 5. Le chapitre 4 fait référence à la méthode introduite dans le chapitre 3. Par ailleurs, la partie 4.2 se réfère à des résultats présentés dans l'annexe. Ces derniers nous paraissent pouvoir être aisément admis par le lecteur ne désirant pas s'investir dans la théorie des polynômes à plusieurs variables.

1.3 Publications

Les travaux réalisés ont fait l'objet des publications suivantes dans la période 2002-2004 :

Articles de revue

- Marc Castella, Jean-Christophe Pesquet et Athina P. Petropulu, *A Family of Frequency- and Time-Domain Contrasts for Blind Separation of Convolutive Mixtures of Temporally Dependent Signals*. Accepté pour publication dans *IEEE Transactions on Signal Processing*.
- Marc Castella, Pascal Bianchi, Antoine Chevreuil et Jean-Christophe Pesquet, *A blind source separation framework for detecting CPM sources mixed by a convolutive MIMO filter*. Soumis à *Signal Processing*.

Conférences

- Marc Castella, Jean-Christophe Pesquet et Athina P. Petropulu, *New contrasts for blind separation of non iid sources in the convolutive case*. EUSIPCO 2002, Vol.2, pp.107-110, Toulouse, France.
- Marc Castella et Jean-Christophe Pesquet, *Source separation of a class of non linear time series*. IEEE-EURASIP Workshop on Nonlinear Signal and Image Processing (NSIP) 2003, Grado, Italy (4 pages).
- Marc Castella, Antoine Chevreuil et Jean-Christophe Pesquet, *Séparation aveugle d'un mélange convolutif de sources non linéaires par une approche hiérarchique*. GRETSI 2003, Paris, France (4 pages).
- Marc Castella, Eric Moreau et Jean-Christophe Pesquet, *A quadratic MISO contrast for blind equalization*. ICASSP 2004, pp.681-684, Montréal, Canada.
- Marc Castella, Pascal Bianchi, Antoine Chevreuil et Jean-Christophe Pesquet, *Blind MIMO detection of convolutively mixed CPM sources*. EUSIPCO 2004, Vienne, Autriche (4 pages).
- Marc Castella et Jean-Christophe Pesquet, *An iterative blind source separation method for convolutive mixtures of images*. ICA 2004, Grenade, Espagne (8 pages).
- Marc Castella, Saloua Rhoui, Eric Moreau et Jean-Christophe Pesquet, *Source separation by quadratic contrast functions : a blind approach based on any higher-order statistics*. Accepté pour publication, ICASSP 2005, Philadelphie, USA.

Autre

- Journée Image de l'Action Spécifique «Séparation de Sources et Analyse en Composantes Indépendantes» du CNRS : quelques résultats sur la séparation de mélanges convolutifs 2D, Paris, 4 décembre 2003, communication orale.

1.4 Problématique et état de l'art

1.4.1 Le problème de la séparation de sources

1.4.1-a) Problème et notations

Nous rappelons ici la problématique de la séparation aveugle de sources telle qu'elle est considérée dans ce document, c'est-à-dire dans le cas des mélanges linéaires convolutifs. Les notations proposées seront celles utilisées dans l'ensemble du document ; elles sont résumées par le schéma de la figure 1.2.

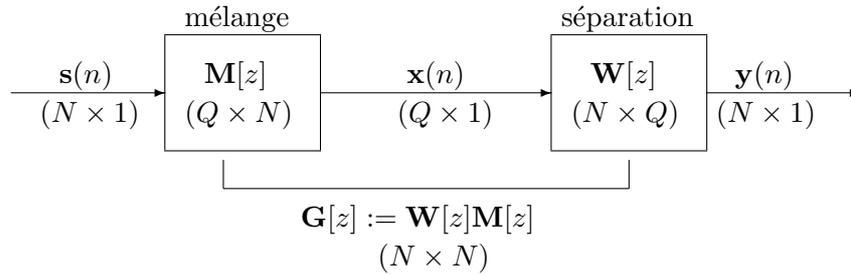


FIG. 1.2 – Schéma du système considéré pour la séparation de sources.

Les $N \in \mathbb{N}^*$ signaux source, supposés non connus et non observables sont regroupés dans le vecteur $\mathbf{s}(n) := (s_1(n), \dots, s_N(n))^T$. Le vecteur $\mathbf{x}(n) := (x_1(n), \dots, x_Q(n))^T$ contient quant à lui les $Q \in \mathbb{N}^*$ signaux observés. Ces *observations* résultent du mélange des sources par un filtre de réponse impulsionnelle $(\mathbf{M}(n))_{n \in \mathbb{Z}}$, ce qui donne l'équation de mélange :

$$\forall n \in \mathbb{Z} \quad \mathbf{x}(n) = \sum_{k \in \mathbb{Z}} \mathbf{M}(k) \mathbf{s}(n - k) \quad (1.2)$$

Le problème de la séparation de sources tel que nous l'envisageons, consiste à reconstituer $\mathbf{s}(n)$ à partir du seul vecteur des observations et sans aucune connaissance *a priori* sur le système de mélange, en dehors de la structure décrite par l'équation ci-dessus. Plus précisément, il est connu que, sous certaines conditions, il existe un filtre inverse du mélange. La question consiste alors à rechercher un filtre de réponse impulsionnelle $(\mathbf{W}(n))_{n \in \mathbb{Z}}$ qui permette d'inverser le mélange. La sortie du filtre *séparateur* s'écrit :

$$\forall n \in \mathbb{Z} \quad \mathbf{y}(n) = \sum_{k \in \mathbb{Z}} \mathbf{W}(k) \mathbf{x}(n - k) \quad (1.3)$$

Dans ce document, la transformée en z d'un filtre de réponse impulsionnelle $(h(n))_{n \in \mathbb{Z}}$ sera notée par des crochets : $h[z]$. Par commodité, nous noterons également $h[z]s(n)$ la sortie de ce même filtre attaqué en entrée par un signal $s(n)$. Les mêmes conventions s'appliqueront pour les filtres éventuellement multi-entrées / multi-sorties et les équations (1.2) et (1.3) peuvent alors se récrire de manière plus concise :

$$\mathbf{x}(n) = \mathbf{M}[z]\mathbf{s}(n) \quad \mathbf{y}(n) = \mathbf{W}[z]\mathbf{x}(n) \quad (1.4)$$

L'étude de la séparation peut, par changement de variable, être ramenée à un problème ayant pour seul paramètre le filtre global $\mathbf{G}[z] := \mathbf{W}[z]\mathbf{M}[z]$ résultant de la mise en cascade

du filtre de mélange et du séparateur. Sa réponse impulsionnelle est donnée par :

$$\forall n \in \mathbb{Z} \quad \mathbf{G}(n) = \sum_{k \in \mathbb{Z}} \mathbf{W}(n-k) \mathbf{M}(k) \quad (1.5)$$

Dans ce cas, la sortie globale peut s'exprimer en fonction des sources et du filtre global :

$$\mathbf{y}(n) = \mathbf{G}[z] \mathbf{s}(n) \quad (1.6)$$

ce qui s'écrit aussi dans le domaine temporel :

$$\forall n \in \mathbb{Z} \quad \mathbf{y}(n) = \sum_{k \in \mathbb{Z}} \mathbf{G}(k) \mathbf{s}(n-k) \quad (1.7)$$

Compte tenu du cadre aveugle choisi, qui implique l'impossibilité d'observer les sources et l'absence de toute connaissance sur $\mathbf{M}[z]$ en dehors du modèle décrit ci-dessus, des hypothèses fortes sont nécessaires sur les statistiques des sources afin d'être en mesure de procéder à leur séparation. L'hypothèse fondamentale suivante sera faite tout au long de ce document :

H.1 Les composantes $(s_1(n))_{n \in \mathbb{Z}}, \dots, (s_N(n))_{n \in \mathbb{Z}}$ du vecteur $(\mathbf{s}(n))_{n \in \mathbb{Z}}$ sont des processus aléatoires stationnaires, centrés, mutuellement indépendants et de variance unité.

Les travaux présentés dans ce document reposent principalement sur des statistiques d'ordre quatre. Ainsi, la stationnarité et l'indépendance peuvent, dans ce cadre, se limiter au quatrième ordre.

Il est connu que la seule hypothèse H.1 ne permet pas par la seule donnée des observations $(\mathbf{x}(n))_{n \in \mathbb{Z}}$, d'obtenir une sortie globale $(\mathbf{y}(n))_{n \in \mathbb{Z}}$ qui soit la copie conforme des sources $(\mathbf{s}(n))_{n \in \mathbb{Z}}$. Nécessairement, le problème ne peut être résolu qu'à certaines indéterminations près :

- l'ordre des sources : il est purement conventionnel et ne peut donc pas être retrouvé sans une connaissance supplémentaire sur le mélange.
- un filtrage scalaire de chacune des sources : un filtrage SISO de chacune des sources laisse en effet inchangée l'hypothèse H.1, ce qui explique cette indétermination dans le cas général. Sous des conditions supplémentaires, cette ambiguïté peut cependant se réduire :
 - *sources i.i.d.* : l'ambiguïté de filtrage se limite à un décalage temporel et un facteur d'amplitude près.
 - *sources temporellement décorrélées* : le filtrage scalaire résiduel peut alors se limiter à un filtre passe-tout. Dans le cas d'un mélange MIMO et de plus RIF, il existe un filtre global MIMO RIF passe-tout sur chacune des composantes. Il est donc possible dans ce cas, de conserver une simple ambiguïté passe-tout RIF, ce qui est équivalent à un simple retard et facteur d'amplitude. Plus généralement, le raisonnement qui vient d'être fait peut s'appliquer dès que les statistiques d'ordre deux des sources sont supposées connues.

Nous considérerons que le problème de la séparation de sources est résolu lorsque $(\mathbf{y}(n))_{n \in \mathbb{Z}}$ correspond à $(\mathbf{s}(n))_{n \in \mathbb{Z}}$ aux indéterminations près ci-dessus. En d'autres termes, la séparation est réalisée lorsque le filtre global $\mathbf{G}[z]$ se réduit à un filtrage scalaire de chacune des sources, suivi d'une permutation. Dans ce cas, le filtre global peut s'écrire :

$$\mathbf{G}[z] = \mathbf{P} \mathbf{D}[z] \quad (1.8)$$

où \mathbf{P} représente une matrice de permutation et $\mathbf{D}[z]$ est diagonale. On dit parfois de $\mathbf{G}[z]$ qu'elle est «diagonale à une permutation près». L'hypothèse suivante assure l'inversibilité du filtre au sens que nous venons de définir :

H.2 Il existe un filtre $\mathbf{W}^0[z]$ de taille $N \times Q$ tel que le filtre global $\mathbf{W}^0[z]\mathbf{M}[z]$ soit diagonal. Cette hypothèse est supposée toujours vérifiée et des conditions permettant de la satisfaire seront précisées par la suite.

1.4.1-b) Le modèle de séparation de sources avec bruit additif sur les capteurs

Le modèle de la séparation de sources présenté dans le paragraphe précédent est celui qui correspond à un cas idéal sans bruit. Il est cependant possible de compléter ce modèle en considérant un bruit additif $(\mathbf{b}(n))_{n \in \mathbb{Z}}$ sur les capteurs, ce qui donne l'équation de mélange :

$$\forall n \in \mathbb{Z} \quad \mathbf{x}(n) = \sum_{k \in \mathbb{Z}} \mathbf{M}(k)\mathbf{s}(n-k) + \mathbf{b}(n), \quad (1.9)$$

qui s'écrit aussi :

$$\mathbf{x}(n) = \mathbf{M}[z]\mathbf{s}(n) + \mathbf{b}(n). \quad (1.10)$$

La sortie du séparateur s'écrit dans ce cas :

$$\mathbf{y}(n) = \mathbf{W}[z]\mathbf{x}(n) \quad (1.11)$$

et donc en fonction du filtre global :

$$\mathbf{y}(n) = \mathbf{G}[z]\mathbf{s}(n) + \mathbf{b}_w(n) \text{ avec : } \mathbf{b}_w(n) := \mathbf{W}[z]\mathbf{b}(n) \quad (1.12)$$

Ce jeu de notations est rappelé sur la figure 1.3.

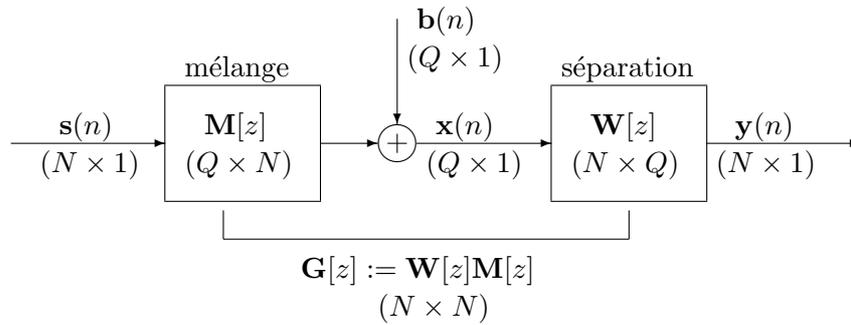


FIG. 1.3 – Schéma du système considéré pour la séparation de sources.

Dans le cadre des travaux présentés, le bruit n'est pas à proprement parler pris en compte au cours de l'étude de la méthode et il apparaît simplement comme une perturbation. S'il est de plus supposé gaussien et indépendant, et dans la mesure où les méthodes en question sont basées sur l'emploi de cumulants d'ordre quatre, le bruit n'a en théorie aucune contribution dans l'estimation des critères de contraste et donc du séparateur. Cependant, son influence transparait au niveau des statistique d'ordre deux qui sont implicitement utilisées (préblanchiment, décorrélation des sorties, ...) et des méthodes plus évoluées pourraient être envisagées, dans lesquelles le bruit serait une partie intégrante du modèle considéré.

1.4.1-c) Sources i.i.d., linéaires et non linéaires

Contrairement à la majorité des travaux existant, ce travail se place principalement dans le cadre de sources *non i.i.d.*. Plus précisément, la difficulté réside dans l'étude du cas de sources *non linéaires* dont nous rappelons la définition :

Définition 1 Une série temporelle $((s(n))_{n \in \mathbb{Z}}$ est dite :

- linéaire lorsqu'elle peut s'écrire comme la filtrée d'un processus i.i.d.
- non linéaire dans tous les autres cas.

D'après cette définition, nous constatons qu'un processus i.i.d. est un cas particulier de processus linéaire. Par contraposée, tout processus non linéaire est non i.i.d.

Compte tenu de l'ambiguïté de filtrage scalaire, il apparaît que le cas des sources linéaires se ramène au cas des sources i.i.d. [71]. Dans ce cas en effet, chaque source $(s_i(n))_{n \in \mathbb{Z}}$ peut s'écrire comme la filtrée d'un processus i.i.d. $(a_i(n))_{n \in \mathbb{Z}}$ par un filtre $f_i[z]$:

$$s_i(n) = f_i[z]a_i(n) \quad (1.13)$$

Dès lors, le modèle de la figure 1.3 s'écrit de façon équivalent avec $(a_i(n))_{n \in \mathbb{Z}}, i \in \{1, \dots, N\}$ comme sources et $\mathbf{M}[z]\text{Diag}(f_1[z], \dots, f_N[z])$ comme matrice de mélange. Ceci correspond au cas i.i.d. et il est alors connu dans ce cas que la blancheur temporelle des sources permet de limiter l'ambiguïté de filtrage scalaire à une ambiguïté de retard et d'amplitude [24]. Ainsi, les méthodes de séparation permettent de remonter aux processus générateurs $(a_i(n))_{n \in \mathbb{Z}}, i \in \{1, \dots, N\}$ et les sources peuvent éventuellement être reconstruites par une méthode de déconvolution SISO.

Pour les raisons évoquées ci-dessus, ce travail considérera principalement le cas de sources non i.i.d. et non linéaires. Des exemples de signaux non linéaires seront donnés au fur et à mesure (voir en particulier les parties 2.3 et 4.1.1).

1.4.2 Travaux antérieurs

Nous faisons ici un rapide tour d'horizon des méthodes de séparation de sources proposées dans la littérature. Nous commençons par donner quelques références concernant les premières solutions proposées dans le cas instantané avant de nous concentrer sur les méthodes convolutives. On pourra consulter avec profit [64] pour des compléments de bibliographie.

1.4.2-a) Le cas instantané

Les premières études en séparation de sources ont d'abord concerné les mélanges instantanés. Dans ce cas, les filtres de mélange $\mathbf{M}[z]$ et séparant $\mathbf{W}[z]$ se réduisent à de simples matrices, et, conformément à l'équation (1.4) les modèles de mélange et de séparation s'écrivent comme les simples produits matriciels, respectivement $\mathbf{x}(n) = \mathbf{M}\mathbf{s}(n)$ et $\mathbf{y}(n) = \mathbf{W}\mathbf{x}(n)$: le problème ainsi posé, plus connu sous le nom d'analyse en composantes indépendantes (ou ICA, acronyme de l'anglais Independent Component Analysis) a donné lieu à d'important travaux [49, 6, 23, 15, 78, 75]. Des algorithmes de séparation de sources ont été proposés dès la fin des années quatre-vingts par la communauté des réseaux de neurones, utilisant un fonctionnement neuromimétique [49, 6]. Parallèlement, le problème de l'analyse en composantes indépendantes a été étudié en tant que tel et donné lieu à des solutions efficaces [23, 15]. On pourra consulter [45, 56] pour une introduction à l'ICA et [14] pour une revue des solutions proposées.

Les méthodes auxquelles nous venons de faire référence utilisent l'hypothèse selon laquelle les sources ont une distribution de probabilité *non gaussienne*. Ainsi, la plupart d'entre elles appartiennent à la catégorie des méthodes exploitant des statistiques d'ordre supérieur ([56] offre une introduction à cette notion). Sous certaines conditions, les statistiques d'ordre deux peuvent cependant suffire. Dans ce cas, les méthodes tirent alors parti soit de la *diversité spectrale* des sources [78, 8], soit de leur *non stationnarité* [77].

1.4.2-b) Le cas convolutif : approches aux ordres supérieurs

Afin de mener à la séparation, les méthodes dans le cas convolutif peuvent également utiliser les trois voies citées ci-dessus pour les mélanges instantanés : non gaussianité, diversité spectrale ou non stationnarité. Nous présentons d'abord les méthodes utilisant des statistiques d'ordre supérieur [99, 24, 68, 57, 88, 93]. Dans ce cas, différentes approches ont été proposées.

Les approches temporelles conjointes Nous parlerons d'approche conjointe lorsque toutes les sources sont extraites simultanément et traitées en parallèle, d'une manière complètement symétrique. Ceci s'oppose aux méthodes séquentielles mentionnées plus loin.

Dans la continuité des développements du cas instantané, des approches temporelles ont permis de généraliser au cas de sources i.i.d. mélangées convolutivement, des contrastes connus dans le cas instantané [24, 68, 91]. Des adaptations ont été proposées, menant à l'idée de contrastes «non symétriques» [69]. Le cas des sources non i.i.d. n'a été abordé que récemment dans [60] qui propose certains critères de contrastes identiques à ceux obtenus par le biais du domaine fréquentiel et que nous exposons dans le chapitre 2.

L'ensemble de ces approches conjointes suppose la possibilité de préblanchir les sources —le paragraphe 2.1 donnera des explications théoriques supplémentaires. Cette nécessité de préblanchir constitue l'inconvénient majeur de ces méthodes, à la fois en raison du caractère délicat de la tâche et de la difficulté à effectuer ensuite une optimisation dans le sous-ensemble des filtres dits para-unitaires. Aussi, plusieurs solutions ont été envisagées pour prendre en compte cette contrainte, souvent basées sur des paramétrisations par des angles de rotations dites de Jacobi [27, 26].

Par ailleurs, des approches de type information mutuelle permettent d'établir des classes importantes de contrastes [76, 5]. Des travaux ont cependant illustré que ces méthodes, développées dans le cas de sources i.i.d., ne donnent pas toujours des résultats satisfaisants dans le cas non linéaire [71].

Les approches fréquentielles Le point de vue fréquentiel présente l'avantage de transformer un problème convolutif en une collection de problèmes instantanés. De telles méthodes se placent en général dans le cadre de signaux i.i.d. et consistent à réaliser des séparations instantanées sur un ensemble de fréquences distinctes [22, 13, 57, 30, 31, 29]. La difficulté majeure est alors de résoudre le problème de l'indétermination de permutation entre les différentes séparations instantanées du domaine fréquentiel.

Les dernières approches évoquées utilisent pour la séparation instantanée à chaque fréquence, des statistiques d'ordre supérieur des transformées de Fourier des observations. Il est possible de développer des méthodes exploitant directement les multi-spectres (en général bi- ou trispectre) des observations [20]. Dans ce cas toutefois, le nombre de points nécessaires pour l'estimation des trispectres peut se révéler important. En revanche, l'utilisation de quantités spectrales permet de saisir la nature non linéaire des signaux sources

et d'étendre certains résultats au cas non i.i.d. [74]. Nous reviendrons sur ce thème au cours du chapitre 2.

Les approches séquentielles Les approches séquentielles se distinguent des approches précédentes par le fait que les sources ne sont pas extraites simultanément, mais les unes après les autres.

Un certain nombre d'approches séquentielles reposent sur l'existence de méthodes d'extraction d'une source : nous les appellerons méthodes MISO dans la mesure où, à partir de plusieurs entrées (sources ou capteurs), une seule source est extraite. Dans ce cadre, des critères de contraste —inspirés des méthodes de déconvolution SISO [85]— ont été exhibés [88, 93] et des procédures itératives menant à la filtrée d'une des sources [47] ont été proposées. L'attrait principal de telles approches est qu'elles demeurent valables dans le cas de sources non linéaires [88].

Le passage d'une procédure MISO à l'extraction MIMO de toutes les sources s'effectue en itérant successivement l'étape d'extraction. Avant de relancer la procédure, il est cependant nécessaire de retirer la contribution de la source précédemment extraite sur les capteurs [87].

Enfin, signalons l'existence de méthodes dites hiérarchiques [92]. Dans ces dernières, la séparation provient de la maximisation successive d'une série de critères dans lesquels sont introduits à chaque étape un terme supplémentaire de pénalisation qui exprime une décorrélation spatiale des signaux de sortie.

1.4.2-c) Le cas convolutif : approches à l'ordre deux et autres méthodes

Nous donnons dans ce paragraphe quelques éléments et pointeurs sur d'autres méthodes, en particulier celles utilisant des statistiques d'ordre deux et donc soit la diversité spectrale soit la non stationnarité des sources.

Utilisation de la diversité spectrale Dans le cas où les sources présentent des statistiques d'ordre deux distinctes et où les filtres sont RIF, il a été prouvé que la séparation est possible à l'aide des seules statistiques d'ordre deux [43]. Les algorithmes opérant dans le domaine fréquentiel présentent souvent l'inconvénient d'être confrontés à des problèmes d'ambiguïtés de permutation et de facteur multiplicatif entre chacune des fréquences considérées [98, 33]. Des algorithmes, proposant des critères développés dans le domaine fréquentiel, mais optimisés par rapport aux paramètres temporels du filtre séparant ont été proposés [52, 53]. Dans ce cas, les difficultés des approches fréquentielles liées aux permutations n'apparaissent pas. Par ailleurs, une approche itérative, de type MISO, a été proposée utilisant les seules statistiques d'ordre deux [54]. Pour terminer, rappelons l'existence de méthodes de type sous-espace qui utilisent uniquement des statistiques d'ordre 2 [2, 39] et qui visent à identifier les coefficients du système de mélange.

Utilisation de la non stationnarité Des critères proches de ceux évoqués ci-dessus ont été proposés [79, 73]. Ils permettent dans ce cas d'utiliser la non stationnarité des signaux en vue d'effectuer la séparation.

Méthodes algébriques Evoquons l'existence de critères déterministes [25, 97] fondés sur l'appartenance des sources à un alphabet fini et qui mènent à des méthodes dites algébriques.

Chapitre 2

Approches conjointes

2.1 Préblanchiment

Dans l'ensemble de ce chapitre sur les approches conjointes, les hypothèses suivantes seront faites sur le bruit et les sources :

H.3 Le bruit $(\mathbf{b}(n))_{n \in \mathbb{Z}}$ est gaussien, centré, stationnaire, indépendant des sources. Ses caractéristiques à l'ordre deux sont supposées connues.

H.4 Chacune des sources $(s_1(n))_{n \in \mathbb{Z}}, \dots, (s_N(n))_{n \in \mathbb{Z}}$ est temporellement décorrélée (mais non nécessairement i.i.d.) et de puissance unité, *i.e.* :

$$\forall k \in \mathbb{Z} \quad \mathbb{E}\{\mathbf{s}(n)\mathbf{s}(n-k)^H\} = \delta_k \mathbf{Id}_N \quad (2.1)$$

Notons que cette dernière hypothèse n'est aucunement restrictive : des sources non blanches à l'ordre deux peuvent être considérées comme provenant d'un filtrage SISO de sources blanches. L'indétermination SISO inhérente à la séparation de sources ne permet pas en général de distinguer ces sources réelles non blanches à l'ordre deux du processus blanc à l'ordre deux qui leur est associé. En d'autres termes, il est possible de passer des sources non décorrélées à leur version décorrélée par un simple blanchiment temporel. Par ailleurs, il existe des catégories de signaux satisfaisant l'hypothèse H.4, parmi lesquels on peut évoquer :

- les signaux de communication après codage correcteur d'erreur [80].
- le cas de signaux obtenus par sous-échantillonnage d'un processus à temps continu linéaire, décorrélé et stationnaire strict : soit en effet $\underline{s}(t)$ un tel signal dont le trispectre est défini par

$$\underline{H}(-\omega_1 - \omega_2 - \omega_3) \underline{H}^*(-\omega_1) \underline{H}(\omega_2) \underline{H}^*(-\omega_3)$$

où \underline{H} est la réponse fréquentielle d'un certain filtre linéaire à temps continu. Après échantillonnage à une période $T > 0$, nous obtenons un processus à temps discret dont le trispectre, par la formule de Poisson, s'écrit :

$$\frac{1}{T^3} \sum_{(l_1, l_2, l_3) \in \mathbb{Z}^3} \underline{H}\left(-\omega_1 - \omega_2 - \omega_3 - 2\pi \frac{l_1 + l_2 + l_3}{T}\right) \underline{H}^*\left(-\omega_1 - 2\pi \frac{l_1}{T}\right) \underline{H}\left(\omega_2 + 2\pi \frac{l_2}{T}\right) \underline{H}^*\left(-\omega_3 - 2\pi \frac{l_3}{T}\right).$$

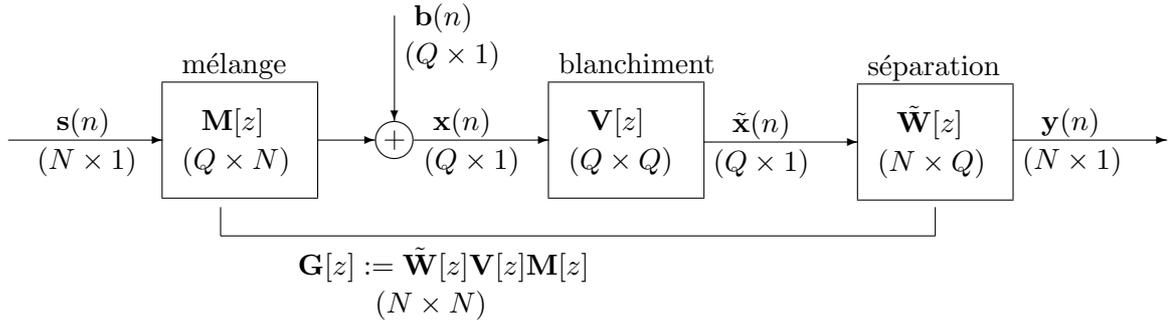
Si H n'est pas à bande limitée (ou à bande limitée de fréquence de coupure supérieure à la fréquence de Nyquist), le processus résultant n'est pas linéaire car l'expression ci dessus ne pourra pas s'écrire sous la forme $H(-\omega_1 - \omega_2 - \omega_3)H^*(-\omega_1)H(\omega_2)H^*(-\omega_3)$ où H est une fonction 2π -périodique.

- les signaux CPM (modulation à phase continue) dans le cas particulier d'un indice de modulation $1/2$. Ces signaux serviront plus particulièrement d'exemple d'application dans le chapitre suivant et seront présentés à cette occasion.
- les exemples donnés dans la partie 2.3 (ARCH, volatilité stochastique, ...).

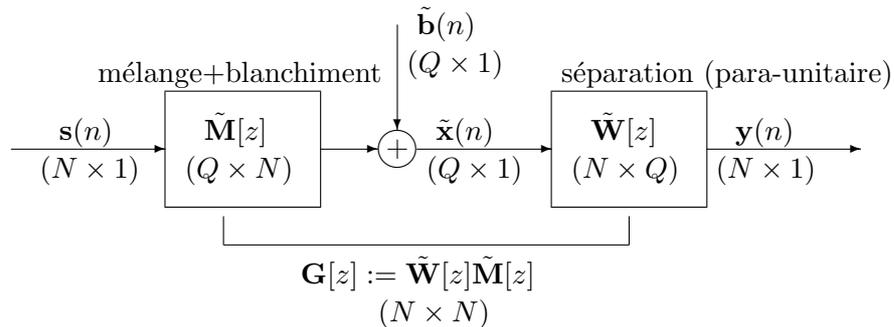
Inspiré par les méthodes d'ICA, les approches conjointes de la séparation de sources procèdent fréquemment en deux étapes : un préblanchiment à l'ordre deux des données précède la recherche d'un filtre séparateur dans un sous-ensemble restreint de filtres. En notant $\mathbf{V}[z]$ un filtre blanchisseur, le filtre séparant se décompose en (voir figure 2.1(a)) :

$$\mathbf{W}[z] = \tilde{\mathbf{W}}[z]\mathbf{V}[z] \quad (2.2)$$

Si l'on note $\tilde{\mathbf{b}}(n) = \mathbf{V}[z]\mathbf{b}(n)$ et $\tilde{\mathbf{M}}[z] = \mathbf{V}[z]\mathbf{M}[z]$, il est alors possible d'étudier de façon équivalente le modèle donné par la figure 2.1(b). L'autocorrélation du bruit étant supposée



(a) Système complet avec préblanchiment



(b) Système équivalent avec mélange para-unitaire

FIG. 2.1 – Blanchiment et filtre para-unitaire.

connue, un filtre blanchisseur est tel que les données blanchies satisfont :

$$\forall k \in \mathbb{Z} \quad \mathbf{E}\{\tilde{\mathbf{x}}(n)\tilde{\mathbf{x}}(n-k)^H\} - \mathbf{E}\{\tilde{\mathbf{b}}(n)\tilde{\mathbf{b}}(n-k)^H\} = \delta_k \mathbf{Id}_Q \quad (2.3)$$

Or, d'après le calcul élémentaire suivant (où l'on a utilisé H.3 et H.4) :

$$\begin{aligned}
\forall k \in \mathbb{Z} \quad E\{\tilde{\mathbf{x}}(n)\tilde{\mathbf{x}}(n-k)^H\} &= \sum_{(l_1, l_2) \in \mathbb{Z}^2} \widetilde{\mathbf{M}}(l_1) E\{\mathbf{s}(n-l_1)\mathbf{s}(n-k-l_2)^H\} \widetilde{\mathbf{M}}(l_2)^H \\
&\quad + \sum_{l_1 \in \mathbb{Z}} \widetilde{\mathbf{M}}(l_1) E\{\mathbf{s}(n-l_1)\tilde{\mathbf{b}}(n-k)^H\} \\
&\quad + \sum_{l_2 \in \mathbb{Z}} E\{\tilde{\mathbf{b}}(n)\mathbf{s}(n-k-l_2)^H\} \widetilde{\mathbf{M}}(l_2)^H \\
&\quad + E\{\tilde{\mathbf{b}}(n)\tilde{\mathbf{b}}(n-k)^H\} \\
&= \sum_{l \in \mathbb{Z}} \widetilde{\mathbf{M}}(l)\widetilde{\mathbf{M}}(l-k)^H + E\{\tilde{\mathbf{b}}(n)\tilde{\mathbf{b}}(n-k)^H\} \quad (2.4)
\end{aligned}$$

il apparaît que (2.3) se traduit par une condition sur $\widetilde{\mathbf{M}}[z]$. Rappelons en effet la définition suivante d'un filtre para-unitaire :

Définition 2 *Un filtre $\mathbf{H}[z]$ est dit para-unitaire lorsque l'une des deux conditions équivalentes ci-dessous est vérifiée :*

$$\begin{aligned}
(i) \quad \mathbf{H}[z]\mathbf{H}\left[\frac{1}{z^*}\right]^H &= \mathbf{Id} \\
(ii) \quad \sum_{l \in \mathbb{Z}} \mathbf{H}(l)\mathbf{H}(l-k)^H &= \delta_k \mathbf{Id} \quad (2.5)
\end{aligned}$$

D'après cette définition, nous constatons que le blanchiment des données est équivalent à la para-unitarité du filtre $\widetilde{\mathbf{M}}[z]$. Un aspect important est qu'un filtre para-unitaire admet un inverse para-unitaire. Ainsi, il est suffisant de se restreindre à la recherche de $\widetilde{\mathbf{W}}[z]$ dans l'ensemble des filtres para-unitaires.

Il existe des méthodes classiques permettant de réaliser un préblanchiment ; on pourra se reporter à [12] ou [7] pour plus de détails. Nous ne considérerons pas cette étape de la séparation et supposerons qu'elle a été réalisée de façon parfaite. Ainsi, nous nous ramenons au modèle équivalent de la figure 2.1(b) et recherchons le filtre séparant $\widetilde{\mathbf{W}}[z]$ dans l'ensemble des filtres para-unitaires. Afin de ne pas alourdir les notations, nous supprimerons dans la suite de ce chapitre le signe tilde \sim . Nous supposerons de plus que le nombre de capteurs est égal au nombre de sources, c'est-à-dire que $Q = N$. Le modèle supposé correspond alors à celui de départ (figure 1.3) où le filtre de mélange est supposé para-unitaire et ayant autant de sorties que d'entrées. Cette configuration restera valable dans l'ensemble de ce chapitre.

2.2 Généralisation des contrastes

Les approches développées dans cette thèse sont basées sur la notion de fonction de contraste. L'intérêt de ce concept est de ramener le problème de la séparation de sources à un problème d'optimisation : une fonction de contraste —appelée aussi contraste— est un critère qui, par sa maximisation, permet d'obtenir la séparation des sources. La définition d'un contraste peut cependant légèrement différer selon le cadre dans lequel on se place (instantané, convolutif, MIMO, MISO) et les propriétés que l'on souhaite imposer au critère en question (symétrie vis à vis des ambiguïtés de la séparation de sources, ...). La définition que nous adoptons dans ce chapitre sur les approches conjointes est la suivante :

Définition 3 Une fonction de contraste (ou contraste) est une fonction réelle des sorties du système global et qui satisfait les deux conditions suivantes :

- (i) $\mathcal{C}((\mathbf{y}(n))_{n \in \mathbb{Z}}) \leq \mathcal{C}((\mathbf{s}(n))_{n \in \mathbb{Z}})$,
- (ii) si $\mathcal{C}((\mathbf{y}(n))_{n \in \mathbb{Z}}) = \mathcal{C}((\mathbf{s}(n))_{n \in \mathbb{Z}})$, alors la séparation des sources est effective, où $(\mathbf{y}(n))_{n \in \mathbb{Z}}$ est la sortie du système global, résultant d'un mélange par un filtre quelconque d'un tirage du vecteur $(\mathbf{s}(n))_{n \in \mathbb{Z}}$ de sources indépendantes.

Notons qu'un contraste est une fonction du filtre séparant $\mathbf{W}[z]$, mais que nous pourrions le considérer de façon équivalente comme fonction du filtre global $\mathbf{G}[z]$, ou, par extension, du vecteur des sorties globales $(\mathbf{y}(n))_{n \in \mathbb{Z}}$.

2.2.1 Une approche fréquentielle

Nous adoptons dans tout ce paragraphe 2.2.1 un point de vue fréquentiel. La matrice de la réponse fréquentielle du filtre de fonction de transfert $\mathbf{M}[z]$ sera notée $\mathbf{M}(\omega)$. Rappelons que nous supposons les observations blanchies et par conséquent $\mathbf{M}[z]$ est para-unitaire, ce qui se traduit sur la réponse fréquentielle par l'unitarité de $\mathbf{M}(\omega)$:

$$\forall \omega \quad \mathbf{M}(\omega)\mathbf{M}(\omega)^H = \mathbf{Id}_N \quad (2.6)$$

Comme expliqué précédemment, le filtre séparant $\mathbf{W}[z]$ —de réponse fréquentielle $\mathbf{W}(\omega)$ — est recherché dans la classe des filtres para-unitaires. Il en résulte que le filtre global $\mathbf{G}[z]$ —de réponse fréquentielle $\mathbf{G}(\omega)$ — est également para-unitaire. Notons au passage une conséquence du fait que seuls des filtres para-unitaires sont considérés : l'ambiguïté de filtrage scalaire indiquée dans le paragraphe 1.4.1-a) est dans ce cas une ambiguïté de filtrage scalaire *pas-se-tout* et l'équation (1.8) traduisant le fait que le filtre global est séparant s'écrit :

$$\forall \omega \quad \mathbf{G}(\omega) = \mathbf{P}e^{j\Phi(\omega)} \quad (2.7)$$

où $\Phi(\omega) = \text{Diag}(\phi_1(\omega), \dots, \phi_N(\omega))$ est une matrice diagonale réelle et \mathbf{P} une matrice de permutation. Ce fait apparaîtra dans les preuves et les énoncés établissant la validité des fonctions de contrastes ici introduites.

Nous citons d'abord les idées fondatrices d'une méthode de séparation fréquentielle. Le critère de diagonalisation conjointe introduit à ce propos servira de base à la construction des contrastes ensuite démontrés. Enfin, l'effet d'une discrétisation en fréquences sera considéré.

2.2.1-a) Un critère de diagonalisation conjointe

Ce paragraphe rappelle des idées présentes dans [19] et [20] et qui ont été à l'origine d'un algorithme de séparation de sources dans le cas i.i.d. Le principe est basé sur la diagonalisation conjointe de matrices, ce qui nous amènera à définir le critère (2.12). Cette même expression sera utilisée dans le paragraphe suivant afin de déduire des contrastes fréquentiels valables dans le cas de sources non i.i.d.

Considérons l'inter-trispectre des observations $(x_i(n))_{n \in \mathbb{Z}}$ prises sur les capteurs $(i, j, l_1, l_2) \in \{1, \dots, N\}^4$. Ce dernier s'écrit :

$$C_{ijl_1l_2}^4(\omega_1, \omega_2, \omega_3) = \sum_{(\tau_1, \tau_2, \tau_3) \in \mathbb{Z}^3} \text{Cum}\{x_i(n), x_j^*(n + \tau_1), x_{l_1}(n + \tau_2), x_{l_2}^*(n + \tau_3)\} e^{-j(\omega_1\tau_1 + \omega_2\tau_2 + \omega_3\tau_3)}. \quad (2.8)$$

L'indépendance et la gaussianité du bruit permet d'éliminer l'influence de ce dernier au cours du calcul. L'inter-trispectre peut alors s'exprimer en fonction du filtre de mélange et du trispectre $\Gamma_p^4(\omega_1, \omega_2, \omega_3)$ de chacune des sources $(s_p(n))_{n \in \mathbb{Z}, p \in \{1, \dots, N\}}$:

$$C_{ijl_1l_2}^4(\omega_1, \omega_2, \omega_3) = \sum_{p=1}^N \Gamma_p^4(\omega_1, \omega_2, \omega_3) M_{ip}(-\omega_1 - \omega_2 - \omega_3) M_{jp}^*(-\omega_1) M_{l_1p}(\omega_2) M_{l_2p}^*(-\omega_3). \quad (2.9)$$

Si l'on définit $\mathbf{C}_{l_1l_2}^4(\omega_1, \omega_2, \omega_3)$ la matrice dont le (i, j) -ème élément vaut $C_{ijl_1l_2}^4(\omega_1, \omega_2, \omega_3)$ et la matrice diagonale

$$\mathbf{D}_{l_1l_2}(\omega_1, \omega_2, \omega_3) := \text{Diag}\left(\{\Gamma_p^4(\omega_1, \omega_2, \omega_3) M_{l_1p}(\omega_2) M_{l_2p}^*(-\omega_3)\}_{p=1, \dots, N}\right), \quad (2.10)$$

l'équation ci-dessus peut s'écrire :

$$\mathbf{C}_{l_1l_2}^4(\omega_1, \omega_2, \omega_3) = \mathbf{M}(-\omega_1 - \omega_2 - \omega_3) \mathbf{D}_{l_1l_2}(\omega_1, \omega_2, \omega_3) \mathbf{M}(-\omega_1)^H. \quad (2.11)$$

Le problème de la séparation de sources peut alors être abordé de la façon suivante : pour un couple de fréquences (ω_2, ω_3) , on cherche $\mathbf{W}(-\omega_1 - \omega_2 - \omega_3)$ et $\mathbf{W}(-\omega_1)^H$, inverses estimées respectives des matrices unitaires $\mathbf{M}(-\omega_1 - \omega_2 - \omega_3)$ et $\mathbf{M}(-\omega_1)^H$, qui mènent à une diagonalisation conjointe (ou plus précisément une décomposition conjointe en valeurs singulières) de l'ensemble des matrices $(\mathbf{C}_{l_1l_2}^4(\omega_1, \omega_2, \omega_3))_{l_1, l_2}$. Ceci est l'idée à la base d'un algorithme de séparation [20], valable uniquement dans le cas de sources i.i.d. Plus précisément, [20] suggère pour tout ω_1 , de rechercher $\mathbf{W}(-\omega_1 - \omega_2 - \omega_3)$ et $\mathbf{W}(-\omega_1)^H$ comme étant les matrices unitaires qui maximisent :

$$\mathcal{I}(\omega_1, \omega_2, \omega_3) := \sum_{l_1, l_2} \text{on}(\mathbf{W}(-\omega_1 - \omega_2 - \omega_3) \mathbf{C}_{l_1l_2}^4(\omega_1, \omega_2, \omega_3) \mathbf{W}(-\omega_1)^H) \quad (2.12)$$

où $\text{on}(\mathbf{M})$ représente la somme des modules au carré des termes sur la diagonale de la matrice \mathbf{M} et (ω_2, ω_3) est choisi de façon appropriée.

Nous n'irons pas plus loin dans l'approche développée dans [20] mais la fonction \mathcal{I} va jouer un rôle clé dans la déduction des contrastes fréquentiels qui suivent. Notons que $\mathcal{I}(\omega_1, \omega_2, \omega_3)$ est une fonction de $\mathbf{W}(\omega)$, c'est-à-dire du système séparant. Par souci de concision des notations, nous ne rendrons pas explicite cette dépendance ni pour $\mathcal{I}(\omega_1, \omega_2, \omega_3)$, ni pour les autres critères qui seront déduits de $\mathcal{I}(\omega_1, \omega_2, \omega_3)$.

La propriété suivante de $\mathcal{I}(\omega_1, \omega_2, \omega_3)$ sera utilisée par la suite dans la démonstration de la validité des contrastes proposés.

Lemme 1 *On a :*

$$\mathcal{I}(\omega_1, \omega_2, \omega_3) = \sum_{i,j} |\Gamma_j^4(\omega_1, \omega_2, \omega_3)|^2 |G_{ij}(-\omega_1 - \omega_2 - \omega_3)|^2 |G_{ij}(-\omega_1)|^2 \quad (2.13)$$

$$= \sum_{il_1l_2} |\tilde{C}_{il_1l_2}^4(\omega_1, \omega_2, \omega_3)|^2 \quad (2.14)$$

où $(\tilde{C}_{il_1l_2}^4(\omega_1, \omega_2, \omega_3))_{i, l_1, l_2}$ correspond à l'inter-trispectre des sorties du système global. De plus, nous avons l'inégalité suivante :

$$\mathcal{I}(\omega_1, \omega_2, \omega_3) \leq \sum_j |\Gamma_j^4(\omega_1, \omega_2, \omega_3)|^2. \quad (2.15)$$

Preuve: D'après (2.11), (2.12) et sachant que $\mathbf{G}(\omega) = \mathbf{W}(\omega)\mathbf{M}(\omega)$ pour tout ω , nous avons

$$\mathcal{I}(\omega_1, \omega_2, \omega_3) = \sum_{l_1, l_2} \text{on}(\mathbf{G}(-\omega_1 - \omega_2 - \omega_3)\mathbf{D}_{l_1 l_2}(\omega_1, \omega_2, \omega_3)\mathbf{G}(-\omega_1)^H) \quad (2.16)$$

et

$$\begin{aligned} \text{on}(\mathbf{G}(-\omega_1 - \omega_2 - \omega_3)\mathbf{D}_{l_1 l_2}(\omega_1, \omega_2, \omega_3)\mathbf{G}(-\omega_1)^H) = \\ \sum_i \left| \sum_j G_{ij}(-\omega_1 - \omega_2 - \omega_3)\Gamma_j^4(\omega_1, \omega_2, \omega_3)M_{l_1 j}(\omega_2)M_{l_2 j}^*(-\omega_3)G_{ij}^*(-\omega_1) \right|^2. \end{aligned} \quad (2.17)$$

On obtient par conséquent :

$$\begin{aligned} \mathcal{I}(\omega_1, \omega_2, \omega_3) = \sum_{i, j_1, j_2} \left(\Gamma_{j_1}^4(\omega_1, \omega_2, \omega_3)\Gamma_{j_2}^{4*}(\omega_1, \omega_2, \omega_3)G_{ij_1}(-\omega_1 - \omega_2 - \omega_3)G_{ij_2}^*(-\omega_1 - \omega_2 - \omega_3) \right. \\ \left. G_{ij_1}^*(-\omega_1)G_{ij_2}(-\omega_1) \left(\sum_{l_1} M_{l_1 j_1}(\omega_2)M_{l_1 j_2}^*(\omega_2) \right) \left(\sum_{l_2} M_{l_2 j_1}^*(-\omega_3)M_{l_2 j_2}(-\omega_3) \right) \right). \end{aligned} \quad (2.18)$$

En utilisant le fait que les réponses fréquentielles considérées sont unitaires ($\forall \omega, \mathbf{M}(\omega)\mathbf{M}(\omega)^H = \mathbf{Id}$), il vient :

$$\sum_{l_1} M_{l_1 j_1}(\omega_2)M_{l_1 j_2}^*(\omega_2) = \sum_{l_2} M_{l_2 j_1}^*(-\omega_3)M_{l_2 j_2}(-\omega_3) = \delta_{j_1 - j_2} \quad (2.19)$$

ce qui donne immédiatement la première égalité de (2.14). De plus on a :

$$\begin{aligned} \sum_{i l_1 l_2} |\tilde{C}_{i l_1 l_2}^4(\omega_1, \omega_2, \omega_3)|^2 = \sum_{i, j_1, j_2} \left(\Gamma_{j_1}^4(\omega_1, \omega_2, \omega_3)\Gamma_{j_2}^{4*}(\omega_1, \omega_2, \omega_3)G_{ij_1}(-\omega_1 - \omega_2 - \omega_3) \right. \\ \left. G_{ij_2}^*(-\omega_1 - \omega_2 - \omega_3)G_{ij_1}^*(-\omega_1)G_{ij_2}(-\omega_1) \right. \\ \left. \left(\sum_{l_1} G_{l_1 j_1}(\omega_2)G_{l_1 j_2}^*(\omega_2) \right) \left(\sum_{l_2} G_{l_2 j_1}^*(-\omega_3)G_{l_2 j_2}(-\omega_3) \right) \right). \end{aligned} \quad (2.20)$$

Cette formule étant similaire à (2.18), en procédant comme précédemment, on montre qu'elle est égale à l'expression (2.14) trouvée pour $\mathcal{I}(\omega_1, \omega_2, \omega_3)$.

Maintenant, il découle de la para-unitarité de $\mathbf{G}(\omega)$ que pour tout ω et j , $\sum_i |G_{ij}(\omega)|^2 = 1$, ce qui entraîne que pour tout i , $|G_{ij}(\omega)| \leq 1$. Ceci mène à :

$$\sum_i |G_{ij}(-\omega_1 - \omega_2 - \omega_3)|^2 |G_{ij}(-\omega_1)|^2 \leq \sum_i |G_{ij}(-\omega_1 - \omega_2 - \omega_3)|^2 = 1. \quad (2.21)$$

Cette inégalité, combinée avec (2.14) donne (2.15). ■

Remarque 1: Le critère $\mathcal{I}(\omega_1, \omega_2, \omega_3)$ ne réalise pas un contraste. En effet, par des raisonnements similaires à ceux qui seront donnés dans la démonstration de la proposition 1, on obtient que $\mathcal{I}(\omega_1, \omega_2, \omega_3)$ est maximum uniquement dans le cas où le système global «sépare», mais pour seulement un couple de fréquences. Ceci n'est pas suffisant pour assurer la propriété de séparation et dans les paragraphes suivants, nous serons donc amenés à intégrer ce critère sur différents ensembles de fréquences.

D'autre part, l'équation (2.14) montre de façon simple que, en général, on ne peut espérer identifier la phase du système à partir de la seule maximisation de $\mathcal{I}(\omega_1, \omega_2, \omega_3)$, car la valeur de ce critère ne dépend que des modules de $G_{ij}(-\omega_1 - \omega_2 - \omega_3)$ et $G_{ij}(-\omega_1)$. Le lemme montre aussi que le critère optimisé ne dépend en fait pas des valeurs $M_{l_1 i}(\omega_2)M_{l_2 i}^*(-\omega_3)$ qui apparaissent dans la définition de $\mathbf{D}_{l_1 l_2}(\omega_1, \omega_2, \omega_3)$. Ces valeurs sont cependant importantes dans la méthode décrite dans [20] afin de résoudre l'ambiguïté de permutation.

Remarque 2: Il est intéressant de noter que si les quantités dans (2.14) ne dépendaient pas de la fréquence, ce qui reviendrait à considérer des mélanges instantanés de sources i.i.d., le critère considéré se réduirait à celui utilisé dans l'algorithme de séparation JADE [15]. Cela signifie que notre approche étend le travail de [15] au cas convolutif.

2.2.1-b) Contrastes fréquentiels

Dans ce paragraphe, nous proposons de nouveaux contrastes et démontrons leur validité. Ces derniers sont définis à partir du critère \mathcal{I} défini en (2.12). À ce titre, ils font intervenir les trispectres des sources et des sorties globales (d'après (2.14)) et peuvent donc être qualifiés de contrastes fréquentiels. Une hypothèse technique est nécessaire au préalable :

H.5 Pour presque tout $(\omega, \nu) \in [0, 2\pi]^2$, il existe $\alpha_{\omega, \nu} \in [-2\pi, 2\pi[$ tel que, pour au moins $N - 1$ sources :

$$\Gamma_j^4\left(\omega, \frac{\nu + \alpha_{\omega, \nu}}{2}, \frac{\nu - \alpha_{\omega, \nu}}{2}\right) \neq 0. \quad (2.22)$$

L'hypothèse adoptée ici est assez faible dans la mesure où elle nous permet de considérer des sources non i.i.d. Pour que cette hypothèse soit vérifiée, au plus une source peut être gaussienne.

Des choix possibles pour $\alpha_{\omega, \nu}$ sont $\alpha_{\omega, \nu} = \nu$, $\alpha_{\omega, \nu} = -\nu$ et $\alpha_{\omega, \nu} = 0$. Pour ces choix, seules des tranches 2D simples du trispectre sont mises en jeu.

Nous allons maintenant prouver que

$$\bar{\mathcal{I}} := \int_0^{2\pi} \int_0^{2\pi} \mathcal{I}\left(\omega, \frac{\nu + \alpha_{\omega, \nu}}{2}, \frac{\nu - \alpha_{\omega, \nu}}{2}\right) d\omega d\nu \quad (2.23)$$

est un contraste, et ainsi, sa maximisation permet de séparer les sources à un filtrage scalaire passe-tout près de chacune d'entre elles.

Proposition 1 *On a la majoration :*

$$\bar{\mathcal{I}} \leq \sum_j \int_0^{2\pi} \int_0^{2\pi} \left| \Gamma_j^4\left(\omega, \frac{\nu + \alpha_{\omega, \nu}}{2}, \frac{\nu - \alpha_{\omega, \nu}}{2}\right) \right|^2 d\omega d\nu =: \bar{\mathcal{I}}_{\max} \quad (2.24)$$

Sous l'hypothèse H.5, la borne supérieure $\bar{\mathcal{I}}_{\max}$ est atteinte si et seulement si la séparation est réalisée, c'est-à-dire si et seulement si $\mathbf{G}(\omega)$ vérifie (2.7).

Preuve: De la formule (2.15) dans le lemme 1 on déduit que :

$$\int_0^{2\pi} \int_0^{2\pi} \mathcal{I}\left(\omega, \frac{\nu + \alpha_{\omega, \nu}}{2}, \frac{\nu - \alpha_{\omega, \nu}}{2}\right) d\omega d\nu \leq \int_0^{2\pi} \int_0^{2\pi} \sum_j \left| \Gamma_j^4\left(\omega, \frac{\nu + \alpha_{\omega, \nu}}{2}, \frac{\nu - \alpha_{\omega, \nu}}{2}\right) \right|^2 d\omega d\nu \quad (2.25)$$

ce qui établit la majoration (2.24).

De plus, ce majorant est atteint si et seulement si, pour presque tout $(\omega, \nu) \in [0, 2\pi]^2$, il y a égalité entre les expressions intégrées dans chaque membre ci-dessus. D'après la preuve du lemme 1 (équation (2.21) notamment) il est alors clair que ce majorant est atteint si et seulement si, pour tout j et presque tout $(\omega, \nu) \in [0, 2\pi]^2$,

$$\left| \Gamma_j^4\left(\omega, \frac{\nu + \alpha_{\omega, \nu}}{2}, \frac{\nu - \alpha_{\omega, \nu}}{2}\right) \right|^2 \left(1 - \sum_i |G_{ij}(-\omega - \nu)|^2 |G_{ij}(-\omega)|^2\right) = 0. \quad (2.26)$$

Soit maintenant j un indice quelconque tel que (2.22) soit vérifiée. On a alors pour presque tout $(\omega, \nu) \in [0, 2\pi]^2$,

$$\sum_i |G_{ij}(-\omega - \nu)|^2 |G_{ij}(-\omega)|^2 = 1. \quad (2.27)$$

ce qui conduit à :

$$\frac{1}{4\pi^2} \sum_i \int_0^{2\pi} \int_0^{2\pi} |G_{ij}(-\omega - \nu)|^2 |G_{ij}(-\omega)|^2 d\omega d\nu = 1. \quad (2.28)$$

Un simple changement de variable donne par ailleurs :

$$\frac{1}{2\pi} \int_0^{2\pi} |G_{ij}(-\omega - \nu)|^2 d\nu = \frac{1}{2\pi} \int_0^{2\pi} |G_{ij}(-\omega)|^2 d\omega =: \|G_{ij}\|^2 \quad (2.29)$$

et (2.28) est équivalente à

$$\sum_i \|G_{ij}\|^4 = 1. \quad (2.30)$$

De plus, l'hypothèse d'unitarité entraîne que

$$\forall j, \quad \sum_i \|G_{ij}\|^2 = 1 \quad (2.31)$$

où, évidemment, $\|G_{ij}\| \leq 1$. Si pour tout i , $\|G_{ij}\| < 1$, nous aurions $\sum_i \|G_{ij}\|^4 < 1$. Ceci montre que l'égalité ne peut arriver dans (2.24) que s'il existe un indice i_j tel que

$$\|G_{i_j j}\| = 1. \quad (2.32)$$

En raison de (2.31), ceci entraîne que pour tout $i \neq i_j$, $\|G_{ij}\| = 0$, i.e. $G_{ij}(\omega) = 0$, pour presque tout $\omega \in [0, 2\pi[$. En procédant de même pour toutes les (au moins $N - 1$) valeurs de j telles que (2.22) soit vérifiée, on établit que sur la j ème colonne correspondante de $\mathbf{G}(\omega)$, il n'y a qu'un seul élément non nul, sur la ligne i_j . Comme la condition d'unitarité introduit aussi la contrainte

$$\forall i, \quad \sum_j \|G_{ij}\|^2 = 1 \quad (2.33)$$

et que (2.32) est satisfaite, deux de ces éléments non nuls de $\mathbf{G}(\omega)$ ne peuvent être situés sur la même ligne.

Quand (2.22) n'est pas vérifiée pour l'un des indices, disons j_0 , on déduit facilement qu'il n'y a qu'un seul élément non nul sur la j_0 ème colonne de $\mathbf{G}(\omega)$ et qu'il est situé sur la ligne restant vacante.

En d'autres termes, on a prouvé que $\mathbf{G}(\omega) = \mathbf{P} \mathbf{\Lambda}(\omega)$ où \mathbf{P} est une permutation et $\mathbf{\Lambda}(\omega)$ est une matrice diagonale. Comme $\mathbf{G}(\omega)$ est unitaire, les éléments diagonaux de $\mathbf{\Lambda}(\omega)$ sont nécessairement de module un, ce qui achève de montrer que $\mathbf{G}(\omega)$ satisfait (2.7). ■

Sous une condition légèrement plus restrictive, une autre forme de contraste peut être obtenue :

Proposition 2 *Supposons que :*

H.6 *Pour presque tout $(\omega, \nu) \in [0, 2\pi]^2$, il existe un ensemble $\mathcal{E}(\omega, \nu) \subset [-2\pi, 2\pi[$ tel que, pour au moins $N - 1$ sources :*

$$\int_{\mathcal{E}(\omega, \nu)} \left| \Gamma_j^4\left(\omega, \frac{\nu + \alpha}{2}, \frac{\nu - \alpha}{2}\right) \right|^2 d\alpha \neq 0. \quad (2.34)$$

Alors, on a la majoration

$$\begin{aligned} \mathcal{I}_{\mathcal{E}} &:= \int_0^{2\pi} \int_0^{2\pi} \left(\int_{\mathcal{E}(\omega, \nu)} \mathcal{I}\left(\omega, \frac{\nu + \alpha}{2}, \frac{\nu - \alpha}{2}\right) d\alpha \right) d\omega d\nu \\ &\leq \sum_j \int_0^{2\pi} \int_0^{2\pi} \left(\int_{\mathcal{E}(\omega, \nu)} |\Gamma_j^4\left(\omega, \frac{\nu + \alpha}{2}, \frac{\nu - \alpha}{2}\right)|^2 d\alpha \right) d\omega d\nu \end{aligned} \quad (2.35)$$

et, de plus, $\mathcal{I}_{\mathcal{E}}$ est un contraste.

La preuve de ce résultat est omise car elle est très similaire à la preuve de la proposition 1.

Remarquons que (2.34) signifie qu'il existe un ensemble $\mathcal{F}_j(\omega, \nu)$ inclus dans $\mathcal{E}(\omega, \nu)$ et de mesure non nulle tel que

$$\forall \alpha \in \mathcal{F}_j(\omega, \nu), \quad \Gamma_j^4\left(\omega, \frac{\nu + \alpha}{2}, \frac{\nu - \alpha}{2}\right) \neq 0. \quad (2.36)$$

Par conséquent, si H.6 est vérifiée, l'hypothèse H.5 l'est clairement aussi. Si de plus pour tout j , $\Gamma_j^4(\omega_1, \omega_2, \omega_3)$ est continue sur $[-\pi, \pi]^3$, nous avons même équivalence entre les deux hypothèses. Dans ce cas en effet, (2.22) implique que (2.34) est satisfaite.

En particulier, si les cumulants des sources forment une famille sommable et si, pour tout j , $\Gamma_j^4(\omega_1, \omega_2, \omega_3)$ est une fonction rationnelle en $(e^{i\omega_1}, e^{i\omega_2}, e^{i\omega_3})$, les conditions H.5 et H.6 sont équivalentes à la condition suivante :

H.7 Pour $N - 1$ sources au moins, il existe des retards temporels $(\tau_{1j}, \tau_{2j}, \tau_{3j})$ tels que

$$\text{Cum}\{s_j(n), s_j^*(n + \tau_{1j}), s_j(n + \tau_{2j}), s_j^*(n + \tau_{3j})\} \neq 0. \quad (2.37)$$

Preuve: La sommabilité de la famille de cumulants des sources assure la continuité de $\Gamma_j^4(\omega_1, \omega_2, \omega_3)$ sur $[-\pi, \pi]^3$ et ainsi l'équivalence entre H.5 et H.6. Nous démontrons par conséquent uniquement l'équivalence entre H.5 et la condition citée ci-dessus.

Si nous supposons (2.22), $\Gamma_j^4(\omega_1, \omega_2, \omega_3)$ n'est pas la fonction identiquement nulle et par suite (2.37) est vérifiée.

Réciproquement, supposons H.5 fautive : il existe alors $\mathcal{N} \subset [0, 2\pi]^2$ de mesure non nulle tel que, pour tout $(\omega, \nu) \in \mathcal{N}$ et pour tout $\alpha \in [-2\pi, 2\pi]$, au moins deux sources vérifient :

$$\Gamma_j^4\left(\omega, \frac{\nu + \alpha}{2}, \frac{\nu - \alpha}{2}\right) = 0. \quad (2.38)$$

Ainsi, il existe un ensemble de mesure non nulle qui annule la fonction ci-dessus pour deux indices j distincts (au moins). Or, $\Gamma_j^4(\omega_1, \omega_2, \omega_3)$ étant une fonction rationnelle en $(e^{i\omega_1}, e^{i\omega_2}, e^{i\omega_3})$, $\Gamma_j^4\left(\omega, \frac{\nu + \alpha}{2}, \frac{\nu - \alpha}{2}\right)$ est une fraction rationnelle en $(e^{i\omega}, e^{i\frac{\nu}{2}}, e^{i\frac{\alpha}{2}})$ et comme (2.38) est vérifiée sur un ensemble de mesure non nulle, cette fraction rationnelle est nulle. Donc Γ_j^4 est identiquement nulle et (2.37) n'est pas vérifié, ce qui achève la démonstration. ■

Il est également clair que lorsque $\mathcal{E}(\omega, \nu)$ est un sous ensemble fixé \mathcal{E} de $[-2\pi, 2\pi[$ (indépendant de (ω, ν)), on a

$$\mathcal{I}_{\mathcal{E}} = \int_{\mathcal{E}} \bar{\mathcal{I}}_{\alpha} d\alpha \quad (2.39)$$

où $\bar{\mathcal{I}}_{\alpha}$ est un critère similaire à (2.23). Un exemple de tel contraste est considéré dans le paragraphe suivant.

Intégration sur tout l'espace fréquentiel Un critère qui pourrait apparaître plus intuitif est donné par :

$$\int_{-\pi}^{\pi} \int_{-\pi}^{\pi} \int_{-\pi}^{\pi} \mathcal{I}(\omega_1, \omega_2, \omega_3) d\omega_1 d\omega_2 d\omega_3. \quad (2.40)$$

En utilisant les 2π -périodicités de $\mathcal{I}(\omega_1, \omega_2, \omega_3)$, et après quelques changements de variables, l'intégrale ci-dessus peut s'exprimer de la manière suivante :

$$\begin{aligned} \int_{-\pi}^{\pi} \int_{-\pi}^{\pi} \int_{-\pi}^{\pi} \mathcal{I}(\omega_1, \omega_2, \omega_3) d\omega_1 d\omega_2 d\omega_3 &= \frac{1}{2} \int_{-\pi}^{\pi} \int_{-2\pi}^{2\pi} \int_{-2\pi+|\nu|}^{2\pi-|\nu|} \mathcal{I}\left(\omega, \frac{\nu+\alpha}{2}, \frac{\nu-\alpha}{2}\right) d\alpha d\nu d\omega \\ &= \frac{1}{2} \left(\int_{-\pi}^{\pi} \int_0^{2\pi} \int_{-2\pi+\nu}^{2\pi-\nu} \mathcal{I}\left(\omega, \frac{\nu+\alpha}{2}, \frac{\nu-\alpha}{2}\right) d\alpha d\nu d\omega \right. \\ &\quad \left. + \int_{-\pi}^{\pi} \int_0^{2\pi} \int_{-\nu}^{\nu} \mathcal{I}\left(\omega, \frac{\nu+\alpha-2\pi}{2}, \frac{\nu-\alpha-2\pi}{2}\right) d\alpha d\nu d\omega \right) \\ &= \frac{1}{2} \int_{-\pi}^{\pi} \int_0^{2\pi} \int_{-2\pi+\nu}^{2\pi+\nu} \mathcal{I}\left(\omega, \frac{\nu+\alpha}{2}, \frac{\nu-\alpha}{2}\right) d\alpha d\nu d\omega \\ &= \frac{1}{2} \int_{-\pi}^{\pi} \int_0^{2\pi} \int_{-2\pi}^{2\pi} \mathcal{I}\left(\omega, \frac{\nu+\alpha}{2}, \frac{\nu-\alpha}{2}\right) d\alpha d\nu d\omega. \end{aligned} \quad (2.41)$$

Cette expression est égale à $\mathcal{I}_{\mathcal{E}}/2$ quand $\mathcal{E}(\omega, \nu) = [-2\pi, 2\pi[$. En utilisant (2.14), il en découle :

$$\frac{1}{2} \mathcal{I}_{[-2\pi, 2\pi[} = \int_{-\pi}^{\pi} \int_{-\pi}^{\pi} \int_{-\pi}^{\pi} \mathcal{I}(\omega_1, \omega_2, \omega_3) d\omega_1 d\omega_2 d\omega_3 \quad (2.42)$$

$$= \sum_{i_1 l_2} \int_{-\pi}^{\pi} \int_{-\pi}^{\pi} \int_{-\pi}^{\pi} |\tilde{C}_{i_1 l_2}^4(\omega_1, \omega_2, \omega_3)|^2 d\omega_1 d\omega_2 d\omega_3. \quad (2.43)$$

En utilisant l'égalité de Parseval, on obtient également une expression de ce critère dans le domaine temporel :

$$\frac{1}{16\pi^3} \mathcal{I}_{[-2\pi, 2\pi[} = \sum_{i_1 l_2} \sum_{\tau_1, \tau_2, \tau_3} |\text{Cum}\{y_i(n), y_i^*(n + \tau_1), y_{l_1}(n + \tau_2), y_{l_2}^*(n + \tau_3)\}|^2. \quad (2.44)$$

Sources i.i.d. Lorsque les sources sont i.i.d., l'hypothèse H.5 se réduit à l'hypothèse plus habituelle : au plus l'un des cumulants d'ordre quatre des sources est nul. De plus, d'après (2.13), $\mathcal{I}(\omega_1, \omega_2, \omega_3)$ devient une fonction des deux variables ω_1 et $\omega_2 + \omega_3$ dans ce cas. Par conséquent, les contrastes $\bar{\mathcal{I}}$ ne dépendent plus du choix de $\alpha_{\omega, \nu}$, qui peut être fait arbitrairement. Enfin, on a

$$\bar{\mathcal{I}} = \frac{1}{4\pi} \mathcal{I}_{[-2\pi, 2\pi[}. \quad (2.45)$$

Ceci montre que le critère correspondant est de nouveau équivalent à un contraste basé sur les intercumulants (2.44). Dans ce cas, il est facile de montrer que (2.24) devient

$$\frac{1}{4\pi^2} \bar{\mathcal{I}} = \sum_j \left(|\gamma_j^4|^2 \sum_i \|G_{ij}\|^4 \right) \leq \sum_j |\gamma_j^4|^2. \quad (2.46)$$

où $\gamma_j^4 := \text{Cum}\{s_j(n), s_j^*(n), s_j(n), s_j^*(n)\}$ est l'autocumulant d'ordre 4 (et sans aucun retard) de la $j^{\text{ème}}$ source.

On peut noter que, contrairement à l'approche de [20], la séparation peut être obtenue en maximisant la fonction de contraste proposée, sans supposer que les valeurs $|M_{l_1 i}(\omega_2)M_{l_2 i}(-\omega_3)|$ sont distinctes. A cet égard, cette approche ne fait pas jouer un rôle critique à certaines valeurs de ω_2 et ω_3 .

2.2.1-c) Contrastes discrétisés

L'utilisation pratique et l'estimation numérique des contrastes fréquentiels introduits précédemment peuvent être délicates. Aussi, nous proposons d'étudier dans ce paragraphe leurs versions équivalentes après discrétisation des fréquences.

Considérons le critère suivant :

$$\bar{\mathcal{I}}^{(K)} := \frac{4\pi^2}{K^2} \sum_{p=0}^{K-1} \sum_{q=0}^{K-1} \mathcal{I}\left(\nu_p, \frac{\nu_{q+r_{p,q}}}{2}, \frac{\nu_{q-r_{p,q}}}{2}\right) \quad (2.47)$$

où pour tout $p \in \{0, \dots, K-1\}$, $\nu_p = 2\pi p/K$ et pour tout $(p, q) \in \{0, \dots, K-1\}^2$, $r_{p,q} \in \{-K+1, \dots, K-1\}$.

$\bar{\mathcal{I}}^{(K)}$ est obtenu à partir de l'équation (2.23) après discrétisation sur $K \in \mathbb{N}^*$ points fréquentiels. Pour clarifier le liens qui existe entre $\bar{\mathcal{I}}^{(K)}$ et le critère $\bar{\mathcal{I}}$ dans (2.23), supposons que $\alpha_{\omega, \nu} = 2\pi a_{\omega, \nu}/P$ où $P \in \mathbb{N}^*$ et $a_{\omega, \nu} \in \{-P+1, \dots, P-1\}$. Supposons aussi que la fonction

$$(\omega, \nu) \mapsto \mathcal{I}\left(\omega, \frac{\nu + \alpha_{\omega, \nu}}{2}, \frac{\nu - \alpha_{\omega, \nu}}{2}\right) \quad (2.48)$$

est intégrable au sens de Riemann. Alors, en choisissant $K = kP$ avec $k \in \mathbb{N}$ et, pour tout $(p, q) \in \{0, \dots, K-1\}^2$, $r_{p,q} = k a_{\nu_p, \nu_q}$, on obtient directement : $\lim_{k \rightarrow \infty} \bar{\mathcal{I}}^{(K)} = \bar{\mathcal{I}}$.

Nous allons étudier maintenant ce qui se passe lorsque nous ne sommes pas dans les conditions asymptotiques. L'hypothèse H.5 doit d'abord modifiée comme suit :

H.8 Pour tout $(p, q) \in \{0, \dots, K-1\}^2$, il existe $r_{p,q} \in \{-K+1, \dots, K-1\}$ tel que, pour au moins $N-1$ sources :

$$\Gamma_j^4\left(\nu_p, \frac{\nu_{q+r_{p,q}}}{2}, \frac{\nu_{q-r_{p,q}}}{2}\right) \neq 0. \quad (2.49)$$

Nous pouvons alors donner le résultat suivant, qui peut être vu comme la forme équivalente de la proposition 1 dans le cas discrétisé :

Proposition 3 *Sous l'hypothèse H.8, la borne supérieure de $\bar{\mathcal{I}}^{(K)}$ est atteinte si et seulement si, pour tout $p \in \{0, \dots, K-1\}$, $\mathbf{G}(\nu_p) = \mathbf{P}^{(K)} e^{i\Phi_p^{(K)}}$, où $\mathbf{P}^{(K)}$ est une matrice de permutation et où, pour tout $p \in \{0, \dots, K-1\}$, $\Phi_p^{(K)}$ est une matrice diagonale réelle. De plus, si le système global a pour réponse impulsionnelle $(g_{ij}(k))_k$ de longueur finie, inférieure ou égale à un entier K , ce que l'on note¹ :*

$$g_{ij}(k) = 0 \quad \text{pour } k < 0 \text{ ou } k \geq K \quad (2.50)$$

la borne supérieure est atteinte si et seulement si la séparation est réalisée. On a alors pour tout ω , $\mathbf{G}(\omega) = \mathbf{P} e^{i(\Phi + \mathbf{D}\omega)}$ où \mathbf{P} est une matrice de permutation, Φ est une matrice diagonale réelle et \mathbf{D} est une matrice diagonale à éléments entiers.

¹L'extension au cas non causal RIF est immédiate.

Preuve: Nous allons procéder de manière similaire à la preuve de la proposition 1. D'après (2.15),

$$\bar{\mathcal{I}}^{(K)} \leq \frac{4\pi^2}{K^2} \sum_{p=0}^{K-1} \sum_{q=0}^{K-1} \sum_j \left| \Gamma_j^4 \left(\nu_p, \frac{\nu_{q+r_{p,q}}}{2}, \frac{\nu_{q-r_{p,q}}}{2} \right) \right|^2 \quad (2.51)$$

et la borne supérieure est atteinte si et seulement si, pour tout j tel que (2.49) soit vérifiée et pour tout $(p, q) \in \{0, \dots, K-1\}^2$,

$$\sum_i |G_{ij}(-\nu_p - \nu_q)|^2 |G_{ij}(-\nu_p)|^2 = 1. \quad (2.52)$$

En utilisant l'égalité de Parseval (pour la TFD), ceci mène à

$$\sum_i \left| \sum_{k=0}^{K-1} |g_{ij}^{(K)}(k)|^2 \right|^2 = 1, \quad (2.53)$$

où $(g_{ij}^{(K)}(k))_{0 \leq k < K}$ correspond à la TFD inverse de $(G_{ij}(\nu_p))_{0 \leq p < K}$ i.e. à une version périodisée de la (i, j) ème réponse impulsionnelle globale $(g_{ij}(k))_k$:

$$g_{ij}^{(K)}(k) := \sum_p g_{ij}(k + pK). \quad (2.54)$$

De plus le fait que, pour tout $p \in \{0, \dots, K-1\}$, la matrice $\mathbf{G}(\nu_p)$ soit une matrice unitaire introduit les contraintes :

$$\forall j, \quad \sum_i \sum_{k=0}^{K-1} |g_{ij}^{(K)}(k)|^2 = 1, \quad (2.55)$$

$$\forall i, \quad \sum_j \sum_{k=0}^{K-1} |g_{ij}^{(K)}(k)|^2 = 1. \quad (2.56)$$

En utilisant les mêmes arguments qu'à la fin de la preuve de la proposition 1, ceci nous permet de conclure que $\mathbf{G}(\nu_p)$ est de la forme $\mathbf{P}^{(K)} e^{i\Phi_p^{(K)}}$.

Sous l'hypothèse (2.50), $(g_{ij}(k))_k$ et sa version périodisée sont identiques pour $k \in \{0, \dots, K-1\}$ (c'est-à-dire $g_{ij}^{(K)}(k) = g_{ij}(k), \forall k \in \{0, \dots, K-1\}$) et par conséquent, pour tout (i, j) fixé dans $\{0, \dots, N-1\}^2$, nous avons l'équivalence :

$$\left(\forall p \in \{0, \dots, K-1\}, \quad G_{i,j}(\nu_p) = 0 \right) \iff \left(\forall k \in \{0, \dots, K-1\}, \quad g_{i,j}(k) = 0 \right). \quad (2.57)$$

Ceci signifie que $\bar{\mathcal{I}}^{(K)}$ est maximal si et seulement si la séparation est réalisée. De plus, comme le système global a été supposé RIF, l'indétermination passe-tout se réduit à une indétermination de phase et de retard (voir paragraphe 2.4.1-b)). ■

Il est également possible de définir une version discrétisée de $\mathcal{I}_{\mathcal{E}}$:

$$\mathcal{I}_{\mathcal{E}^K}^{(K)} := \frac{8\pi^3}{K^3} \sum_{p=0}^{K-1} \sum_{q=0}^{K-1} \sum_{r \in \mathcal{E}_{p,q}^K} \mathcal{I} \left(\nu_p, \frac{\nu_{q+r}}{2}, \frac{\nu_{q-r}}{2} \right) \quad (2.58)$$

où $\mathcal{E}_{p,q}^K \subset \{-K+1, \dots, K-1\}$. Des résultats analogues à ceux obtenus pour le critère précédent s'appliquent alors. On peut en particulier facilement démontrer :

Proposition 4 *Supposons :*

H.9 Pour tout $(p, q) \in \{0, \dots, K-1\}^2$, il existe un ensemble $\mathcal{E}_{p,q}^K \subset \{-K+1, \dots, K-1\}$ tel que, pour au moins $N-1$ sources,

$$\sum_{r \in \mathcal{E}_{p,q}^K} \left| \Gamma_j^4 \left(\nu_p, \frac{\nu_{q+r}}{2}, \frac{\nu_{q-r}}{2} \right) \right|^2 \neq 0. \quad (2.59)$$

Sous l'hypothèse H.9, la borne supérieure de $\mathcal{I}_{\mathcal{E}^K}^{(K)}$ est atteinte si et seulement si, pour tout $p \in \{0, \dots, K-1\}$, $\mathbf{G}(\nu_p) = \mathbf{P}^{(K)} e^{i\Phi_p^{(K)}}$, où $\mathbf{P}^{(K)}$ est une matrice de permutation et, pour tout $p \in \{0, \dots, K-1\}$, $\Phi_p^{(K)}$ est une matrice diagonale réelle.

De plus, si le système global a pour réponse impulsionnelle $(g_{ij}(k))_k$ de longueur finie inférieure ou égale à K , la borne supérieure est atteinte si et seulement si pour tout ω , $\mathbf{G}(\omega) = \mathbf{P} e^{i(\Phi + \mathbf{D}\omega)}$.

Il est intéressant de remarquer que la proposition 3 est en fait un cas particulier de la proposition 4 dans la mesure où nous avons :

$$\bar{\mathcal{I}}^{(K)} = \frac{K}{2\pi} \mathcal{I}_{\mathcal{E}^K}^{(K)} \quad (2.60)$$

quand, pour tout $(p, q) \in \{0, \dots, K-1\}^2$, $\mathcal{E}_{p,q}^K$ se réduit au singleton $\{r_{p,q}\}$. Ainsi, après discrétisation, $\mathcal{I}_{\mathcal{E}^K}^{(K)}$ apparaît comme la forme la plus générale de contraste fréquentiel que l'on peut considérer.

Des commentaires supplémentaires peuvent être faits à propos de l'hypothèse (2.50). En pratique, la connaissance *a priori* que nous pouvons avoir est que le système mélangeant para-unitaire a une réponse impulsionnelle de longueur finie inférieure ou égale à L . Ainsi, ses sorties peuvent être séparées en utilisant un système à réponse impulsionnelle finie, dont la réponse impulsionnelle est de la même longueur maximale et le système global résultant vérifie alors (2.50) avec $K = 2L - 1$. La difficulté principale dans ce cas est de trouver une paramétrisation des matrices $\mathbf{G}(\nu_p)$, $p \in \{0, \dots, K-1\}$ cohérente avec à la fois les hypothèses de réponse impulsionnelle finie et de para-unitarité. Nous verrons dans la partie 2.4.1, qu'une possibilité pour cela est de considérer une représentation de l'inverse estimé du système mélangeant par une structure en treillis sans perte.

2.2.2 Retour sur les contrastes temporels

2.2.2-a) Généralisation de contrastes connus

Les contrastes précédents requièrent l'estimation de nombreux cumulants croisés, comme on peut le voir au travers de l'équation (2.44), par exemple. De ce fait, il faut s'attendre à des performances limitées au niveau de la qualité de l'estimation statistique. De plus, le calcul d'un critère mettant en jeu de nombreux cumulants peut représenter une charge de calcul importante. Pour faire face à ces difficultés, nous adoptons la même approche que dans [90] et nous déduisons des contrastes qui demeurent valables même dans le cas non i.i.d. Cette approche est basée sur la propriété suivante [69, 90] :

Propriété 1 (majoration d'un contraste) Soit \mathcal{C}_1 une fonction réelle des sorties du système global et soit \mathcal{C}_2 un contraste. Si les conditions suivantes sont réunies :

- (i) pour tout signal source $(\mathbf{s}(n))_{n \in \mathbb{Z}}$ dont les composantes sont indépendantes : $\mathcal{C}_1((\mathbf{s}(n))_{n \in \mathbb{Z}}) = \mathcal{C}_2((\mathbf{s}(n))_{n \in \mathbb{Z}})$,

(ii) pour toutes les sorties globales $(\mathbf{y}(n))_{n \in \mathbb{Z}}$ possibles : $\mathcal{C}_1((\mathbf{y}(n))_{n \in \mathbb{Z}}) \leq \mathcal{C}_2((\mathbf{y}(n))_{n \in \mathbb{Z}})$, alors, \mathcal{C}_1 est un contraste.

De manière similaire à l'expression de $\mathcal{I}(\omega_1, \omega_2, \omega_3)$ dans le lemme 1, définissons :

$$\mathcal{J}(\omega_1, \omega_2, \omega_3) := \sum_{i l_1 l_2} \eta_{i l_1 l_2} |\tilde{\mathcal{C}}_{i l_1 l_2}^4(\omega_1, \omega_2, \omega_3)|^2 \quad (2.61)$$

où, les coefficients de pondération $\eta_{i l_1 l_2}$ sont définis par :

$$\forall (i, l_1, l_2) \in \{1, \dots, N\}^3 \quad \eta_{i l_1 l_2} \begin{cases} = 1 & \text{quand } i = l_1 = l_2 \\ \leq 1 & \text{sinon.} \end{cases} \quad (2.62)$$

Par analogie avec (2.23), nous considérons maintenant :

$$\bar{\mathcal{J}} := \int_0^{2\pi} \int_0^{2\pi} \mathcal{J}\left(\omega, \frac{\nu + \alpha_{\omega, \nu}}{2}, \frac{\nu - \alpha_{\omega, \nu}}{2}\right) d\omega d\nu. \quad (2.63)$$

D'après (2.14), nous avons $\mathcal{J}(\omega_1, \omega_2, \omega_3) \leq \mathcal{I}(\omega_1, \omega_2, \omega_3)$ et par conséquent $\bar{\mathcal{J}} \leq \bar{\mathcal{I}}$. De plus, lorsque les sorties du système global sont égales aux sources à une permutation et un filtrage passe-tout près, les termes dans (2.61) et (2.14) pour lesquels l'un des indices i, l_1, l_2 diffère des deux autres, s'annulent. Par conséquent, $\bar{\mathcal{J}}$ atteint la valeur maximale de $\bar{\mathcal{I}}$ et dans ce cas : $\bar{\mathcal{J}} = \bar{\mathcal{I}} = \bar{\mathcal{I}}_{\max}$. En invoquant la propriété 1, nous pouvons alors conclure :

Proposition 5 *Sous l'hypothèse H.5, $\bar{\mathcal{J}}$ est un contraste.*

En utilisant les mêmes arguments, on montre que :

Proposition 6 *Sous l'hypothèse H.6,*

$$\mathcal{J}_{\mathcal{E}} := \int_0^{2\pi} \int_0^{2\pi} \left(\int_{\mathcal{E}(\omega, \nu)} \mathcal{J}\left(\omega, \frac{\nu + \alpha}{2}, \frac{\nu - \alpha}{2}\right) d\alpha \right) d\omega d\nu \quad (2.64)$$

est un contraste.

Le choix des coefficients de pondération $\eta_{i l_1 l_2}$ offre toute une gamme de contrastes possibles. Cependant, il semble judicieux de fixer à zéro certains des coefficients de pondération : les critères obtenus ont alors moins de termes et présentent donc une expression plus simple que celle de $\bar{\mathcal{I}}$ ou $\mathcal{I}_{\mathcal{E}}$ correspondante. Comme premier cas particulier, on peut choisir :

$$\eta_{i l_1 l_2} = \begin{cases} 1 & \text{si } l_1 = l_2 \\ 0 & \text{sinon.} \end{cases} \quad (2.65)$$

En appliquant l'égalité de Parseval, on a alors

$$\frac{1}{16\pi^3} \mathcal{J}_{[-2\pi, 2\pi[} = \sum_{i l} \sum_{\tau_1, \tau_2, \tau_3} |\text{Cum}\{y_i(n), y_i^*(n + \tau_1), y_l(n + \tau_2), y_l^*(n + \tau_3)\}|^2. \quad (2.66)$$

Cette expression nous rappelle un contraste introduit dans [67] pour les mélanges instantanés et dans [90] pour le cas convolutif i.i.d. Un autre choix possible pour les coefficients de pondération est

$$\eta_{i l_1 l_2} = \begin{cases} 1 & \text{si } l_1 = i \\ 0 & \text{sinon.} \end{cases} \quad (2.67)$$

Le contraste s'écrit alors :

$$\frac{1}{16\pi^3} \mathcal{J}_{[-2\pi, 2\pi[} = \sum_{il} \sum_{\tau_1, \tau_2, \tau_3} |\text{Cum}\{y_i(n), y_i^*(n + \tau_1), y_i(n + \tau_2), y_i^*(n + \tau_3)\}|^2, \quad (2.68)$$

qui peut être vu comme une extension du contraste dans [58] au cas convolutif. Enfin, il nous faut mentionner la possibilité suivante :

$$\eta_{il_1 l_2} = \begin{cases} 1 & \text{si } l_1 = l_2 = i \\ 0 & \text{sinon,} \end{cases} \quad (2.69)$$

qui mène à l'expression la plus simple du contraste :

$$\frac{1}{16\pi^3} \mathcal{J}_{[-2\pi, 2\pi[} = \sum_i \sum_{\tau_1, \tau_2, \tau_3} |\text{Cum}\{y_i(n), y_i^*(n + \tau_1), y_i(n + \tau_2), y_i^*(n + \tau_3)\}|^2. \quad (2.70)$$

Cette équation est la plus simple expression connue d'un contraste temporel, basé sur des cumulants d'ordre quatre, et valable de plus dans le cas général de signaux non i.i.d. (i.e. sans aucune hypothèse sur les sources en dehors de leur indépendance mutuelle). Elle a déjà également été obtenue par Liu et Inouye par une approche entièrement temporelle [60]. En raison de la somme infinie sur τ_1, τ_2, τ_3 , ce contraste demeure cependant délicat à utiliser, tout comme l'étaient les contrastes fréquentiels.

2.2.2-b) Simplification maximale des contrastes

En nous appuyant sur certaines propriétés des sources, nous proposons d'exploiter les contrastes précédents pour les simplifier encore plus : nous cherchons à réduire autant que possible le nombre de termes dans les sommes des équations (2.66), (2.68) ou (2.70).

En évoquant à nouveau la propriété 1, on remarque aisément que, dans les expressions (2.66), (2.68) ou (2.70), nous pouvons faire disparaître tous les termes qui s'annulent lorsque la séparation est effectuée. Nous pouvons ainsi affirmer la proposition suivante :

Proposition 7 *Si le support des cumulants d'ordre quatre des sources est contenu dans un support \mathcal{S} , on obtient toujours un contraste en remplaçant la somme infinie sur τ_1, τ_2, τ_3 dans (2.66), (2.68) ou (2.70) par une somme sur \mathcal{S} .*

Il faut souligner ici que nous considérons le support des cumulants des sources, et non pas celui des sorties globales qui est en général plus large et dépend de l'ordre du système de mélange. Le domaine de sommation dépend ainsi des sources considérées, et pour une large classe de sources, il peut être supposé fini.

En particulier, si $s_i(n)$ est i.i.d., seul l'autocumulant sans aucun retard est non nul. On peut ainsi se limiter à $\tau_1 = \tau_2 = \tau_3$ dans la somme (2.70) et nous obtenons le contraste déjà trouvé dans [24] :

$$\zeta_0 = \sum_i |\text{Cum}\{y_i(n), y_i^*(n), y_i(n), y_i^*(n)\}|^2. \quad (2.71)$$

Dans la partie 2.3, d'autres exemples de signaux seront donnés, qui présentent un support fini du champ d'autocumulants. En particulier, les sources de type multiplicatif

telles que celles décrites dans le paragraphe 2.3.2-a), vérifient :

$$\begin{aligned} \text{Cum}\{s_i(n), s_i^*(n + \tau_1), s_i(n + \tau_2), s_i^*(n + \tau_3)\} \neq 0 \\ \Downarrow \\ \begin{cases} (\tau_1 = 0, \tau_2 = \tau_3 \text{ et } |\tau_2| \leq M) \\ \text{ou :} \\ (\tau_2 = 0, \tau_1 = \tau_3 \text{ et } |\tau_1| \leq M) \\ \text{ou :} \\ (\tau_3 = 0, \tau_1 = \tau_2 \text{ et } |\tau_1| \leq M) \end{cases} \end{aligned} \quad (2.72)$$

On obtient dans ce cas le contraste suivant :

$$\zeta_M = \sum_i \left[|\text{Cum}\{y_i(n), y_i^*(n), y_i(n), y_i^*(n)\}|^2 + 6 \sum_{\tau=1}^M |\text{Cum}\{y_i(n), y_i^*(n), y_i(n + \tau), y_i^*(n + \tau)\}|^2 \right]. \quad (2.73)$$

Il a été observé au cours des simulations que, pour un certain nombre de sources non linéaires, le critère ζ_0 semblait mener à la séparation. En conséquence, une question que l'on peut naturellement se poser à ce stade est de savoir si ζ_0 est également un contraste dans le cas non i.i.d. Comme on peut s'y attendre, la réponse est en général négative :

Proposition 8 *Le critère ζ_0 n'est pas nécessairement un contraste pour la séparation aveugle de sources non i.i.d.*

Ce fait important, qui souligne l'intérêt des contrastes (2.73) est prouvé par le contre-exemple ci-après :

Preuve: Considérons des sources multiplicatives telles que celles qui seront décrites au paragraphe 2.3.2-b) et telles que $\epsilon < 1/4$. Ces sources vérifient :

$$\text{Cum}\{s(n), s(n), s(n), s(n)\} = \epsilon \quad (2.74)$$

$$\forall p \neq 0 \quad \text{Cum}\{s(n), s(n), s(n + p), s(n + p)\} = \begin{cases} \frac{1}{4} \text{ si } p \in \{-1, 1\} \\ 0 \text{ si } |p| > 1. \end{cases} \quad (2.75)$$

Supposons que la matrice de réponse fréquentielle du filtre global soit donnée par :

$$\mathbf{G}(\omega) = \frac{1}{2} \begin{pmatrix} 1 - e^{-i\omega} & 1 + e^{-i\omega} \\ -1 - e^{-i\omega} & -1 + e^{-i\omega} \end{pmatrix}. \quad (2.76)$$

En utilisant la multilinéarité des cumulants et l'indépendance des sources, on obtient aisément les autocumulants sans retards :

$$\text{Cum}\{y_1(n), y_1^*(n), y_1(n), y_1^*(n)\} = \text{Cum}\{y_2(n), y_2^*(n), y_2(n), y_2^*(n)\} = \frac{4\epsilon + 3}{16} \quad (2.77)$$

et par suite, nous avons la valeur du contraste pour le filtre \mathbf{G} donné ci-dessus :

$$\zeta_0 = 2 \left(\frac{4\epsilon + 3}{16} \right)^2 \quad (2.78)$$

Pour $\epsilon < \frac{1}{4}$, cette expression est plus grande que $2\epsilon^2$, qui est la plus grande valeur de ζ_0 pour un filtre global séparant. Ainsi, dans ce cas, le maximum de ζ_0 ne correspond pas à une solution séparante et ce ne peut donc pas être un contraste. ■

Nous allons maintenant introduire quelques signaux permettant d'appliquer les idées qui viennent d'être exposées dans cette partie.

2.3 Exemples de signaux non linéaires

Cette partie présente quelques signaux non linéaires qui ont été utilisés dans nos simulations. Conformément aux hypothèses de ce chapitre, les signaux en questions sont décorrélés.

Les domaines de l'économétrie et de la finance ont été rapidement confrontés à des phénomènes que les approches linéaires ne peuvent expliquer, tels que la persistance d'une forte variance des prix, ... Ils ont donc éprouvé la nécessité de considérer des modèles non linéaires et offrent ainsi un grand nombre d'exemples de signaux non linéaires, dont une classe importante est constituée par les processus conditionnellement gaussiens. Nous introduisons très brièvement la notion de volatilité avant de définir les modèles dits de volatilité stochastiques. Un signal inspiré du modèle précédent et construit de manière *ad hoc* pour servir de contre-exemple est présenté. Enfin, nous décrivons les modèles ARCH. Les propriétés utiles de chacun de ces processus sont rappelées à chaque fois.

2.3.1 Signaux de volatilité stochastique

Les données financières ou économiques présentent fréquemment des corrélations dans leurs amplitudes. Ainsi, si $(s(n))_{n \in \mathbb{Z}}$ est la grandeur étudiée, l'analyse empirique de certaines données financières montre que $s(n)^2$ et $s(n+m)^2$ (où $m \in \mathbb{N}^*$) sont corrélées. Ce fait se traduit par un phénomène de regroupement des valeurs prises par série, en des périodes de grande ou de faible agitation. Ce phénomène peut être pris en compte au travers de modèles de type «multiplicatifs» qui s'écrivent comme le produit de deux autres signaux :

$$\forall n \in \mathbb{Z} \quad s(n) = a(n)\xi(n) \quad (2.79)$$

Le cas le plus simple, qui sera celui que nous considérerons, est celui où $(\xi(n))_{n \in \mathbb{Z}}$ est i.i.d. centré et de variance 1. $(\xi(n))_{n \in \mathbb{Z}}$ est de plus indépendant de $(a(n))_{n \in \mathbb{Z}}$. Plusieurs conséquences découlent de ces hypothèses :

- $(s(n))_{n \in \mathbb{Z}}$ est décorrélé et de moyenne nulle.
- La variance de $(s(n))_{n \in \mathbb{Z}}$ est égale à celle de $(a(n))_{n \in \mathbb{Z}}$.
- dans le cas où $(\xi(n))_{n \in \mathbb{Z}}$ est de plus gaussien, la loi conditionnelle de $s(n)$ connaissant $a(n)$ est une loi normale centrée de variance $a(n)^2$.

Ainsi, la non linéarité de $(s(n))_{n \in \mathbb{Z}}$ est liée à la dépendance temporelle introduite par le processus $(a(n))_{n \in \mathbb{Z}}$ qui contrôle la variance conditionnelle de $(s(n))_{n \in \mathbb{Z}}$. Dans la littérature financière, $a(n)$ est appelée *volatilité*² : elle constitue une mesure de la variabilité d'une grandeur et fournit de ce fait une indication sur son risque boursier.

La suite $(a(n))_{n \in \mathbb{Z}}$ peut être une fonction du temps, ou une variable aléatoire. Le modèle dit de volatilité stochastique est celui où $a(n)$ s'écrit $a(n) = e^{\frac{\Delta(n)}{2}}$ et où $\Delta(n)$ est un modèle AR engendré par un bruit indépendant de $(\xi(n))_{n \in \mathbb{Z}}$ [86].

Toutefois, pour nos applications, nous considérerons uniquement le cas où $(a(n))_{n \in \mathbb{Z}}$ est un modèle MA (paragraphe 2.3.2) ainsi que le cas des signaux ARCH (paragraphe 2.3.3). Auparavant, nous examinons l'allure du champs de cumulants d'une source donnée par (2.79).

²Il existe en réalité plusieurs notions de volatilité (implicite, historique, instantanée, ...) correspondant à des définitions distinctes. La volatilité peut parfois être simplement assimilée à l'écart-type des observations effectuées (voir [86] pour une présentation détaillée ou [72] pour une simple introductions aux liens entre marchés financiers et mathématiques).

Champ de cumulants Si l'on note $\sigma_a^2 := E\{a(n)^2\}$, il vient en utilisant l'indépendance de $(a(n))_{n \in \mathbb{Z}}$ et de $(\xi(n))_{n \in \mathbb{Z}}$:

$$\begin{aligned} \kappa_s^4(p_1, p_2, p_3) &= M_s^4(p_1, p_2, p_3) - \sigma_a^4(\delta_{p_1} \delta_{p_2-p_3} + \delta_{p_2} \delta_{p_1-p_3} + \delta_{p_3} \delta_{p_1-p_2}) \\ &= M_a^4(p_1, p_2, p_3) M_\xi^4(p_1, p_2, p_3) - \sigma_a^4(\delta_{p_1} \delta_{p_2-p_3} + \delta_{p_2} \delta_{p_1-p_3} + \delta_{p_3} \delta_{p_1-p_2}). \end{aligned} \quad (2.80)$$

où l'on a noté

$$\kappa_s^4(p_1, p_2, p_3) := \text{Cum}\{s(n), s^*(n+p_1), s(n+p_2), s^*(n+p_3)\} \quad \text{et} \quad (2.81)$$

$$M_s^4(p_1, p_2, p_3) := E\{s(n)s^*(n+p_1)s(n+p_2)s^*(n+p_3)\}. \quad (2.82)$$

Or, $(\xi(n))_{n \in \mathbb{Z}}$ étant i.i.d., $M_\xi^4(p_1, p_2, p_3)$ est non nul uniquement quand (p_1, p_2, p_3) vérifie l'une des conditions :

$$\left\{ \begin{array}{l} (p_1 = 0 \text{ et } p_2 = p_3) \\ \text{ou :} \\ (p_2 = 0 \text{ et } p_1 = p_3) \\ \text{ou :} \\ (p_3 = 0 \text{ et } p_1 = p_2) \end{array} \right. \quad (2.83)$$

L'ensemble des triplets (p_1, p_2, p_3) satisfaisant (2.83) prend la forme de trois droites concourantes en l'origine ; nous noterons \mathcal{S}_* l'ensemble des tels triplets. Il apparaît donc que $\kappa_s^4(p_1, p_2, p_3)$ est nul en dehors de l'ensemble \mathcal{S}_* . Compte tenu de l'invariance des cumulants par permutation, il est alors suffisant de considérer les cumulants pour les triplets du type $(0, p, p)$ avec $p \in \mathbb{N}$.

2.3.2 Cas particuliers où $a(n)$ est MA

Nous supposons maintenant que $(a(n))_{n \in \mathbb{Z}}$ est un modèle MA :

$$a(n) = \sum_{i=0}^M b(n)\varepsilon(n-i), \quad \text{où } (\varepsilon(n))_{n \in \mathbb{Z}} \text{ est i.i.d.} \quad (2.84)$$

Ce choix a été dicté par le souhait d'obtenir un modèle dont le champ de cumulants soit de support borné. Par ailleurs, cela nous a permis de construire un exemple de signal dont l'autocumulant d'ordre quatre en $(0, 0, 0)$ est nul, sans pour autant que le champs d'autocumulants soit nul (ni que le signal soit gaussien).

2.3.2-a) Support fini des cumulants

Compte tenu de (2.84), $a(n)$ et $a(n-p)$ sont indépendants dès que $|p| > M$. Par conséquent :

$$\begin{aligned} \forall p, |p| > M \quad M_a^4(0, p, p) &= E\{a(n)^2 a(n-p)^2\} \\ &= E\{a(n)^2\} E\{a(n-p)^2\} = \sigma_a^4 \end{aligned} \quad (2.85)$$

Comme par ailleurs :

$$\begin{aligned} \forall p \neq 0 \quad \kappa_s^4(0, p, p) &= M_a^4(0, p, p) M_\xi^4(0, p, p) - \sigma_a^4 \\ &= M_a^4(0, p, p) E\{\xi(n)^2\} E\{\xi(n-p)^2\} - \sigma_a^4 \\ &= M_a^4(0, p, p) - \sigma_a^4 \end{aligned} \quad (2.86)$$

on en déduit que $\kappa_s^4(0, p, p) = 0$ si $|p| > M$. Par suite, les cumulants de $(s(n))_{n \in \mathbb{Z}}$ sont non nuls uniquement sur $\mathcal{S}_* \cap \{(p_1, p_2, p_3) \mid |p_1| \leq M, |p_2| \leq M, |p_3| \leq M\}$ et seul un nombre fini de cumulants de $(s(n))_{n \in \mathbb{Z}}$ sont donc non nuls.

2.3.2-b) Une source de cumulants d'ordre quatre nul en (0,0,0)

Nous reprenons le modèle multiplicatif de l'équation (2.79) et considérons le cas où $(a(n))_{n \in \mathbb{Z}}$ est un modèle MA d'ordre 1 donné par :

$$a(n) = \varepsilon(n) + \varepsilon(n-1) \quad (2.87)$$

où $\varepsilon(n)$ est un processus i.i.d. prenant ses valeurs dans $\{-1, 0, +1\}$ avec des probabilités respectives $1/4, 1/2$ et $1/4$. $\xi(n)$ est choisi comme étant un processus i.i.d. prenant avec une même probabilité $1/4$ ses valeurs dans l'ensemble :

$$\left\{ \pm \sqrt{1 + \sqrt{\frac{2(3+\epsilon)}{5}}} - 1; \pm \sqrt{1 - \sqrt{\frac{2(3+\epsilon)}{5}}} - 1 \right\}$$

où l'on suppose que $\epsilon \in]-1/2, 2]$.

On peut aisément vérifier que $a(n)$ et $\xi(n)$ sont de variance unité et que $E\{a(n)^4\} = 5/2$ et $E\{\xi(n)^4\} = \frac{2(3+\epsilon)}{5}$. Ainsi :

$$\text{Cum}\{s(n), s(n), s(n), s(n)\} = E\{a(n)^4\}E\{\xi(n)^4\} - 3 = \epsilon \quad (2.88)$$

et pour $p \neq 0$, on obtient aisément :

$$\text{Cum}\{s(n), s(n), s(n+p), s(n+p)\} = \begin{cases} \frac{1}{4} & \text{si } p \in \{-1, 1\} \\ 0 & \text{si } |p| > 1. \end{cases} \quad (2.89)$$

Ainsi, nous avons obtenu un processus de champs de cumulants d'ordre quatre non nul et ayant une valeur arbitrairement proche de zéro (et éventuellement nulle) de son cumulants d'ordre quatre en (0,0,0).

2.3.3 Signaux ARCH

2.3.3-a) Définition

Les modèles auto-régressifs à hétéroscédasticité conditionnelle (ARCH en anglais) ont été introduits par Engle en 1982 [34]. On pourra consulter [40, 86] pour leur étude plus précise et [11] pour leurs applications possibles en finance.

Les signaux ARCH correspondent à un modèle de signaux multiplicatifs (2.79), mais dans lequel n'intervient qu'une seule source aléatoire :

Définition 4 On dit que le processus $(s(n))_{n \in \mathbb{Z}}$ suit un modèle ARCH(M) lorsqu'il est défini par :

$$\forall n \in \mathbb{Z} \quad s(n) = a(n)\xi(n) \quad (2.90)$$

où $(\xi(n))_{n \in \mathbb{Z}}$ est i.i.d., centré, gaussien et de variance 1 et $(a(n))_{n \in \mathbb{Z}}$ est positif et donné par l'équation suivante dans laquelle $(\alpha_i)_{0 \leq i \leq M}$ sont des coefficients réels :

$$\forall n \in \mathbb{Z} \quad a(n)^2 = \alpha_0 + \sum_{i=1}^M \alpha_i s(n-i)^2 \quad (2.91)$$

On constate dans ce modèle que la variance est une fonction des valeurs passées du processus. L'idée sous-jacente à un tel modèle est que les chocs aléatoires peuvent avoir une plus forte probabilité à certains moments qu'à d'autres. Sur les données financières, il permet d'expliquer les regroupements de volatilité, *i.e.* la persistance de fortes (ou faibles) variations de la grandeur étudiée sur plusieurs périodes.

2.3.3-b) Etude du cas particulier d'un processus ARCH(1)

Dans le cas d'un processus ARCH(1), une étude directe permet d'obtenir rapidement quelques propriétés. Nous les présentons ici avant de citer les théorèmes généraux dans le paragraphe suivant.

Dans le cas ARCH(1), l'équation (2.91) devient :

$$\forall n \in \mathbb{Z} \quad a(n)^2 = \alpha_0 + \alpha_1 s(n-1)^2 \quad (2.92)$$

et on en déduit immédiatement que le processus $s(n) = a(n)\xi(n)$ vérifie :

$$\mathbb{E}\{s(n)^2\} = \alpha_0 + \alpha_1 \mathbb{E}\{s(n-1)^2\}. \quad (2.93)$$

Si $\alpha_1 < 1$, l'équation de récurrence (2.93) admet l'unique solution stationnaire :

$$\mathbb{E}\{s(n)^2\} = \frac{\alpha_0}{1 - \alpha_1} \quad (2.94)$$

De plus, on a :

$$\mathbb{E}\{s(n)^4\} = \mathbb{E}\{a(n)^4\} \mathbb{E}\{\xi(n)^4\} \quad (2.95)$$

et $\xi(n)$ étant gaussien :

$$\mathbb{E}\{s(n)^4\} = 3\mathbb{E}\{a(n)^4\} = 3\mathbb{E}\{(\alpha_0 + \alpha_1 s(n-1)^2)^2\} \quad (2.96)$$

$$= 3(\alpha_0^2 + 2\alpha_0\alpha_1 \mathbb{E}\{s(n-1)^2\} + \alpha_1^2 \mathbb{E}\{s(n-1)^4\}) \quad (2.97)$$

$$= 3 \frac{\alpha_0^2(1 + \alpha_1)}{1 - \alpha_1} + 3\alpha_1^2 \mathbb{E}\{s(n-1)^4\}. \quad (2.98)$$

En supposant à la fois $0 < \alpha_1 < 1$ et $3\alpha_1^2 < 1$ (ce qui se résume à la condition $0 < \alpha_1 < \frac{1}{\sqrt{3}}$), nous pouvons donc obtenir la solution stationnaire à l'ordre quatre :

$$\mathbb{E}\{s(n)^4\} = \frac{3\alpha_0^2(1 + \alpha_1)}{(1 - \alpha_1)(1 - 3\alpha_1^2)}. \quad (2.99)$$

De plus, par les équations (2.94) et (2.99), nous obtenons le cumulants d'ordre quatre du processus ARCH(1) $s(n)$:

$$\text{Cum}\{s(n), s(n), s(n), s(n)\} = \frac{\mathbb{E}\{s(n)^4\}}{\mathbb{E}\{s(n)^2\}^2} - 3 = \frac{6\alpha_1^2}{1 - 3\alpha_1^2}. \quad (2.100)$$

On constate en particulier que le coefficient de kurtosis est positif ce qui implique que la distribution de $s(n)$ est leptokurtique (elle a une queue épaisse).

2.3.3-c) Conditions d'existence d'une solution stationnaire et des moments d'ordre quatre

Les théorèmes suivants sont adaptés de [40] et permettent d'assurer respectivement l'existence d'une solution stationnaire et l'existence des moments d'ordre quatre pour le processus défini par (2.90)- (2.91).

Théorème 1 *Sous les conditions de régularité*

$$\alpha_0 > 0, \alpha_i \geq 0, i = 1, \dots, M \text{ et } \sum_{i=1}^M \alpha_i < 1, \quad (2.101)$$

il existe une solution stationnaire au sens large p.s. pour le processus ARCH(M) décrit par les équations (2.90) et (2.91). Ce processus est non corrélé, de moyenne nulle et de variance $\frac{\alpha_0}{1 - \sum_{i=1}^M \alpha_i}$.

Théorème 2 *Pour un processus ARCH(M) défini par les équations (2.90) et (2.91) avec les paramètres $\alpha_i \geq 0, i = 1, \dots, M$ et $\sum_{i=1}^M \alpha_i < 1$, une condition nécessaire et suffisante pour l'existence de moments d'ordre quatre est :*

$$3\alpha^T(\mathbf{Id} - \psi^{-1})\alpha < 1 \quad (2.102)$$

où $\alpha = (\alpha_1, \dots, \alpha_M)^T$ et ψ est la matrice $\psi = (\psi_{ij})_{(i,j) \in \{1, \dots, M\}^2}$ avec :

$$\psi_{ij} = \alpha_{i+j} + \alpha_{i-j} \quad (2.103)$$

où, dans l'équation (2.103) uniquement, on a adopté la convention $\alpha_p = 0$ si $p \leq 0$ ou $p \geq M$.

Remarque 3: On retrouve les résultats du paragraphe précédent par application des théorèmes ci-dessus. Un processus ARCH(1) défini par l'équation (2.92) :

- admet une solution stationnaire si $\alpha_0 > 0$ et $\alpha_1 < 1$.
- a des moments d'ordre quatre finis si en outre $3\alpha_1^2 < 1$.

2.4 Mise en œuvre de la séparation

2.4.1 Paramétrisation par une structure en treillis sans perte

L'ensemble des résultats obtenus dans la partie 2.2 sont valables à condition d'avoir effectué un préblanchiment et de travailler avec un filtre séparateur para-unitaire. Nous donnons ici une paramétrisation des filtres para-unitaires qui permet de se ramener à un problème d'optimisation sans contraintes. Les techniques utilisées pour la maximisation sont alors brièvement exposées.

2.4.1-a) Paramétrisation

Les filtres para-unitaires peuvent être paramétrisés par une structure dite en treillis sans perte³. On consultera [96] pour plus de précisions. La paramétrisation est la suivante pour un filtre $\mathbf{W}[z]$:

³Il arrive que les filtres para-unitaires soient aussi appelés filtres sans perte.

$$\mathbf{W}[z] = \left(\prod_{k=1}^{N-1} \mathbf{R}_{kN}(\theta_{kN}^L) \right) \mathbf{D}[z] \dots \left(\prod_{k=1}^{N-1} \mathbf{R}_{kN}(\theta_{kN}^2) \right) \mathbf{D}[z] \left(\prod_{k=1}^{N-1} \left(\prod_{j=k+1}^N \mathbf{R}_{kj}(\theta_{kj}^1) \right) \right). \quad (2.104)$$

Dans cette expression, les angles de rotation $\theta_{k,j}^l$ peuvent être choisis dans $[-\pi/2, \pi/2]$ et l'élément (n, m) de chaque matrice de rotation est défini par :

$$\forall \varphi \in [0, \pi/2], \quad [\mathbf{R}_{kj}(\varphi)]_{n,m} = \begin{cases} \cos(\varphi) & \text{si } n = m = k \text{ ou } n = m = j \\ -\sin(\varphi) & \text{si } (n, m) = (k, j) \\ \sin(\varphi) & \text{si } (n, m) = (j, k) \\ 1 & \text{si } n = m \notin \{j, k\} \\ 0 & \text{sinon} \end{cases} \quad (2.105)$$

tandis que $\mathbf{D}[z] = \text{Diag}(1, \dots, 1, z^{-1})$ correspond à un retard sur la dernière composante. Afin de fournir une paramétrisation complète de l'ensemble des filtres para-unitaires de réponse impulsionnelle finie, un terme multiplicatif $\text{Diag}(\pm 1, \dots, \pm 1, \pm z^{-d})$, $d \in \mathbb{Z}$, doit être inclus dans (2.104).

La paramétrisation ci-dessus sera utilisée pour le filtre de séparation. Le filtre de mélange sera choisi également avec une telle paramétrisation. Ceci offre le double avantage de permettre l'étude de la séparation indépendamment de l'étape de préblanchiment (qui peut poser quelques difficultés dans la pratique), et de pouvoir connaître la longueur de la réponse impulsionnelle du filtre para-unitaire séparant —un filtre para-unitaire de réponse impulsionnelle finie est en effet inversible par un filtre para-unitaire dont la réponse impulsionnelle est de même longueur. Il résulte de ceci que le filtre global $\mathbf{G}[z]$, dont on sait qu'il est para-unitaire, sera en outre de réponse impulsionnelle finie dans notre mise en œuvre. Examinons la conséquence de ce fait sur l'ambiguïté de filtrage qui reste après séparation.

2.4.1-b) Une conséquence sur l'ambiguïté de filtrage

Nous avons déjà vu que le fait de travailler dans l'ensemble des filtres para-unitaires transformait l'ambiguïté de filtrage SISO résiduel en une ambiguïté de filtrage *passé-tout*. Le fait de travailler avec des filtres à réponse impulsionnelle finie (ce qui est le cas en utilisant la paramétrisation ci-dessus) entraîne une restriction supplémentaire.

Dans le cas où le filtre global $\mathbf{G}[z]$ est para-unitaire et de réponse impulsionnelle finie, l'ambiguïté de filtrage résiduelle après séparation se réduit à un retard et une phase et le filtre SISO correspondant est du type : $e^{i\varphi} z^{-d}$.

Preuve: Notons en effet $h[z]$ le filtre SISO résiduel à l'issue de la séparation. Comme ce dernier est passé-tout, il vérifie $|h[z]| = 1$ pour tout z du cercle unité. Ayant supposé la réponse impulsionnelle finie, il existe de plus $k \in \mathbb{Z}$ tel que $z^k h[z]$ soit un polynôme, qui vérifie $|z^k h[z]| = 1$ pour tout z du cercle unité. Il s'ensuit que le polynôme $z^k h[z]$ est nécessairement de la forme $e^{i\varphi} z^l$ avec $l \in \mathbb{N}$ et $\varphi \in \mathbb{R}$. On obtient donc le résultat précédemment annoncé avec $d = -l + k$. ■

Le fait souligné ici est d'importance pour les simulations : il nous sera en effet possible dans ce cas de calculer une erreur quadratique moyenne sur les sources reconstituées après suppression des seules ambiguïtés de phase et de retard.

2.4.2 Optimisation des contrastes

2.4.2-a) Algorithmes classiques : optimisation sans contrainte

Une fois donnée la paramétrisation (2.104) du filtre séparant, la maximisation du contraste se ramène à un problème classique d'optimisation sans contrainte (les angles de rotations peuvent être choisis dans $[-\pi/2, \pi/2]$, mais il n'y a nulle obligation). Une introduction aux différentes méthodes d'optimisation possible peut être trouvée dans de nombreux ouvrages ([63] par exemple); les méthodes d'optimisation qui ont été utilisées sont :

- descente successive selon les coordonnées : comme la paramétrisation choisie est fonction d'angles de Jacobi, cette méthode appartient à l'ensemble des méthodes dites de *Jacobi*.
- *gradient* : étant une méthode extrêmement simple, elle a été programmée en premier afin de constater la validité des résultats.
- *gradient conjugué* : elle offre une vitesse de convergence bien plus rapide que la méthode du gradient sans nécessiter le calcul du hessien de la fonction optimisée.
- *méthodes de quasi-Newton* : elles nécessitent le calcul du hessien de la fonction optimisée, mais pas son inversion. La toolbox de Matlab a offert la possibilité d'utiliser aisément ces méthodes dont la rapidité de convergence est nettement améliorée comparativement au gradient.

L'ensemble des méthodes évoquées ci-dessus ne permettent généralement qu'une optimisation locale du critère considéré. Or, comme expliqué dans la partie 2.5, nous avons constaté que les contrastes présentaient de nombreux maxima locaux. Afin de proposer un début de solution, nous avons envisagé l'utilisation d'algorithmes génétiques. Nous proposons une courte présentation de ces derniers dans le paragraphe suivant.

2.4.2-b) Algorithmes génétiques

Les algorithmes génétiques offrent une possibilité d'optimisation globale. Leur mise en œuvre ainsi que leur principe de base, que nous présentons ici, sont relativement simples. L'efficacité de telles méthodes est reconnue dans des cas où d'autres méthodes se révèlent inefficace. Une toolbox Matlab est disponible sur internet [1], accompagnée d'une introduction plus approfondie aux algorithmes génétiques, ainsi que d'une bibliographie.

Qu'est-ce qu'un algorithme génétique ? Un algorithme génétique est un « algorithme stochastique itératif » qui, à partir d'une population initiale, opère sur des ensembles de points codés à l'aide de trois opérateurs : croisement, mutation, sélection. Les deux premiers sont des opérateurs d'exploration de l'espace, tandis que le dernier fait évoluer la population vers les optima d'un problème. Cette méthode est directement inspirée de la théorie de Darwin sur l'évolution des espèces : en permettant aux seuls individus bien adaptés à l'environnement de se reproduire, la nature assure la pérennité de leurs meilleures caractéristiques, lesquelles se recombinent entre elles (chaque enfant reçoit de bonnes caractéristiques à la fois de son père et de sa mère) pour former à chaque génération de nouveaux individus toujours mieux adaptés à leur environnement.

Descriptif du fonctionnement d'un algorithme génétique Les algorithmes génétiques considèrent une population d'individus, qui représentent autant de solutions possibles au problème considéré. Chaque individu est représenté par ses *chromosomes* par le

biais d'un codage défini entre l'ensemble des solutions envisageables *a priori* et un ensemble de chaînes de caractères définies sur un certain alphabet. Dans un cas simple, il est par exemple possible de coder les solutions possibles sur un certain nombre de bits.

Une population initiale est créée, par exemple par tirage aléatoire, et est amenée à vivre. Les individus de cette population se reproduisent en suivant trois règles pour générer la population de la génération suivante :

croisement Etant donnés deux parents de la population actuelle, un croisement donne au chromosome fils des caractéristiques partagées de ses deux parents.

mutation Les mutations consistent à opérer des modifications aléatoires sur les gènes des différents chromosomes afin d'introduire une plus grande diversité et ainsi mieux explorer l'espace des solutions.

sélection Au sein de la population fille, obtenue par reproduction (croisement et mutation), sont sélectionnés les individus les mieux adaptés au problème posé au sens d'un certain critère d'adéquation à définir. Ce critère se définit en général comme une fonction de la grandeur à optimiser. Le résultat de la sélection donne alors la population de la génération suivante.

Le processus ainsi décrit est réitéré sur un certain nombre de générations. Un critère d'arrêt permet l'arrêt du processus, et la, ou les solution(s) retournée(s) est alors celle(s) correspondant à l'individu (ou aux individus) qui, au sein de la dernière génération, répond(ent) le mieux au problème posé.

Avantages des algorithmes génétiques Les algorithmes génétiques présentent un certain nombre de caractéristiques propres. On peut en particulier noter les points suivants :

- les algorithmes génétiques travaillent en parallèle sur une population de points et non pas sur un seul point.
- les algorithmes génétiques ne nécessitent aucune information sur la dérivée, ni de toute autre nature concernant la fonction à maximiser (convexité, continuité, ...).
- les algorithmes génétiques utilisent des règles de transitions aléatoires et non pas déterministes.
- les algorithmes génétiques travaillent sur un codage de l'ensemble des paramètres, pas sur les paramètres eux-mêmes (à moins de se placer dans le cas particulier de gènes réels).

Compte tenu de ces faits, les algorithmes génétiques sont adaptés à certains problèmes particuliers : on peut évoquer entre autres les problèmes de nature discrète par essence, ceux où l'ensemble des solutions envisageable est de grande dimension. Egalement dans le cas de problèmes à solution non unique ou d'optimisation multi-objectifs, les algorithmes génétiques se révèlent potentiellement intéressant pour trouver un ensemble de solutions Pareto-optimales⁴ En revanche, dans le cas de fonctions simples, convexes, dérivables, la performance des algorithmes génétiques est en général bien inférieure à celle des méthodes plus classiques.

D'après ces observations rapides, il semble donc que, dans les problèmes de séparation de sources, les algorithmes génétiques pourraient être utiles afin de résoudre la difficulté liée à l'existence de maxima locaux. En revanche, un algorithme classique, et même un simple gradient, devrait être bien plus efficace pour converger vers le point cherché une

⁴Dans le cadre de la maximisation multi-objectifs, une solution est dite Pareto-optimale lorsque toute autre solution permettant de maximiser plus l'un des critères, en diminue nécessairement un autre.

fois ramené dans le voisinage d'une solution globale. Nous avons donc proposé d'évaluer rapidement les résultats que peut fournir une telle méthode, parfois au prix d'un coût de calcul élevé. Le fait d'obtenir des résultats positifs inviterait à poursuivre l'effort dans la recherche de méthodes conjointes de séparation de sources.

2.5 Simulations

En plus de confirmer la validité des résultats théoriques aussi bien dans le cas i.i.d. que non i.i.d., le but de ces simulations a été de comparer les différents contrastes afin d'offrir des éléments de choix. Intuitivement, un contraste devrait être d'autant plus délicat à estimer qu'il prend en compte un plus grand nombre de cumulants avec des retards plus importants. On peut s'attendre dans ce cas à une moins bonne estimation statistique et à une moindre qualité de la séparation. Cependant, la prise en compte de retards supplémentaires est aussi susceptible d'apporter une information supplémentaire qui peut améliorer la séparation. Rappelons les différents contrastes utilisés dans ces simulations et apportons quelques précisions à leur sujet :

- contrastes fréquentiels :

$$\bar{\mathcal{I}}^{(K)} = \frac{4\pi^2}{K^2} \sum_{p=0}^{K-1} \sum_{q=0}^{K-1} \mathcal{I}\left(\nu_p, \frac{\nu_{q+r_{p,q}}}{2}, \frac{\nu_{q-r_{p,q}}}{2}\right) \quad (2.106)$$

avec $r_{p,q} = 0$ (noté I sur les figures)

$$\mathcal{I}_{\mathcal{E}^K}^{(K)} = \frac{8\pi^3}{K^3} \sum_{p=0}^{K-1} \sum_{q=0}^{K-1} \sum_{r \in \mathcal{E}_{p,q}^K} \mathcal{I}\left(\nu_p, \frac{\nu_{q+r}}{2}, \frac{\nu_{q-r}}{2}\right) \quad (2.107)$$

avec $\mathcal{E}_{p,q}^K = \{-K+1, \dots, K-1\}$ (noté I_{E^K} sur les figures)

Le nombre K de fréquences discrètes a été choisi égal à deux fois la longueur du filtre de mélange. Pour le calcul de ces critères, une estimée du trispectre a été obtenue par une technique du tri-périodogramme lissé. L'optimisation de tels contrastes a été réalisée par une procédure de type Jacobi.

- contrastes temporels :

$$\zeta_{AC,M} = \sum_i \sum_{\substack{|\tau_1| \leq M \\ |\tau_2| \leq M \\ |\tau_3| \leq M}} |\text{Cum}\{y_i(n), y_i^*(n + \tau_1), y_i(n + \tau_2), y_i^*(n + \tau_3)\}|^2 \quad (2.108)$$

$$\zeta_M = \sum_i \left[|\text{Cum}\{y_i(n), y_i^*(n), y_i(n), y_i^*(n)\}|^2 + 6 \sum_{\tau=1}^M |\text{Cum}\{y_i(n), y_i^*(n), y_i(n + \tau), y_i^*(n + \tau)\}|^2 \right] \quad (2.109)$$

$$\zeta_0 = \sum_i |\text{Cum}\{y_i(n), y_i^*(n), y_i(n), y_i^*(n)\}|^2 \quad (2.110)$$

Le contraste $\zeta_{AC,M}$ est la forme la plus générale, valable théoriquement pour tout type de signal lorsque $M \rightarrow +\infty$, tandis que ζ_M est valable pour les signaux multiplicatifs et que ζ_0 n'est théoriquement valable que dans le cas de signaux i.i.d. La méthode d'optimisation employée a été un gradient en ce qui concerne ces contrastes temporels.

Les signaux source considérés étaient tous centrés et de variance unité. Plus précisément, il s'agissait de :

- signaux i.i.d. : nous avons considéré des signaux réels de type PAM4 (rappelons qu'ils prennent avec une probabilité 1/4 l'une des quatre valeurs $\{\pm\frac{1}{\sqrt{5}}, \pm\frac{3}{\sqrt{5}}\}$) et des signaux distribués selon une loi uniforme. L'ensemble des critères considérés constituent des contrastes pour ces sources.
- des signaux de volatilité stochastique (voir équation (2.79)), avec $a(n)$ pris comme un modèle MA. Les sources de contre-exemple données dans le paragraphe 2.3.2-b) en sont un cas particulier particulièrement intéressant. Pour de telles sources, les critères fréquentiels, $\zeta_{AC,M}$ et ζ_M (avec M suffisamment grand, mais fini) sont des contrastes. En revanche, ζ_0 n'est pas un contraste.
- de signaux ARCH, pour lesquels en théorie, les mêmes contrastes que ci-dessus sont valables, avec cependant M infini.
- des signaux obtenus par sous-échantillonnage d'un facteur 2 d'une source linéaire, provenant elle-même du filtrage d'une source i.i.d. binaire par le filtre de réponse impulsionnelle : (0.1759 0.0350 0.3609 0.1890 -0.0545 0.3231 0.3087 0.4688 0.1204 -0.6041). Dans ce cas, les critères fréquentiels ainsi que $\zeta_{AC,M}$ avec $M \geq 5$ sont des contrastes ; tandis que ζ_M n'en est pas un quel que soit M .

2.5.1 Propriétés locales des contrastes

Afin d'isoler les différents problèmes susceptibles d'intervenir, nous avons utilisé, dans toute cette partie sur les propriétés locales, des algorithmes d'optimisation locale initialisés à des points proches de la vraie valeur des paramètres. Plus précisément, l'algorithme d'optimisation a été initialisé de façon systématique en prenant pour valeur initiale de chaque angle, la vraie valeur cherchée à laquelle a été ajoutée une variable aléatoire uniformément répartie dans $[-\pi/20, \pi/20]$: ceci représente un écart maximal de 10% de l'intervalle de variation des angles.

Après séparation, les seules indéterminations restantes (phase et retard d'après le paragraphe 2.4.1-b)) ont été éliminées par une méthode de corrélation. Ceci a permis de considérer l'erreur quadratique moyenne sur chacune des sources reconstituées afin de quantifier la qualité de la séparation.

Tous les résultats et valeurs proviennent du moyennage sur 100 réalisations des filtres et des sources, tirés aléatoirement par une méthode de Monte-Carlo.

Résultats et commentaires L'erreur quadratique moyenne de reconstruction des sources (EQM) est donnée en fonction du nombre d'échantillons et pour différents contrastes utilisés sur les figures 2.2, 2.3, 2.4 et 2.5. Le type de sources utilisées sont respectivement des sources i.i.d. PAM4, les sources du contre-exemple donné en 2.3.2-b), des sources ARCH(9) et des sources obtenues par sous-échantillonnage d'un processus linéaire. Le tableau 2.1 donne également la valeur de l'EQM pour les intervalles de confiance à 90%, pour des sources i.i.d. et du contre-exemple. La figure 2.6 donne la valeur de l'EQM moyenne en fonction de la longueur du filtre, tandis que la figure 2.7 permet de voir l'influence d'un bruit additif.

D'une façon générale, on constate qu'il est avantageux de considérer le contraste qui comporte le moins de cumulants possibles. En particulier, les contrastes fréquentiels, qui utilisent théoriquement le champ de cumulants complet, offrent de moins bons résultats que les contrastes temporels. Parmi les critères temporels possibles, les meilleurs résultats

sont obtenus en utilisant le critère comprenant le moins de cumulants, mais dont on a l'assurance théorique que c'est un contraste pour les sources en question : dans le cas de sources i.i.d. (Figure 2.2), il s'agit de ζ_0 tandis que dans le cas des sources du contre-exemple du paragraphe 2.3.2-b), ζ_1 offre les meilleurs résultats. De même sur la figure 2.5, on sait que $\zeta_{AC,5}$ est bien un contraste; il fournit des résultats corrects, tandis que les autres critères donnent des résultats de moindre qualité, en accord avec l'impossibilité que nous avons d'affirmer qu'il s'agit de contrastes. Enfin, si l'on regarde les figures 2.3 et 2.5, nous constatons clairement que ζ_0 n'est pas un contraste pour ce type de sources.

Notons finalement que $\mathcal{I}_{\mathcal{E}^K}^{(K)}$ mène à de meilleurs résultats que $\bar{\mathcal{I}}^{(K)}$, ce qui montre que le fait de prendre en compte plus de points fréquentiels améliore la séparation.

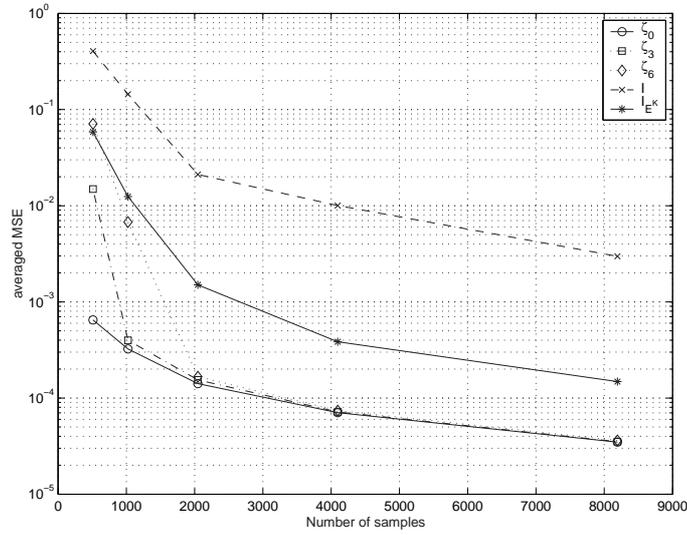


FIG. 2.2 – Comparaison des contrastes dans le cas de deux sources i.i.d. PAM4 (moyenne de l'EQM en fonction du nombre d'échantillons). Le filtre de mélange est de longueur 5.

		Nombre d'échantillons			
		2048	4096	8192	16384
sources PAM4 (ζ_0 , Fig. 2.2)	EQM moyenne	1.4540e-04	7.4188e-05	3.5643e-05	1.7824e-05
	inter. conf. 90%	2.8466e-04	1.4084e-04	6.5007e-05	3.3668e-05
sources non i.i.d. ($\epsilon = 0.5$, ζ_1 Fig. 2.3)	EQM moyenne	0.1245	0.0222	0.0067	0.0030
	inter. conf. 90%	0.4278	0.0375	0.0133	0.0056

TAB. 2.1 – EQM moyenne et pour des intervalles de confiance à 90% en fonction du nombre d'échantillons.

2.5.2 Propriétés globales des contrastes

2.5.2-a) Existence de maxima locaux

Afin de se faire une idée plus précise sur l'allure et le comportement global des fonctions de contraste, nous avons considéré le cas de 2 sources et d'un filtre de mélange de longueur 2. Le filtre séparant est alors également un filtre de longueur 2, et est paramétré par deux

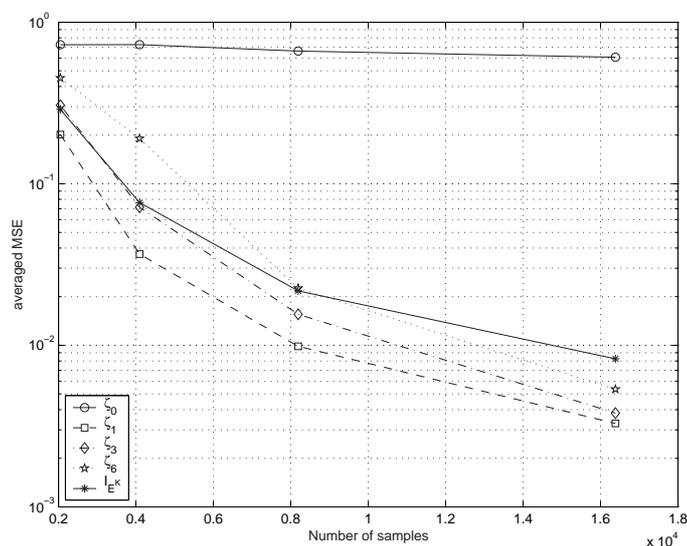
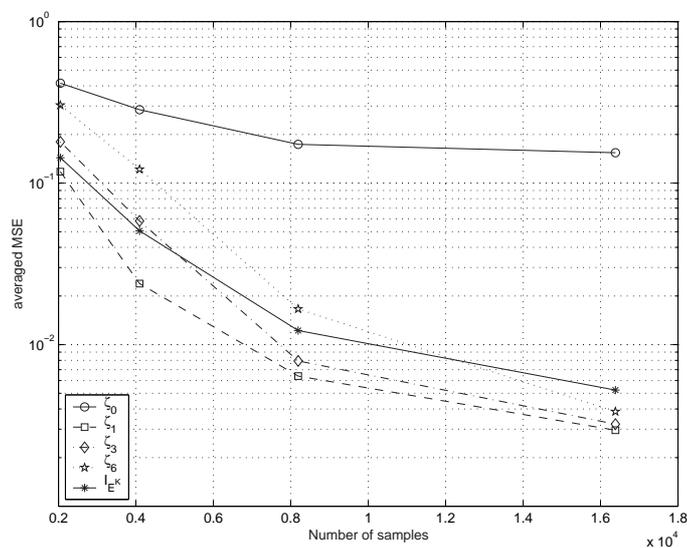
(a) $\epsilon = 0$ (b) $\epsilon = 0.5$

FIG. 2.3 – Comparaison des contrastes dans le cas de deux sources multiplicatives du contre-exemple en 2.3.2-b) et pour deux valeurs du paramètre ϵ (EQM moyenne en fonction du nombre d'échantillons). Le filtre de mélange est de longueur 5.

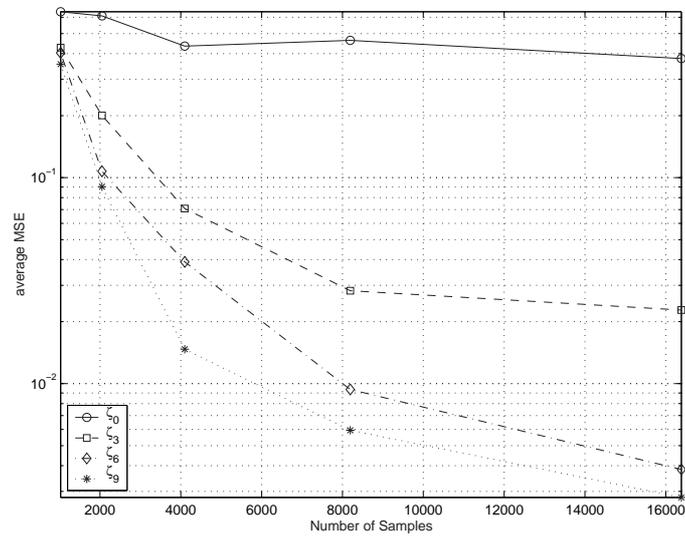


FIG. 2.4 – Comparaison des contrastes pour deux sources ARCH(9). Le filtre de mélange est de longueur 5.

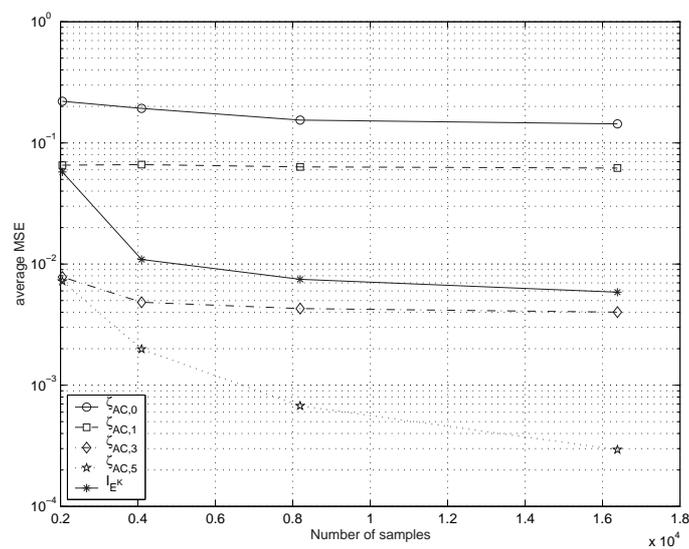


FIG. 2.5 – Comparaison des contrastes pour deux sources non linéaire obtenues par sous-échantillonnage d'un processus linéaire. Le filtre de mélange est de longueur 3.

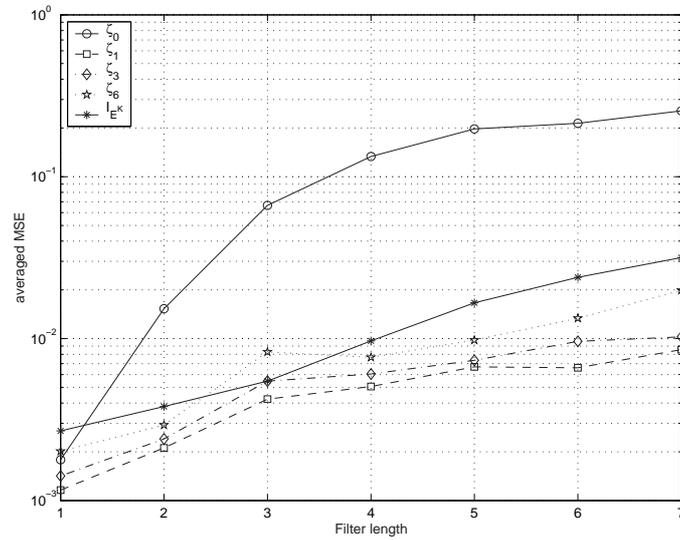


FIG. 2.6 – EQM moyenne en fonction de la longueur du filtre de mélange pour des signaux du contre-exemple en 2.3.2-b) avec $\epsilon = 0.5$. Le nombre de sources était de 2 et le nombre d'échantillons de 8192.

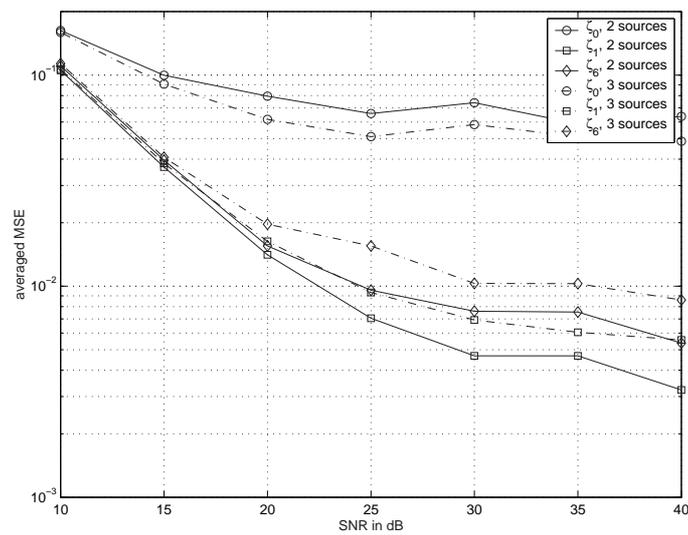


FIG. 2.7 – EQM moyenne en fonction du SNR pour des signaux du contre-exemple en 2.3.2-b) $\epsilon = 0.5$; Le nombre de sources était de 2 ou 3, le nombre d'échantillons de 8192 et la longueur du filtre de mélange de 5.

angles θ_1 et θ_2 . Nous avons tracé sur la figure 2.8 l'allure du contraste ζ_0 en fonction de ces deux angles. L'allure des contrastes ζ_M avec $M < 5$ est par ailleurs tout à fait semblable à celle de ζ_0 . Nous constatons sans équivoque l'existence de maxima locaux du contraste.

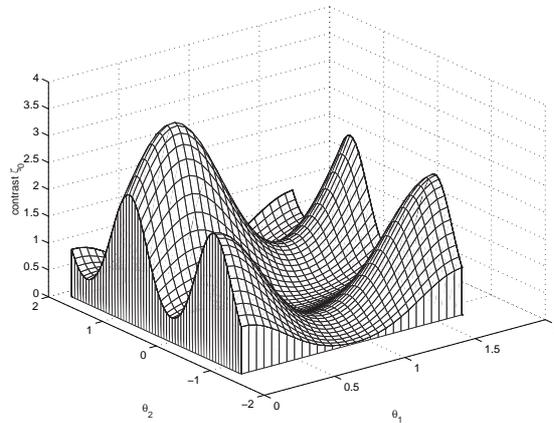


FIG. 2.8 – Représentation du contraste ζ_0 comme fonction des angles de rotation. Les sources sont PAM4 i.i.d. et ont 2048 échantillons.

2.5.2-b) Difficultés d'estimation

Afin de confirmer l'intérêt de considérer, dans la mesure du possible, un nombre minimal de cumulants, nous avons regardé, dans le cas de sources i.i.d., la variance de l'estimation des contrastes $\zeta_0, \zeta_2, \zeta_3$ et ζ_4 . Le résultat a été obtenu par moyennage sur 2500 réalisations de Monte-Carlo et est donné sur la figure 2.9. Nous constatons que plus le nombre de cumulants intervenant dans le contraste est élevé, plus la variance d'estimation est importante. De plus, les deux maxima de la figure 2.9 indiquent les valeurs des paramètres du séparateur pour lesquelles la variance d'estimation du contraste est la plus grande et correspondent précisément aux minima du contraste. Ainsi, les problèmes d'estimation sont susceptibles de faire apparaître des maxima locaux parasites supplémentaires.

2.5.2-c) Apport des algorithmes génétiques

Pour affirmer l'intérêt des approches conjointes et pour démontrer qu'elles sont utilisables malgré la présence de maxima locaux parasites, nous avons envisagé l'utilisation d'algorithmes génétiques. L'utilisation de telles méthodes se fait cependant au prix d'un coût de calcul accru. Une toolbox Matlab disponible gratuitement [1] a été utilisée. Les résultats sont donnés dans le tableau 2.2 : nous constatons que l'utilisation d'un algorithme génétique avant une phase d'optimisation classique augmente considérablement les chances de succès de la séparation.

2.6 Conclusion

Nous nous sommes intéressés dans ce chapitre à une approche conjointe de la séparation de sources. Après avoir présenté le problème, nous avons exposé comment il est possible,

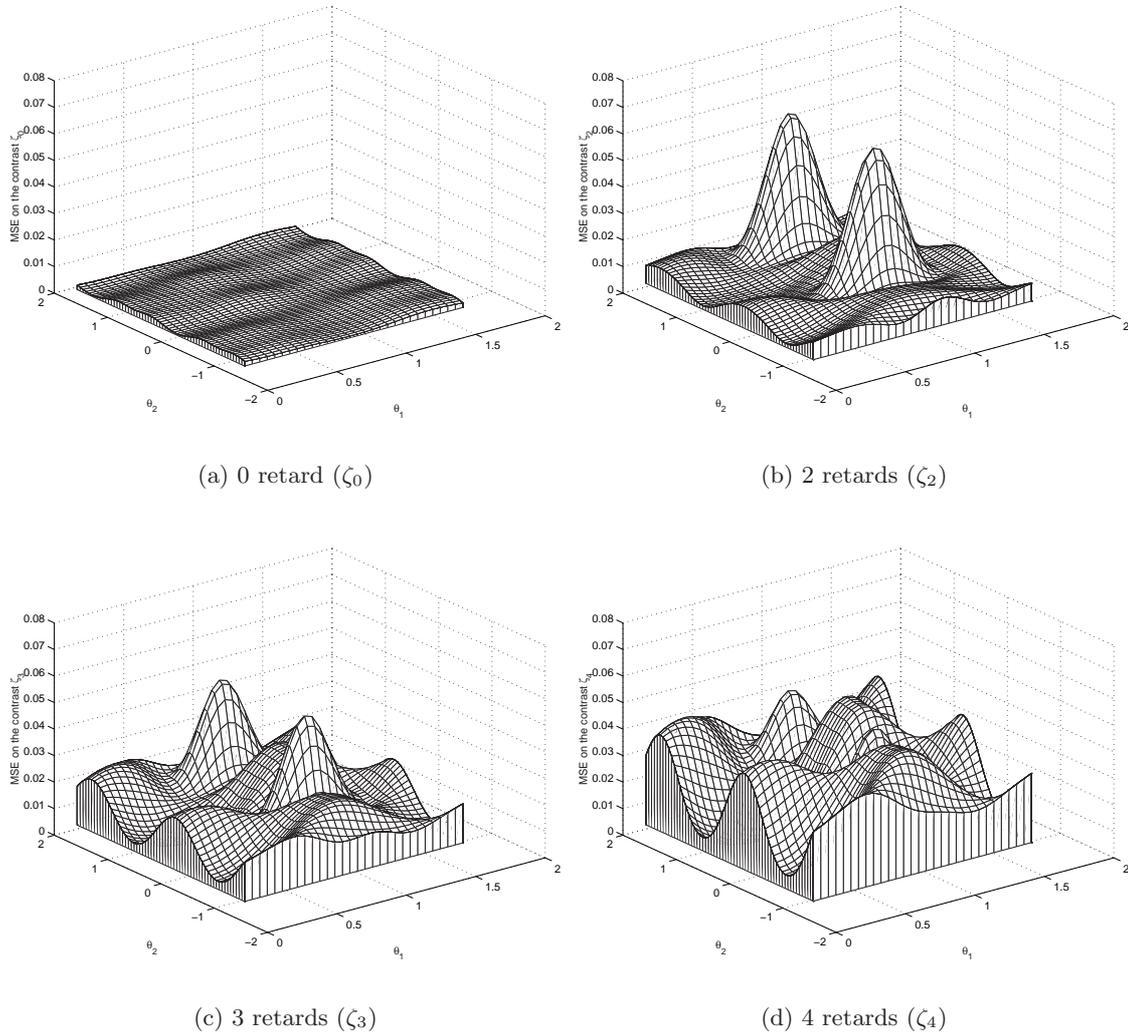


FIG. 2.9 – Erreur d'estimation des contrastes. Les sources sont PAM4 i.i.d. avec 2048 échantillons.

Méthode	% de succès
Gradient initialisé aléatoirement	51
Alg. gén. (gènes binaires) + gradient	92
Alg. gén. (gènes réels) + gradient	96

TAB. 2.2 – Pourcentage de succès de la séparation pour deux sources ARCH(5) mélangées par un filtre de longueur 3 (La séparation a été considérée comme réussie lorsque l'EQM moyenne sur les deux sources était inférieure à $1e-2$).

d'utiliser les statistiques d'ordre deux pour un préblanchiment et de se restreindre ensuite à la recherche d'un inverse dans l'ensemble des filtres para-unitaires.

En se basant sur une approche spectrale, nous avons pu démontrer la validité de contrastes fréquentiels. La relation de Parseval a permis de déterminer des contrastes à base de cumulants qui leur correspondent dans le domaine temporel. Le problème de l'estimation de ces contrastes se pose dans leur utilisation concrète : il convient en effet de limiter au maximum le nombre de termes estimés afin d'obtenir des performances convenables. Par la prise en compte d'informations sur la nature des sources et plus précisément leur champs de cumulants nous avons alors :

- mis en évidence une classe très générale de contrastes. Leur expression, simplifiée au maximum facilite leur utilisation.
- montré que le critère standard (2.71) n'est en général pas un contraste pour des processus non linéaires.

Les simulations ont confirmé l'intérêt de la famille de contrastes proposés. D'autre part, l'applicabilité de l'ensemble des résultats a été montrée, au prix d'une charge de calcul éventuellement importante. L'étape de préblanchiment n'a cependant pas été considérée ; cette dernière limite l'utilisation pratique de telles méthodes.

Chapitre 3

Approches séquentielles

3.1 Extension des contrastes MISO

3.1.1 Présentation des approches séquentielles

Dans ce chapitre, nous abordons le problème de la séparation de sources par des méthodes séquentielles. Les méthodes conjointes cherchent à déterminer *simultanément* l'ensemble des paramètres d'un filtre qui sépare *toutes* les sources, c'est-à-dire un filtre MIMO $\mathbf{W}[z]$ tel que le filtre $\mathbf{G}[z] := \mathbf{W}[z]\mathbf{M}[z]$ soit le produit d'une permutation et d'un filtre diagonal (voir notations du paragraphe 1.4.1-a)). Dans le chapitre précédent, ceci était réalisé par la maximisation d'un critère de contraste, précédée d'un préblanchiment.

Les approches séquentielles procèdent au contraire par étape et les sources sont extraites les unes après les autres. Le fondement de ces approches réside en la possibilité d'extraire l'une des N sources du mélange à partir du vecteur des observations. Les autres sources peuvent alors être extraites en itérant le processus d'extraction précédent sous certaines précautions : une approche classique, appelée déflation, consiste à retrancher la contribution de la source extraite au vecteur des observations et à recommencer à l'identique le processus d'extraction jusqu'à la dernière source. D'autres méthodes existent également, parmi lesquelles nous pouvons évoquer la possibilité d'ajouter d'un terme de pénalisation au critère utilisé [92]. Dans ce chapitre nous expliquerons comment il est également envisageable d'introduire des contraintes sur l'ensemble dans lequel s'effectue l'optimisation.

Par construction, les méthodes séquentielles reviennent à chercher et à construire le filtre séparant $\mathbf{W}[z]$ ligne par ligne : cette étape constitue le point central qui fonde ces méthodes. Nous serons donc amenés à étudier l'une quelconque de ces étapes et, dans ce but, nous noterons $\mathbf{w}[z]$ une des lignes du filtre $\mathbf{W}[z]$ ($\mathbf{w}[z]$ est un filtre MISO donc de taille $1 \times Q$). La ligne correspondante du filtre global $\mathbf{G}[z]$ est de taille $1 \times N$ et notée $\mathbf{g}[z] := \mathbf{w}[z]\mathbf{M}[z]$. Ainsi $\mathbf{g}[z]$ constitue un filtre MISO dont la sortie $(y(n))_{n \in \mathbb{Z}}$ s'écrit :

$$y(n) = \mathbf{w}[z]\mathbf{x}(n) = \mathbf{g}[z]\mathbf{s}(n) \quad (3.1)$$

et dans le cas d'un bruit additif $\mathbf{b}(n)$ sur les capteurs :

$$y(n) = \mathbf{w}[z]\mathbf{x}(n) = \mathbf{w}[z]\left(\mathbf{M}[z]\mathbf{s}(n) + \mathbf{b}(n)\right) = \mathbf{g}[z]\mathbf{s}(n) + \tilde{\mathbf{b}}(n). \quad (3.2)$$

où $\tilde{\mathbf{b}}(n) = \mathbf{w}[z]\mathbf{b}(n)$ est le résultat du filtrage du bruit par le filtre ligne extracteur. Nous allons considérer l'extraction d'une source et étudier en fonction du filtre global $\mathbf{g}[z]$ les

critères dont la maximisation permet d'atteindre ce but. De tels critères sont appelés critère de contraste. Afin de les distinguer des critères de contraste «conjoints» définis au chapitre précédent, nous nous autoriserons à parler, si toutefois il pouvait y avoir ambiguïté, de critère de contraste MISO. Dans le paragraphe suivant, des résultats existant ayant trait aux contrastes MISO sont présentés puis étendus.

3.1.2 Généralisation des contrastes existants

Dans le cas MISO, la propriété essentielle qui permet de définir une fonction de contraste est que sa maximisation permet l'extraction d'une source. Cependant, il convient de faire certaines distinctions selon la nature des sources. Dans le cas le plus simple de sources i.i.d., la définition suivante d'un contraste peut-être proposée :

Définition 5 (contraste MISO, cas i.i.d.) Une fonction de contraste MISO (ou contraste MISO) est une fonction réelle de la sortie $(y(n))_{n \in \mathbb{Z}}$ du filtre ligne global $\mathbf{g}[z]$ qui vérifie :

- (i) $\mathcal{C}((y(n))_{n \in \mathbb{Z}}) \leq \max_{i=1}^N \mathcal{C}((s_i(n))_{n \in \mathbb{Z}})$
- (ii) Si $\mathcal{C}((y(n))_{n \in \mathbb{Z}}) = \max_{i=1}^N \mathcal{C}((s_i(n))_{n \in \mathbb{Z}})$, alors le filtre MISO global est du type $\mathbf{g}[z] = (0, \dots, 0, \alpha z^{-d}, 0, \dots, 0)$, où α est un scalaire non nul, $d \in \mathbb{Z}$ et la position de l'élément non nul z^{-d} correspond à un indice i de l'une des sources pour lesquelles $\mathcal{C}((s_i(n))_{n \in \mathbb{Z}})$ est maximal.

Remarque 4: Dans la définition ci-dessus, il a implicitement été supposé qu'une fonction de contraste dépendait des statistiques du signal de sortie et d'elles seules. En raison de la stationnarité toujours supposée (hypothèse H.1), ceci entraîne une invariance par translation temporelle, c'est-à-dire $\mathcal{C}((y(n))_{n \in \mathbb{Z}}) = \mathcal{C}((y(n-l))_{n \in \mathbb{Z}})$ pour tout $l \in \mathbb{Z}$. Dans le chapitre 5, nous nous intéresserons à des familles de contrastes paramétrées par un autre signal fixé (signal de référence). Dans ce cas, l'invariance du contraste par translation temporelle ne sera pas nécessairement vérifiée et nous devons adapter la définition correspondante.

Dans le cas i.i.d., il a été démontré [62, 93] que le critère

$$\frac{|\kappa_y(0, 0, 0)|}{(\mathbb{E}\{|y(n)|^2\})^2}, \quad (3.3)$$

est un contraste au sens de la définition ci-dessus.

Le cas de sources non i.i.d. requiert un peu plus de technicité et une définition plus souple de la notion de contraste. Nous ferons l'hypothèse suivante sur les sources :

H.10 Pour tout $j \in \{1, \dots, N\}$, la $j^{\text{ème}}$ source est centrée, stationnaire et sa fonction d'autocorrélation, notée $(\gamma_j(k))_{k \in \mathbb{Z}} := (\mathbb{E}\{s_j(n)s_j^*(n-k)\})_{k \in \mathbb{Z}}$, est définie positive. De plus, chaque source est de puissance unité ($\gamma_j(0) := \mathbb{E}\{|s_j(n)|^2\} = 1$).

Alors, pour tout indice $j \in \{1, \dots, N\}$ et tout filtre SISO $h[z]$, nous posons :

$$\|h\|_j := \left(\sum_{(k,l) \in \mathbb{Z}^2} h(k)h(l)^* \gamma_j(l-k) \right)^{\frac{1}{2}}. \quad (3.4)$$

Compte tenu de H.10, $\|\cdot\|_j$ définit bien une norme. Il est alors possible de définir la norme ℓ^2 pondérée du filtre global $\mathbf{g}[z] := (g_1[z], \dots, g_N[z])$:

$$\|\mathbf{g}\| := \left(\sum_{j=1}^N \|g_j\|_j^2 \right)^{\frac{1}{2}}. \quad (3.5)$$

Enfin, nous introduisons les ensembles suivants :

$$\forall i \in \{1, \dots, N\} \quad \mathcal{G}_i := \{\mathbf{g}[z] = (g_1[z], \dots, g_N[z]) \mid \|g_j\|_j = \delta_{i-j}, \forall j\} \quad (3.6)$$

Pour tout $i \in \{1, \dots, N\}$, \mathcal{G}_i est le sous-ensemble des filtres globaux de norme unité dont toutes les composantes sont nulles, sauf la $i^{\text{ème}}$ dont la norme- i vaut 1. Ainsi, un filtre global MISO appartenant à \mathcal{G}_i donne en sortie une filtrée scalaire de la $i^{\text{ème}}$ source. A partir de ces éléments, il est maintenant possible de donner la définition d'un contraste dans le cas de sources non i.i.d. et non linéaires :

Définition 6 (contraste MISO, cas non i.i.d.) \mathcal{C} est un contraste lorsque :

- (i) $\mathcal{C}((\mathbf{g}[z]\mathbf{s}(n))_{n \in \mathbb{Z}}) \leq \max_{i=1}^N \sup_{\tilde{\mathbf{g}} \in \mathcal{G}_i} \mathcal{C}((\tilde{\mathbf{g}}[z]\mathbf{s}(n))_{n \in \mathbb{Z}}) < +\infty$
- (ii) $\mathcal{C}((\mathbf{g}[z]\mathbf{s}(n))_{n \in \mathbb{Z}}) = \max_{i=1}^N \sup_{\tilde{\mathbf{g}} \in \mathcal{G}_i} \mathcal{C}((\tilde{\mathbf{g}}[z]\mathbf{s}(n))_{n \in \mathbb{Z}})$ si et seulement si il existe un scalaire $\alpha \neq 0$ et $\tilde{\mathbf{g}} \in \mathcal{G}_{i_0}$ permettant d'écrire $\mathbf{g} = \alpha \tilde{\mathbf{g}}$, où i_0 est tel que $\sup_{\tilde{\mathbf{g}} \in \mathcal{G}_{i_0}} \mathcal{C}((\tilde{\mathbf{g}}[z]\mathbf{s}(n))_{n \in \mathbb{Z}}) = \max_{i=1}^N \sup_{\tilde{\mathbf{g}} \in \mathcal{G}_i} \mathcal{C}((\tilde{\mathbf{g}}[z]\mathbf{s}(n))_{n \in \mathbb{Z}})$.

Remarque 5: Dans les définitions 5 et 6, le contraste \mathcal{C} a été noté comme dépendant de la sortie globale. Cependant, pour des sources et un mélange donnés, la sortie s'exprime en réalité en fonction des observations selon l'équation (3.1) et le contraste dépend ainsi des paramètres du filtre séparant.

En revanche, il est plus commode pour l'étude théorique de considérer la dépendance du contraste en fonction du filtre MISO global $\mathbf{g}[z]$. C'est ce dernier point de vue qui prévaut dans les développements qui suivent : le critère considéré sera noté $J(\mathbf{g})$ et étudié en fonction de sa dépendance vis-à-vis du filtre global.

Remarque 6: Dans les définitions 5 et 6 a été introduit un scalaire non nul α : il correspond à l'indétermination de multiplication des sources par un facteur multiplicatif. En général, les fonctions de contraste ne sont considérées que sur l'ensemble des filtres globaux de norme unité, ce qui équivaut à une normalisation en puissance ($E\{|y(n)|^2\} = 1$) de la sortie globale. Dans ce cas, nous pouvons imposer $|\alpha| = 1$ dans les deux définitions.

Il a été démontré dans [88] que le critère défini par (3.3) est un contraste pour des sources non i.i.d. au sens de la définition ci-dessus. Comme le critère (3.3) est inchangé par multiplication des coefficients du filtre global par un scalaire non nul, il est possible de considérer ce critère uniquement sur l'ensemble des filtres $\mathbf{g}[z]$ de norme unité. Dans ce cas, puisque $\|\mathbf{g}\|^2 = E\{|y(n)|^2\}$ et que ces deux quantités valent 1, il est possible d'affirmer que la fonction

$$\mathbf{g}[z] \mapsto |\kappa_y(0, 0, 0)| \quad (3.7)$$

est un contraste sur l'ensemble des filtres tels que $\|\mathbf{g}\| = 1$; c'est-à-dire que cette fonction est un contraste sous une contrainte de puissance unité en sortie. En nous basant sur ce résultat, nous introduisons le critère J défini par :

$$J(\mathbf{g}) := \sum_{(k_1, k_2, k_3) \in \mathcal{S}} f(\kappa_y(k_1, k_2, k_3)) \quad (3.8)$$

où $f : \mathbb{C} \rightarrow \mathbb{R}$ est une fonction convexe telle que $f(0) = 0$, et \mathcal{S} est un sous-ensemble de \mathbb{Z}^3 . Définissons également :

$$\mathcal{M}_i := \sup_{\mathbf{g} \in \mathcal{G}_i} J(\mathbf{g}). \quad (3.9)$$

\mathcal{M}_i est la borne supérieure du critère J lorsque seule la $i^{\text{ème}}$ source est présente en sortie du filtre global. Nous supposons également par la suite que la condition technique suivante est vérifiée (sa validité en pratique sera vérifiée dans le paragraphe 3.4.2) :

H.11 La borne supérieure définie à l'équation (3.9) est atteinte pour un filtre noté $\mathbf{g}_i^\sharp \in \mathcal{G}_i$. Nous pouvons alors affirmer :

Proposition 9 *Si f et \mathcal{S} ont été choisis de sorte à satisfaire :*

$$\forall j \in \{1, \dots, N\} \quad 0 < \mathcal{M}_j < +\infty, \quad (3.10)$$

alors J est un contraste sur l'ensemble des filtres \mathbf{g} de norme unité ($\|\mathbf{g}\| = 1$).

Preuve: Cette proposition se démontre par une méthode très proche de celle que l'on trouve dans [88]. Nous présentons ici cette démonstration, que l'on généralisera ensuite dans la proposition 11.

D'après l'équation (3.2) la sortie globale $y(n)$ peut s'écrire :

$$y(n) = \sum_{j=1}^N g_j[z] s_j(n) + \tilde{b}(n) = \sum_{j=1}^N \|g_j\|_j \tilde{g}_j[z] s_j(k) + \tilde{b}(n) \quad (3.11)$$

où l'on a posé :

$$\tilde{g}_j[z] := \begin{cases} \frac{g_j[z]}{\|g_j\|_j} & \text{si } \|g_j\|_j \neq 0, \\ 0 & \text{sinon.} \end{cases} \quad (3.12)$$

Ainsi, chacun des filtres scalaires $\tilde{g}_j[z]$ est de norme- j unité. En utilisant la multilinéarité des cumulants, la gaussianité du bruit et son indépendance, il vient :

$$\kappa_y(k_1, k_2, k_3) = \sum_{j=1}^N \|g_j\|_j^4 \kappa_{\tilde{g}_j[z] s_j}(k_1, k_2, k_3) + \left(1 - \sum_{j=1}^N \|g_j\|_j^4\right) \cdot 0. \quad (3.13)$$

Par convexité de f , puisque $\sum_{j=1}^N \|g_j\|_j^4 \leq 1$ et $f(0) = 0$, l'inégalité de Jensen permet d'écrire :

$$f(\kappa_y(k_1, k_2, k_3)) \leq \sum_{j=1}^N \|g_j\|_j^4 f(\kappa_{\tilde{g}_j[z] s_j}(k_1, k_2, k_3)) \quad (3.14)$$

et ainsi :

$$J(\mathbf{g}) \leq \sum_{j=1}^N \|g_j\|_j^4 \sum_{(k_1, k_2, k_3) \in \mathcal{S}} f(\kappa_{\tilde{g}_j[z] s_j}(k_1, k_2, k_3)) \quad (3.15)$$

Comme $\tilde{g}_i \in \mathcal{G}_i$, en utilisant (3.9) - (3.10), il vient :

$$J(\mathbf{g}) \leq \sum_{j=1}^N \|g_j\|_j^4 \mathcal{M}_j \quad (3.16)$$

$$\leq \sum_{j=1}^N \|g_j\|_j^2 \mathcal{M}_j \quad (3.17)$$

$$\leq \left(\max_{j \in \{1, \dots, N\}} \mathcal{M}_j \right) \sum_{j=1}^N \|g_j\|_j^2 = \left(\max_{j \in \{1, \dots, N\}} \mathcal{M}_j \right) \|\mathbf{g}\|^2 = \max_{j \in \{1, \dots, N\}} \mathcal{M}_j. \quad (3.18)$$

Si $J(\mathbf{g}) = \max_{j \in \{1, \dots, N\}} \mathcal{M}_j$, les inégalités ci-dessus deviennent des égalités. En particulier, l'égalité entre (3.16) et (3.17) entraîne que $\|g_j\|_j$ ne peut valoir que 0 ou 1. De plus, en considérant l'égalité

entre (3.17) et (3.18), il vient que $\|g_{i_0}\|_{i_0} = 1$ et $\|g_j\|_j = 0$ pour tout $j \neq i_0$, où i_0 est tel que $\mathcal{M}_{i_0} = \max_j \mathcal{M}_j$.

Réciproquement, $\mathbf{g} = \mathbf{g}_{i_0}^\#$ atteint la borne supérieure définie par (3.9) car dans ce cas, on a $J(\mathbf{g}) = \mathcal{M}_{i_0}$. \blacksquare

Plusieurs remarques s'imposent au sujet du résultat précédent :

- Contrairement à des résultats existants dans le cas i.i.d. [93] ou non i.i.d. [88], nous ne supposons pas que J est une fonction convexe du *module* des cumulants, mais une fonction convexe des cumulants eux-mêmes.
- Pour des sources dont les cumulants d'ordre quatre sont non nuls en $(0, 0, 0)$ (c'est-à-dire $\forall i, \kappa_{s_i}(0, 0, 0) \neq 0$), nous obtenons le contraste (3.7) comme cas particulier de (3.8) avec les choix $\mathcal{S} = \{(0, 0, 0)\}$ et $f(\cdot) = |\cdot|$.
- Dans le cas de sources de cumulants négatifs en $(0, 0, 0)$, il est possible d'obtenir un contraste en choisissant moins l'identité pour f . Ce choix n'est pas équivalent au choix $f(\cdot) = |\cdot|$ qui vient d'être évoqué.

En effet, considérons des sources qui, pour tout j , vérifient $\kappa_{s_j}(0, 0, 0) < 0$. Dans le cas de sources i.i.d. nous pouvons garantir que la sortie globale vérifie $\kappa_y(0, 0, 0) < 0$ et donc $|\kappa_y(0, 0, 0)| = -\kappa_y(0, 0, 0)$. Ainsi, les deux critères sont égaux dans ce cas. Dans le cas de sources non linéaires en revanche, rien ne permet d'affirmer que $\kappa_y(0, 0, 0) < 0$ en sortie globale et les deux critères ne sont donc pas identiques.

- La condition (3.10) est (en général) satisfaite pour des choix simples de \mathcal{S} et f . Pour tout support \mathcal{S} fini, il est possible de trouver une condition suffisante pour que \mathcal{M}_i soit fini. Cette condition est satisfaite pour les signaux que nous considérerons (i.i.d. ou CPM) et s'énonce :

Proposition 10 *Pour tout i fixé dans $\{1, \dots, N\}$, la borne supérieure \mathcal{M}_i est finie dès que les trois conditions suivantes sont réunies :*

1. \mathcal{S} est de cardinal fini.
2. La famille des autocumulants de la source $(s_i(n))_{n \in \mathbb{Z}}$ est sommable.
3. Il existe $\rho > 0$ tel que la densité spectrale de la source $(s_i(n))_{n \in \mathbb{Z}}$ vérifie $\Gamma_i[e^{i\omega}] \geq \rho > 0, \forall \omega$.

Preuve: Comme dans (3.8) la somme est effectuée sur \mathcal{S} dont le cardinal est supposé fini, il suffit de prouver que $f(\kappa_y(k_1, k_2, k_3))$ est fini, c'est-à-dire encore que $|\kappa_y(k_1, k_2, k_3)|$ est bornée car f est convexe sur \mathbb{C} et donc continue. Soit $\mathbf{g}[z] = (0, \dots, 0, g_i[z], 0, \dots, 0)$ un élément de \mathcal{G}_i . On peut écrire :

$$\begin{aligned} \kappa_y(k_1, k_2, k_3) &= \sum_{(l_1, l_2, l_3, l_4) \in \mathbb{Z}^4} g_i(l_1)g_i^*(l_2)g_i(l_3)g_i^*(l_4) \text{Cum}\{s_i(k-l_1), \\ &\quad s_i^*(k-k_1-l_2), s_i(k-k_2-l_3), s_i^*(k-k_3-l_4)\} \\ &= \sum_{(l_1, l_2, l_3, l_4) \in \mathbb{Z}^4} g_i(l_1)g_i^*(l_2)g_i(l_3)g_i^*(l_4) \kappa_{s_i}(l_1-k_1-l_2, l_1-k_2-l_3, l_1-k_3-l_4) \\ &= \sum_{(l, p_1, p_2, p_3) \in \mathbb{Z}^4} g_i(l)g_i^*(l-p_1)g_i(l-p_2)g_i^*(l-p_3) \kappa_{s_i}(p_1-k_1, p_2-k_2, p_3-k_3) \\ &= \sum_{(p_1, p_2, p_3) \in \mathbb{Z}^3} \kappa_{s_i}(p_1-k_1, p_2-k_2, p_3-k_3) \sum_{l \in \mathbb{Z}} g_i(l)g_i^*(l-p_1)g_i(l-p_2)g_i^*(l-p_3). \end{aligned}$$

L'inégalité de Hölder généralisée permet alors d'obtenir :

$$\sum_{l \in \mathbb{Z}} |g_i(l)| |g_i^*(l-p_1)| |g_i(l-p_2)| |g_i^*(l-p_3)| \leq \sum_{l \in \mathbb{Z}} |g_i(l)|^4 \leq \|g_i\|_{\mathcal{L}^2}^4 \quad (3.19)$$

où $\|g_i\|_{\mathcal{L}^2} = \sqrt{\sum_{l \in \mathbb{Z}} |g_i(l)|^2}$. Ainsi, nous avons la majoration :

$$|\kappa_y(k_1, k_2, k_3)| \leq \|g_i\|_{\mathcal{L}^2}^4 \sum_{(p_1, p_2, p_3) \in \mathbb{Z}^3} |\kappa_{s_i}(p_1 - k_1, p_2 - k_2, p_3 - k_3)| \quad (3.20)$$

Par ailleurs, l'hypothèse sur la densité spectrale de la source $(s_i(n))_{n \in \mathbb{Z}}$ permet d'écrire :

$$\|g_i\|_i^2 = \frac{1}{2\pi} \int_{-\pi}^{\pi} |G_i[e^{i\omega}]|^2 \Gamma_i[e^{i\omega}] d\omega \geq \frac{\rho}{2\pi} \int_{-\pi}^{\pi} |G_i[e^{i\omega}]|^2 d\omega \geq \frac{\rho}{2\pi} \|g_i\|_{\mathcal{L}^2}^2 \quad (3.21)$$

Puisque $\mathbf{g} \in \mathcal{G}_i$, il vient donc $\|g_i\|_{\mathcal{L}^2} \leq \sqrt{2\pi\rho^{-1/2}} \|g_i\|_i = \sqrt{2\pi\rho^{-1/2}}$. Cette inégalité, combinée avec la sommabilité des cumulants permet d'affirmer que le majorant de l'équation (3.20) est fini et donc que \mathcal{M}_i est fini. ■

- D'après la définition (3.9), une condition suffisante pour assurer $\mathcal{M}_j > 0$ est qu'il existe $\mathbf{g} \in \mathcal{G}_j$ tel que $J(\mathbf{g}) > 0$. Dès lors, on constate aisément que, dans le cas particulier où $\mathcal{S} = \{(0, 0, 0)\}$ et où f est moins l'identité, le contraste donné par (3.8) est valable si les sources sont de kurtosis négatif. En effet, si l'on considère $\mathbf{g} \in \mathcal{G}_j$ défini par un filtre identité sur la $j^{\text{ème}}$ composante et des zéros ailleurs ($\mathbf{g}[z] = (0, \dots, 0, z^0, 0, \dots, 0)$), alors $J(\mathbf{g}) = -\kappa_{s_j}(0, 0, 0) > 0$, ce qui implique que (3.10) est vérifiée et que $\mathbf{g} \mapsto -\kappa_y(0, 0, 0)$ est un contraste.

3.2 De l'extraction d'une source à l'extraction MIMO

Le paragraphe précédent a mis en avant une famille de critères dont l'optimisation permet d'obtenir l'extraction d'une source. Ainsi la maximisation d'un critère semblable à (3.8) permet d'après la proposition 9 d'obtenir un filtre MISO séparant. Afin d'obtenir un filtre séparant MIMO, cette même étape doit être répétée N fois. Naturellement, afin de ne pas retrouver plusieurs fois la même source, les maximisations successives doivent être réalisées sous certaines conditions ou contraintes.

Dans l'approche par déflation, la contribution sur les capteurs de la source précédemment extraite est soustraite par une méthode de moindres carrés. Cette contribution étant supprimée, la source en question ne peut être obtenue une seconde fois. Nous rappelons le principe de la déflation dans le paragraphe suivant. Cette approche présente l'inconvénient d'accumuler les erreurs d'estimation au cours des itérations du procédé de déflation. En conséquence, la qualité des résultats tend à se détériorer au fur et à mesure que les sources sont extraites. Nous allons proposer une autre méthode, qui consiste en l'introduction de contraintes supplémentaires à chaque maximisation successive du ou des critère(s) considéré(s). Ceci est expliqué dans le paragraphe 3.2.2.

3.2.1 Procédé de déflation

Parmi les procédés d'extraction séquentielle des sources, une méthode classique, appelée déflation, consiste à supprimer la contribution de la source précédemment extraite sur les capteurs. Le procédé a été introduit dans [62] pour le cas convolutif et développé dans [32] pour le cas instantané. On le retrouve également dans [88, 93, 87] et nous en rappelons ici brièvement le fonctionnement.

Le principe de base consiste à exploiter la possibilité d'extraire une source afin d'en déduire une procédure séquentielle, dont les étapes sont les suivantes :

- recherche d'un filtre $\mathbf{w}[z]$ de taille $1 \times Q$ tel que $y(n) = \mathbf{w}[z]\mathbf{x}(n)$ soit une version filtrée de l'une des sources, par exemple la première. On note dans ce cas $y(n) = g_1[z]s_1(n)$.
- à partir de $y(n)$, détermination et soustraction de la contribution de $s_1(n)$ sur les capteurs. Plus précisément, on recherche un filtre $\mathbf{t}[z] = (t_1[z], \dots, t_Q[z])^T$ de taille $Q \times 1$ qui pour tout $j \in \{1, \dots, Q\}$ minimise le critère :

$$\mathbf{E}\{|x_j(n) - t_j[z]y(n)|^2\} = \mathbf{E}\{|(M_{j1}[z] - t_j[z]g_1[z])s_1(n) + \sum_{i=2}^N M_{ji}[z]s_i(n)|^2\}$$

ce qui s'écrit par indépendance des sources :

$$= \mathbf{E}\{|(M_{j1}[z] - t_j[z]g_1[z])s_1(n)|^2\} + \mathbf{E}\{|\sum_{i=2}^N M_{ji}[z]s_i(n)|^2\}$$

Le critère ci-dessus est minimal quand $M_{j1}[z] = t_j[z]g_1[z]$. Par suite, les critères ci-dessus pour $j \in \{1, \dots, Q\}$ sont minimaux si $\mathbf{t}[z]y(n)$ correspond à la contribution de $s_1(n)$ sur le réseau de capteurs. Il est donc possible de créer l'erreur de prédiction $\tilde{\mathbf{x}}(n) := \mathbf{x}(n) - \mathbf{t}[z]y(n)$, qui est le signal résultant du mélange des $N - 1$ sources $s_2(n), \dots, s_N(n)$ par un filtre composé des $N - 1$ dernières colonnes de $\mathbf{M}[z]$.

- itération des deux étapes précédentes après avoir remplacé $\mathbf{x}(n)$ par un nouveau vecteur d'observations $\tilde{\mathbf{x}}(n)$. Le processus s'arrête lorsque toutes les sources ont été extraites.

Malgré quelques abus de langage, nous réservons le nom de déflation au processus qui consiste à retrancher par moindres carrés la contribution d'une source, dans le but d'en déduire une méthode de séparation séquentielle (ce terme provient originellement de techniques de calcul numérique sur les matrices).

Remarque 7: Notons que le choix de l'ordre des filtres $\mathbf{t}[z]$ n'est pas évident dans le cas non i.i.d. : en effet, la source extraite n'est reconstruite qu'à un filtrage près, de type MA. Afin de soustraire sa contribution sur les capteurs, il convient alors d'utiliser un filtre $\mathbf{t}[z]$ de type AR et donc de réponse impulsionnelle infinie en théorie. Cette difficulté de choix de l'ordre du filtre n'apparaîtra pas dans l'approche proposée aux paragraphes 3.2.2 et 3.4.

Une autre procédure séquentielle est présentée dans le paragraphe suivant.

3.2.2 Maximisation stricte des contrastes sur un sous-ensemble

Nous proposons ici une approche différente de la déflation : au lieu de soustraire les contributions des sources précédemment extraites, ces dernières nous servent à introduire des contraintes supplémentaires sur l'ensemble des filtres MISO sur lequel l'optimisation est menée à l'étape suivante. Définissons les sous-ensembles suivants de l'ensemble des filtres $1 \times N$:

$$\forall P \in \{1, \dots, N\} \quad \mathcal{F}_P := \{\mathbf{g}[z] \mid g_1 = \dots = g_{P-1} = 0 \text{ et } \|\mathbf{g}\| = 1\} \quad (3.22)$$

En particulier, \mathcal{F}_1 correspond à la sphère unité, tandis que les $\mathcal{F}_P, P > 1$ sont les projections de \mathcal{F}_1 sur les sous-espaces des filtres MISO dont les $P - 1$ premières composantes sont identiquement nulles. Nous allons considérer (pour tout $P \in \{1, \dots, N\}$) les restrictions $J|_{\mathcal{F}_P}$ de J sur \mathcal{F}_P .

La proposition 9 affirme d'ores et déjà que que $J_{|\mathcal{F}_1}$ est un contraste : supposons que par sa maximisation on puisse extraire la première source. Il est clair que la maximisation de $J_{|\mathcal{F}_2}$ ne peut mener une seconde fois à l'extraction de la première source. Cependant, puisque l'on considère le critère J sur un sous-ensemble de l'ensemble des filtres de norme unité, $J_{|\mathcal{F}_2}$ présente des maxima autres que ceux de $J_{|\mathcal{F}_1}$ et donc, parmi les critères $J_{|\mathcal{F}_P}, P \geq 1$, seul $J_{|\mathcal{F}_1}$ constitue un contraste *a priori*. La proposition suivante permet de surmonter cette difficulté et de généraliser le résultat de la proposition 9 :

Proposition 11 *Si (3.10) est vérifiée, alors, pour tout $P \in \{1, \dots, N\}$, $J_{|\mathcal{F}_P}$ est un contraste au sens où sa maximisation globale¹ mène à l'extraction d'une source. De plus, l'indice de la source extraite par la maximisation de $J_{|\mathcal{F}_P}$ est supérieur ou égal à P .*

Preuve: Fixons P quelconque dans $\{1, \dots, N\}$ et considérons $\mathbf{g} \in \mathcal{F}_P$. Alors, $g_1 = \dots = g_{P-1} = 0$ et l'équation (3.11) s'écrit :

$$y(n) = \sum_{j=P}^N \|g_j\|_j \tilde{g}_j[z] s_j(k) + \tilde{b}(n), \quad (3.23)$$

où $\tilde{g}_j[z]$ est défini dans (3.12). La preuve de la proposition 11 se déroule alors de manière identique à celle de la proposition 9, dans laquelle on aura limité la sommation à des indices $j \in \{P, \dots, N\}$. ■

Nous pouvons maintenant remarquer que la proposition précédente permet d'extraire les sources les unes après les autres. La validité de la méthode correspondante, appelée hiérarchique, est affirmée dans la proposition qui suit. Dans la mesure où l'ordre d'extraction des sources est purement conventionnel, nous supposons toujours que la $P^{\text{ème}}$ source est extraite à l'étape P de la méthode.

Proposition 12 (procédure de séparation hiérarchique) *La procédure qui consiste à maximiser successivement $J_{|\mathcal{F}_1}, \dots, J_{|\mathcal{F}_N}$ mène à l'obtention d'un filtre MIMO global séparant.*

Preuve: Puisque \mathcal{F}_1 est l'ensemble des filtre globaux de norme unité, d'après la proposition 9, la maximisation de $J_{|\mathcal{F}_1}$ mène à un filtre MISO séparant. La source extraite correspond à l'une des sources pour lesquelles le paramètre $\mathcal{M}_i, i \in \{1, \dots, N\}$ est maximal. D'après l'ordre choisi pour les sources, nous supposons en toute généralité que la première source a été extraite.

Supposons maintenant nous trouver à l'étape $P \in \{1, \dots, N\}$ où, par maximisation successive de $J_{|\mathcal{F}_1}, \dots, J_{|\mathcal{F}_{P-1}}$, nous avons extrait les $P - 1$ premières sources. La proposition 11 assure que la maximisation de $J_{|\mathcal{F}_P}$ mène à la séparation d'une source distincte des $P - 1$ déjà extraites. Par convention sur l'ordre des sources, il s'agit de la $P^{\text{ème}}$ et la proposition se déduit donc par récurrence. ■

Dans le paragraphe 3.4.1 nous traiterons de la faisabilité de l'optimisation contrainte introduite dans les propositions 11 et 12. Auparavant, nous étudions les propriétés locales des critères introduits.

3.3 Etude locale des contrastes

Comme il sera expliqué dans les paragraphes 3.4.2 et 3.4.4, l'optimisation est en général réalisée par une méthode d'optimisation itérative locale. Ainsi, une étude locale donnera-t-elle une meilleure compréhension du comportement des algorithmes. De plus, l'analyse du

¹Maximisation globale est à comprendre au sens de maximisation stricte, *i.e.* relative à la recherche du maximum absolu, par opposition à la maximisation locale.

paragraphe 3.3.2 permettra de justifier une méthode de post-optimisation qui permet de compenser l'accumulation d'erreurs inhérente aux approches séquentielles de la séparation de source : nous constaterons par la suite que ceci permet une nette amélioration de la qualité des résultats.

3.3.1 Utilisation des maxima locaux pour la séparation

L'approche précédente se fonde sur des maximisations globales de $J_{|\mathcal{F}_1}, \dots, J_{|\mathcal{F}_N}$. Il serait intéressant d'avoir en plus des informations sur de possibles maxima locaux parasites de ces critères. Il a été démontré dans [88, 93] que tout maximum local du critère (3.3) est un filtre séparant (ce résultat s'applique naturellement au critère (3.7)). Comme nous considérons ici ces mêmes critères sous contrainte, ces résultats ne sont pas applicables en l'état et nous proposons maintenant de les généraliser.

Inspiré des idées de [88], nous nous concentrons dans les paragraphes 3.3.1 et 3.3.2 sur le choix particulier $\mathcal{S} = \{(0, 0, 0)\}$. Ce choix apparaît justifié aussi par le fait supplémentaire que l'estimation du contraste se trouve d'autant plus facilitée qu'il y a moins de termes à estimer dans J . D'après notre expérience —notamment les résultats du chapitre précédent— ceci devrait mener à de meilleures performances statistiques.

Enfin, une hypothèse technique supplémentaire est nécessaire dans le paragraphe 3.3.1 (mais pas dans 3.3.2) : en plus de la convexité de f , nous supposons que la fonction $f : \mathbb{R} \rightarrow \mathbb{R}$ est strictement monotone (croissante ou décroissante) sur chacun des intervalles $]-\infty, 0]$ et $[0, +\infty[$. En particulier, si f est moins l'identité (choix que l'on considérera dans le chapitre suivant, paragraphe 4.1), la condition est remplie. De même, les fonctions habituellement utilisées telles que la valeur absolue ou le module au carré satisfont la condition requise.

Les résultats de [88] sur les maxima locaux du critère (3.3) proviennent indirectement du lemme suivant :

Lemme 2 Soient $m \in \mathbb{N}^*$, des réels c_1, \dots, c_m et la fonction :

$$\psi : \mathbb{R}^m \rightarrow \mathbb{R}, \quad \mathbf{a} = \begin{pmatrix} a_1 \\ \vdots \\ a_m \end{pmatrix} \mapsto \sum_{k=1}^m c_k a_k^4 \quad (3.24)$$

Notons $\mathcal{S} = \{\mathbf{a} \in \mathbb{R}^m \mid \mathbf{a}^T \mathbf{a} = 1\}$ la sphère unité de \mathbb{R}^m et $\mathbf{e}_k, k \in \{1, \dots, m\}$ les vecteurs de la base canonique de \mathbb{R}^m . On peut affirmer que :

1. si $\forall k, c_k > 0$:
 - (i) les maxima locaux de ψ sur \mathcal{S} sont les vecteurs \mathbf{a}^{\max} tels que $\mathbf{a}^{\max} = \pm \mathbf{e}_k$.
 - (ii) les minima locaux de ψ sur \mathcal{S} sont à chercher parmi les vecteurs \mathbf{a}^{\min} dont les composantes vérifient : $\forall k, (a_k^{\min})^2 = \frac{\lambda}{4c_k}$ où $\lambda = \frac{4}{\sum_{l=1}^m \frac{1}{c_l}}$.
2. si $\forall k, c_k < 0$:
 - (i) les minima locaux de ψ sur \mathcal{S} sont les vecteurs \mathbf{a}^{\min} tels que $\mathbf{a}^{\min} = \pm \mathbf{e}_k$.
 - (ii) les maxima locaux de ψ sur \mathcal{S} sont à chercher parmi les vecteurs \mathbf{a}^{\max} dont les composantes vérifient : $\forall k, (a_k^{\max})^2 = \frac{\lambda}{4c_k}$ où $\lambda = \frac{4}{\sum_{l=1}^m \frac{1}{c_l}}$.
3. si les c_k ne sont pas tous strictement de même signe (i.e. certains sont > 0 , d'autres ≤ 0 ou inversement) :
 - (i) les maxima locaux de ψ sur \mathcal{S} sont les vecteurs $\mathbf{a}^{\max} = \pm \mathbf{e}_{k_0}$ où k_0 est tel que $c_{k_0} > 0$.
 - (ii) les minima locaux de ψ sur \mathcal{S} sont les vecteurs $\mathbf{a}^{\min} = \pm \mathbf{e}_{k_0}$ où k_0 est tel que $c_{k_0} < 0$.

Ce lemme peut se démontrer par une étude de tous les cas où le Lagrangien du problème s'annule. La démonstration complète se trouve dans [32]. Au moyen d'une preuve semblable à celle dans [88], nous pouvons alors prouver le résultat suivant :

Proposition 13 *Supposons que $\mathcal{S} = \{(0, 0, 0)\}$. Alors, pour tout $P \in \{1, \dots, N\}$, tout maximum local positif de $J_{|\mathcal{F}_P}$ correspond à un filtre séparant. De plus l'indice de la source extraite par tout maximum local de $J_{|\mathcal{F}_P}$ est supérieur ou égal à P .*

Preuve: Soit P fixé dans $\{1, \dots, N\}$ et soit $\mathbf{g} \in \mathcal{F}_P$; on peut alors écrire (3.23) et il s'ensuit :

$$f(\kappa_{\mathbf{y}}(0, 0, 0)) = f\left(\sum_{j=P}^N \|g_j\|_j^4 \kappa_{\tilde{g}_j[z]s_j(k)}(0, 0, 0)\right). \quad (3.25)$$

Supposons que $\hat{\mathbf{g}}$ soit un maximum local appartenant à \mathcal{F}_P de l'expression ci-dessus. Deux cas doivent être distingués :

- si l'argument de la fonction f est non nul en $\hat{\mathbf{g}}$, l'hypothèse technique précédente sur f , associée avec sa continuité assure que $\hat{\mathbf{g}}$ est soit un maximum local, soit un minimum local sur \mathcal{F}_P de :

$$\mathbf{g} \mapsto \sum_{j=P}^N \|g_j\|_j^4 \kappa_{\tilde{g}_j[z]s_j(k)}(0, 0, 0). \quad (3.26)$$

(où l'expression ci-dessus est l'argument de f dans (3.25)).

- si l'argument de la fonction f dans (3.25) s'annule en $\hat{\mathbf{g}}$, le même résultat demeure lorsque f est monotone sur \mathbb{R} . Si en revanche un changement du sens de variation de f intervient en 0, en raison de sa convexité, f est nécessairement décroissante sur $]-\infty, 0]$ et croissante sur $[0, +\infty[$. Ainsi, 0 est un minimum global de f sur \mathbb{R} et $\hat{\mathbf{g}}$ serait aussi un minimum global de l'expression dans (3.25). Ceci entre manifestement en contradiction avec le fait que $\hat{\mathbf{g}}$ est un maximum local de cette même expression, démontrant ainsi que ce cas ne peut se produire.

Nous déduisons en particulier de cette discussion que le vecteur $(\|\hat{g}_P\|_P, \dots, \|\hat{g}_N\|_N)$ est un extremum sur la sphère unité $(\sum_{j=P}^N \|g_j\|_j^2 = 1)$ de la fonction :

$$\begin{aligned} \mathbb{R}^{N-P+1} &\rightarrow \mathbb{R} \\ (\|g_P\|_P, \dots, \|g_N\|_N) &\mapsto \sum_{j=P}^N \|g_j\|_j^4 \tilde{C}_j \end{aligned} \quad (3.27)$$

où les $\tilde{C}_j := \kappa_{\tilde{g}_j[z]s_j(k)}(0, 0, 0)$ sont des constantes dépendant de $\hat{\mathbf{g}}$. Cette fonction correspond à celle du lemme 2. De plus, puisque le maximum local de $J_{|\mathcal{F}_P}$ est supposé positif, il ne peut s'agir d'un minimum local de (3.27) dans le cas 1 du lemme. En effet, si nous étions dans ce dernier cas, les \tilde{C}_j seraient positifs et donc aussi la fonction (3.27). Dans le cas où f est croissante sur \mathbb{R}^+ ceci correspond à un minimum local de (3.25) ce qui contredit l'hypothèse de maximum local. Dans le cas où f est décroissante sur \mathbb{R}^+ , on obtient un maximum local négatif de (3.25) puisque $f(0) = 0$. Notons que ce maximum serait strictement négatif compte tenu de la forme des solutions de 1-(ii), ce qui est contraire à l'hypothèse de maximum positif. Ces deux possibilités mènent donc bien à une contradiction.

Pour des raisons similaires, nous ne pouvons être dans le cas 2-(ii) du lemme. Ainsi le lemme permet de conclure que toutes les composantes $\|\hat{g}_j\|_j, j = P, \dots, N$ sont nulles sauf une. Comme de plus $\hat{\mathbf{g}} \in \mathcal{F}_P$, ceci prouve que $\hat{\mathbf{g}}$ est séparant. ■

La proposition ci-dessus démontre qu'il est possible d'extraire une source par une méthode de maximisation locale. Par une preuve identique à celle de la proposition 12, nous pouvons en déduire une méthode de séparation de l'ensemble des sources. De même que dans le paragraphe 3.2.2 nous supposons par souci de simplicité (et sans perte de généralité) que les sources sont ordonnées selon leur ordre d'extraction.

Proposition 14 *La procédure qui consiste en la recherche successive des maxima locaux positifs des critères $J_{|\mathcal{F}_1}, J_{|\mathcal{F}_2}, \dots, J_{|\mathcal{F}_N}$ mène à l'obtention d'un filtre MIMO global séparant.*

Les résultats obtenus dans le cadre de ce paragraphe ont fait apparaître que certains contrastes peuvent présenter des caractéristiques plus intéressantes que celles imposées par les définitions 5 ou 6. Le fait qu'un contraste soit séparant pour tout maximum local est en effet d'un intérêt bien plus important pour la mise en œuvre de la méthode que de savoir la propriété vérifiée pour le maximum global. Ce résultat local permet de s'affranchir d'une recherche de maximum global et nous pouvons remarquer à ce propos que de façon générale, il serait souhaitable dans l'étude des critères de séparation, de ne pas se concentrer uniquement sur leurs maxima globaux. L'intérêt des propriétés *locales* sera encore souligné dans la suite de nos investigations, qui concerne principalement une réciproque de la proposition 13.

3.3.2 Analyse au voisinage d'un filtre MISO séparant

Les résultats précédents affirment que tout maximum local positif des critères considérés (et en particulier tout maximum local positif de $J_{|\mathcal{F}_1}$) correspond à un filtre séparant. Cependant, rien n'interdit *a priori* qu'un point séparant obtenu par l'une des maximisations de $J_{|\mathcal{F}_P}$ pour $P > 1$ ne soit pas un maximum local de $J_{|\mathcal{F}_1}$: c'est précisément cette question qui nous préoccupe dans ce paragraphe et nous montrons que cette situation ne peut pas se produire. En particulier nous caractérisons, autour de tout maximum local de $J_{|\mathcal{F}_P}$, $P \geq 1$, une boule dans laquelle ce maximum local est maximum local de $J_{|\mathcal{F}_1}$ également. Le résultat est valable à la fois dans le cas i.i.d. et non i.i.d. et permettra de proposer dans le chapitre 3.4.5 une amélioration de la méthode de séparation dans le but de compenser les accumulations d'erreurs.

Soit $\hat{\mathbf{g}}[z]$ un filtre MISO de taille $1 \times N$ qui maximise le critère $J_{|\mathcal{F}_P}$ où P est fixé quelconque dans $\{1, \dots, N\}$. La proposition 13 (ou 11) assure que $\hat{\mathbf{g}}[z]$ permet l'extraction d'une source, la $P^{\text{ème}}$. De plus, $\hat{\mathbf{g}}[z]$ étant un maximum local de $J_{|\mathcal{F}_P}$, il existe $\rho > 0$ tel que :

$$\forall \mathbf{g} \in \mathcal{F}_P, \quad \|\mathbf{g} - \hat{\mathbf{g}}\| < \rho \Rightarrow J(\mathbf{g}) \leq J(\hat{\mathbf{g}}). \quad (3.28)$$

Proposition 15 *Si l'on pose :*

$$\widehat{\mathcal{M}}_P := J(\hat{\mathbf{g}}) \quad \text{et} \quad \epsilon_P := \frac{2\widehat{\mathcal{M}}_P}{\max_{j \neq P} \mathcal{M}_j + \widehat{\mathcal{M}}_P}, \quad (3.29)$$

alors on a :

$$\forall \mathbf{g} \in \mathcal{F}_1, \quad \|\mathbf{g} - \hat{\mathbf{g}}\| < \min(\sqrt{\epsilon_P}, \rho) \Rightarrow J(\mathbf{g}) \leq J(\hat{\mathbf{g}}) \quad (3.30)$$

et pour tout \mathbf{g} appartenant à

$$\mathcal{F}_1 \cap \{\mathbf{g} \mid \|\mathbf{g} - \hat{\mathbf{g}}\| < \min(\sqrt{\epsilon_P}, \frac{\rho}{2})\} \quad (3.31)$$

l'égalité $J(\mathbf{g}) = J(\hat{\mathbf{g}})$ ne peut être vraie que si \mathbf{g} extrait la même source que $\hat{\mathbf{g}}$ (c'est-à-dire $\|g_j\|_j = \delta_{j-P}$).

Preuve:

D'après (3.14), nous avons :

$$f(\kappa_{y(k)}(0, 0, 0)) \leq \sum_{j \neq P} \|g_j\|_j^4 f(\kappa_{\tilde{g}_j[z]s_j(n)}(0, 0, 0)) + \|g_P\|_P^4 f(\kappa_{\tilde{g}_P[z]s_P(n)}(0, 0, 0)) \quad (3.32)$$

et, en utilisant la définition (3.9) :

$$f(\kappa_{y(k)}(0, 0, 0)) \leq \max_{j \neq P} \mathcal{M}_j \sum_{j \neq P} \|g_j\|_j^4 + \|g_P\|_P^4 f(\kappa_{\tilde{g}_P[z]s_P(n)}(0, 0, 0)). \quad (3.33)$$

Soit $\tilde{\mathbf{g}} := (0, \dots, 0, \tilde{g}_P, 0, \dots, 0)$. Comme seule la $P^{\text{ème}}$ composante de $\hat{\mathbf{g}}$ et $\tilde{\mathbf{g}}$ est non nulle, nous avons $\|\tilde{\mathbf{g}} - \hat{\mathbf{g}}\| = \|\tilde{g}_P - \hat{g}_P\|_P$ et cette quantité est majorée par $2\|\mathbf{g} - \hat{\mathbf{g}}\|$ comme le montre la suite d'inégalités ci-dessous :

$$\begin{aligned} \|\tilde{\mathbf{g}} - \hat{\mathbf{g}}\| &\leq \|\tilde{g}_P - g_P\|_P + \|g_P - \hat{g}_P\|_P \\ &\leq \|g_P\|_P \left(\frac{1}{\|g_P\|_P} - 1 \right) + \|g_P - \hat{g}_P\|_P \\ &\leq \left| 1 - \|g_P\|_P \right| + \|g_P - \hat{g}_P\|_P = \left| \|\hat{g}_P\|_P - \|g_P\|_P \right| + \|g_P - \hat{g}_P\|_P \\ &\leq 2\|g_P - \hat{g}_P\|_P \leq 2\|\mathbf{g} - \hat{\mathbf{g}}\|. \end{aligned} \quad (3.34)$$

Si \mathbf{g} appartient à l'ensemble décrit par (3.31), alors $\|\mathbf{g} - \hat{\mathbf{g}}\| < \rho/2$ et ainsi $\|\tilde{\mathbf{g}} - \hat{\mathbf{g}}\| < \rho$, ce qui entraîne

$$f(\kappa_{\tilde{g}_P[z]s_P(n)}(0, 0, 0)) = J(\tilde{\mathbf{g}}) \leq J(\hat{\mathbf{g}}) = \widehat{\mathcal{M}}_P \quad (3.35)$$

et donc, d'après (3.33) :

$$J(\mathbf{g}) = f(\kappa_{y(k)}(0, 0, 0)) \leq \max_{j \neq P} \mathcal{M}_j \sum_{j \neq P} \|g_j\|_j^4 + \|g_P\|_P^4 \widehat{\mathcal{M}}_P. \quad (3.36)$$

Introduisons $\epsilon := \sum_{j \neq P} \|g_j\|_j^2 \in [0, 1]$. Nous avons :

$$\sum_{j \neq P} \|g_j\|_j^4 \leq \left(\sum_{j \neq P} \|g_j\|_j^2 \right)^2 = \epsilon^2 \quad (3.37)$$

et ainsi :

$$J(\mathbf{g}) \leq \epsilon^2 \max_{j \neq P} \mathcal{M}_j + (1 - \epsilon)^2 \widehat{\mathcal{M}}_P =: \phi(\epsilon). \quad (3.38)$$

On constate alors immédiatement que $\phi(\epsilon) < \phi(0)$ lorsque $0 < \epsilon < \epsilon_P$ (où ϵ_P est défini par (3.29)) et $\phi(\epsilon)$ admet en zéro le maximum local $\phi(0) = \widehat{\mathcal{M}}_P$.

Maintenant, pour tout \mathbf{g} appartenant à l'ensemble décrit par (3.31), $\|\mathbf{g} - \hat{\mathbf{g}}\|^2 = \sum_{j=1}^N \|g_j - \hat{g}_j\|_j^2 < \epsilon_P$. Comme $\hat{g}_j[z] = 0$ pour tout $j \neq P$, ceci implique évidemment $\epsilon < \epsilon_P$. Ainsi, si $\mathbf{g}[z]$ appartient à l'ensemble (3.31) et $\epsilon > 0$, en se rappelant (3.38), nous obtenons $J(\mathbf{g}) < \widehat{\mathcal{M}}_P = J(\hat{\mathbf{g}})$. Si $\epsilon = 0$, \mathbf{g} vérifie $\|g_j\|_j = \delta_{j-P}$ et $J(\mathbf{g}) \leq J(\hat{\mathbf{g}})$. En particulier, si $J(\mathbf{g}) = J(\hat{\mathbf{g}})$, nous sommes nécessairement dans ce dernier cas et la proposition 15 est donc établie. ■

Selon la proposition 14, maximiser $J_{|\mathcal{F}_P}$ mène théoriquement à une solution séparante. En pratique cependant, les ensembles \mathcal{F}_P sont estimés et seule une approximation de la solution réelle cherchée peut être obtenue (voir le paragraphe 3.4.5 pour une discussion plus approfondie). D'après le résultat précédent, et à condition que l'approximation obtenue soit suffisamment proche de la vraie solution (c'est-à-dire à condition qu'elle soit dans l'ensemble (3.31)) une maximisation locale de $J_{|\mathcal{F}_1}$ devrait améliorer la solution.

Remarque 8: Un filtre $\hat{\mathbf{g}}$ qui maximise l'une des fonctions $J_{|\mathcal{F}_P}$ est nécessairement séparable et, d'après les notations choisies, $\hat{\mathbf{g}} \in \mathcal{G}_P$. Un cas plus particulièrement intéressant est celui où $\hat{\mathbf{g}}$ est un maximum global de $J_{\mathcal{F}_P}$. Dans ce cas en effet, nous avons $\widehat{\mathcal{M}}_P = \mathcal{M}_P$ et $\rho = \infty$ et l'ensemble (3.31) devient l'intersection de \mathcal{F}_1 avec la boule de centre $\hat{\mathbf{g}}$ et de rayon $\sqrt{\epsilon_P}$.

Remarque 9: L'introduction du paramètre ρ dans la proposition 15 et dans la définition de l'ensemble (3.31) peut sembler une limitation importante de l'intérêt de ces résultats. Il est cependant possible de proposer une version légèrement différente de la dernière partie de cette proposition :

Proposition 16 *Soit $\hat{\mathbf{g}}[z]$ un filtre MISO de taille $1 \times N$ qui soit maximum global ou local de l'une des fonctions $J_{|\mathcal{F}_1}, \dots, J_{|\mathcal{F}_N}$, et qui permette d'extraire la $j_0^{\text{ème}}$ source. Alors, tout autre maximum local de $J : \mathbf{g}[z] \mapsto J(\mathbf{g})$ sur l'ensemble*

$$\{\mathbf{g} \mid \|\mathbf{g} - \hat{\mathbf{g}}\| < 1 \text{ et } \|\mathbf{g}\| = 1\} \quad (3.39)$$

extraît également la $j_0^{\text{ème}}$ source.

Preuve: Soit $\mathbf{g} = (g_1, \dots, g_N)$ un maximum local de J appartenant à l'ensemble (3.39). \mathbf{g} étant un maximum local de $J_{|\mathcal{F}_1}$, ses composantes sont toutes nulles sauf une, par exemple la $j_1^{\text{ème}}$, ce qui s'écrit $\|g_j\|_j = \delta_{j-j_1}$. De plus, $\hat{\mathbf{g}}$ est tel que $\|\hat{g}_j\|_j = \delta_{j-j_0}$ et donc :

$$\|\mathbf{g} - \hat{\mathbf{g}}\|^2 = \sum_{j=1}^N \|g_j - \hat{g}_j\|_j^2 = \sum_{j \neq j_0} \|g_j\|_j^2 + \|g_{j_0} - \hat{g}_{j_0}\|_{j_0}^2 \quad (3.40)$$

De $\|\mathbf{g} - \hat{\mathbf{g}}\| < 1$ et de l'égalité ci-dessus, il s'ensuit une contradiction puisque dans le cas où $j_1 \neq j_0$, (3.40) conduit à $\|\mathbf{g} - \hat{\mathbf{g}}\| \geq 1$. Par conséquent, on a nécessairement $j_1 = j_0$, ce qui démontre la proposition 16. ■

Remarque 10: Dans le cas où $\mathcal{S} \neq \{(0, 0, 0)\}$, la proposition 13 n'est plus valable et $J_{|\mathcal{F}_P}(\mathbf{g})$ peut présenter des maxima locaux parasites non séparants. En revanche, lorsque le maximum global a été trouvé, le filtre correspondant est nécessairement séparable. Pour affiner le résultat, il est alors justifié de procéder à une post-optimisation locale autour de ce point uniquement lorsque toutes les sources ont des statistiques identiques : en effet, le maximum global atteint par $J_{|\mathcal{F}_P}(\mathbf{g})$ est alors égal au maximum global (et donc aussi local) de $J_{|\mathcal{F}_1}(\mathbf{g})$.

3.4 Vers une méthode effective de séparation MIMO

3.4.1 Caractérisation des ensembles \mathcal{F}_i

Afin de mettre en œuvre les procédures de séparation données par les propositions précédentes, nous montrons que les ensembles $\mathcal{F}_i, i \in \{1, \dots, N\}$ peuvent être caractérisés à partir des seules données observées. La proposition qui suit montre que $\mathbf{g}[z]$ appartient à l'ensemble souhaité \mathcal{F}_i si et seulement si le signal de sortie correspondant est de variance unité et vérifie des contraintes de décorrélation.

Considérons la $P^{\text{ème}}$ étape ($P \in \{1, \dots, N\}$) de la procédure de séparation de la proposition 12 ou 14 et supposons que les $P - 1$ étapes précédentes ont été réalisées avec succès. Nous disposons alors de $P - 1$ sorties globales $y_1(n), \dots, y_{P-1}(n)$ qui sont des filtrées SISO des sources $s_1(n), \dots, s_{P-1}(n)$ respectivement. Ceci sera noté de la façon suivante : $y_i(n) = g_i^\# [z] s_i(n), \forall i \in \{1, \dots, P - 1\}$. Soit $\mathbf{g}[z] = (g_1[z], \dots, g_N[z])$ un filtre de taille $1 \times N$ et $y(n) := \mathbf{g}[z] \mathbf{s}(n)$.

Proposition 17 *Il y a équivalence entre les deux propositions :*

1. $\mathbf{g} \in \mathcal{F}_P$

$$2. \begin{cases} \text{C.1 } \mathbb{E}\{|y(n)|^2\} = 1 \\ \text{C.2 } \forall i \in \{1, \dots, P-1\} \quad \forall k \in \mathbb{Z} \quad \mathbb{E}\{y_i(n)y^*(n-k)\} = 0 \end{cases}$$

Preuve: Il est aisé de constater d'abord que C.1 est équivalent à $\|\mathbf{g}\| = 1$. De plus, soit i fixé quelconque dans $\{1, \dots, P-1\}$. C.2 est équivalent à l'annulation de l'interspectre $S_i[e^{z\omega}]$ entre $y_i(n)$ et $y(n)$, ce qui peut s'écrire :

$$S_i[e^{z\omega}] = g_i^\sharp[e^{z\omega}](g_i[e^{z\omega}])^* \Gamma_i[e^{z\omega}] = 0 \quad (3.41)$$

Ainsi, sous hypothèse d'analyticité de $g_i^\sharp[z]$ et de $\Gamma_i[z]$ dans un anneau contenant le cercle unité, nous obtenons que $\mathbb{E}\{y_i(n)y^*(n-k)\}$ est identiquement nul si et seulement si $g_i = 0$. Par conséquent, les deux conditions données par 2 sont équivalentes à $\|\mathbf{g}\| = 1$ et $\forall i \in \{1, \dots, P-1\}, g_i = 0$, ce qui prouve l'équivalence des deux propositions. ■

D'après la proposition 17, la procédure de séparation de la proposition 12 (resp. proposition 14) consiste en N maximisations globales (resp. locales) successives de J sous les contraintes C.1 et C.2, où P prend successivement les valeurs $1, 2, \dots, N$. Remarquons au passage que la contrainte C.2 est linéaire, ce qui facilite sa prise en compte.

Un algorithme similaire a été proposé dans [46] uniquement dans le cas de sources i.i.d. Nous avons ici démontré sa validité dans le cas non i.i.d. et, contrairement à [46], nous n'avons aucunement besoin d'une opération de préblanchiment. Ceci est d'une grande importance si l'on veut en pratique obtenir une méthode de séparation efficace : il est connu en effet, qu'une opération de préblanchiment peut altérer les performances de la séparation.

3.4.2 Optimisation pour des filtres RIF

L'optimisation des contrastes que nous avons étudiés peut s'effectuer à l'aide d'un algorithme d'optimisation itératif classique. Dès lors, la contrainte C.1 peut facilement être obtenue en effectuant une renormalisation du filtre séparant après chaque itération de la méthode d'optimisation. Nous nous intéressons donc maintenant plus particulièrement à la contrainte C.2 et la réécrivons comme suit :

$$\text{C.3 } \forall i, 1 \leq i < P \quad \forall k \in \mathbb{Z} \quad \mathbb{E}\{y_i(n+k)(\mathbf{w}[z]\mathbf{x}(n))^*\} = 0$$

Nous pouvons constater que C.3 donne un nombre infini de contraintes car l'égalité ci-dessus doit être assurée pour tout $k \in \mathbb{Z}$. En effet, pour des sources temporellement corrélées, dont l'autocorrélation ne s'annule jamais, il n'y a aucune raison pour que l'égalité ci-dessus soit automatiquement vérifiée lorsque $|k|$ est plus grand qu'une certaine valeur finie. Afin d'obtenir une meilleure compréhension du problème, nous allons considérer par la suite des filtres RIF.

Supposons à compter de maintenant (et jusqu'à la fin du chapitre) que le filtre de mélange $\mathbf{M}[z]$ est de longueur finie L (c'est-à-dire $\deg \mathbf{M}[z] = L-1$, où \deg est le degré maximal des polynômes en z^{-1} dans la matrice $\mathbf{M}[z]$). Plus précisément, $\mathbf{M}[z]$ se compose de N colonnes $\mathbf{M}[z] := (\mathbf{m}_1[z] \quad \dots \quad \mathbf{m}_N[z])$ de degré respectif $\deg \mathbf{m}_i[z] =: L_i - 1$ (et nous avons donc $L = \max_i L_i$). Cette hypothèse assure que le filtre $\mathbf{M}[z]$ est stable. Précisons des conditions sous lesquelles le filtre peut être inversé :

H.12 $\mathbf{M}[z]$ est irréductible *i.e.* $\text{Rang}(\mathbf{M}[z]) = N, \forall z \in \mathbb{C}^* \cup \{\infty\}$.

H.13 $\mathbf{M}[z]$ est à colonnes réduites *i.e.* $\text{Rang}(\mathbf{m}_1(L_1 - 1) \ \dots \ \mathbf{m}_N(L_N - 1)) = N$.

Sous H.12, l'existence d'un filtre inverse RIF est assurée. Si H.13 est satisfaite, une indication supplémentaire sur le degré d'un filtre séparant est apportée : pour tout D tel que $D \geq \sum_{i=1}^N (L_i - 1)$, il existe un filtre séparant $\mathbf{W}[z]$ de longueur D (voir [39] et ses références). Même si nous supposons les deux hypothèses vérifiées, H.13 n'est faite que par commodité afin de connaître le degré de l'inverse du filtre de mélange, tandis que H.12 est une hypothèse tout à fait raisonnable puisque, sous certaines conditions techniques, elle est presque toujours vraie dès qu'il y a plus de capteurs que de sources ($Q > N$).

En conséquence des hypothèses ci-dessus, l'optimisation de J est réalisée sur l'ensemble des filtre $\mathbf{w}[z]$ de longueur finie $D \in \mathbb{N}^*$, et qui satisfont C.1 et C.3. Si l'on note :

$$\underline{\mathbf{w}} := (\mathbf{w}(0) \ \mathbf{w}(1) \ \dots \ \mathbf{w}(D - 1)) \quad (3.42)$$

on constate aisément que C.3 peut donner, comme déjà expliqué, un nombre infini de contraintes sur $\underline{\mathbf{w}}$. Toutefois, en raisonnant des hypothèses ci-dessus, nous sommes assurés de l'existence d'un $\underline{\mathbf{w}} \neq 0$ séparateur. Les contraintes d'orthogonalité données par C.3 doivent donc nécessairement être redondantes, car sinon, elles impliqueraient $\underline{\mathbf{w}} = 0$.

Un problème analogue a été rencontré dans [46], quand bien même, dans ces travaux, les sources étaient supposées décorréelées et que C.3 pouvait par conséquent se réduire à un nombre fini de contraintes. Afin d'obtenir un vecteur $\underline{\mathbf{w}}$ non nul, il avait été proposé d'ajuster la longueur du filtre séparant à chaque étape de la procédure hiérarchique (*i.e.* D dépendait de P et différait pour chaque ligne du filtre $\mathbf{W}[z]$). L'étude du cas non i.i.d. nous laisse penser que cette approche n'est pas la meilleure, car dans ce dernier cas, une réduction du nombre de contraintes est dans tous les cas nécessaire afin d'obtenir un filtre $\mathbf{w}[z]$ de réponse impulsionnelle finie.

En pratique, nous considérons plus d'équations de contraintes d'orthogonalité que strictement nécessaire et nous sélectionnons à l'aide d'une décomposition en valeurs singulières les contraintes numériquement les plus significatives. Afin de réaliser cette projection, nous déterminons dans le paragraphe suivant la dimension du sous-espace des vecteurs $\underline{\mathbf{w}}$ telles que C.3 est vérifiée.

3.4.3 Dimension des ensembles \mathcal{F}_i

Dans ce paragraphe, nous discutons de la dimension minimale de l'ensemble des filtres séparants $\mathbf{w}[z]$ de longueur donnée D et tels que le filtre global $\mathbf{g}[z] = \mathbf{w}[z]\mathbf{M}[z]$ appartienne à \mathcal{F}_P (et ainsi la sortie globale satisfasse C.1 et C.2). Si l'on se rappelle que le $i^{\text{ème}}$ élément de $\mathbf{g}[z]$ s'écrit $g_i[z] = \mathbf{w}[z]\mathbf{m}_i[z]$, où $\mathbf{m}_i[z]$ représente la $i^{\text{ème}}$ colonne du filtre de mélange, C.2 peut se réécrire :

$$\text{C.4 } \forall i \in \{1, \dots, P - 1\} \quad \mathbf{w}[z]\mathbf{m}_i[z] = 0 .$$

Définissons $\mathcal{T}_D(\mathbf{m}_i)$ comme la matrice suivante de taille $QD \times (L_i + D - 1)$:

$$\mathcal{T}_D(\mathbf{m}_i) := \begin{pmatrix} \mathbf{m}_i(0) & \mathbf{m}_i(1) & \dots & \mathbf{m}_i(L_i - 1) & 0 & \dots & 0 \\ 0 & \mathbf{m}_i(0) & \mathbf{m}_i(1) & \dots & \mathbf{m}_i(L_i - 1) & 0 & 0 \\ \vdots & \ddots & \ddots & \ddots & & \ddots & \vdots \\ 0 & \dots & 0 & \mathbf{m}_i(0) & \mathbf{m}_i(1) & \dots & \mathbf{m}_i(L_i - 1) \end{pmatrix} \quad (3.43)$$

C.4 est équivalente à :

$$\underline{\mathbf{w}} (\mathcal{T}_D(\mathbf{m}_1) \ \dots \ \mathcal{T}_D(\mathbf{m}_{P-1})) = 0 \quad (3.44)$$

Puisque $\mathbf{M}[z]$ est irréductible et à colonnes réduites, $(\mathcal{T}_D(\mathbf{m}_1) \dots \mathcal{T}_D(\mathbf{m}_{P-1}))$ est de rang maximal, c'est-à-dire $((P-1)(D-1) + \sum_{i=1}^{P-1} L_i)$. Le nombre de conditions d'orthogonalité qu'impose l'équation (3.44) est donc $((P-1)(D-1) + \sum_{i=1}^{P-1} L_i)$. C'est le nombre de contraintes linéaires données par C.2 en théorie. Comme en fait $L = \max_i L_i$, nous obtenons que C.2 correspond au plus à $(P-1)(D+L-1)$ équations de contrainte linéairement indépendantes.

3.4.4 Implantation de la méthode de séparation

Nous donnons maintenant des détails concernant l'implantation pratique de la séparation par la méthode hiérarchique². Le problème se ramène à l'optimisation de $J_{|\mathcal{F}_P}$ pour $P = 1, \dots, N$ et donc à des optimisations sous les contraintes C.1 et C.2. Un algorithme de gradient a été utilisé pour ces maximisations successives. Pour l'étape $P = 1$, la contrainte C.2 disparaît et l'optimisation est identique à ce qui a été présenté dans [93] : la contrainte C.1 est imposée grâce à une renormalisation du vecteur $\underline{\mathbf{w}}$ effectuée à chaque itération du gradient.

Considérons à présent la gestion de la contrainte C.2 supplémentaire qui apparaît dès que $P > 1$. En définissant :

$$\underline{\mathbf{x}}(n) := (\mathbf{x}(n)^T, \mathbf{x}(n-1)^T, \dots, \mathbf{x}(n-D+1)^T)^T \quad (3.45)$$

on peut écrire :

$$y(n) = \underline{\mathbf{w}} \underline{\mathbf{x}}(n) \quad (3.46)$$

Dans le cas $P = 2$, C.2 (ou C.3) exprime la décorrélation entre la première source reconstituée $y_1(n)$ et celle que l'on cherche. Avec les notations ci-dessus, ceci s'écrit :

$$\forall k \in \mathbb{Z} \quad \underline{\mathbf{w}}^* \mathbf{E}\{y_1(n+k) \underline{\mathbf{x}}^*(n)\} = 0 \quad (3.47)$$

Sachant que les conditions ci-dessus sont redondantes, nous ne les considérons que pour $k \in \{-k_{\max}, \dots, k_{\max}\}$. Nous définissons ainsi la matrice \mathbf{A}_1 , composée des intercorrélations entre $y_1(n)$ et les composantes de $\underline{\mathbf{x}}(n)$, et telle que la condition ci-dessus, prise pour $k \in \{-k_{\max}, \dots, k_{\max}\}$ s'écrive alors

$$\underline{\mathbf{w}}^* \mathbf{A}_1 = \mathbf{0}. \quad (3.48)$$

En pratique, nous avons choisi $k_{\max} = L+D-1$, qui correspond à l'ensemble des valeurs de k à considérer de sorte que (3.47) soit vérifié pour tout k dans l'hypothèse de sources i.i.d. Ainsi, l'équation (3.48) traduit $2k_{\max} + 1$ conditions d'orthogonalité, ce qui est plus que le nombre $D+L-1$ de conditions à imposer d'après le paragraphe précédent. La matrice \mathbf{A}_1 n'est donc pas de rang plein et nous réduisons par conséquent le nombre de conditions d'orthogonalité en sélectionnant dans une matrice $\tilde{\mathbf{A}}_1$ l'espace colonne correspondant aux plus grandes valeurs singulières de \mathbf{A}_1 . La recherche du second filtre séparateur s'effectue ainsi dans l'espace défini par $\underline{\mathbf{w}}^* \tilde{\mathbf{A}}_1 = \mathbf{0}$ et il est aisé d'en déduire un algorithme de gradient projeté. Dans le cas où $P > 2$, nous construisons de même des matrices $\tilde{\mathbf{A}}_i, i \in \{1, \dots, P\}$ et le vecteur $\underline{\mathbf{w}}$ est recherché dans l'espace défini par $\underline{\mathbf{w}}^* \tilde{\mathbf{A}}_1 = \dots = \underline{\mathbf{w}}^* \tilde{\mathbf{A}}_P = \mathbf{0}$.

²Une séparation hiérarchique a déjà été introduite dans le cas de sources i.i.d. [92]. La différence consiste en l'introduction de termes de pénalisation du critère, qui permettent de prendre en compte de manière pondérée les contraintes de décorrélation, au lieu de les imposer comme il est fait ici.

3.4.5 Amélioration des performances par une post-optimisation locale

On peut reprocher aux méthodes de séparation hiérarchiques (propositions 12 et 14), de ne pas traiter les sources de façon symétrique : la première est reconstruite après une maximisation dans l'ensemble \mathcal{F}_1 des filtres $1 \times N$ de norme unité, tandis que les filtres MISO séparateurs suivants sont cherchés dans des sous-ensembles $\mathcal{F}_P, P \in \{2, \dots, N\}$, qui satisfont les inclusions successives : $\mathcal{F}_N \subsetneq \dots \subsetneq \mathcal{F}_2 \subsetneq \mathcal{F}_1$. Comme les ensembles $\mathcal{F}_i, i \in \{2, \dots, N\}$ dépendent des estimations précédentes, et comme les erreurs d'estimation s'accumulent, il faut s'attendre à ce que l'extraction des sources soit de moindre qualité pour celles extraites en dernier. Cette même critique peut être adressée aux approches par déflation qui soustraient la contribution des sources précédemment extraites sur les capteurs.

Pour limiter ces effets, il a été proposé dans [94, 66] et [87], de réaliser après chaque extraction de la procédure une maximisation locale *non contrainte* de $J_{|\mathcal{F}_1}$ mais avec une *initialisation appropriée* de l'algorithme d'optimisation afin d'éviter d'extraire deux fois la même source. Lorsque $\mathcal{S} = \{(0, 0, 0)\}$, la proposition 15 fournit une justification théorique à cette pratique.

L'initialisation adéquate peut être donnée par l'algorithme décrit dans la proposition 12 (ou 14). Une autre solution consiste à identifier le filtre ligne séparateur correspondant après l'extraction d'une source par déflation. Ceci peut se faire par moindres carrés.

3.5 Simulations

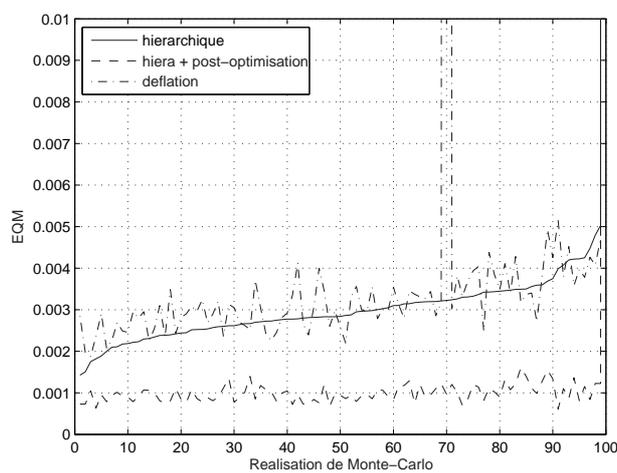
Les simulations présentées dans ce paragraphe se concentrent principalement sur le cas de sources i.i.d. réelles dont la distribution est soit uniforme, soit de type PAM4. Le cas non i.i.d. est présenté également mais a été étudié plus attentivement dans le cadre des applications exposées au chapitre 4. Nous renvoyons en particulier le lecteur aux simulations du paragraphe 4.1.3. Enfin, le choix de se placer dans le cas où le nombre de capteurs est strictement supérieur au nombre de sources se justifie par le souhait d'assurer l'existence d'un filtre MIMO et RIF qui permette d'inverser le mélange.

3.5.1 Sources i.i.d.

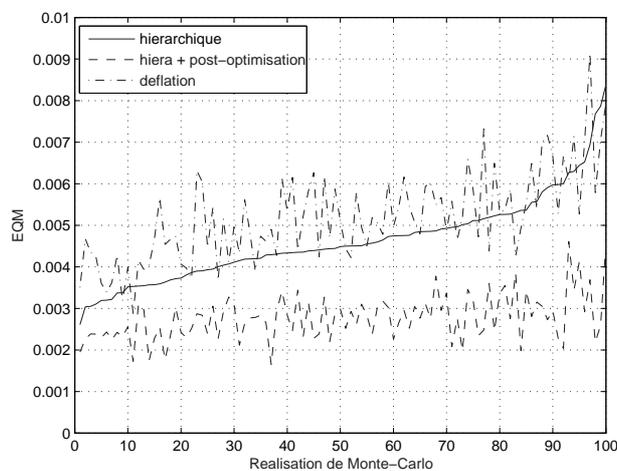
La méthode de séparation déduite des propositions 12 et 14, et dont des éléments de mise en œuvre ont été donnés au paragraphe 3.4 a été considérée sur des signaux i.i.d. Le contraste MISO classique $J(\mathbf{g}) = |\kappa_y(0, 0, 0)|$ correspondant à $\mathcal{S} = \{(0, 0, 0)\}$ et $f(\cdot) = |\cdot|$ a été choisi dans ce cadre. Une comparaison a été menée entre la méthode dite hiérarchique, cette même méthode à laquelle a été ajoutée une étape de post-optimisation sans contrainte de décorrélation et la méthode par déflation.

Méthode séparation	Source 1	Source 2	Source 3
hiérarchique	0.0027	0.0046	0.0066
hiéra. + post-opt.	0.0027	0.0027	0.0029
déflation	0.0027	0.0052	0.0077

TAB. 3.1 – Comparaison des EQM moyennes sur 1000 réalisations données par différentes méthodes de séparation séquentielles. Cas de 3 sources i.i.d. à distribution uniforme, 6 capteurs, et d'un filtre de mélange de longueur 3.



(a) Sources i.i.d. PAM4.



(b) Sources i.i.d. uniformes.

FIG. 3.1 – EQM moyenne sur les sources reconstituées. Comparaison sur 100 réalisations de Monte-Carlo (3 sources, 6 capteurs, filtre de longueur 3, 4096 échantillons). Les réalisations ont été classées par ordre croissant de l'EQM pour la méthode hiérarchique.

La figure 3.1 trace pour 100 réalisations de Monte-Carlo la moyenne des EQM sur les trois sources reconstituées par chacune des méthodes. Ces réalisations ont été classées de sorte à donner aux graphiques une meilleure lisibilité. Il apparaît que les performances de la méthode hiérarchique sont légèrement meilleures que celles de la déflation. L'ajout de l'étape de post-optimisation permet d'améliorer encore la qualité du résultat.

Les valeurs des erreurs quadratiques moyennes (moyennées sur les 1000 réalisations) sont données au niveau du tableau 3.1. Nous constatons alors que l'amélioration des performances est plus particulièrement observable sur la deuxième et surtout troisième source extraite. La méthode par déflation entraîne en effet une accumulation sensible des erreurs au fur et à mesure que les sources sont extraites. Cet inconvénient demeure dans la méthode hiérarchique, mais de façon atténuée. En revanche, le fait de procéder à une post-optimisation permet de rétablir une symétrie de traitement entre la première source extraite et les suivantes. Ainsi, l'erreur sur la troisième source extraite n'est-elle pas sensiblement plus importante que celle sur la première.

3.5.2 Sources non i.i.d.

Le cas de sources non i.i.d. requiert l'introduction d'un critère de séparation insensible aux ambiguïtés de filtrage SISO. Aussi, pour la $P^{\text{ème}}$ source extraite, nous avons considéré le critère suivant, défini à partir de la $P^{\text{ème}}$ ligne du filtre global, notée $\mathbf{g}[z] = (g_1[z], \dots, g_N[z])$:

$$\tau_P := 1 - \frac{\max_{j \geq P} \|g_j\|_j^2}{\sum_j \|g_j\|_j^2} \quad (3.49)$$

Il est immédiat de constater que, pour P quelconque, $0 \leq \tau_P < 1$. La séparation de la $P^{\text{ème}}$ source est réussie lorsque la $P^{\text{ème}}$ ligne du filtre global ne contient qu'une seule composante non nulle. Cette composante doit être de plus distincte des composantes obtenues aux lignes précédentes, que l'on a supposé être les composantes $1, \dots, P-1$: le contraire signifierait que la même source a été extraite plus d'une fois. Ainsi, nous constatons que des valeurs de τ_P proches de zéro correspondent à des extractions de bonne qualité de la $P^{\text{ème}}$ source.

Les résultats donnés par la méthode hiérarchique ont été examinés. En particulier, les valeurs moyennes des critères τ_1 à τ_4 ont été tracées à la figure 3.2 en fonction du nombre d'échantillons disponibles pour des sources ARCH(1) présentées au paragraphe 2.3.3. Nous constatons la validité de la méthode. D'autres résultats dans le cas non i.i.d. accompagnés d'une étude plus approfondie seront donnés à l'occasion de l'étude de l'application à des signaux CPM.

3.6 Conclusion

Ce chapitre a concerné l'étude des méthodes séquentielles basées sur des contrastes MISO. Après avoir rappelé les notions nécessaires, nous avons obtenu des formes généralisées de fonctions de contraste. Le passage à la séparation MIMO a été obtenu au travers de contraintes supplémentaires ajoutées à l'issue de chaque extraction. Ces contraintes imposent et traduisent une décorrélation entre la source en cours d'extraction et celles précédemment extraites. Les difficultés pratiques de mise en œuvre de la méthode ont également été exposées. Par ailleurs, ce chapitre a été l'occasion d'élargir et de clarifier des résultats relatifs à une étude locale des critères considérés. Outre l'assurance de ne

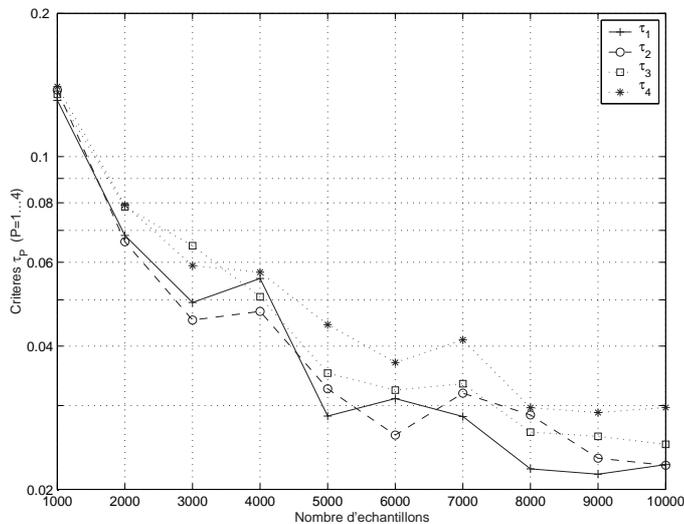


FIG. 3.2 – Séparation de 4 sources de types ARCH(1) : valeur moyenne sur 100 réalisations des critères $\tau_i, i \in \{1, \dots, 4\}$ en fonction du nombre d'échantillons. Le filtre de mélange comportait 8 capteurs, était de longueur 2 et a été tiré aléatoirement selon une loi gaussienne normalisée à chaque réalisation.

pas obtenir de maxima locaux non séparant, qui était déjà connue dans le cas MISO, nous avons justifié la possibilité d'utiliser une optimisation locale non contrainte pour affiner les résultats et limiter le problème d'accumulation des erreurs.

Chapitre 4

Applications des approches séquentielles

Le chapitre précédent a présenté d'un point de vue général la séparation de signaux sources par des méthodes séquentielles. Nous considérons dans ce chapitre deux applications intéressantes de la méthode qui a été proposée. Nous étudions d'abord l'utilisation des méthodes de séparation dans le cadre des télécommunications et plus particulièrement des signaux modulés par une modulation à phase continue (CPM). Ensuite, nous établissons de quelle manière les méthodes développées peuvent se généraliser au cas des images et des signaux multidimensionnels.

4.1 Séparation et détection d'un mélange de signaux modulés CPM

4.1.1 Introduction aux modulations CPM

4.1.1-a) Généralités

Les modulations à phase continue, couramment appelées modulations CPM (continuous phase modulation) présentent des propriétés qui les rendent particulièrement attractives, parmi lesquelles la propriété importante de module constant de leur enveloppe. De plus, elles présentent une très bonne efficacité spectrale et réalisent un bon compromis entre efficacité de puissance et complexité des modulateurs/démodulateurs. Pour ces raisons, les modulations CPM sont utilisées dans le système européen de téléphones mobiles de seconde génération (GSM), ainsi que dans de nombreux systèmes de communications militaires, ...

La démodulation de signaux CPM ayant subi une transmission selon un canal à trajets multiples trouve un intérêt pratique dans l'application à l'écoute passive. Dans le cas d'un émetteur unique, la plupart des méthodes consistent à estimer conjointement les paramètres du canal et la séquence de symboles [70, 21, 51]. En réalité, d'après [10], il est possible d'estimer la séquence de symboles à l'aide d'un égaliseur —par exemple l'algorithme de module constant (CMA) de Godard [38]— suivi d'un algorithme de détection classique. Nous allons illustrer que cette méthode est généralisable dans la configuration d'émetteurs multiples. Nous rappelons d'abord les propriétés essentielles des CPM dont nous aurons l'utilité. Elles sont exposées plus en détail dans [9]; d'autres informations peuvent être trouvées dans [81].

4.1.1-b) Modèle de signal

Soit $(a_n)_{n \in \mathbb{Z}}$ une suite i.i.d. de symboles de communication. Nous supposons que, pour tout n , a_n prend de manière équiprobable sa valeur dans l'alphabet binaire $\{+1, -1\}$. L'enveloppe complexe (à temps continu) du signal CPM associé est :

$$\underline{s}(t) := e^{i\psi(t)}, \quad (4.1)$$

où la phase $\psi(t)$ de $\underline{s}(t)$ vaut :

$$\psi(t) := \pi h \left(\sum_{k \in \mathbb{Z}} a_k \phi(t - kT_s) \right). \quad (4.2)$$

Dans l'expression ci-dessus, $h \in]0, 1[$ est un paramètre fixe appelé indice de modulation et T_s désigne la période symbole. Enfin, ϕ est une fonction continue, croissante telle que, si l'on fixe $L \in \mathbb{N}^*$ un entier naturel non nul :

$$\begin{aligned} \forall t \leq 0, \quad \phi(t) &= 0 \\ \forall t \in [0, LT_s] \quad \phi(t) &\in [0, 1] \\ \forall t \geq LT_s \quad \phi(t) &= 1 \end{aligned} \quad (4.3)$$

Dans la pratique, la fonction ϕ est une primitive de ce qui est appelé la fonction de mise en forme $g(t)$: cette dernière est continue, positive, de support $[0, LT_s]$ (L est donc la longueur du filtre de mise en forme) et normalisée de sorte que $\int_0^{LT_s} g(t) dt = 1$. Nous avons par conséquent la relation : $\phi(t) := \int_0^t g(u) du$. Selon la longueur du filtre de mise en forme, on distingue :

- si $L = 1$, la modulation est dite modulation CPM à réponse complète. Dans ce cas, le support de la fonction de mise en forme correspond à $[0, T_s]$.
- si $L > 1$, la modulation est dite modulation CPM à réponse partielle.

Dans la suite de l'exposé, seules les modulations CPM à réponse complète seront envisagées.

4.1.1-c) Représentation de Laurent (CPM à réponse complète)

Dans le cas de CPM à réponse complète, il a été démontré dans [59] que le signal $\underline{s}(t)$ peut s'écrire comme une modulation linéaire de pseudo-symboles. Nous reprenons ici la démonstration donnée dans [9].

D'après (4.1) et (4.2),

$$\underline{s}(t) = e^{i\pi h \left(\sum_{k \in \mathbb{Z}} a_k \phi(t - kT_s) \right)} \quad (4.4)$$

Supposons $t \in [nT_s, (n+1)T_s[$. En utilisant les propriétés (4.3) de la fonction ϕ , il vient :

$$\forall t \in [nT_s, (n+1)T_s[\quad \underline{s}(t) = \underline{s}(nT_s) e^{i\pi h a_n \phi(t - nT_s)} \quad (4.5)$$

Par ailleurs, en exploitant le fait que $a_n = \pm 1$, on vérifie aisément que :

$$e^{i\pi h a_n \phi(t - nT_s)} = e^{i\pi h a_n} \frac{\sin(\pi h \phi(t - nT_s))}{\sin(\pi h)} + \frac{\sin(\pi h (1 - \phi(t - nT_s)))}{\sin(\pi h)} \quad (4.6)$$

et donc :

$$\forall t \in [nT_s, (n+1)T_s[$$

$$\underline{s}(t) = \underline{s}(nT_s)e^{i\pi h a_n} \frac{\sin(\pi h \phi(t - nT_s))}{\sin(\pi h)} + \underline{s}(nT_s) \frac{\sin(\pi h(1 - \phi(t - nT_s)))}{\sin(\pi h)} \quad (4.7)$$

Introduisons la suite $(s(n))_{n \in \mathbb{Z}} = (\underline{s}((n+1)T_s))_{n \in \mathbb{Z}}$, appelée suite de pseudo-symboles. Cette suite est liée à $(a_n)_{n \in \mathbb{Z}}$ et vérifie :

$$\begin{cases} \forall n \in \mathbb{Z}, & s(n+1) = e^{i\pi h a_n} s(n) \\ s(0) \text{ est une variable aléatoire centrée, de module 1, uniformément distribuée} \\ \text{sur l'ensemble de ses états possibles, et indépendante de } (a_n)_{n \in \mathbb{Z}^+}. \end{cases} \quad (4.8)$$

De (4.7), on déduit alors successivement

$$\forall t \in [nT_s, (n+1)T_s[$$

$$\underline{s}(t) = s(n) \frac{\sin(\pi h \phi(t - nT_s))}{\sin(\pi h)} + s(n-1) \frac{\sin(\pi h(1 - \phi(t - nT_s)))}{\sin(\pi h)} \quad (4.9)$$

puis l'équation suivante, appelée représentation de Laurent :

$$\forall t \in \mathbb{R} \quad \underline{s}(t) = \sum_{k \in \mathbb{Z}} s(k)c(t - kT_s) \quad (4.10)$$

où $c(t)$ est la fonction de mise en forme de Laurent définie par :

$$c(t) = \begin{cases} 0 & \text{si } t \notin]0, 2T_s[, \\ \frac{\sin(\pi h \phi(t))}{\sin(\pi h)} & \text{si } 0 \leq t \leq T_s, \\ \frac{\sin(\pi h(1 - \phi(t - T_s)))}{\sin(\pi h)} & \text{si } T_s \leq t \leq 2T_s, \end{cases} \quad (4.11)$$

Il apparaît donc que le signal $\underline{s}(t)$ a la structure d'une modulation linéaire, mais où la suite de pseudo-symboles —qui remplace la suite de symbole originelle— n'est pas i.i.d. mais seulement stationnaire. Plus précisément, $(s(n))_{n \in \mathbb{Z}}$ est strictement stationnaire et n'est pas un processus linéaire.

4.1.2 Séparation de CPM dans le cas où la période symbole est égale à la période d'échantillonnage

4.1.2-a) Formulation du problème et lien avec les chapitres précédents

Nous considérons le cas d'une transmission de N signaux de communication distincts modulés en CPM. Nous notons $s_j(t)$, $j \in \{1, \dots, N\}$ chacun de ces N signaux de communications; h_j , $j \in \{1, \dots, N\}$ désigneront les indices de modulation respectivement associés et $(a_n^j)_{n \in \mathbb{Z}}$, $j \in \{1, \dots, N\}$ seront les symboles correspondants aux différents utilisateurs. Ces signaux sont transmis au travers d'un canal de propagation linéaire et invariant dans le temps, qui, par exemple, peut résulter d'un modèle de propagation multi-trajets.

Au niveau des capteurs, et après échantillonnage au rythme de la période symbole T_s , les observations $\mathbf{x}(n)$ suivent alors un modèle de mélange convolutif du type :

$$\mathbf{x}(n) = \sum_{k \in \mathbb{Z}} \mathbf{M}(n-k)\mathbf{s}(n) + \mathbf{b}(n) \quad (4.12)$$

Il s'agit là du modèle exposé au chapitre 1.4.1, où les sources sont des suites de pseudo-symboles vérifiant les propriétés de l'équation (4.8). Le problème est de reconstituer chacune des communications, ce qui revient à l'étude de la séparation du mélange de pseudo-symboles, puis à l'application d'une démodulation. Un schéma de principe est donné à la figure 4.1.

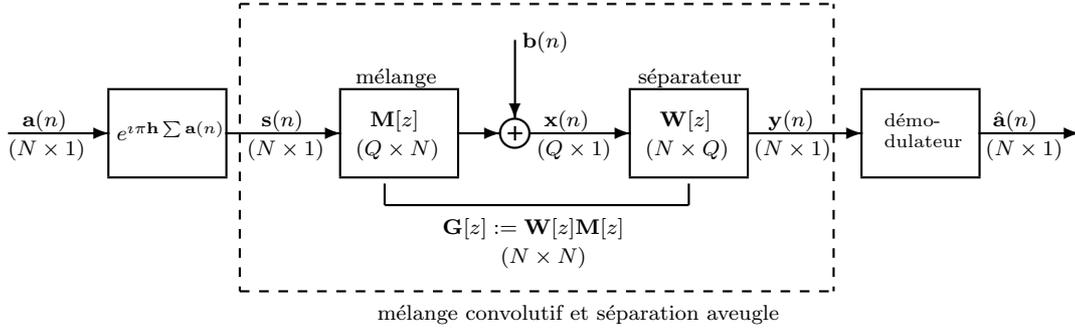


FIG. 4.1 – Système global à temps discret.

Dans cette application, nous considérons uniquement des approches séquentielles. Nous supposons être à la $P^{\text{ème}}$ étape de la séparation, qui a été achevée avec succès par maximisation de $J|_{\mathcal{F}_P}$ et éventuellement post-optimisation (voir le chapitre précédent). Ainsi, la $P^{\text{ème}}$ sortie $y_P(n)$ est une filtrée de la $P^{\text{ème}}$ source, ce que nous écrivons :

$$y_P(n) = \sum_{k \in \mathbb{Z}} g_P(k) s_P(n - k) \quad (4.13)$$

Pour atteindre notre objectif, qui est de détecter les symboles $(a_k^P)_{k \in \mathbb{Z}}$, il serait possible d'effectuer une égalisation du signal $(y_P(n))_{n \in \mathbb{Z}}$ et d'appliquer ensuite un algorithme de décision dure. Nous allons montrer dans la suite que cette étape peut être grandement simplifiée.

4.1.2-b) Contraste choisi et critère de module constant

Le choix effectué pour le critère J de l'équation (3.8) est $\mathcal{S} = \{(0, 0, 0)\}$ et f égale à moins l'identité. Ainsi, nous avons :

$$J(\mathbf{g}) = -\kappa_y(0, 0, 0) \quad (4.14)$$

Ce choix nous permet d'établir un lien entre J et le critère dit de module constant défini par :

$$J_{\text{CM}}(\mathbf{g}) := \mathbb{E}\{(|y(n)|^2 - 1)^2\} \quad (4.15)$$

Un tel lien a déjà été remarqué dans le cadre SISO [84]. Il est facile de généraliser ce résultat :

Lemme 3 *Si le filtre global vérifie $\|\mathbf{g}\| = 1$, alors :*

$$J(\mathbf{g}) = 1 - J_{\text{CM}}(\mathbf{g}) \quad (4.16)$$

De plus, pour tout $i \in \{1, \dots, N\}$, la borne supérieure définie dans (3.9) est atteinte pour le filtre identité et $\forall i \in \{1, \dots, N\}, \mathcal{M}_i = 1$.

Preuve: En supposant $\|\mathbf{g}\|^2 = \mathbb{E}\{|y(n)|^2\} = 1$ et en utilisant la circularité de $y(n)$ qui entraîne $\mathbb{E}\{y(n)^2\} = 0$, nous pouvons écrire :

$$\begin{aligned}\kappa_y(0, 0, 0) &= \mathbb{E}\{|y(n)|^4\} - 2\mathbb{E}\{|y(n)|^2\}^2 - |\mathbb{E}\{y(n)^2\}|^2 \\ &= \mathbb{E}\{|y(n)|^4\} - 2\mathbb{E}\{|y(n)|^2\} + 1 - 1 \\ &= \mathbb{E}\{(|y(n)|^2 - 1)^2\} - 1\end{aligned}\tag{4.17}$$

ce qui donne (4.16). D'autre part, il est clair d'après (4.15) et (4.16) que $J(\mathbf{g})$ est majoré par 1. Enfin, si $\mathbf{g} = (0, \dots, 0, z^0, 0, \dots, 0) \in \mathcal{G}_i$, nous avons $J_{CM}(\mathbf{g}) = 0$ et ainsi $J(\mathbf{g}) = 1$, ce qui prouve que la borne supérieure est atteinte et vaut $\mathcal{M}_i = 1$. ■

Le lemme ci-dessus justifie l'hypothèse H.11 qui affirme que la borne supérieure (3.9) est atteinte. En général, les autres travaux [88] ont considéré ce point comme une hypothèse technique supplémentaire.

4.1.2-c) Caractérisation de l'ambiguïté de filtrage scalaire

L'ensemble des filtres numériques qui, attaqués en entrée par un signal CPM, donnent une sortie de module constant a été caractérisé dans [10]. Ainsi, le lien établi entre J et le critère J_{CM} permet de caractériser l'ambiguïté de filtrage scalaire sur chacune des sources. On notera que les résultats sont distincts dans les cas où l'indice de modulation vaut $h_P = \frac{1}{2}$ et $h_P \neq \frac{1}{2}$.

Proposition 18 (cas $h_P \neq \frac{1}{2}$) *Supposons $h_P \neq \frac{1}{2}$. La suite $(g_P(k))_{k \in \mathbb{Z}}$ peut alors s'écrire comme*

$$g_P(k) = e^{i\varphi_P} \check{g}_P(k - k_P), \quad \forall k \in \mathbb{Z},\tag{4.18}$$

où k_P est un entier, $\varphi_P \in [-\pi, \pi)$ et où $(\check{g}_P(k))_{k \in \mathbb{Z}}$ coïncide avec l'une des suites suivantes :

$$\left\{ \begin{array}{l} \check{g}_P(0) = \frac{\sin \theta_P}{\sin(\pi h_P)} \\ \check{g}_P(1) = \frac{\sin(\pi h_P - \theta_P)}{\sin(\pi h_P)} \\ \check{g}_P(2) = 0 \end{array} \right. \quad \text{ou} \quad \left\{ \begin{array}{l} \check{g}_P(0) = \frac{\sin \theta_P}{\sin(\pi h_P)} \\ \check{g}_P(1) = \frac{e^{-i\theta_P}}{i \tan(\pi h_P)} \\ \check{g}_P(2) = \frac{i \cos \theta_P}{\sin(\pi h_P)} \end{array} \right.\tag{4.19}$$

et $\check{g}_P(k) = 0$ pour tout k différent de 0, 1 ou 2. Ici, θ_P prend n'importe quelle valeur dans l'intervalle $[-\pi, \pi[$.

Si maintenant $h_P = \frac{1}{2}$, le résultat suivant a été présenté dans [10] et est une conséquence directe de [41] :

Proposition 19 (cas $h_P = \frac{1}{2}$) *Supposons $h_P = \frac{1}{2}$. La suite $(g_P(k))_{k \in \mathbb{Z}}$ peut alors s'écrire comme :*

$$g_P(k) = e^{i\varphi_P} \check{g}_P(k - k_P), \quad \forall k \in \mathbb{Z},\tag{4.20}$$

où k_P est un entier, $\varphi_P \in [-\pi, \pi[$ et où $(\check{g}_P(k))_{k \in \mathbb{Z}}$ coïncide avec l'une des séquences suivantes :

$$\left\{ \begin{array}{l} \check{g}_P(0) = \sin \theta_P \\ \check{g}_P(K) = i^{-(K-1)} \cos \theta_P \end{array} \right.\tag{4.21}$$

où $\check{g}_P(k) = 0$ pour tout k différent de 0 et K . Ici, $\theta_P \in [-\pi, \pi[$ et K est un entier strictement positif.

A partir de maintenant, nous nous concentrons uniquement sur le cas $h_P \neq \frac{1}{2}$.

4.1.2-d) Commentaires

Nous dénommerons respectivement filtre de type I et type II les familles de filtres précédemment définies dans (4.19). Ainsi, on peut dire que le filtre $(g_P(k))_{k \in \mathbb{Z}}$ correspond à un filtre de type I ou II à un retard et un déphasage près.

Dans le cas où $(g_P(k))_{k \in \mathbb{Z}}$ correspond à un filtre de type I (à un retard et un déphasage près), le signal de sortie $(y_P(n))_{n \in \mathbb{Z}}$ est tel que pour tout entier n ,

$$y_P(n) = e^{i\varphi_P} \left(\frac{\sin \theta_P}{\sin(\pi h_P)} s_i(n - k_P) + \frac{\sin(\pi h_P - \theta_P)}{\sin(\pi h_P)} s_P(n - k_P - 1) \right) \quad (4.22)$$

où k_P , ϕ_P et θ_P sont les paramètres définis dans la proposition 18. Supposons maintenant que θ_P soit tel que $0 \leq \theta_P \leq \pi h_P$. Dans ce cas, θ_P peut s'écrire $\theta_P = \pi h_P \phi_P(\tau_P)$, où τ_P est un élément de l'intervalle de temps $[0, T_s]$ et où ϕ_P est la primitive de la fonction de mise en forme telle que définie à l'aide de (4.3). En utilisant la décomposition de Laurent (4.10) (ou bien l'équation (4.9)), on constate que l'équation (4.22) entraîne que pour tout entier n , $y_P(n) = e^{i\varphi_P} \underline{s}_P(nT_s - k_P T_s - \tau_P)$, où $\underline{s}_P(t)$ représente la source CPM à temps continu. En d'autres termes, le filtre $\sum_k g_P(k) z^{-k}$ peut s'interpréter comme un filtre interpolateur (à un déphasage φ_P près). Ainsi, une étape de synchronisation suivie d'un algorithme de détection de CPM classique sont les seuls éléments nécessaires pour reconstituer les symboles émis $(a_k^P)_{k \in \mathbb{Z}}$.

Comme le paramètre θ_P ne vérifie pas nécessairement $0 \leq \theta_P \leq \pi h_P$ et comme $(g_P(k))_{k \in \mathbb{Z}}$ peut également être un filtre de type II, l'estimation des symboles peut demander une procédure plus compliquée. Cependant, il est intéressant de noter que dans tous les cas, l'utilisation d'un égaliseur SISO permettant de récupérer $(s_P(n))_{n \in \mathbb{Z}}$ n'est pas indispensable. La proposition 18 donne une paramétrisation du filtre $(g_P(k))_{k \in \mathbb{Z}}$ et permet donc de remplacer une étape d'égalisation potentiellement coûteuse par une estimation plus simple des paramètres inconnus. Par exemple, un estimateur du maximum de vraisemblance peut être considéré en chaque point θ_P d'une grille discrète afin d'estimer simultanément θ_P et la suite de symboles. D'autres méthodes plus évoluées ont également été proposées parmi lesquelles on peut citer le «Per-Survivor Processing» [82].

4.1.3 Simulations

4.1.3-a) Méthodologie et description générale

Nous avons étudié d'abord la méthode de séparation en elle-même et ensuite envisagé la détection des différentes suites de symboles. Pour l'étude de la seule séparation, ont été pris en compte les critères

$$\tau_P := 1 - \frac{\max_{j \geq P} \|g_j\|_j^2}{\sum_j \|g_j\|_j^2} \quad (P \text{ fixé, } 1 \leq P \leq N) \quad (4.23)$$

où $\mathbf{g}[z] = (g_1[z], \dots, g_N[z])$ représente la $P^{\text{ème}}$ ligne du filtre global. Comme expliqué au paragraphe 3.5.2, nous avons $0 \leq \tau_P \leq 1$ et $\tau_P = 0$ uniquement dans le cas d'une séparation parfaitement réussie. Nous avons par ailleurs utilisé la moyenne $\tau = \frac{1}{N} \sum_{P=1}^N \tau_P$ comme critère de réussite de la séparation MIMO.

Compte tenu du contexte applicatif, nous avons considéré un modèle de canal classique en télécommunications. Les filtres de mélange ont été tirés selon un modèle de propagation multi-trajets réduit à trois trajets. Les facteurs d'atténuation ont été fixés à 1 pour le

premier trajet et tirés aléatoirement dans $[-1, 1]$ (loi uniforme) pour les autres trajets. Les retards de propagation ont été supposés nuls pour le premier trajet et tirés uniformément dans $[0, 3T_s]$ pour les autres trajets. Enfin, l'effet des angles d'arrivée a été modélisé par un déphasage complexe entre les capteurs, lequel a été tiré selon une loi uniforme sur $[0, 2\pi]$.

4.1.3-b) Validité de la méthode

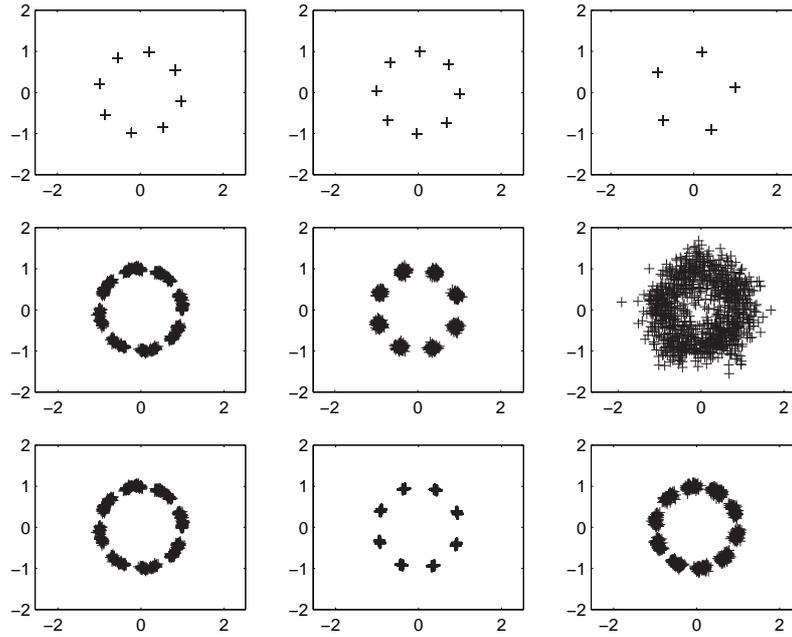


FIG. 4.2 – Sources originales (1^{ère} ligne) et reconstituées, sans post-optimisation locale (2^{ème} ligne) et avec (3^{ème} ligne) (Indices de modulations : 0,25 ; 0,75 et 0,4 ; 4 capteurs, filtre de mélange de longueur 4, 1000 échantillons.)

La figure 4.2 représente les constellations de $N = 3$ sources ainsi que les constellations obtenues après séparation de leur mélange sur $Q = 4$ capteurs. Nous constatons clairement que la méthode de séparation réussit particulièrement bien pour les deux premières sources (colonnes 1 et 2), contrairement à la troisième source dont la restauration semble de moindre qualité.

En outre, la procédure post-optimisation expliquée au paragraphe 3.4.5 mène à des résultats de qualité bien meilleure comme le montre la dernière ligne de la figure 4.2. Cet exemple illustre que, au cours des étapes successives, l'estimation des ensembles $\mathcal{F}_i, i \in \{2, \dots, N\}$ devient de moindre qualité. Il s'agit d'un point qu'il convient de traiter avec soin.

On remarquera que les sources extraites sont de module unité, témoignant de ce que l'algorithme, basé sur une méthode d'optimisation locale, a en réalité convergé vers une solution de module un : il s'agit donc d'un maximum global du critère J . Il est intéressant de noter aussi que les sources sont bien reconstituées à un filtrage près de type I tel que défini à l'équation (4.19) : ceci s'observe particulièrement bien à la dernière ligne, dernière colonne de la figure 4.2, où l'on distingue clairement 10 groupes de points au lieu des 5 groupes représentés en haut de la même colonne et correspondant à un indice de

modulation de 0,4. Le même phénomène s'est produit dans la seconde colonne, mais le paramètre θ étant proche de zéro (ou π), les différents groupes de points ne peuvent être clairement distingués.

4.1.3-c) Difficulté du cas des sources CPM

Nous avons illustré précédemment que la méthode de séparation hiérarchique était susceptible d'accumuler des erreurs. De ce fait, comme les ensembles $\mathcal{F}_i, i \in \{1, \dots, N\}$ ne sont pas connus exactement mais seulement estimés, des difficultés peuvent survenir lorsque l'on essaie d'imposer le nombre requis de contraintes. Nous avons tracé sur la figure 4.3 deux exemples typiques de distribution des valeurs singulières de la matrice de contrainte (*i.e.* le vecteur des paramètres recherchés appartient au noyau de cette matrice, qui avait été notée \mathbf{A}_1 dans le paragraphe 3.4.4). Ces exemples correspondent aux mélanges de deux sources i.i.d. et de deux sources CPM respectivement. Sur la figure 4.3(a), les

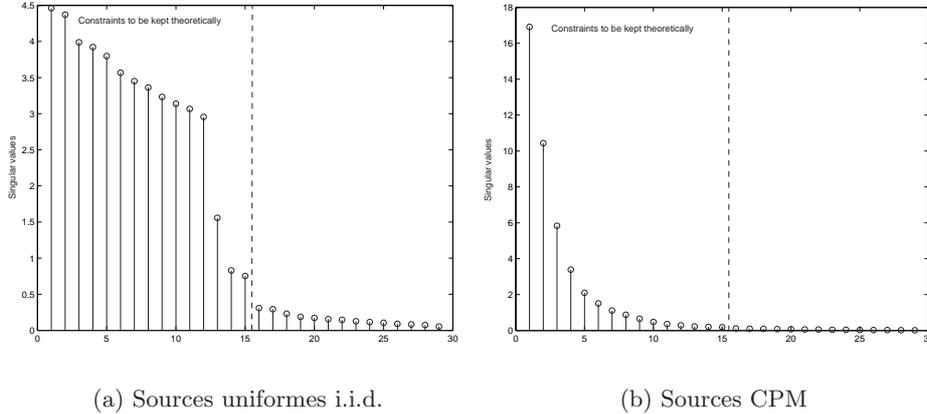


FIG. 4.3 – Deux exemples typiques des distributions des valeurs singulières de la matrice de contrainte (obtenus avec un filtre de mélange identique, 10000 échantillons et deux types de sources distincts).

valeurs singulières peuvent être séparées en deux groupes et l'espace vectoriel engendré par les vecteurs correspondant aux valeurs singulières les plus petites permet d'aboutir à une estimation fiable de l'ensemble \mathcal{F}_i cherché. En comparaison, ce phénomène ne se produit pas pour des sources CPM et il n'apparaît pas de saut dans la distribution des valeurs singulières de la figure 4.3(b). Dès lors, il ne peut être exclu que l'estimation des ensembles \mathcal{F}_i soit fortement compromise. Ainsi, le nombre idéal de contraintes à prendre en compte résulte d'un compromis et le nombre théorique déterminé au paragraphe C.4 n'apparaît plus que comme une indication. Prendre en compte plus de contraintes limitera le risque de retrouver plusieurs fois sur une source déjà extraite, mais limitera fortement la qualité de la séparation. En revanche, choisir moins de contraintes mènera à une séparation de meilleure qualité, mais à un risque accru d'extraire une source déjà extraite.

4.1.3-d) Etude sur deux filtres fixés

Nous illustrons ici sur deux filtres fixés les possibilités offertes par la méthode hiérarchique de séparation sur des sources CPM. Les performances sont également comparées à

celles offertes par une méthode de déflation.

Le premier filtre considéré comporte 4 capteurs pour 3 sources en entrée et sa longueur est de 4. D'après le paragraphe 3.4, le filtre séparant a été choisi de longueur 9 et le nombre de contraintes d'orthogonalité à imposer en théorie est de 12 par source précédemment extraite. Nous avons représenté sur la figure 4.4 les fonctions de répartition du critère de séparation τ pour respectivement 4 et 9 conditions d'orthogonalité par source extraite. Nous constatons sur la figure 4.4(a) que, si le nombre de contraintes imposées est suffisamment faible, la méthode hiérarchique offre des performances très satisfaisante en comparaison de la déflation. Une post-optimisation améliore encore les performances de la séparation. En revanche, lorsque le nombre de contraintes d'orthogonalité est plus important, la qualité de la séparation est moindre, comme indiqué par la figure 4.4(b).

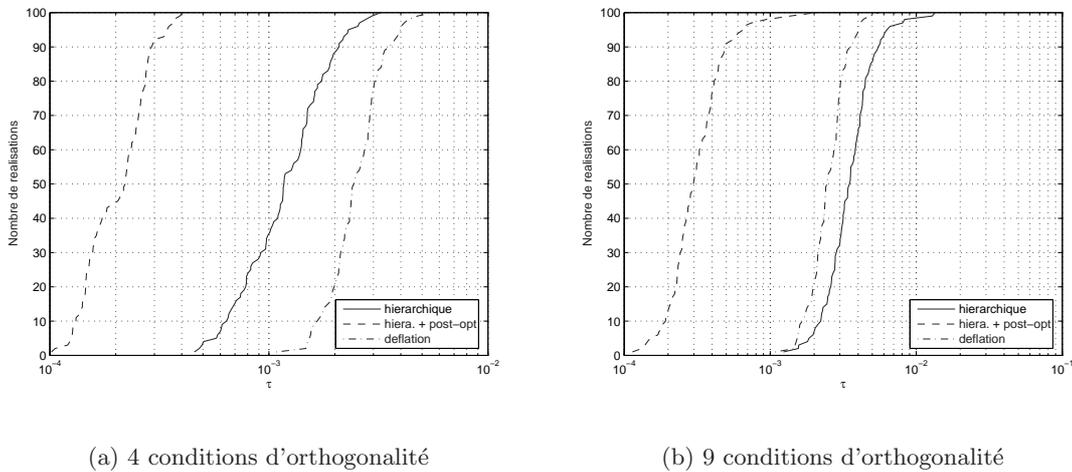
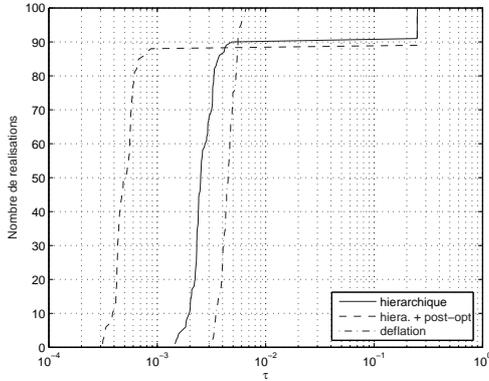
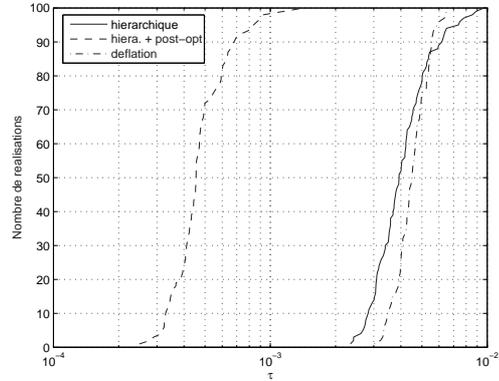


FIG. 4.4 – Fonction de répartition du critère $\frac{\sum_{i=1}^3 \tau_i}{3}$ pour un filtre 4×3 fixé. (4000 échantillons)

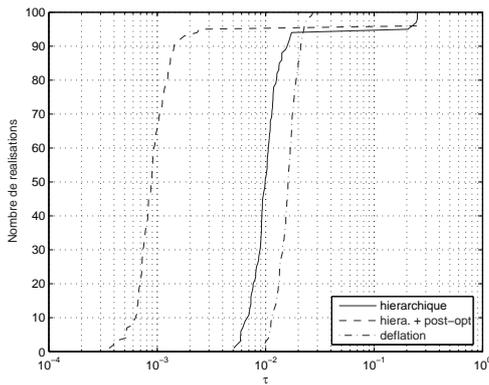
Afin de mieux cerner les limites de la méthode hiérarchique, il est intéressant de se placer dans le cas d'un mélange de 4 sources sur 5 capteurs et de longueur 4 (figure 4.5 et tableau 4.1). Nous remarquons que la méthode hiérarchique offre, pour 4000 échantillons disponibles, de bons résultats lorsque 9 conditions d'orthogonalité sont prises en compte par source (figure 4.5(b)). Un nombre plus faible de conditions d'orthogonalité divise en revanche les réalisations en deux catégories (figure 4.5(a)) : une majorité d'entre elles donne une séparation meilleure que dans le cas d'une déflation, tandis que dans une minorité de cas, la séparation a échoué. L'apparition de ces cas d'échec complet de la séparation se remarque aussi dans le cas d'une diminution du nombre d'échantillons (figure 4.5(d)). L'observation des tableaux 4.1 et 4.2 montre par ailleurs que la difficulté provient de l'estimation de la dernière source. Les premières sources sont en effet correctement estimées, avec une accumulation d'erreur inférieure à celle de la déflation. Dans le cas d'un nombre trop faible de contraintes, il existe en revanche une minorité de cas où la dernière source est l'une de celles précédemment extraite. Ceci se traduit à la fois par une très forte dégradation de la valeur moyenne du critère de séparation de la dernière source, et par une dégradation plus mesurée de la valeur médiane.



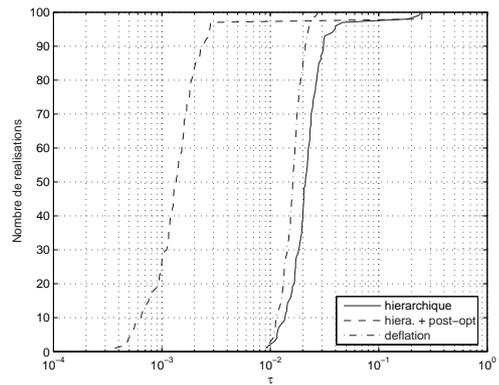
(a) 6 conditions d'orthogonalité, 4000 échantillons



(b) 9 conditions d'orthogonalité, 4000 échantillons



(c) 6 conditions d'orthogonalité, 1000 échantillons



(d) 9 conditions d'orthogonalité, 1000 échantillons

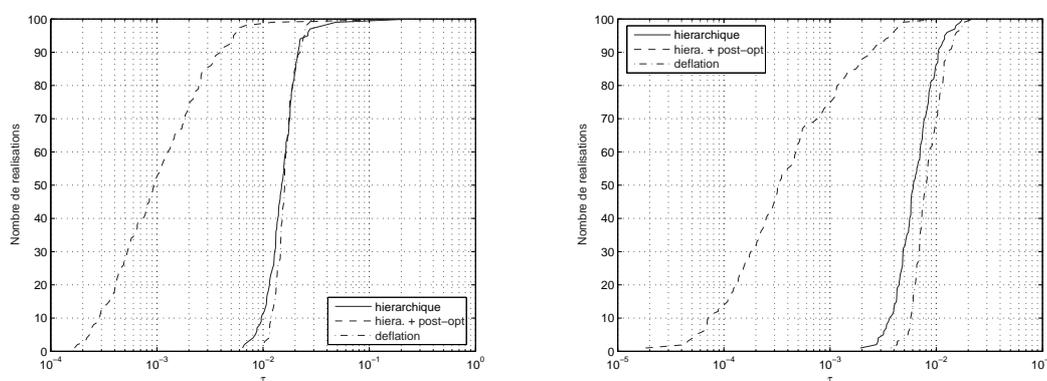
FIG. 4.5 – Fonction de répartition du critère $\frac{\sum_{i=1}^4 \tau_i}{4}$ pour un filtre 5×4 fixé.

	6 cond. orth.			9 cond. orth.		
	hiéra	hiéra+ post-opt.	déflation	hiéra.	hiéra + post-opt.	déflation
1 ^{ère} source : τ_1	0.0004	0.0004	0.0004	0.0004	3.695e-3	0.0004
2 ^{ème} source : τ_2	0.0004	0.0002	0.0034	0.0011	3.734e-3	0.0034
3 ^{ème} source : τ_3	0.0044	0.0003	0.0063	0.0068	3.350e-3	0.0063
4 ^{ème} source : τ_4	0.1042	0.1209	0.0082	0.0088	9.136e-3	0.0082

TAB. 4.1 – Valeurs moyennes des critères $\tau_i, i \in \{1, \dots, 4\}$ pour un filtre 5×4 fixé, 4000 échantillons, 100 réalisations de Monte-Carlo.

4.1.3-e) Etude de Monte-Carlo avec canal aléatoire

Comme indiqué précédemment, le nombre de conditions d'orthogonalité sélectionnées a un impact important sur la qualité de séparation. Aussi, de nombreuses possibilités permettent de faire le choix des valeurs singulières à conserver ou non. A la figure 4.6 sont donnés les résultats pour des filtres tirés aléatoirement et où seules les contraintes correspondant à des valeurs singulières supérieures à un seuil ont été conservées. Nous constatons que la méthode hiérarchique est susceptible de fournir de bon résultats sous réserve de bien effectuer le choix des conditions d'orthogonalité. Remarquons aussi que l'étape de post-optimisation améliore considérablement les résultats. Enfin, ces simulations témoignent également du fait que les résultats précédents ne sont pas particuliers aux filtres choisis.



(a) 4 sources, limite des valeurs singulières conservées : 1

(b) 3 sources, limite des valeurs singulières conservées : 0.7

FIG. 4.6 – Fonction de répartition du critère $\frac{\sum_{i=1}^N \tau_i}{N}$ pour des filtres 4×3 et 5×4 générés aléatoirement. Le nombre d'échantillons a été fixé à 1000.

4.1.3-f) Cas réel bruité et taux d'erreurs

Nous avons envisagé la détection des symboles dans le cas favorable du mélange de la figure 4.4(a) avec 4 conditions d'orthogonalité. Il convient dans le cadre d'une application de prendre en compte le bruit. Aussi, le tableau 4.3 donne pour différents rapports signal à bruit les critères de séparation moyen obtenus. Le tableau 4.4 donne quant à lui les taux d'erreurs estimés à partir de 1000 réalisations. Nous constatons que les résultats sont de bonne qualité, ce qui permet d'envisager des applications concrètes.

	6 cond. orth.			9 cond. orth.		
	hiéra	hiéra+ post-opt.	déflation	hiéra.	hiéra + post-opt.	déflation
1 ^{ère} source : τ_1	0.0004	0.0004	0.0004	0.0004	3.721e-4	0.0004
2 ^{ème} source : τ_2	0.0003	0.0001	0.0031	0.0011	1.052e-4	0.0031
3 ^{ème} source : τ_3	0.0041	0.0002	0.0063	0.0064	2.300e-4	0.0063
4 ^{ème} source : τ_4	0.0051	0.0011	0.0079	0.0077	9.770e-4	0.0079

TAB. 4.2 – Valeurs médiane des critères $\tau_i, i \in \{1, \dots, 4\}$ pour un filtre 5×4 fixé, 4000 échantillons, 100 réalisations de Monte-Carlo.

SNR en dB	Hiéra.				Hiéra. + post-opt.			
	25	20	15	10	25	20	15	10
1 ^{ère} source : τ_1	8.3e-4	1.6e-3	3.5e-3	7.1e-3	8.3e-4	1.6e-3	3.5e-3	7.1e-3
2 ^{ème} source : τ_2	1.0e-3	2.4e-3	4.5e-3	7.1e-3	9.4e-4	2.3e-3	4.6e-3	7.8e-3
3 ^{ème} source : τ_3	1.3e-2	1.6e-2	2.0e-2	3.5e-2	1.2e-3	3.1e-3	8.3e-3	2.3e-2

TAB. 4.3 – Valeur moyenne des critères $\tau_i, i \in \{1, \dots, 3\}$ pour un filtre de mélange 4×3 fixé et pour différentes valeurs du SNR. Le nombre de réalisations de Monte-Carlo est de 100.

SNR en dB	Hiéra.			Hiéra. + post-opt.		
	∞	25	15	∞	25	15
1 ^{ère} source : τ_1	1.4e-5	5.8e-5	2.6e-3	1.4e-5	5.8e-5	2.6e-3
2 ^{ème} source : τ_2	2.9e-5	8.1e-5	4.6e-3	1.8e-5	8.3e-5	4.2e-3
3 ^{ème} source : τ_3	1.4e-5	2.4e-5	1.9e-3	2.0e-6	1.9e-5	0.7e-3

TAB. 4.4 – Estimation du BER pour un filtre de mélange 4×3 fixé et pour différents SNR. Le nombre de réalisations de Monte-Carlo est de 1000.

4.2 Séparation de mélanges convolutifs d'images

Nous considérons dans cette partie une application de la séparation de sources aux mélanges linéaires d'images. Deux difficultés se sont principalement présentées au cours de cette étude.

Nous avons d'abord dû étudier l'inversibilité des mélanges MIMO convolutifs 2D. Ce dernier problème peut se ramener à des propriétés des matrices polynomiales à deux indéterminées. Même si leur accès n'est pas aisé dans la littérature, les résultats requis existent quel que soit le nombre d'indéterminées : ils sont présentés dans l'annexe A mais les points utiles peuvent être admis sans nuire à la compréhension du texte. Après résolution des difficultés propres au cas 2D, nous avons pu étendre les algorithmes de séparation 1D au cas des images.

L'autre problème auquel nous avons été confrontés est que le contenu spectral des images se révèle en général non uniforme, notamment dans le cas d'images naturelles. Il est donc apparu indispensable de vérifier que nos méthodes s'appliquent dans le cas de telles sources. Par delà l'aspect théorique, il convient d'être conscient des limites qu'imposent des sources au spectre très inégalement réparti.

Enfin, des simulations illustreront la validité des points théoriques qui ont été soulevés.

4.2.1 Bref aperçu des méthodes de séparation/restauration en image

La séparation de sources a soulevé un certain intérêt dans le domaine du traitement d'images. En général, le modèle exploité reste toutefois celui des mélanges instantanés et de l'analyse en composantes indépendantes. Parmi les nombreuses applications de l'analyse en composantes indépendantes dans le domaine des images, on peut citer l'extraction de composantes modélisant le fonctionnement neuronal du cerveau [44], la séparation du rayonnement fossile dit CMB (Cosmic Microwave Background) [16], ...

Le problème de la déconvolution d'images appartient à un domaine différent de celui de la séparation de source et qui a été exploré depuis plus longtemps. Dans le cas SISO, de nombreuses méthodes ont été proposées. Elles font généralement appel à une régularisation ou une information *a priori* (voir [55] pour une revue des méthodes existantes). Par ailleurs, il est dans certain cas possible d'utiliser l'acquisition d'une même image sur différents canaux : cette configuration de type SIMO (une seule et même image se trouve sur plusieurs capteurs) a également donné lieu à un certain nombre de contributions, notamment dans le cas aveugle (voir [37, 95] et les références qui y sont citées). A notre connaissance, le cas MIMO convolutif n'avait en revanche encore jamais été abordé pour des images.

4.2.2 Inversibilité des mélanges

Afin de ne pas laisser d'ambiguïté de notation, nous reformulons brièvement le problème de la séparation de mélanges convolutifs de sources dans le cas bidimensionnel¹. Les N sources, mutuellement indépendantes, sont notées $(s_i(n_1, n_2))_{(n_1, n_2) \in \mathbb{Z}^2}$ pour $i \in \{1, \dots, N\}$ et sont mélangées par un filtre MIMO de réponse impulsionnelle $(\mathbf{M}(n_1, n_2))_{(n_1, n_2) \in \mathbb{Z}^2}$ suivant le modèle :

$$\mathbf{x}(n_1, n_2) = \sum_{(k_1, k_2) \in \mathbb{Z}^2} \mathbf{M}(k_1, k_2) \mathbf{s}(n_1 - k_1, n_2 - k_2) \quad (4.24)$$

¹Compte tenu du fait que les résultats de l'annexe A sont valables dans le cas général multidimensionnel, une extension à des signaux nD (avec $n > 2$) est immédiate sur le plan théorique. Pour des raisons de lisibilité, ceci est laissé au soin du lecteur.

Le vecteur $(\mathbf{x}(n_1, n_2))_{(n_1, n_2) \in \mathbb{Z}^2}$ est composé des observations sur les $Q \geq N$ capteurs et, à partir de sa seule connaissance, le problème de la séparation recherche un séparateur $(\mathbf{W}(n_1, n_2))_{(n_1, n_2) \in \mathbb{Z}^2}$ tel que la sortie globale permette de restituer les sources (aux ambiguïtés près : filtrage scalaire, permutation) et s'écrive :

$$\mathbf{y}(n_1, n_2) = \sum_{(k_1, k_2) \in \mathbb{Z}^2} \mathbf{W}(k_1, k_2) \mathbf{x}(n_1 - k_1, n_2 - k_2) \quad (4.25)$$

De même que dans le cas unidimensionnel, on définit la transformée en z bidimensionnelle du système de mélange par :

$$\mathbf{M}[z_1, z_2] = \sum_{(k_1, k_2) \in \mathbb{Z}^2} \mathbf{M}(k_1, k_2) z_1^{-k_1} z_2^{-k_2} \quad (4.26)$$

La transformée en z du système séparant est définie de façon analogue. Dans le cas général de sources non i.i.d., on souhaite alors trouver $\mathbf{W}[z_1, z_2]$ tel que le filtre global $\mathbf{G}[z_1, z_2] := \mathbf{W}[z_1, z_2] \mathbf{M}[z_1, z_2]$ soit «diagonal à une permutation près».

Dans le cas de filtres unidimensionnels et RIF, il a été possible d'assurer l'existence d'un tel inverse. Ce résultat se généralise au cas bidimensionnel :

Propriété 2 *Supposons que $\mathbf{M}[z_1, z_2]$ soit RIF. Alors, $\mathbf{M}[z_1, z_2]$ admet un inverse RIF si et seulement si le nombre de capteurs est supérieur ou égal au nombre de sources ($Q \geq N$) et si de plus :*

$$\forall (z_1, z_2) \in \mathbb{C}^{*2} \quad \text{Rang}(\mathbf{M}[z_1, z_2]) = N \quad (4.27)$$

Cette propriété découle directement de résultats sur les matrices polynomiales à plusieurs indéterminées [35] (voir le paragraphe A.3 de l'annexe pour plus d'informations).

Dans le cas d'un filtre de mélange obtenu aléatoirement, il est intéressant de savoir si son inversibilité constitue une caractéristique vraisemblable. Des éléments de réponse ont été donnés dans [83] et en particulier, la propriété suivante s'applique (voir le paragraphe 1.3.3-c) de l'annexe A) :

Propriété 3 *Supposons que le filtre $\mathbf{M}[z_1, z_2]$ de taille $Q \times N$ soit RIF. Si ses coefficients sont tirés selon une lois de probabilité à densité continue, et si $Q > N$, alors $\mathbf{M}[z_1, z_2]$ est presque sûrement inversible à gauche par un filtre RIF.*

Pour terminer, signalons également notre intérêt à connaître l'ordre minimal d'un inverse en cas d'existence. Cette question a déjà été en partie abordée et, dans certains cas, des majorants de l'ordre minimal ont été donnés [83]. La question de l'obtention de bornes minimales semble toutefois encore ouverte.

4.2.3 Application des méthodes de séparation séquentielles dans le cas des images

Nous supposons dans ce paragraphe que le filtre de mélange admet un inverse à gauche de réponse impulsionnelle finie. Nous allons alors montrer que les procédures de séparation utilisées dans le cas unidimensionnel peuvent aisément être utilisées pour la séparation d'images.

Considérons dans un premier temps l'extraction d'une source à partir des observations $(\mathbf{x}(n_1, n_2))_{(n_1, n_2) \in \mathbb{Z}^2}$. De même que dans le cas unidimensionnel, nous noterons $\mathbf{w}[z_1, z_2]$

une ligne du filtre séparant $\mathbf{W}[z_1, z_2]$. $(y(n_1, n_2))_{(n_1, n_2) \in \mathbb{Z}^2}$ est la sortie correspondante, ce que l'on écrit : $y(n_1, n_2) = \mathbf{w}[z_1, z_2] \mathbf{x}(n_1, n_2)$.

Dans le cas où le système séparant $\mathbf{w}[z_1, z_2]$ est RIF et causal², il existe des entiers positifs D_1 et D_2 qui caractérisent la longueur du filtre séparant³, c'est-à-dire tels que :

$$w(k_1, k_2) = 0 \text{ pour } (k_1, k_2) \notin \{0, \dots, D_1 - 1\} \times \{0, \dots, D_2 - 1\} \quad (4.28)$$

$\mathbf{w}[z_1, z_2]$ est donc défini par ses coefficients $\mathbf{w}(k_1, k_2)$ pour $(k_1, k_2) \in \{0, \dots, D_1 - 1\} \times \{0, \dots, D_2 - 1\}$. Il est alors possible de définir un vecteur ligne $\underline{\mathbf{w}}$ et, pour tout (n_1, n_2) , un vecteur colonne $\underline{\mathbf{x}}(n_1, n_2)$ qui contiennent les termes respectifs $w_j(k_1, k_2)$ et $x_j(n_1 - k_1, n_2 - k_2)$ lorsque (k_1, k_2) varie dans $\{0, \dots, D_1 - 1\} \times \{0, \dots, D_2 - 1\}$ et j varie dans $\{1, \dots, Q\}$. La sortie du séparateur peut alors s'exprimer ainsi :

$$\forall (n_1, n_2) \quad y(n_1, n_2) = \underline{\mathbf{w}} \underline{\mathbf{x}}(n_1, n_2) \quad (4.29)$$

Dans le cas unidimensionnel, nous avons de la même façon :

$$\forall n \quad y(n) = \underline{\mathbf{w}} \underline{\mathbf{x}}(n) \quad (4.30)$$

La méthode de séparation justifiée au chapitre 3 par la proposition 9 et les suivantes (propositions 11, 12, 13 et 14) utilise une optimisation d'un critère de contraste ne dépendant que de $y(n)$. L'optimisation était effectuée par rapport aux paramètres du vecteur $\underline{\mathbf{w}}$ à l'aide d'un gradient, en prenant soin de le renormaliser à chaque itération de sorte que la sortie globale soit de puissance unité.

Il ressort donc que la procédure de séparation 1D ne requiert pour son utilisation que les réalisations du vecteur $\underline{\mathbf{x}}(n)$. Compte tenu de l'écriture similaire des équations (4.29) et (4.30), nous constatons que les procédures de séparation 1D utilisées dans les méthodes séquentielles sont utilisables pour la séparation 2D à la condition d'adapter au préalable les procédés de construction de $\underline{\mathbf{w}}$ et de $\underline{\mathbf{x}}(n)$.

On peut constater que le problème de séparation bidimensionnelle perd sa spécificité une fois effectuée la traduction pour une résolution par des procédures 1D. En particulier, toute éventuelle structure satisfaite par des filtres bidimensionnels (telle que l'isotropie, la symétrie, ...) est ignorée par les procédures 1D. Naturellement, dans le cas où une telle information *a priori* est disponible, il est tout à fait judicieux d'essayer de la prendre en compte et de l'incorporer au modèle du filtre.

4.2.4 Problèmes liés à l'occupation spectrale des sources

Les images naturelles sont fortement corrélées et leur spectre est principalement concentré sur les basses fréquences. Aussi, nous considérons ici le cas limite où les sources n'occupent pas toute la bande spectrale existante mais où leur spectre est positif sur un ensemble Ω et s'annule sur son complémentaire $\bar{\Omega}$. Pour des raisons de facilité de notation, nous revenons temporairement (uniquement dans les paragraphes 4.2.4 et 4.2.5-a)) à des notations unidimensionnelles. Le cas étudié revient donc à supposer :

²La causalité n'est ici supposée que par facilité de notation et peut être obtenue par un simple décalage de la réponse impulsionnelle puisque cette dernière est supposée finie.

³Des valeurs de D_1, D_2 qui assurent l'existence d'un inverse sont données dans certains cas et en fonction des différents paramètres dans [83]. Ces bornes ne sont cependant pas minimales et nous renvoyons le lecteur à la partie simulations pour connaître les ordres de grandeur menant à des résultats satisfaisants.

H.14 Pour tout $i \in \{1, \dots, N\}$, la densité spectrale de puissance de la $i^{\text{ème}}$ source vérifie :

$$\begin{cases} \forall \omega \in \Omega & \Gamma_i(\omega) > 0 \\ \forall \omega \in \bar{\Omega} & \Gamma_i(\omega) = 0 \end{cases} \quad (4.31)$$

Cette hypothèse remplace et affaiblit l'hypothèse H.10 faite au chapitre précédent, paragraphe 3.1.2. Si l'on considère le filtre global $\mathbf{g}[z] := (g_1[z], \dots, g_N[z])$, et que pour $i \in \{1, \dots, N\}$ quelconque, on pose :

$$\|g_i\|_i := \left(\sum_{k,l} g_i(k) g_i^*(l) \gamma_i(l-k) \right)^{\frac{1}{2}} \quad (4.32)$$

il apparaît que $\|\cdot\|_i$ ne définit une norme que lorsque $\bar{\Omega}$ est de mesure nulle. Les résultats du chapitre précédent s'appliquent uniquement dans ce cas, car il est alors possible d'écrire pour tout i et pour tout filtre $g_i[z]$:

$$g_i[z] = \|g_i\|_i \tilde{g}_i[z] \quad (4.33)$$

où l'on pose de même qu'à l'équation (3.12) :

$$\tilde{g}_i[z] = \begin{cases} \frac{g_i[z]}{\|g_i\|_i} & \text{si } \|g_i\|_i \neq 0, \\ 0 & \text{si } \|g_i\|_i = 0. \end{cases} \quad (4.34)$$

Cette approche n'est pas valable en l'état dans le cas où $\bar{\Omega}$ est de mesure non nulle, la division par $\|g_i\|_i$ n'étant pas toujours possible ci-dessus. En conséquence, nous proposons de décomposer le filtre global en :

$$\mathbf{g}[z] = \mathbf{g}^\Omega[z] + \mathbf{g}^{\bar{\Omega}}[z] \text{ avec :} \quad (4.35)$$

$$\mathbf{g}^\Omega[e^{i\omega}] = \begin{cases} \mathbf{g}[e^{i\omega}] & \text{si } \omega \in \Omega, \\ 0 & \text{si } \omega \in \bar{\Omega}. \end{cases} \quad \text{et :} \quad \mathbf{g}^{\bar{\Omega}}[e^{i\omega}] = \begin{cases} 0 & \text{si } \omega \in \Omega, \\ \mathbf{g}[e^{i\omega}] & \text{si } \omega \in \bar{\Omega}. \end{cases} \quad (4.36)$$

Compte tenu de l'hypothèse H.14, nous avons pour tout $n \in \mathbb{Z}$:

$$\mathbf{g}^\Omega[z] \mathbf{s}(n) = \mathbf{g}[z] \mathbf{s}(n) \quad \text{et :} \quad \mathbf{g}^{\bar{\Omega}}[z] \mathbf{s}(n) = 0 \quad (4.37)$$

et de plus :

$$\|\mathbf{g}\| = \|\mathbf{g}^\Omega\|. \quad (4.38)$$

Nous pouvons donc reprendre les démonstrations du chapitre précédent à condition de ne considérer que le filtre $\mathbf{g}^\Omega[z] := (g_1^\Omega[z], \dots, g_N^\Omega[z])$. Ainsi, les contrastes habituels sont maximum si et seulement si $\mathbf{g}^\Omega[z]$ est séparant, ce que nous résumons dans la proposition suivante :

Proposition 20 *Les résultats du chapitre 3 (propositions 11, 12, 13, 14, 15 et 16) demeurent valables si l'on remplace l'hypothèse H.10 par l'hypothèse H.14 à condition de remplacer également les ensembles $\mathcal{G}_i, i \in \{1, \dots, N\}$ et $\mathcal{F}_i, i \in \{1, \dots, N\}$ par $\mathcal{G}_i \cap \mathfrak{G}, i \in \{1, \dots, N\}$ et $\mathcal{F}_i \cap \mathfrak{G}, i \in \{1, \dots, N\}$ respectivement, où \mathfrak{G} représente l'ensemble des filtres globaux tels que $\mathbf{g}^{\bar{\Omega}}[z] = 0$.*

Remarque 11: L'extension de cette propriété au cas multidimensionnel est immédiate après adaptation des notations.

Remarque 12: Par utilisation de la proposition ci-dessus, nous savons que les procédures séquentielles donnent un filtre tel que $\mathbf{g}^\Omega[z]$ soit séparant. En revanche, nous n'avons aucune information concernant $\mathbf{g}^{\bar{\Omega}}[z]$, qui peut être quelconque. En raison de (4.37), ceci n'a toutefois aucune influence sur la sortie du séparateur, dans le cas idéal où la densité spectrale des sources s'annule strictement sur $\bar{\Omega}$. Cependant, on peut imaginer que, dans le cas où la densité spectrale ne s'annule pas strictement, l'absence de contrôle de $\mathbf{g}^{\bar{\Omega}}[z]$ puisse être indésirable et amplifier fortement des composantes de bruit.

4.2.5 Simulations

Nous illustrons d'abord l'effet d'une bande limitée des signaux sources. Ensuite, les simulations ont été réalisées sur des images offrant une difficulté croissante pour la séparation : deux images i.i.d. d'abord, le mélange d'une image i.i.d. et d'une image naturelle ensuite, et enfin le mélange de deux images naturelles.

4.2.5-a) Séparation de sources de bande non pleine

Afin de distinguer plus nettement les seuls effets liés à la considération de sources qui n'occupent pas toute la bande spectrale, nous avons mené quelques simulations sur des signaux unidimensionnels pour des filtres de longueur $L = 3$. Des sources ont été générées par filtrage de sorte que seule une certaine proportion de la bande spectrale soit occupée : après tirage de signaux selon une loi i.i.d. uniforme centrée, les coefficients des transformées de Fourier discrètes respectives ont été artificiellement mis à zéro pour des fréquences supérieures à un certain seuil.

La figure 4.7 représente un cas typique obtenu dans le cas d'une séparation de deux signaux de 16384 échantillons et qui n'occupent que 30% de la bande spectrale. Sont représentés sur la figure 4.7 les périodogrammes des deux sources présentes dans le mélange ainsi que les composantes du filtre ligne global, $g_1[e^{i\omega}]$ et $g_2[e^{i\omega}]$ respectivement.

Nous constatons que le filtre global $\mathbf{g}[z] = (g_1[z], g_2[z])$ est séparant dans la mesure où, sur la bande occupée par les sources, nous avons sensiblement :

$$\forall \omega \in \Omega \quad |g_1[e^{i\omega}]| = |g_1^\Omega[e^{i\omega}]| \approx 1 \quad (4.39)$$

$$|g_2[e^{i\omega}]| = |g_2^\Omega[e^{i\omega}]| \approx 0. \quad (4.40)$$

En revanche, pour $\omega \in \bar{\Omega}$, nous constatons des valeurs relativement élevées de $g_1^{\bar{\Omega}}[e^{i\omega}]$ et $g_2^{\bar{\Omega}}[e^{i\omega}]$. Ceci confirme la discussion théorique du paragraphe 4.2.4 : $\mathbf{g}^\Omega[z]$ est séparant tandis que $\mathbf{g}^{\bar{\Omega}}[z]$ est quelconque.

Le tableau 4.5 donne, pour divers pourcentages de la bande occupée et différents nombres d'échantillons, la valeur du critère τ_1 défini à l'équation (4.23) et obtenu par moyennage sur 10 réalisations distinctes des sources et du canal (tiré aléatoirement, avec 3 capteurs, 2 sources). Le graphique 4.8 représente ce même critère τ_1 en fonction du nombre d'échantillons et du pourcentage de bande occupée ; il correspond donc aux valeurs du tableau 4.5 et en donne une représentation visuelle. Nous constatons que, sous réserve d'augmenter de façon importante le nombre d'échantillons, la séparation est effective même pour une très faible occupation de la bande spectrale. Cependant, compte

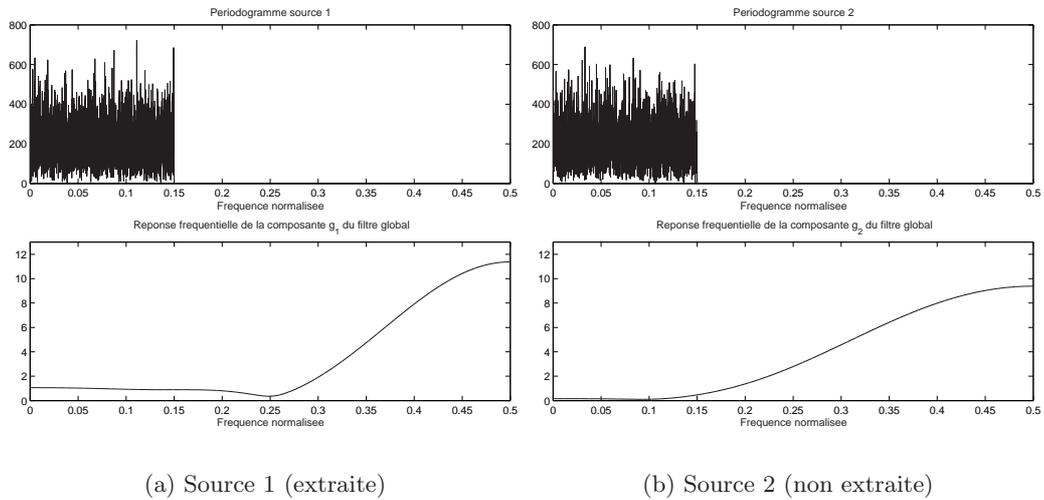


FIG. 4.7 – Exemple typique de résultat de séparation dans le cas de sources n’occupant pas toute la bande : observations au niveau spectral (16384 échantillons).

tenu du procédé utilisé pour la génération des sources, il a été observé que celles-ci se rapprochaient de la gaussianité pour une faible occupation de la bande spectrale. Ce point explique que des résultats corrects de la séparation ne puissent être obtenus que difficilement dans le cas d’une faible bande occupée.

échant \ % bande		% bande				
		0.9	0.7	0.5	0.3	0.1
	2^{11}	0.0061	0.0220	0.0710	0.2389	0.2870
	2^{12}	0.0030	0.0111	0.0513	0.1614	0.2429
	2^{13}	0.0020	0.0078	0.0185	0.1640	0.2579
	2^{14}	0.0045	0.0048	0.0132	0.0480	0.2418
	2^{15}	0.0004	0.0009	0.0049	0.0185	0.2199
	2^{16}	0.0003	0.0006	0.0020	0.0126	0.1260
	2^{17}	0.0008	0.0004	0.0012	0.0059	0.1179

TAB. 4.5 – Valeur du critère τ_1 en fonction du nombre d’échantillons et du pourcentage occupé de la bande.

Il ressort donc de ces quelques simulations que la difficulté de séparation peut être sérieusement accrue dans le cas des images naturelles dont le spectre est concentré dans les basses fréquences. Nous constaterons en particulier ce même phénomène d’amplification des hautes fréquences de la source séparée.

4.2.5-b) Cas d’images sources i.i.d.

L’étude a été réalisée sur des ensembles de 100 réalisations de Monte-Carlo. Pour chacune des réalisations, les coefficients du filtre de mélange ont été tirés selon une loi gaussienne, centrée de variance unité. Compte tenu de la taille du filtre MIMO — $N = 2$ sources et $Q = 3$ capteurs—, ceci assure d’après la propriété 2 l’existence presque sûre d’un

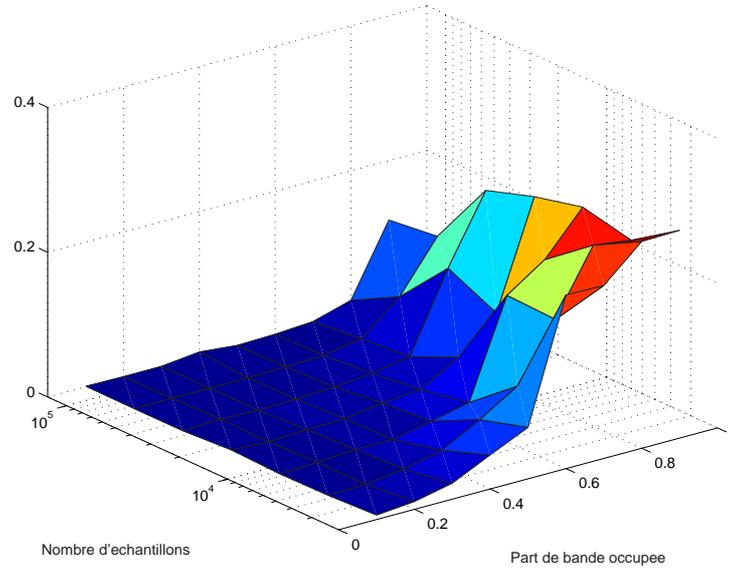


FIG. 4.8 – Valeur du critère τ_1 en fonction du nombre d'échantillons et du pourcentage occupé de la bande.

inverse MIMO RIF. Les sources étaient des sources de loi uniforme, centrée et de variance unité. La longueur du système mélangeant était $L_1 = L_2 = 2$ et celle du séparateur était $D_1 = D_2 = 5$. Les sources étant i.i.d., elles ont été reconstituées à un retard et un facteur multiplicatif près ; après élimination de ces ambiguïtés par corrélation, nous avons donc considéré l'EQM sur chacune des sources comme critère de séparation. Les résultats sont donnés sur la figure 4.9 pour des images de taille 256×256 pixels. La moyenne de l'EQM sur toute les réalisations était de 5.5×10^{-3} . Ces résultats confirment l'applicabilité des méthodes de séparation dans le cas convolutif multidimensionnel.

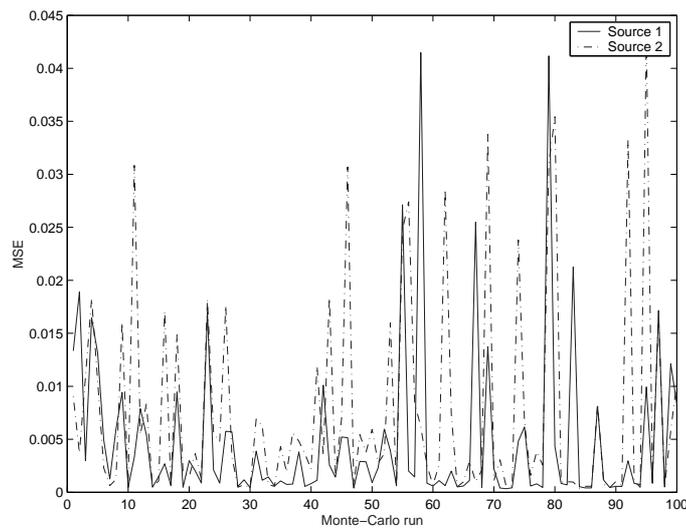


FIG. 4.9 – EQM après séparation de deux sources i.i.d. uniformes 2D.

4.2.5-c) Images naturelles

Le cas de la séparation d'un bruit et d'une image naturelle, mélangés par un filtre de longueur $L_1 = 3, L_2 = 1$ est donné en figure 4.10. Le nombre de pixels était de 256×256 et la longueur du filtre séparant était $D_1 = 4, D_2 = 1$. Un autre exemple de séparation de deux images naturelles de taille 256×256 est donné par la figure 4.11. Le filtre de mélange est de longueur $L_1 = 2, L_2 = 1$ et le séparateur de longueur $D_1 = 2, D_2 = 1$.

Les sorties directes de l'algorithme de séparation, représentées sur les figures 4.10(c) et 4.11(c) illustrent clairement le fait que les sources ne sont reconstituées qu'à un filtrage (ici passe-haut) près. D'autres filtres résiduels, dont certains au comportement plutôt passe-bas, ont également été trouvés comme points stationnaires des algorithmes de séparation utilisés. Ainsi, afin de supprimer l'ambiguïté de filtrage SISO et de reconstituer complètement l'image de départ, il convient d'envisager l'utilisation d'autres méthodes de traitement d'image. Nous nous sommes cependant placé dans des cas simplifiés.

Par une étape de déflation, il est possible de supprimer la contribution d'une des sources au mélange. Comme nous nous sommes restreint à deux sources, il est en réalité possible d'obtenir la contribution séparée de chacune des sources sur chaque capteur.

Dans le cas de la figure 4.10, la contribution du bruit a été retirée des capteurs par moindre carrés, ce qui conduit au résultat de la figure 4.10(d). Nous pouvons constater le bon fonctionnement de la méthode. En revanche, selon le filtre de mélange choisi, nous pourrions ne retrouver que des images semblables à celle du milieu de la figure 4.10(d).

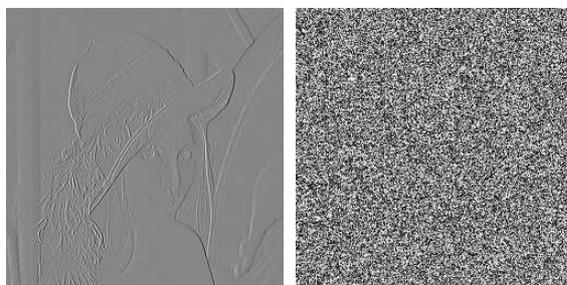
Pour éviter ce cas délicat, le filtre de mélange de la figure 4.11 a été choisi de sorte que chacune des sources ait une contribution non convolutive sur un des capteurs au moins. La même méthode que précédemment assure alors que l'image originale se retrouve sur l'un des capteurs au moins, après soustraction de la contribution de l'autre image. Le résultat de cette opération est donné en figure 4.11(d).



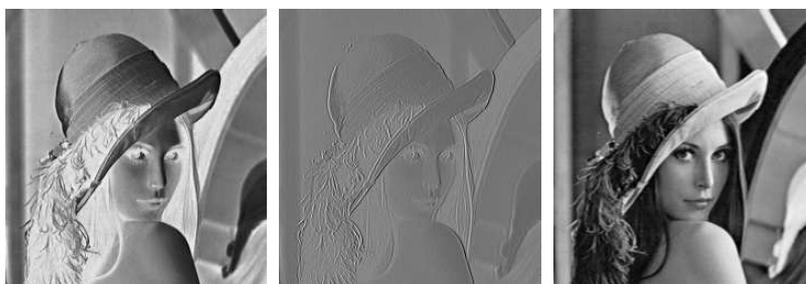
(a) Images sources originales



(b) Observations sur les capteurs



(c) Sortie de l'algorithme de séparation

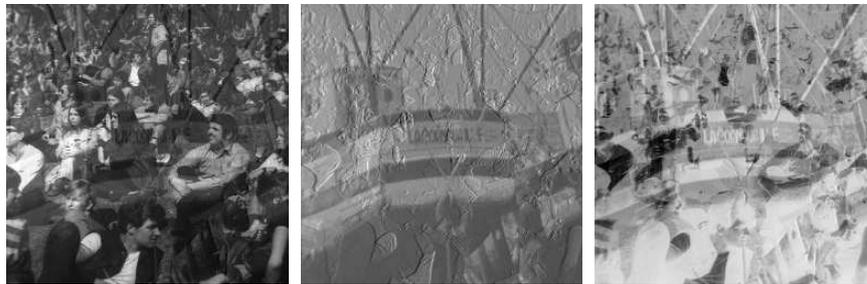


(d) Soustraction du bruit reconstitué aux capteurs

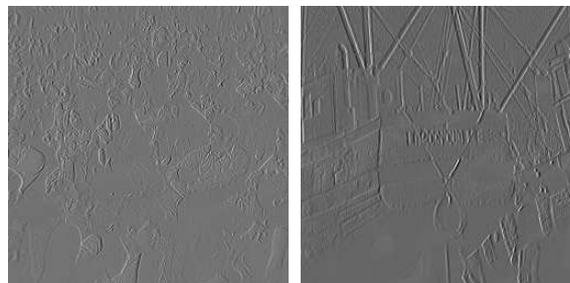
FIG. 4.10 – Exemple de séparation d'une image naturelle et d'un bruit i.i.d. uniforme



(a) Images sources originales



(b) Observations sur les capteurs



(c) Sortie de l'algorithme de séparation



(d) Reconstruction après moindres carrés

FIG. 4.11 – Exemple de séparation de deux images naturelles

Chapitre 5

Fonctions de contraste avec référence

5.1 Présentation

Ce chapitre se place dans le cadre des approches séquentielles qui ont déjà été abordées dans les chapitres 3 et 4. Nous nous concentrerons plus particulièrement sur l'extraction d'une source par un filtre MISO. L'extension à des méthodes MIMO est ensuite possible par utilisation d'un procédé de déflation.

L'intérêt majeur de ce chapitre est de proposer des fonctions de contraste dont la dépendance est *quadratique* en les paramètres. La maximisation du critère est alors fortement simplifiée : il n'est plus nécessaire de recourir à une optimisation itérative et il en résulte une charge de calcul considérablement allégée en comparaison de celle requise par la méthode du chapitre 3. Cet allègement se traduit par une accélération remarquable du processus de séparation.

L'idée centrale qui a permis l'élaboration des contrastes est celle de signal de *référence*. Ainsi, notre approche pourrait, dans un premier temps, être qualifiée de semi-aveugle même si son utilité s'est par la suite avérée dans le cas aveugle. La notion de signal de référence, ainsi qu'un rapide tour d'horizon des travaux en lien avec les nôtres, sont présentés dans un premier temps. Nous démontrons ensuite la validité des critères de contrastes dans le cas i.i.d. et non i.i.d. Enfin, nous détaillons les possibles difficultés de mise en œuvre avant de donner quelques résultats de simulation.

5.1.1 Signal de référence

Notations Dans ce chapitre, nous considérons l'extraction MISO d'une source à partir des observations. Nous reprenons par conséquent des notations identiques à celles du chapitre 3 et nous nous plaçons dans le cas sans bruit : $(y(n))_{n \in \mathbb{Z}}$ est ainsi la sortie du filtre séparateur $\mathbf{w}[z]$ et, en introduisant le filtre ligne global $\mathbf{g}[z] := \mathbf{w}[z]\mathbf{M}[z]$, on a dans le cas non bruité $y(n) = \mathbf{w}[z]\mathbf{x}(n) = \mathbf{g}[z]\mathbf{s}(n)$. Les fonctions de transfert $\mathbf{w}[z]$ et $\mathbf{g}[z]$ correspondent à une même ligne des filtres respectifs $\mathbf{W}[z]$ (filtre séparateur) et $\mathbf{G}[z]$ (filtre global).

Référence L'approche ici développée s'apparente aux approches *semi-aveugles* dans la mesure où nous supposons que nous disposons d'une première approximation notée $\mathbf{w}_r[z]$

5.1.3 Retour sur la définition d'un contraste

5.1.3-a) Dissymétrie possible

Dans le chapitre 3, la définition 5 d'une fonction de contraste supposait une invariance par translation temporelle, c'est-à-dire que la valeur du contraste était la même pour la sortie $(y(n))_{n \in \mathbb{Z}}$ et ses translatées $(y(n-l))_{n \in \mathbb{Z}}$ où l est fixé quelconque dans \mathbb{Z} . Cette invariance, valable de fait pour tous les contrastes considérés jusqu'à présent ne nous est toutefois d'aucune utilité pour la séparation de sources : un critère de séparation qui privilégierait l'obtention d'une version retardée particulière de l'une des sources pourrait en effet tout à fait nous convenir. C'est, en l'occurrence, ce qui advient ici lorsque le contraste considéré dépend non plus des seules statistiques de la sortie, mais des statistiques conjointes de la sortie et du signal de référence. Pour cette raison, nous utiliserons dans ce chapitre la définition suivante :

Définition 7 (contraste MISO, cas i.i.d. sans invariance temporelle) Une fonction de contraste MISO (ou contraste MISO) est une fonction réelle de la sortie $(y(n))_{n \in \mathbb{Z}}$ du filtre ligne global $\mathbf{g}[z]$ qui vérifie :

(i) Il existe $(i_0, l_0) \in \{1, \dots, N\} \times \mathbb{Z}$ tel que pour toute sortie globale $(y(n))_{n \in \mathbb{Z}}$ on ait :

$$\mathcal{C}((y(n))_{n \in \mathbb{Z}}) \leq \mathcal{C}((s_{i_0}(n-l_0))_{n \in \mathbb{Z}}). \quad (5.3)$$

(ii) L'égalité dans l'équation (5.3) ci-dessus n'est possible que si le filtre MISO global est du type $\mathbf{g}[z] = (0, \dots, 0, \alpha z^{-d}, 0, \dots, 0)$, où α est un scalaire non nul, $d \in \mathbb{Z}$ et la position de l'élément non nul αz^{-d} correspond à un indice i_0 qui assure la condition (i).

La définition ci-dessus prend en compte les variations éventuelles de la valeur d'un contraste pour des retards différents de la source reconstruite. En revanche, dans le cas de sources non linéaires (et non i.i.d.) aucune adaptation de la définition n'est nécessaire. En effet, la définition 6 prend en compte le fait qu'une source non i.i.d. ne peut être reconstruite qu'à un filtrage scalaire près, ce qui inclut l'ensemble des décalages temporels possibles.

5.1.3-b) Normalisation en puissance

Nous rappelons le cadre de base mis en place au chapitre 3 pour l'étude des approches séquentielles, et que l'on retrouve ici. Nous conservons l'hypothèse H.10 qui s'écrit pour mémoire :

H.10 Pour tout $j \in \{1, \dots, N\}$, la fonction d'autocorrélation de la $j^{\text{ème}}$ source, notée $(\gamma_j(k))_{j \in \mathbb{Z}} := (\mathbb{E}\{s_j(n)s_j^*(n-k)\})_{k \in \mathbb{Z}}$, est définie positive. De plus, chaque source est de puissance unité ($\gamma_j(0) := \mathbb{E}\{|s_j(n)|^2\} = 1$).

Dès lors, pour tout indice $j \in \{1, \dots, N\}$ et tout filtre SISO $h[z]$, nous définissons la norme :

$$\|h\|_j := \left(\sum_{(k,l) \in \mathbb{Z}^2} h(k)h^*(l)\gamma_j(l-k) \right)^{\frac{1}{2}}. \quad (5.4)$$

et, pour tout filtre global $\mathbf{g}[z] := (g_1[z], \dots, g_N[z])$, nous définissons sa norme ℓ^2 pondérée :

$$\|\mathbf{g}\| := \left(\sum_{j=1}^N \|g_j\|_j^2 \right)^{\frac{1}{2}}. \quad (5.5)$$

Enfin, rappelons que nous notons :

$$\forall i \in \{1, \dots, N\} \quad \mathcal{G}_i := \{ \mathbf{g}[z] = (g_1[z], \dots, g_N[z]) \mid \|g_j\|_j = \delta_{i-j}, \forall j \} \quad (5.6)$$

l'ensemble des filtres globaux de norme unité dont toutes les composantes sont nulles, sauf la $i^{\text{ème}}$.

De même qu'au chapitre 3, nous normaliserons la sortie globale afin qu'elle soit de puissance unité : $E\{|y(n)|^2\} = 1$. Compte tenu de la puissance unité des sources, ceci est équivalent à travailler sur l'ensemble des filtres globaux de norme unité qui a été noté :

$$\mathcal{F}_1 := \{ \mathbf{g}[z] \mid \|\mathbf{g}\| = 1 \}. \quad (5.7)$$

Tous les critères seront considérés sous contrainte que le filtre global soit de norme un ; par conséquent, le scalaire α introduit dans les définitions 6 et 7 vérifie $|\alpha| = 1$. Enfin, nous supposons :

H.15 Le filtre global de référence $\mathbf{t}[z]$ est stable (*i.e.* sa réponse impulsionnelle est absolument sommable) et de norme unité : $\|\mathbf{t}\| = 1$.

Notons que dans le cas de sources i.i.d. de variance 1, la contrainte de norme 1 s'écrit pour le filtre global et le filtre de référence :

$$\|\mathbf{g}\|^2 = \sum_{i=1}^N \sum_{k \in \mathbb{Z}} |g_i(k)|^2 = 1 \quad \text{et} \quad \|\mathbf{t}\|^2 = \sum_{i=1}^N \sum_{k \in \mathbb{Z}} |t_i(k)|^2 = 1. \quad (5.8)$$

5.2 Contrastes MISO dans le cas convolutif

5.2.1 Cas de sources i.i.d.

5.2.1-a) Contraste quadratique à base de cumulants croisés d'ordre 4

Dans ce paragraphe, nous supposons disposer d'un signal de référence $(z(n))_{n \in \mathbb{Z}}$ et nous définissons le cumulante croisé suivant :

$$\kappa_4\{y(n), z(n)\} := \text{Cum}\{y(n), y^*(n), z(n), z^*(n)\} \quad (5.9)$$

Nous considérons alors le critère suivant :

$$\mathcal{C}_z\{y(n)\} := |\kappa_4\{y(n), z(n)\}| \quad (5.10)$$

Nous allons étudier sous quelles conditions sur le signal ou filtre de référence, le critère (5.10) constitue un contraste. Le résultat sera donné par la proposition 21 avant que les conditions soient discutées dans le paragraphe 5.2.1-b). Nous pouvons dans un premier temps affirmer :

Lemme 4 *On suppose vérifiée la contrainte de puissance $E\{|y(n)|^2\} = \|\mathbf{g}\|^2 = 1$. Il existe $(i_0, l_0) \in \{1, \dots, N\} \times \mathbb{Z}$ tel que pour toute sortie globale $(y(n))_{n \in \mathbb{Z}}$ on ait : $\mathcal{C}_z\{y(n)\} \leq \mathcal{C}_z\{s_{i_0}(n - l_0)\}$ (équation (5.3)).*

Preuve: Par multilinéarité des cumulants, on peut écrire :

$$\kappa_4\{y(n), z(n)\} = \sum_{j=1}^N \sum_{k \in \mathbb{Z}} |g_j(k)|^2 |t_j(k)|^2 \kappa_{s_j}(0, 0, 0) \quad (5.11)$$

Définissons

$$\forall (j, k) \in \{1, \dots, N\} \times \mathbb{Z}, \quad \mathcal{M}_j(k) := |t_j(k)|^2 \kappa_{s_j}(0, 0, 0) \quad (5.12)$$

et :

$$\mathcal{M}_{\max} := \max_{j=1}^N \left(\sup_{k \in \mathbb{Z}} |\mathcal{M}_j(k)| \right) \quad (5.13)$$

En raison de la stabilité de $\mathbf{t}[z]$, la famille $|\mathcal{M}_j(k)|, k \in \mathbb{Z}$ est sommable. Ainsi, la borne supérieure qui a été définie est atteinte pour un couple $(i_0, l_0) \in \{1, \dots, N\} \times \mathbb{Z}$. Il vient alors :

$$\mathcal{C}_z\{y(n)\} = |\kappa_4\{y(n), z(n)\}| = \left| \sum_{j=1}^N \sum_{k \in \mathbb{Z}} |g_j(k)|^2 \mathcal{M}_j(k) \right| \quad (5.14)$$

$$\leq \sum_{j=1}^N \sum_{k \in \mathbb{Z}} |g_j(k)|^2 |\mathcal{M}_j(k)| \quad (5.15)$$

$$\leq \mathcal{M}_{\max} \sum_{j=1}^N \sum_{k \in \mathbb{Z}} |g_j(k)|^2. \quad (5.16)$$

En utilisant successivement l'hypothèse de norme unité et le fait que la borne supérieure est atteinte, on en déduit que :

$$\mathcal{C}_z\{y(n)\} \leq \mathcal{M}_{\max} = \mathcal{C}_z\{s_{i_0}(n - l_0)\}, \quad (5.17)$$

d'où le résultat annoncé. ■

Le lemme suivant donne une condition nécessaire et suffisante pour que le majorant ci-dessus soit atteint :

Lemme 5 *La fonction $\mathcal{C}_z\{\cdot\}$ atteint sa borne supérieure (i.e. $\mathcal{C}_z\{y(n)\} = \mathcal{M}_{\max}$) si et seulement si :*

$$\sum_{(j,k) \in \mathfrak{I}} |g_j(k)|^2 = 1 \quad (5.18)$$

où \mathfrak{I} est l'un des ensembles :

$$\mathfrak{I}_+ := \{(j, k) \in \{1, \dots, N\} \times \mathbb{Z} \mid \mathcal{M}_j(k) = \mathcal{M}_{\max}\} \quad (5.19)$$

$$\mathfrak{I}_- := \{(j, k) \in \{1, \dots, N\} \times \mathbb{Z} \mid \mathcal{M}_j(k) = -\mathcal{M}_{\max}\} \quad (5.20)$$

Cette condition revient de façon équivalente à dire que les coefficients $g_j(k)$ sont identiquement nuls en dehors de l'ensemble \mathfrak{I}_+ (ou \mathfrak{I}_-).

Preuve: Supposons $\mathcal{C}_z\{y(n)\} = \mathcal{M}_{\max}$. D'après l'égalité (5.14), il existe $\epsilon \in \{-1, +1\}$ tel que :

$$\sum_{j=1}^N \sum_{k \in \mathbb{Z}} |g_j(k)|^2 \mathcal{M}_j(k) = \epsilon \mathcal{M}_{\max} \quad (5.21)$$

Puisque le filtre $\mathbf{g}[z]$ est de norme unité, l'égalité ci-dessus peut s'écrire :

$$\sum_{j=1}^N \sum_{k \in \mathbb{Z}} |g_j(k)|^2 (\mathcal{M}_j(k) - \epsilon \mathcal{M}_{\max}) = 0 \quad (5.22)$$

Comme $(\mathcal{M}_j(k) - \epsilon \mathcal{M}_{\max})$ est de signe constant, nous en déduisons que pour tout $(j, k) \in \{1, \dots, N\}$ nous avons soit $|g_j(k)|^2 = 0$, soit $\mathcal{M}_j(k) = \epsilon \mathcal{M}_{\max}$. L'égalité (5.18) découle alors immédiatement si l'on se souvient que $\mathbf{g}[z]$ est de norme unité.

Réciproquement, si (5.18) est vérifiée, il découle de l'égalité $\sum_{j=1}^N \sum_{k \in \mathbb{Z}} |g_j(k)|^2 = 1$ que $\forall (j, k) \notin \mathcal{I}, g_j(k) = 0$ et l'égalité $\mathcal{C}_z\{y(n)\} = \mathcal{M}_{\max}$ est évidente. ■

Nous sommes à ce stade en mesure de démontrer la propriété suivante qui caractérise les conditions sous lesquelles la fonction $\mathcal{C}_z\{.\}$ est un contraste :

Proposition 21 *Sous contrainte de norme 1, la fonction $\mathcal{C}_z\{.\}$ est un contraste si et seulement si les ensembles \mathcal{I}_+ et \mathcal{I}_- contiennent chacun au plus un élément.*

Preuve: Supposons d'abord que \mathcal{I}_+ et \mathcal{I}_- contiennent au plus un élément. Le lemme 4 assure la majoration nécessaire pour obtenir la propriété de contraste.

Remarquons que $\mathcal{I}_+ \cup \mathcal{I}_-$ contient au moins un élément qui est le couple (i_0, l_0) du lemme 4. Si $\mathcal{C}_z\{.\}$ atteint sa borne supérieure, d'après le lemme 5, $|g_i(k)| = 0$ pour tout $(i, k) \notin \mathcal{I}$, où \mathcal{I} représente l'un des ensembles \mathcal{I}_+ ou \mathcal{I}_- . Comme ces derniers ensembles ont au plus un élément, et que $\mathcal{I}_+ \cup \mathcal{I}_-$ est non vide, c'est donc que \mathcal{I} (c'est-à-dire soit \mathcal{I}_+ soit \mathcal{I}_- selon le cas) a un et un seul élément (j_0, k_0) . Dès lors, nous avons $|g_{j_0}(k_0)| = 1$ et $g_i(k) = 0$ pour tout $(i, k) \neq (j_0, k_0)$, ce qui prouve que le filtre est séparant. La propriété de contraste en découle donc.

Réciproquement, supposons par exemple que \mathcal{I}_+ contienne plus d'un élément. Tout filtre global satisfaisant

$$\begin{cases} \forall (j, k) \notin \mathcal{I}_+ & g_j(k) = 0, \\ \forall (j, k) \in \mathcal{I}_+ & |g_j(k)|^2 = \frac{1}{\text{card}(\mathcal{I}_+)} \end{cases} \quad (5.23)$$

où card signifie cardinal de l'ensemble, rend le critère maximal sans pour autant être séparant au sens de la définition 7. La condition sur les ensembles \mathcal{I}_+ et \mathcal{I}_- est donc non seulement suffisante, mais aussi nécessaire. ■

5.2.1-b) Discussion des conditions de validité du contraste

Nous discutons ici la condition de validité du critère de contraste avec référence donnée par la proposition 21.

Condition sur le système de référence seul En premier lieu, nous pouvons noter que la condition sur les ensembles \mathcal{I}_+ et \mathcal{I}_- est une condition conjointe sur le système de référence et les caractéristiques des sources. Les valeurs des coefficients $\mathcal{M}_j(k)$ définis à l'équation (5.12) dépendent en effet simultanément des coefficients $t_j(k)$ du système de référence et des autocumulants $\kappa_{s_j}(0, 0, 0)$ des sources. Il peut paraître maladroit d'exprimer la condition de validité du contraste comme dépendant de caractéristiques inconnues des sources ; ce point de vue permet cependant d'exprimer des conditions générales.

Le choix du système de référence est libre et il peut être plus naturel d'exprimer une condition sur lui seul. Nous supposons donc les caractéristiques des sources fixées quelconques, dans la mesure toutefois où l'une des sources au moins a un cumulant d'ordre quatre non nul :

$$\exists j \in \{1, \dots, N\}, \quad \kappa_{s_j}(0, 0, 0) \neq 0. \quad (5.24)$$

En toute généralité et par convention, nous supposons que la source numéro 1 a un cumulant plus grand en valeur absolue que tous les autres cumulants des sources. Ceci revient à écrire :

$$\forall j \in \{1, \dots, N\} \quad |\kappa_{s_j}(0, 0, 0)| \leq |\kappa_{s_1}(0, 0, 0)| \quad (5.25)$$

Il est alors immédiat de vérifier que si le système de référence satisfait :

$$\exists l \in \mathbb{Z} \mid \forall (j, k) \in \{1, \dots, N\} \times \mathbb{Z} \quad (j \neq 1 \text{ ou } k \neq l) \Rightarrow |t_j(k)| < |t_1(l)| \quad (5.26)$$

nous avons pour tout $(j, k) \neq (1, l)$

$$|\mathcal{M}_j(k)| = |t_j(k)| |\kappa_{s_j}(0, 0, 0)| < |t_1(l)| |\kappa_{s_1}(0, 0, 0)| = |\mathcal{M}_1(l)| \quad (5.27)$$

et $\mathfrak{J}_+ \cup \mathfrak{J}_-$ se réduit exactement au singleton $\{(1, l)\}$. \mathfrak{J}_+ et \mathfrak{J}_- ont donc chacun au plus un élément et \mathcal{C}_z est un contraste.

Une interprétation possible Une lecture possible de la condition donnée par l'équation (5.26) est la suivante : parmi les coefficients du système de référence global, l'un d'entre eux, $t_1(l)$ est en même temps le plus grand en valeur absolue et correspond à la première source, c'est-à-dire celle dont le cumuland est le plus grand. D'après la preuve de la proposition 21, le critère de contraste avec référence qui vérifie cette condition mène alors à la séparation de cette source avec un retard temporel l donné par le coefficient le plus grand.

De façon plus générale, nous pouvons interpréter la condition de la proposition 21 comme la nécessité de présence majoritaire de l'une des versions retardée des sources dans le signal de référence. Cette présence majoritaire peut être liée soit à une valeur plus importante du cumuland de la source en question, soit à un système de référence privilégiant la source en question. Ainsi, cette approche donne bien une condition moins restrictive pour la validité des contrastes que si cette dernière est exprimée sur le seul filtre de référence, comme nous venons de le faire.

Choix du système de référence Nous avons jusqu'à présent supposé le système de référence donné. Connaissant maintenant les conditions que ce dernier doit respecter, nous pouvons examiner la façon de le choisir.

Supposons que les sources aient toutes les mêmes caractéristiques et qu'il existe un couple d'indices $(i_0, k_0) \in \{1, \dots, N\} \times \mathbb{Z}$ tel que pour tout $(i, k) \neq (i_0, k_0)$ on ait $|t_i(k)| < |t_{i_0}(k_0)|$. Par le même raisonnement qu'au paragraphe ci-dessus, il est immédiat de remarquer alors que $\mathfrak{J}_+ \cup \mathfrak{J}_-$ est un singleton et que donc $\mathcal{C}_z\{\cdot\}$ est un contraste. Il suffit donc que l'un des coefficients du filtre de mélange majore strictement les autres pour que le contraste soit valable. Plus généralement, nous pouvons énoncer la proposition suivante, dont la preuve est évidente :

Proposition 22 *Une condition suffisante pour que la fonction $\mathcal{C}_z\{\cdot\}$ soit un contraste (sous contrainte de norme unité) est :*

$$\exists (j_0, k_0) \in \{1, \dots, N\} \times \mathbb{Z} \mid \forall (j, k) \neq (j_0, k_0) \quad |\mathcal{M}_j(k)| < |\mathcal{M}_{j_0}(k_0)| \quad (5.28)$$

La condition de validité d'un contraste avec référence apparaît donc comme très peu contraignante. Plus généralement, nous pouvons affirmer :

Proposition 23 *Supposons qu'il existe une source de cumuland d'ordre quatre non nul (équation (5.24)). Si le système de référence est de réponse impulsionnelle finie et ses coefficients sont tirés aléatoirement selon une loi conjointe continue, alors la condition de la proposition 21 sur les ensembles \mathfrak{J}_+ et \mathfrak{J}_- est presque sûrement vérifiée.*

Preuve: En effet, l'hypothèse sur les sources assure que \mathcal{M}_{\max} est non nul. De plus, comme indiqué dans la preuve de la proposition 21, $\mathfrak{J}_+ \cup \mathfrak{J}_-$ est non vide et ainsi il existe un couple (i_0, l_0) (le couple du lemme 4) tel que $\mathcal{M}_{\max} = \epsilon \mathcal{M}_{i_0}(l_0)$ où $\epsilon \in \{-1, +1\}$.

Un couple $(j, k) \neq (i_0, l_0)$ appartient donc à \mathfrak{J}_+ si et seulement si $\mathcal{M}_j(k) = \epsilon \mathcal{M}_{i_0}(l_0)$, ce qui s'écrit aussi $|t_j(k)|^2 \kappa_{s_j}(0, 0, 0) = \epsilon |t_{i_0}(l_0)|^2 \kappa_{i_0}(0, 0, 0)$. Selon les valeurs de ϵ et $\kappa_{s_j}(0, 0, 0)$, ceci est ou bien impossible, ou alors arrive avec une probabilité nulle lorsque les coefficients $(t_j(k))$, $(j, k) \in \{1, \dots, N\} \times \mathbb{Z}$ sont tirés selon une loi continue. Le même raisonnement peut être mené sur \mathfrak{J}_- , ce qui prouve que presque sûrement, les ensembles \mathfrak{J}_+ et \mathfrak{J}_- ont au plus un élément. ■

5.2.1-c) Généralisation pour des cumulants à tous ordres

Nous expliquons dans ce paragraphe comment les résultats du paragraphe précédents peuvent se généraliser à d'autres cumulants d'ordre supérieur ou égal à 3. Par ailleurs, il est possible de choisir librement de conjuguer ou non les signaux dans les cumulants. Par commodité, utilisons le symbole $\mathbf{y}(n)$ pour désigner au choix le signal $y(n)$ ou son conjugué $y^*(n)$. De même, $\mathbf{z}(n)$ désigne au choix $z(n)$ ou $z^*(n)$. Nous pouvons alors définir pour tout $R \geq 3$ la forme générale du cumulants pouvant servir à l'élaboration d'un contraste :

$$\kappa_R\{\mathbf{y}(n), \mathbf{z}(n)\} := \text{Cum}\{\mathbf{y}(n), \mathbf{y}(n), \underbrace{\mathbf{z}(n), \dots, \mathbf{z}(n)}_{R-2 \text{ fois}}\} \quad (5.29)$$

Les propriétés du paragraphe précédent se généralisent alors au critère :

$$\mathbf{C}_z\{\mathbf{y}(n)\} := |\kappa_R\{\mathbf{y}(n), \mathbf{z}(n)\}| \quad (5.30)$$

Comme le cumulants ci-dessus n'est plus nécessairement réel, il convient au lieu des ensembles \mathfrak{J}_+ et \mathfrak{J}_- de considérer l'ensemble :

$$\mathfrak{J} = \{(j, k) \in \{1, \dots, N\} \times \mathbb{Z} \mid |\mathcal{M}_j(k)| = \mathcal{M}_{\max}\} \quad (5.31)$$

Les propositions du paragraphe précédent se généralisent alors aisément et permettent d'affirmer :

Proposition 24 *Sous contrainte de norme 1, la fonction $\mathbf{C}_z\{\cdot\}$ est un contraste si et seulement si l'ensemble \mathfrak{J} contient un élément et un seul.*

La démonstration de cette proposition est quasiment identique à celle déjà faite. Comme le seul cas pratique considéré en simulation est celui des cumulants d'ordre quatre, nous ne démontrons pas ce résultat. Remarquons aussi que pour des signaux sources complexes circulaires, il convient pour obtenir un contraste que le nombre de conjugués et de non conjugués s'équilibre dans (5.29). A défaut, nous aurions en effet $\mathcal{M}_{\max} = 0$ et $\mathfrak{J} = \{1, \dots, N\} \times \mathbb{Z}$.

5.2.2 Cas de sources non i.i.d.

Nous étudions à présent le cas de sources non linéaires (et non i.i.d.). Nous allons d'abord montrer que, sous certaines hypothèses techniques, les critères qui viennent d'être introduits constituent des contrastes pour des sources non i.i.d. au sens de la définition 6. L'étude des hypothèses techniques et la recherche de conditions suffisantes pour qu'elles soient satisfaites nous préoccupera dans un second temps.

5.2.2-a) Validité des contrastes avec référence dans le cas non i.i.d.

La démonstration de la validité des critères de contraste avec référence présente de fortes similitudes avec celle de la proposition 9 du chapitre 3. Pour tout i , nous introduisons ainsi une borne supérieure du critère lorsque seule la $i^{\text{ème}}$ source est présente :

$$\forall i \in \{1, \dots, N\} \quad \mathcal{M}_i := \sup_{\mathbf{g} \in \mathcal{G}_i} \mathcal{C}_z\{\mathbf{g}[z]\mathbf{s}(n)\} \quad (5.32)$$

Contrairement au cas i.i.d. il n'est pas assuré que cette borne supérieure soit atteinte. De même qu'au chapitre 3, nous poserons donc cette hypothèse technique :

H.16 Pour tout $i \in \{1, \dots, N\}$, la borne supérieure définie à l'équation (5.33) est atteinte pour un filtre noté $\mathbf{g}_i^\# \in \mathcal{G}_i$.

Enfin, par analogie avec le cas i.i.d. du paragraphe 5.2.1, nous définissons :

$$\mathcal{M}_{\max} := \max_{j=1}^N \left(\mathcal{M}_j \right) = \max_{j=1}^N \left(\sup_{\mathbf{g} \in \mathcal{G}_j} \mathcal{C}_z\{\mathbf{g}[z]\mathbf{s}(n)\} \right) \quad (5.33)$$

La proposition suivante peut alors être énoncée :

Proposition 25 *Sous contrainte de norme 1 et sous l'hypothèse H.16, la fonction $\mathcal{C}_z\{\cdot\}$ est un contraste si et seulement si l'ensemble*

$$\mathfrak{J}' := \{j \in \{1, \dots, N\} \mid \mathcal{M}_j = \mathcal{M}_{\max}\} \quad (5.34)$$

contient un unique élément.

Preuve: De même qu'à l'équation (3.11), la sortie globale $y(n)$ peut s'écrire :

$$y(n) = \sum_{j=1}^N g_j[z]s_j(n) = \sum_{j=1}^N \|g_j\|_j \tilde{g}_j[z]s_j(n) = \sum_{j=1}^N \|g_j\|_j \tilde{y}_j(n) \quad (5.35)$$

où l'on a posé :

$$\tilde{g}_j[z] := \begin{cases} \frac{g_j[z]}{\|g_j\|_j} & \text{si } \|g_j\|_j \neq 0 \\ 0 & \text{sinon} \end{cases} \quad (5.36)$$

et :

$$\tilde{y}_j(n) := \tilde{g}_j[z]s_j(n) \quad (5.37)$$

Par multilinéarité des cumulants et indépendance des $(\tilde{y}_j(n))_{n \in \mathbb{Z}}, j \in \{1, \dots, N\}$, il vient :

$$\kappa_4\{y(n), z(n)\} = \sum_{j=1}^N \|g_j\|_j^2 \kappa_4\{\tilde{y}_j(n), z(n)\} \quad (5.38)$$

Nous pouvons alors effectuer les majorations suivantes :

$$\mathcal{C}_z\{y(n)\} = |\kappa_4\{y(n), z(n)\}| \leq \sum_{j=1}^N \|g_j\|_j^2 |\kappa_4\{\tilde{y}_j(n), z(n)\}| \quad (5.39)$$

$$\leq \sum_{j=1}^N \|g_j\|_j^2 \mathcal{C}_z\{\tilde{y}_j(n)\} \quad (5.40)$$

et donc, se rappelant (5.32), (5.37) et que chacun des filtres scalaires $\tilde{g}_j[z]$ est de j -norme unité :

$$\mathcal{C}_z\{y(n)\} \leq \sum_{j=1}^N \|g_j\|_j^2 \mathcal{M}_j \quad (5.41)$$

$$\leq \mathcal{M}_{\max} \sum_{j=1}^N \|g_j\|_j^2 \leq \mathcal{M}_{\max} \quad (5.42)$$

La propriété de majoration nécessaire à la validité du contraste est donc prouvée.

Par ailleurs, en cas d'égalité ci-dessus (ce qui se produit pour au moins un filtre global compte tenu de l'hypothèse H.16), en exploitant l'égalité entre (5.41) et (5.42), il vient :

$$\sum_{j=1}^N \|g_j\|_j^2 (\mathcal{M}_{\max} - \mathcal{M}_j) = 0 \quad (5.43)$$

et ainsi :

$$\forall j \in \{1, \dots, N\} \quad \begin{cases} \|g_j\|_j = 0 \\ \mathcal{M}_j = \mathcal{M}_{\max}, \end{cases} \quad \text{ou bien :} \quad (5.44)$$

ce qui donne encore $\forall j \notin \mathcal{J}', \|g_j\|_j = 0$. Si \mathcal{J}' contient un unique élément j_0 , nous avons alors $\|g_j\|_j = 0$ si $j \neq j_0$ et $\|g_{j_0}\|_{j_0} = 1$ compte tenu de la norme unité du filtre global. Ce dernier est donc séparant, ce qui prouve la propriété de contraste.

Dans le cas où \mathcal{J}' contient en revanche plusieurs éléments dont entre autres j_1 et j_2 , le filtre global défini par $\frac{1}{\sqrt{2}}(\mathbf{g}_{j_1}^\# [z] + \mathbf{g}_{j_2}^\# [z])$ atteint le maximum du critère sans être pour autant séparant. La condition sur \mathcal{J}' est donc nécessaire. ■

Remarque 13: L'hypothèse H.16 peut en réalité être affaiblie en :

H.17 Pour tout $i \in \mathcal{J}'$, la borne supérieure définie à l'équation (5.33) est atteinte pour un filtre noté $\mathbf{g}_i^\# \in \mathcal{G}_i$.

Au cours de la démonstration, nous utilisons en effet le fait que la borne supérieure est atteinte uniquement pour le filtre séparant correspondant à un indice dans l'ensemble \mathcal{J}' .

5.2.2-b) Condition suffisante de validité du contraste

Nous cherchons dans ce paragraphe une condition suffisante pour pouvoir appliquer la proposition 25. Comme nous allons le constater, la généralisation de la proposition 22 au cas non i.i.d. est plus délicate et l'interprétation de son équivalent est moins immédiate. Aussi, le lecteur pourra-t-il sauter ce paragraphe au cours d'une première lecture.

Examinons d'abord le cas particulier où la référence consiste en la filtrée scalaire de l'une des sources, par exemple la $j_0^{\text{ème}}$. Soit $j \in \{1, \dots, N\}, j \neq j_0$, et soit un filtre global $\mathbf{g} \in \mathcal{G}_j$ dont seule la $j^{\text{ème}}$ composante est non nulle. La sortie globale $y(n) = \mathbf{g}[z]\mathbf{s}(n)$, qui ne dépend que de $(s_j(n)), n \in \mathbb{Z}$, est alors indépendante de la référence, qui ne dépend que de $(s_{j_0}(n))_{n \in \mathbb{Z}}$. Le contraste (5.10) est donc nul et on en déduit par conséquent que pour tout $j \neq j_0$, $\mathcal{M}_j = 0$. Ainsi, l'ensemble \mathcal{J}' de la proposition 25 ne contient qu'un élément, si bien sûr les cumulants d'ordre quatre des sources sont non nuls. La condition de la proposition 25 est donc vérifiée lorsque la référence est une filtrée scalaire de l'une des sources et ce résultat est conforme à nos attentes.

Nous donnons maintenant une autre condition suffisante, généralisant la proposition 22 :

Proposition 26 Une condition suffisante pour que la condition de la proposition 25 soit vérifiée est qu'il existe $j_0 \in \{1, \dots, N\}$ tel que :

$$\forall j \neq j_0 \quad \|t_{j_0}\|_\infty^2 |\kappa_{s_{j_0}}(0, 0, 0)| > (\|t_{j_0}\|_1^2 - \|t_{j_0}\|_\infty^2) \sup_{(l_1, l_2) \neq (0, 0)} |\kappa_{s_{j_0}}(0, l_1, l_2)| \\ + (\Gamma_j^{\min})^{-1} \|t_j\|_\infty^2 \|\kappa_{s_j}\|_1, \quad (5.45)$$

où pour toute famille de nombres $(a(k))_{k \in \mathcal{K}}$ on a défini les normes :

$$\|a\|_1 := \sum_{k \in \mathcal{K}} |a(k)| \\ \|a\|_\infty := \sup_{k \in \mathcal{K}} |a(k)|$$

et où pour chacune des sources $(s_j(n))_{n \in \mathbb{Z}}, j \in \{1, \dots, N\}$ on a supposé la densité spectrale minorée par $\Gamma_j^{\min} > 0$.

Preuve: La démonstration est quelque peu fastidieuse et se déroule en trois temps : minoration des \mathcal{M}_j , majoration des \mathcal{M}_j et enfin conclusion sur la proposition.

Minoration des \mathcal{M}_j Par multilinéarité des cumulants et indépendance des sources, on a

$$\kappa_4\{\tilde{y}_j(n), z(n)\} = \\ \sum_{k_1, k_2, l_1, l_2} \tilde{g}_j(k_1) \tilde{g}_j^*(k_2) t_j(l_1) t_j^*(l_2) \text{Cum}\{s_j(n - k_1), s_j^*(n - k_2), s_j(n - l_1), s_j^*(n - l_2)\} \\ = \sum_{k_1, k_2, l_1, l_2} \tilde{g}_j(k_1) \tilde{g}_j^*(k_2) t_j(l_1) t_j^*(l_2) \kappa_{s_j}(k_1 - k_2, k_1 - l_1, k_1 - l_2). \quad (5.46)$$

Puisque (5.32) s'écrit aussi $\mathcal{M}_j := \sup_{\|\tilde{g}_j\|=1} \mathcal{C}_z\{\tilde{g}_j[z]s_j(n)\}$, en prenant le filtre particulier de réponse impulsionnelle $(\delta_{n-p})_{n \in \mathbb{Z}}$, on en déduit que pour tout $p \in \mathbb{Z}$:

$$\mathcal{M}_j \geq \left| \sum_{k_1, k_2, l_1, l_2} \delta(k_1 - p) \delta(k_2 - p) t_j(l_1) t_j^*(l_2) \kappa_{s_j}(k_1 - k_2, k_1 - l_1, k_1 - l_2) \right| \\ = \left| \sum_{l_1, l_2} t_j(l_1) t_j^*(l_2) \kappa_{s_j}(0, p - l_1, p - l_2) \right| \\ \geq |t_j(p)|^2 |\kappa_{s_j}(0, 0, 0)| - \left| \sum_{(l_1, l_2) \neq (p, p)} t_j(l_1) t_j^*(l_2) \kappa_{s_j}(0, p - l_1, p - l_2) \right|.$$

De plus, nous pouvons écrire :

$$\left| \sum_{(l_1, l_2) \neq (p, p)} t_j(l_1) t_j^*(l_2) \kappa_{s_j}(0, p - l_1, p - l_2) \right| \\ \leq \sup_{(l_1, l_2) \neq (p, p)} |\kappa_{s_j}(0, p - l_1, p - l_2)| \sum_{(l_1, l_2) \neq (p, p)} |t_j(l_1)| |t_j(l_2)| \\ = \sup_{(l_1, l_2) \neq (0, 0)} |\kappa_{s_j}(0, l_1, l_2)| (\|t_j\|_1^2 - |t_j(p)|^2).$$

Les deux inégalités précédentes mènent, pour tout $p \in \mathbb{Z}$, à :

$$\mathcal{M}_j \geq |t_j(p)|^2 \left(|\kappa_{s_j}(0, 0, 0)| + \sup_{(l_1, l_2) \neq (0, 0)} |\kappa_{s_j}(0, l_1, l_2)| \right) - \|t_j\|_1^2 \sup_{(l_1, l_2) \neq (0, 0)} |\kappa_{s_j}(0, l_1, l_2)| \quad (5.47)$$

et en prenant la borne supérieure sur l'ensemble des $|t_j(p)|, p \in \mathbb{Z}$:

$$\mathcal{M}_j \geq \|t_j\|_\infty^2 \left(|\kappa_{s_j}(0, 0, 0)| + \sup_{(l_1, l_2) \neq (0, 0)} |\kappa_{s_j}(0, l_1, l_2)| \right) - \|t_j\|_1^2 \sup_{(l_1, l_2) \neq (0, 0)} |\kappa_{s_j}(0, l_1, l_2)| \quad (5.48)$$

Majoration des \mathcal{M}_j D'après (5.46), nous pouvons écrire :

$$\begin{aligned} |\kappa_4\{\tilde{y}_j(n), z(n)\}| &\leq \sum_{k_1, k_2, l_1, l_2} |\tilde{g}_j(k_1)| |\tilde{g}_j^*(k_2)| |t_j(l_1)| |t_j^*(l_2)| |\kappa_{s_j}(k_1 - k_2, k_1 - l_1, k_1 - l_2)| \\ &\leq \|t_j\|_\infty^2 \sum_{k_1, k_2} |\tilde{g}_j(k_1)| |\tilde{g}_j(k_2)| \sum_{l_1, l_2} |\kappa_{s_j}(k_1 - k_2, l_1, l_2)| \end{aligned} \quad (5.49)$$

En changeant de variable $k = k_1, p = k_1 - k_2$ et en réarrangeant, il vient :

$$|\kappa_4\{\tilde{y}_j(n), z(n)\}| \leq \|t_j\|_\infty^2 \sum_{l_1, l_2} \sum_p |\kappa_{s_j}(p, l_1, l_2)| \sum_k |\tilde{g}_j(k)| |\tilde{g}_j(k - p)|$$

et comme $\sum_k |\tilde{g}_j(k)| |\tilde{g}_j(k - p)| \leq \sum_k |\tilde{g}_j(k)|^2 \leq (\Gamma_j^{\min})^{-1} \|\tilde{g}_j\|_j^2 = (\Gamma_j^{\min})^{-1}$:

$$|\kappa_4\{\tilde{y}_j(n), z(n)\}| \leq \|t_j\|_\infty^2 \sum_{l_1, l_2} \sum_p |\kappa_{s_j}(p, l_1, l_2)| = (\Gamma_j^{\min})^{-1} \|t_j\|_\infty^2 \|\kappa_{s_j}\|_1 \quad (5.50)$$

La borne supérieure du terme à gauche ci-dessus vérifie donc :

$$\mathcal{M}_j \leq (\Gamma_j^{\min})^{-1} \|t_j\|_\infty^2 \|\kappa_{s_j}\|_1. \quad (5.51)$$

Preuve proposition 26 Une condition suffisante pour que \mathfrak{J}' de l'équation (5.34) se réduise au singleton $\{j_0\}$ est :

$$\begin{aligned} (\Gamma_j^{\min})^{-1} \|t_j\|_\infty^2 \|\kappa_{s_j}\|_1 &< \|t_{j_0}\|_\infty^2 \left(|\kappa_{s_{j_0}}(0, 0, 0)| + \sup_{(l_1, l_2) \neq (0, 0)} |\kappa_{s_{j_0}}(0, l_1, l_2)| \right) \\ &\quad - \|t_{j_0}\|_1^2 \sup_{(l_1, l_2) \neq (0, 0)} |\kappa_{s_{j_0}}(0, l_1, l_2)| \end{aligned} \quad (5.52)$$

Le membre de gauche est en effet celui de la majoration de \mathcal{M}_j dans (5.51) et celui de droite est celui de la minoration de \mathcal{M}_{j_0} dans (5.48). Il s'ensuit donc $\mathcal{M}_j < \mathcal{M}_{j_0}$ pour tout $j \neq j_0$, ce qui assure que l'ensemble \mathfrak{J}' de l'équation est un singleton. On vérifie alors immédiatement l'équivalence entre (5.45) et (5.52). ■

Comme le système de référence a été supposé stable (hypothèse H.15), les différentes normes $\|\cdot\|_1$, $\|\cdot\|$ et $\|\cdot\|_\infty$ sur le système de référence existent et de plus, il existe $p^\sharp \in \mathbb{Z}$ tel que $|t_{j_0}(p^\sharp)| = \|t_{j_0}\|_\infty$. La condition (5.45) de la proposition 26 peut s'interpréter alors comme le fait que $t_{j_0}(p^\sharp)$ est le terme «dominant» dans $(t_j(k))_{1 \leq j \leq N, k \in \mathbb{Z}}$. Cette même condition (5.45) requiert également que $|\kappa_{s_{j_0}}(0, 0, 0)|$ ne soit pas trop petit par rapport à $\sup_{(l_1, l_2) \neq (0, 0)} |\kappa_{s_{j_0}}(0, l_1, l_2)|$, ce qui n'est pas nécessairement vérifiée pour tout signal non i.i.d. comme l'a prouvé le contre-exemple du paragraphe 2.3.2-b). Nous constatons ainsi que la condition obtenue de validité du contraste avec référence peut s'interpréter de façon analogue au cas i.i.d. : il s'agit d'une condition conjointe sur les sources et le système de référence, qui assure la présence majoritaire de l'une des sources dans le signal de référence.

Remarque 14: Dans le cas du paragraphe précédent où les $(s_j(n))_{n \in \mathbb{Z}}$ forment des suites i.i.d., on a $\Gamma_j^{\min} = 1$ et $\kappa_{s_j}(k_1, k_2, k_3) = 0$ pour tout $(k_1, k_2, k_3) \neq (0, 0, 0)$. Par suite, $\|\kappa_{s_j}\|_1 = |\kappa_{s_j}(0, 0, 0)|$ et (5.45) se simplifie en :

$$\forall j \neq j_0 \quad \|t_{j_0}\|_\infty^2 |\kappa_{s_{j_0}}(0, 0, 0)| > \|t_j\|_\infty^2 |\kappa_{s_j}(0, 0, 0)|. \quad (5.53)$$

Nous retrouvons donc bien la condition (5.28) obtenue dans le cas i.i.d.

De plus, puisque (5.45) implique que (5.53) est vérifiée, nous constatons que la condition de validité obtenue dans le cas non i.i.d. est plus restrictive que celle du cas i.i.d. (avec équivalence dans le cas de sources i.i.d.).

5.3 Éléments pour la mise en œuvre

5.3.1 Expression des formes quadratiques : critère et contrainte

De même que dans le paragraphe 3.4.2, nous supposons maintenant que le filtre séparant est RIF. Sa réponse impulsionnelle est donnée par $\mathbf{w}(0), \dots, \mathbf{w}(D-1)$ et nous définissons le vecteur suivant de taille $1 \times QD$:

$$\underline{\mathbf{w}} := (\mathbf{w}(0) \quad \mathbf{w}(1) \quad \dots \quad \mathbf{w}(D-1)) \quad (5.54)$$

ainsi que pour tout $n \in \mathbb{Z}$ le vecteur colonne $\underline{\mathbf{x}}(n) := (\underline{x}_1(n), \dots, \underline{x}_{QD}(n))^T$ de taille $QD \times 1$:

$$\underline{\mathbf{x}}(n) := (\mathbf{x}(n)^T \quad \mathbf{x}(n-1)^T \quad \dots \quad \mathbf{x}(n-D+1)^T)^T \quad (5.55)$$

La sortie du filtre global s'exprime alors :

$$y(n) = \underline{\mathbf{w}} \underline{\mathbf{x}}(n) \quad (5.56)$$

Les observations étant données, $\underline{\mathbf{x}}(n)$ est donné et $y(n)$ s'exprime donc en fonction de $\underline{\mathbf{w}}$ seulement. Dès lors, la méthode proposée pour l'extraction d'une source se ramène, d'après la proposition 21, au problème d'optimisation :

$$\text{maximiser } \mathcal{C}_z\{y(n)\} \quad \text{sous la contrainte : } \mathbb{E}\{|y(n)|^2\} = 1, \quad (5.57)$$

où la variable à optimiser est le vecteur $\underline{\mathbf{w}}$. En effet, à condition de définir les notations qui suivent, les expressions apparaissant ci-dessus peuvent être données explicitement en fonction de $\underline{\mathbf{w}}$. Définissons

$$\mathbf{R} := \mathbb{E}\{\underline{\mathbf{x}}(n)\underline{\mathbf{x}}(n)^H\} \quad (5.58)$$

et $\mathbf{C} = (C_{ij})_{(i,j) \in \{1, \dots, QD\}^2}$ la matrice dont les éléments sont donnés par :

$$\forall (i, j) \in \{1, \dots, QD\}^2 \quad C_{ij} := \text{Cum}\{\underline{x}_i(n), \underline{x}_j^*(n), z(n), z^*(n)\}. \quad (5.59)$$

Les définitions (5.58) et (5.59) permettent alors d'écrire :

$$\mathbb{E}\{|y(n)|^2\} = \underline{\mathbf{w}} \mathbf{R} \underline{\mathbf{w}}^H \quad \kappa_4\{y(n), z(n)\} = \underline{\mathbf{w}} \mathbf{C} \underline{\mathbf{w}}^H \quad (5.60)$$

et donc aussi :

$$\mathcal{C}_z\{y(n)\} := |\kappa_4\{y(n), z(n)\}| = |\underline{\mathbf{w}} \mathbf{C} \underline{\mathbf{w}}^H|. \quad (5.61)$$

Remarquons au passage que les matrices \mathbf{R} ainsi que \mathbf{C} sont hermitiennes, ce qui est d'ailleurs en conformité avec le fait que les quantités ci-dessus sont réelles. A ce stade, le problème (5.57) se réduit au problème d'optimisation suivant :

$$\max |\underline{\mathbf{w}} \mathbf{C} \underline{\mathbf{w}}^H| \quad \text{sous la contrainte : } \underline{\mathbf{w}} \mathbf{R} \underline{\mathbf{w}}^H = 1. \quad (5.62)$$

Il s'agit d'un problème dont le critère est une valeur absolue d'une fonction quadratique et dont la contrainte est quadratique. Nous allons montrer ci-dessous qu'il est aisé d'en obtenir une solution.

5.3.2 Restriction du problème

5.3.2-a) Définie positivité de \mathbf{R} ou positivité seulement ?

Dans le cas où \mathbf{R} est définie positive, l'ensemble des vecteurs $\underline{\mathbf{w}}$ tels que $\underline{\mathbf{w}}\mathbf{R}\underline{\mathbf{w}}^H = 1$ est un compact et, grâce à la continuité, nous sommes assurés de l'existence d'une borne supérieure finie au critère que nous souhaitons maximiser. Cette borne est par ailleurs atteinte pour au moins un vecteur, qui constitue alors une solution au problème donné à l'équation (5.62). L'existence d'une solution au problème est donc acquise lorsque \mathbf{R} est définie positive.

Cependant, la matrice \mathbf{R} n'est en général que positive non définie. C'est en particulier le cas lorsque :

- le mélange est instantané et il y a strictement plus de capteurs que de sources. Dans ce cas en effet, $\underline{\mathbf{x}}(n) = \mathbf{x}(n) = \mathbf{M}(0)\mathbf{s}(n)$ et ainsi $\mathbf{R} = \mathbf{E}\{\mathbf{x}(n)\mathbf{x}(n)^H\} = \mathbf{M}(0)\mathbf{M}(0)^H$ n'est pas de rang plein.
- une des sources est identiquement nulle. C'est, dans l'idéal, ce qui arrive après une étape de déflation. Ce cas sera étudié au paragraphe 5.3.4.
- de façon générale \mathbf{R} n'est pas de rang plein dès que les nombres de capteurs et de sources vérifient $Q > N + 1$ et que la longueur du filtre séparant est choisie comme la longueur minimale assurant l'existence d'un inverse ($D = N(L - 1)$).

Compte tenu de ce fait, rien ne nous assure *a priori* que le critère à maximiser est borné ni que le problème posé admet une solution.

5.3.2-b) Non unicité du filtre séparant

Nous nous plaçons maintenant dans le cas où \mathbf{R} est positive non définie. Compte tenu de sa définition, ceci survient lorsqu'il existe une relation de dépendance linéaire entre les signaux $\underline{x}_i(n), i \in \{1, \dots, QD\}$ contenus dans $\underline{\mathbf{x}}(n)$. Il s'ensuit alors que le filtre séparant n'est pas unique.

Supposons en effet $\underline{\mathbf{w}}_0^H \in \text{Ker } \mathbf{R}$. Alors, $\underline{\mathbf{w}}_0\mathbf{R}\underline{\mathbf{w}}_0^H = \mathbf{E}\{\underline{\mathbf{w}}_0\mathbf{x}(n)\mathbf{x}(n)^H\underline{\mathbf{w}}_0^H\} = 0$ et le signal $\underline{\mathbf{w}}_0\underline{\mathbf{x}}(n)$ est alors identiquement nul. Ceci montre que deux filtres $\underline{\mathbf{w}}$ et $\underline{\mathbf{w}} + \underline{\mathbf{w}}_0$ avec $\underline{\mathbf{w}}_0^H \in \text{Ker } \mathbf{R}$ donnent le même signal de sortie au niveau du séparateur.

5.3.2-c) Contrainte supplémentaire

Compte tenu des remarques précédentes, nous imposons par conséquent dans le problème d'optimisation la contrainte supplémentaire suivante :

$$\underline{\mathbf{w}}^H \in (\text{Ker } \mathbf{R})^\perp \quad (5.63)$$

et le problème d'optimisation à résoudre s'écrit finalement :

$$\max |\underline{\mathbf{w}}\mathbf{C}\underline{\mathbf{w}}^H| \quad \text{sous les contraintes : } \begin{cases} \underline{\mathbf{w}}\mathbf{R}\underline{\mathbf{w}}^H = 1 \text{ et :} \\ \underline{\mathbf{w}}^H \in (\text{Ker } \mathbf{R})^\perp. \end{cases} \quad (5.64)$$

L'ensemble des vecteurs $\underline{\mathbf{w}}$ satisfaisant les deux contraintes ci-dessus est un compact et l'existence d'une solution est ainsi garantie. La discussion du paragraphe précédent assure par ailleurs qu'il est suffisant de considérer cet ensemble pour décrire l'ensemble des sorties possibles au niveau du séparateur.

5.3.3 Maximisation exacte par décomposition en valeurs singulières

Bien que les résultats soient connus, nous expliquons maintenant pour des raisons d'exhaustivité et de clarté, la façon simple dont le problème d'optimisation (5.64) peut se résoudre. La méthode est basée sur le lemme suivant :

Lemme 6 *Soit $p \in \mathbb{N}^*$ et $\tilde{\mathbf{C}} \in \mathbb{C}^{p \times p}$ une matrice symétrique hermitienne ($\tilde{\mathbf{C}}^H = \tilde{\mathbf{C}}$). Les vecteurs de $\mathbb{C}^{p \times 1}$ solutions du problème d'optimisation*

$$\max |\tilde{\mathbf{x}}^H \tilde{\mathbf{C}} \tilde{\mathbf{x}}| \quad \text{sous la contrainte : } \tilde{\mathbf{x}}^H \tilde{\mathbf{x}} = 1 \quad (5.65)$$

sont les vecteurs appartenant à l'intersection de la sphère unité ($\tilde{\mathbf{x}}^H \tilde{\mathbf{x}} = 1$) et de la réunion des sous-espaces propres de $\tilde{\mathbf{C}}$ relatifs aux valeurs propres maximales en valeur absolue.

Voyons la façon dont nous nous ramenons au lemme précédent. \mathbf{R} étant hermitienne, elle est diagonalisable dans une base orthogonale. \mathbf{R} étant positive, mais non nécessairement définie, nous supprimons dans cette diagonalisation les valeurs propres nulles et les colonnes correspondantes et l'écrivons sous la forme $\mathbf{R} = \mathbf{U}\mathbf{S}\mathbf{U}^H$. Dans cette écriture, \mathbf{S} est carrée, diagonale. De plus, sa taille est celle de \mathbf{R} uniquement dans le cas où cette dernière est définie positive. De même, \mathbf{U} est semi-unitaire (carrée unitaire uniquement si \mathbf{R} est définie positive). Nous définissons alors :

$$\mathbf{P} := \mathbf{S}^{\frac{1}{2}}\mathbf{U}^H \quad \text{et : } \mathbf{Q} := \mathbf{U}\mathbf{S}^{-\frac{1}{2}} \quad (5.66)$$

\mathbf{P} et \mathbf{Q} sont pseudo-inverses l'une de l'autre au sens de Moore-Penrose. Par ailleurs, l'espace engendré par \mathbf{Q} est $\text{Im } \mathbf{Q} = \text{Im } \mathbf{R} = (\text{Ker } \mathbf{R})^\perp$. La contrainte $\underline{\mathbf{w}}^H \in (\text{Ker } \mathbf{R})^\perp$ du problème (5.64) est donc prise en compte si l'on écrit :

$$\underline{\mathbf{w}}^H = \mathbf{Q}\tilde{\mathbf{x}} \quad (\text{ce qui revient à poser : } \tilde{\mathbf{x}} = \mathbf{P}\underline{\mathbf{w}}^H). \quad (5.67)$$

On a alors :

$$\underline{\mathbf{w}}\mathbf{R}\underline{\mathbf{w}}^H = \underline{\mathbf{w}}\mathbf{U}\mathbf{S}\mathbf{U}^H\underline{\mathbf{w}}^H = \underline{\mathbf{w}}\mathbf{P}^H\mathbf{P}\underline{\mathbf{w}}^H = \tilde{\mathbf{x}}^H\tilde{\mathbf{x}}, \quad (5.68)$$

et en posant $\tilde{\mathbf{C}} := \mathbf{Q}^H\mathbf{C}\mathbf{Q}$, nous sommes bien ramenés au problème du lemme 6.

Par conséquent, une solution exacte au problème peut être obtenue par une diagonalisation (ou décomposition en valeurs singulières) de la matrice $\tilde{\mathbf{C}}$. L'avantage que l'on en tire est double : d'une part, la maximisation du contraste ne requiert qu'un nombre fini d'opérations. La rapidité de la maximisation est ainsi fortement accrue et il n'est pas nécessaire de recourir à un algorithme itératif d'optimisation qui se révèle bien plus lourd. D'autre part, le maximum *global* est obtenu et la méthode est donc insensible à l'existence de maxima locaux potentiellement non séparants.

Remarque 15: Le problème d'optimisation défini par (5.62) peut aussi être abordé au travers de l'étude du Lagrangien. On tombe alors sur le problème de recherche de valeurs propres généralisées λ et de vecteurs propres généralisés $\underline{\mathbf{w}}^H$, qui s'écrit :

$$\mathbf{C}\underline{\mathbf{w}}^H = \lambda\mathbf{R}\underline{\mathbf{w}}^H \quad (5.69)$$

La résolution de ce problème est équivalente à l'approche qui a été présentée.

5.3.4 Procédé séquentiel de déflation

Disposant maintenant de tous les éléments pour extraire une source, nous considérons le problème d'une séparation MIMO de l'ensemble des sources. La solution la plus simple consiste à utiliser une procédure de déflation telle qu'elle a été présentée dans le paragraphe 3.2.1.

Nous avons expliqué qu'il fallait dans la méthode d'optimisation tenir compte du rang de la matrice \mathbf{R} définie à l'équation (5.58). Examinons l'effet de la suppression d'une source sur le rang de cette matrice. Nous définissons pour cela $\underline{\mathbf{s}}(n)$ le vecteur colonne de dimension $N(L + D - 1)$:

$$\underline{\mathbf{s}}(n) := (\mathbf{s}(n)^T \quad \mathbf{s}(n-1)^T \quad \dots \quad \mathbf{s}(n-L-D+2)^T)^T \quad (5.70)$$

et \mathcal{T}_D la matrice de Sylvester suivante de taille $QD \times N(L + D - 1)$:

$$\mathcal{T}_D(\mathbf{M}) := \begin{pmatrix} \mathbf{M}(0) & \mathbf{M}(1) & \dots & \mathbf{M}(L-1) & 0 & \dots & 0 \\ 0 & \mathbf{M}(0) & \mathbf{M}(1) & \dots & \mathbf{M}(L-1) & 0 & 0 \\ \vdots & \ddots & \ddots & \ddots & & \ddots & \vdots \\ 0 & \dots & 0 & \mathbf{M}(0) & \mathbf{M}(1) & \dots & \mathbf{M}(L-1) \end{pmatrix}. \quad (5.71)$$

Le vecteur $\underline{\mathbf{x}}(n)$ défini à l'équation (5.55) s'exprime alors par :

$$\underline{\mathbf{x}}(n) = \mathcal{T}_D(\mathbf{M})\underline{\mathbf{s}}(n) \quad (5.72)$$

$$(5.73)$$

ce qui mène à l'expression suivante de la matrice \mathbf{R} :

$$\mathbf{R} := E\{\underline{\mathbf{x}}(n)\underline{\mathbf{x}}(n)^H\} = \mathcal{T}_D(\mathbf{M})E\{\underline{\mathbf{s}}(n)\underline{\mathbf{s}}(n)^H\}\mathcal{T}_D(\mathbf{M})^H \quad (5.74)$$

La suppression de la contribution d'une source au cours d'une étape de déflation est équivalente, dans le cas idéal, à considérer que simultanément sont mis à zéros l'une des composantes du vecteur $\mathbf{s}(n)$ et la colonne correspondante du filtre $\mathbf{M}[z]$. Ceci entraîne donc la mise à zéro simultanément de $L + D - 1$ composantes de $\underline{\mathbf{s}}(n)$ ou de $L + D - 1$ colonnes de $\mathcal{T}_D(\mathbf{M})$, ce qui entraîne une chute équivalente du rang de la matrice \mathbf{R} .

Nous en concluons qu'il convient de prendre en compte dans les procédures précédentes le fait que le rang de \mathbf{R} a chuté de $(D+L-1)$ fois le nombre de contributions déjà soustraites des observations. Les valeurs singulières de \mathbf{R} les plus petites sont donc considérées en nombre correspondant comme nulles dans la description de la procédure de maximisation du paragraphe 5.3.3.

5.4 Simulations

Nous présentons les simulations en considérant d'abord l'extraction d'une seule source. La possibilité d'itérer par processus de déflation sera ensuite vérifiée.

5.4.1 Extraction d'une source à l'aide d'une référence

5.4.1-a) Utilisation de la source comme référence

Un premier test a été mené dans le cas idéal où la source réelle sert de référence. Les résultats sont donnés sur la figure 5.2. Les erreurs quadratiques moyennes sur la source

reconstituée ont été tracées en fonction du nombre d'échantillons disponibles. Ces résultats proviennent d'un moyennage sur 1000 réalisations où les filtres de mélanges, de longueur 3, ont des coefficients tirés aléatoirement selon une loi gaussienne centrée et normalisée. Les simulations ont été réalisées pour des systèmes de mélange à 2 entrées/3 sorties et 3 entrées/4 sorties, ainsi que pour deux types de sources distinctes : i.i.d. uniformes et i.i.d. PAM4. L'observation des EQM moyennes de la figure 5.2 est encourageante et nous invite à engager une étude plus approfondie de la méthode et à considérer le problème de l'obtention d'une référence dans le cadre aveugle.

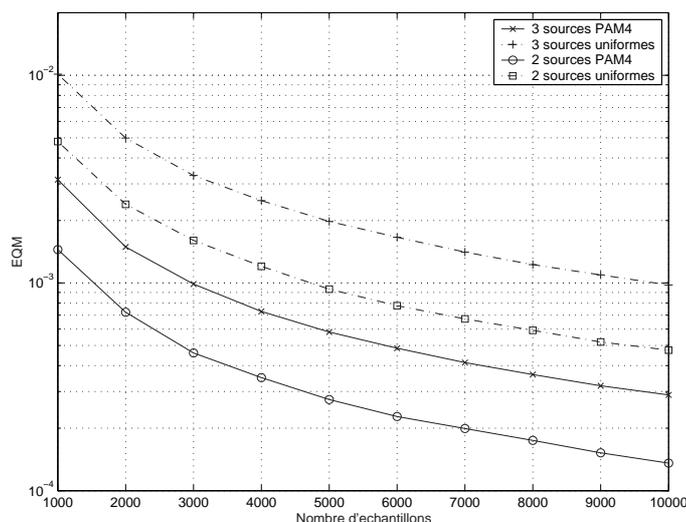


FIG. 5.2 – EQM moyenne sur les sources reconstituées en fonction du nombre d'échantillons. Les valeurs ont été obtenues avec des filtres de mélange de longueur 3, et à l'aide d'un moyennage sur 1000 réalisations.

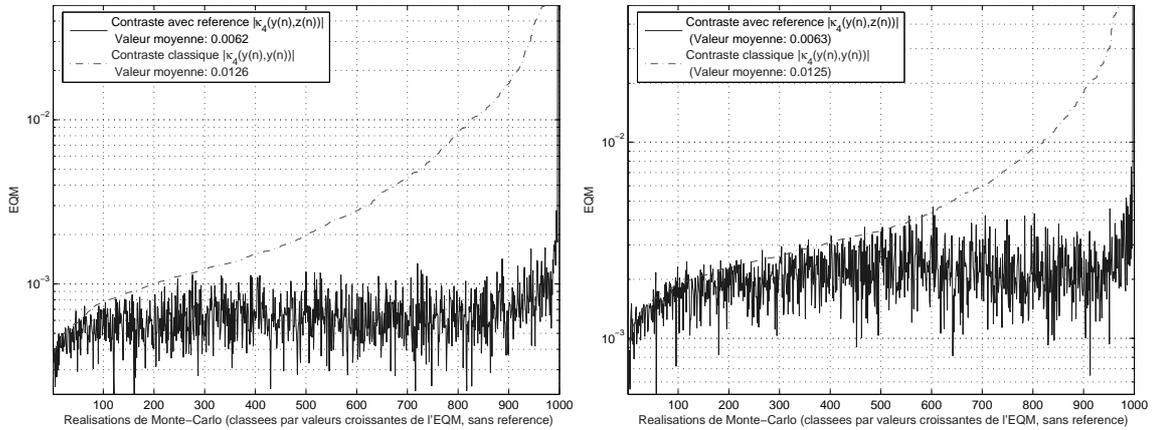
5.4.1-b) Utilisation initiale d'un contraste sans référence

Convergence de l'optimisation complète du contraste sans référence Pour obtenir les résultats précédents (figure 5.2), nous avons artificiellement supposé connue l'une des sources que nous avons utilisée comme signal de référence. Cette méthode n'est naturellement pas applicable en pratique et nous proposons de considérer le cas où la référence utilisée est une source extraite à l'aide d'un contraste MISO tel que ceux présentés au chapitre 3.

Les résultats sont donnés pour les deux types de sources précédemment considérées (i.i.d. uniformes ou PAM4) sur les graphiques de la figure 5.3. Ils représentent, pour chacune des 1000 réalisations de Monte-Carlo, l'EQM sur la source reconstituée pour la méthode sans référence et pour la méthode avec. Pour des raisons de lisibilité, les réalisations ont été classées par ordre croissant de l'EQM obtenue par la méthode sans référence.

Nous constatons que l'utilisation d'un contraste avec référence améliore sensiblement la qualité de la séparation. Ceci illustre un premier intérêt des développements de ce chapitre.

L'apport apparaît toutefois plus nettement si l'on compare les charges de calculs induites par chacune des méthodes : à titre indicatif, pour un mélange avec un filtre 3 sources/4 capteurs, de longueur 3 et 5000 échantillons, la même séparation dure de l'ordre



(a) Sources i.i.d. PAM4.

(b) Sources i.i.d. uniformes.

FIG. 5.3 – EQM sur les sources reconstituées. Comparaison sur 1000 réalisations de Monte-Carlo (3 sources, 4 capteurs, filtres de mélange tirés aléatoirement, 5000 échantillons) du résultat avec le contraste $\mathcal{C}_z\{\cdot\}$ et le contraste classique sans référence $|\kappa_4\{y(n), y(n)\}|$. Les réalisations ont été classées par ordre croissant de l'EQM pour le contraste sans référence.

de 20 secondes pour les 1000 itérations nécessaires à une bonne convergence d'un contraste classique. En comparaison, l'étape d'optimisation du contraste avec référence prend, dans les mêmes conditions, au plus 20 *centièmes* de secondes. Le gain en temps est donc considérable et constitue l'atout majeur de la méthode. De plus, la méthode ne se trouve que faiblement handicapée par un nombre accru d'échantillons : l'estimation des matrices notées \mathbf{C} et \mathbf{R} dans le paragraphe 5.3 ne se réalise en effet qu'une seule fois et l'optimisation s'effectue alors en ne faisant appel qu'à ces seules matrices dont la taille est réduite. Ainsi, pour 10000 et 50000 échantillons, les temps respectifs d'optimisation passent à 44 et 221 secondes pour un contraste classique contre 23 et 93 centièmes de secondes pour le contraste avec référence. Ce point explique par ailleurs le nombre très élevé d'échantillons qui a pu être utilisé dans la suite de nos simulations.

Non convergence de l'optimisation du contraste sans référence Compte tenu des résultats précédents, il est apparu intéressant de regarder l'effet de l'utilisation comme référence du résultat fourni par un contraste classique mais dont l'optimisation itérative a été interrompue avant convergence. Les résultats sont donnés à la figure 5.4. Il apparaît une grande majorité des cas où le contraste par référence offre une performance très honorable malgré une référence manifestement éloignée de l'une des sources. Cependant, pour environ 5% des cas, le résultat du contraste classique n'est pas convenable (en raison d'une optimisation inachevée) et la méthode par référence ne parvient pas non plus à une solution acceptable. L'étude du paragraphe 5.4.2 fournira des éléments d'explication. Le gain en temps de calcul demeure toutefois élevé dans la mesure où l'on s'épargne la nécessité d'attendre la convergence de l'algorithme d'optimisation itératif.

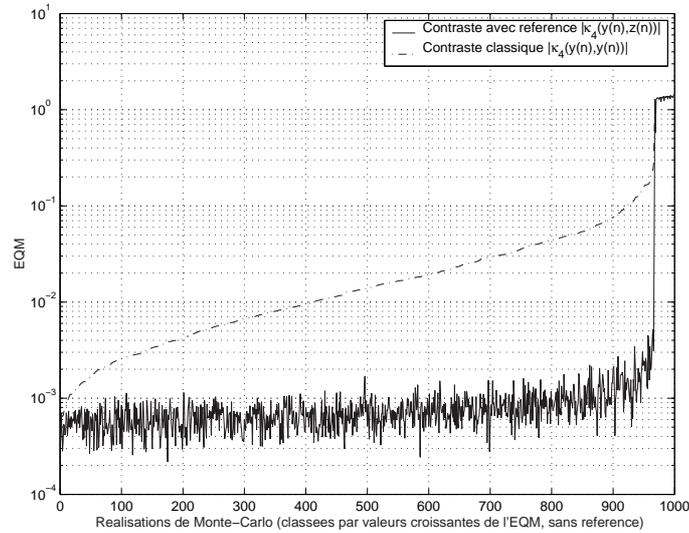


FIG. 5.4 – EQM sur les sources reconstituées (3 sources de type PAM4, 4 capteurs, filtres de mélange tirés aléatoirement, 5000 échantillons). L’optimisation du contraste sans référence n’a pas été menée jusqu’à la convergence et le résultat obtenu a servi de référence. Les réalisations ont été classées par ordre croissant de l’EQM pour le contraste sans référence.

5.4.1-c) Cas de sources non i.i.d.

Les simulations précédentes ont été brièvement reprises pour des sources non i.i.d. afin d’effectuer une vérification de principe de la validité de la méthode dans ce cas. Les résultats sont donnés dans le tableau 5.1 pour des sources CPM d’indices de modulation 0.25, 0.4 et 0.6.

	Contraste sans réf, 10 itérations	Avec réf. colonne préc.	Contraste sans réf, 100 itérations	Avec réf. colonne préc.
Sources CPM, 5000 échantillons	0.1993	0.0052	0.0068	0.0010
Sources CPM, 2000 échantillons	0.1917	0.0123	0.0075	0.0022

TAB. 5.1 – Valeur moyenne du critère τ_1 sur 1000 réalisations. Le mélange était tiré aléatoirement, à coefficients complexes avec 3 sources et 4 capteurs.

5.4.2 Tirage aléatoire du système de référence

Au cours des simulations précédentes, l’obtention d’un signal de référence a constitué la difficulté majeure dans l’utilisation des techniques de ce chapitre. Il est cependant apparu dans les propositions 22 et 23 que les conditions sur le système de référence étaient relativement faibles. En particulier, n’importe quel contraste avec référence obtenu à l’aide d’un système de référence tiré aléatoirement devrait d’après la proposition 23 mener à la séparation. Cette affirmation doit cependant être modulée par le fait que nous ne disposons que de statistiques estimées. Nous examinons maintenant ce point au travers de simulations.

Parallèlement, il nous est venu à l'idée d'itérer le processus de séparation avec référence et d'obtenir ainsi un algorithme de type «point fixe». Plus précisément, la méthode mise en œuvre et testée a été la suivante :

- choisir par tirage aléatoire un vecteur $\underline{\mathbf{w}}^0$ correspondant à un système séparant de référence.
- pour n variant de 1 à n_{\max} :
 - fixer la référence à $z(n) = \underline{\mathbf{w}}^{n-1} \underline{\mathbf{x}}(n)$.
 - optimiser $\mathcal{C}_z\{\cdot\}$, ce qui donne le filtre séparateur $\underline{\mathbf{w}}^n$.
- le filtre séparateur retenu est celui correspondant à $\underline{\mathbf{w}}^{n_{\max}}$ et la source reconstituée s'écrit $\underline{\mathbf{w}}^{n_{\max}} \underline{\mathbf{x}}(n)$.

Les résultats fournis par cette méthode sur 100 réalisations sont livrés dans les tableaux 5.2, 5.3 et 5.4 pour un nombre d'itération maximal $n_{\max} = 5$. Le nombre de sources a été fixé à 3, le nombre de capteurs à 4. Les filtres de mélange étaient de longueur 3 et ont été tirés aléatoirement (loi gaussienne centrée normalisée sur les coefficients).

Nombre d'itérations	EQM moyenne	% de réalisations avec une EQM :		
		$< 5.10^{-2}$	$< 10^{-2}$	$< 10^{-3}$
1	0.2687	18	0	0
2	0.0435	88	73	16
3	0.0087	96	96	81
4	0.0009	100	99	91
5	0.0006	100	100	96

TAB. 5.2 – Optimisation itérée d'un contraste avec référence tirée au hasard : résultat pour 3 sources i.i.d. PAM4 et 5000 échantillons.

Nombre d'itérations	EQM moyenne	% de réalisations avec une EQM :		
		$< 5.10^{-2}$	$< 10^{-2}$	$< 10^{-3}$
1	0.1662	80	5	0
2	0.0069	99	90	69
3	0.0003	100	100	99
4	0.0003	100	100	100
5	0.0003	100	100	100

TAB. 5.3 – Optimisation itérée d'un contraste avec référence tirée au hasard : résultat pour 3 sources i.i.d. PAM4 et 10000 échantillons.

Nombre d'itérations	EQM moyenne	% de réalisations avec une EQM :		
		$< 5.10^{-2}$	$< 10^{-2}$	$< 10^{-3}$
1	0.0191	89	60	4
2	3.5620e-5	100	100	100
3	2.8186e-5	100	100	100
4	2.8186e-5	100	100	100
5	2.8186e-5	100	100	100

TAB. 5.4 – Optimisation itérée d'un contraste avec référence tirée au hasard : résultat pour 3 sources i.i.d. PAM4 et 100000 échantillons.

Nous constatons dans le tableau 5.4 que, pour un nombre d'échantillons grand, le

résultat de séparation est en général bon dès la première maximisation du contraste avec référence. Un examen plus attentif indique toutefois que les cas de séparation insatisfaisante à la première étape sont liés à l'existence de deux valeurs propres relativement proches dans la matrice notée $\tilde{\mathbf{C}}$ au paragraphe 5.3, donnant ainsi une possibilité de se prémunir contre de tels cas. Par ailleurs, l'augmentation du nombre d'échantillons permet de fortement limiter le nombre de cas non séparant dès la première étape, comme l'indique l'évolution des valeurs dans les tableaux 5.2, 5.3 et 5.4.

En conclusion, en présence d'un nombre limité d'échantillons, les erreurs d'estimation des matrices \mathbf{C} et \mathbf{R} peuvent conduire à un écart de performances par rapport au comportement théorique prédit par la proposition 23. Néanmoins, il est tout à fait positif de constater que, dans les cas où la séparation n'est pas effective dès la première maximisation du contraste avec référence, le fait de recommencer cette maximisation permet de converger vers l'une des sources. Ainsi, la méthode proposée semble pouvoir s'appliquer dans un cadre aveugle.

5.4.3 Séparation MIMO par déflation

La séparation MIMO a été testée avec 3 sources, 4 capteurs, des filtres de mélange tirés aléatoirement de longueur 3 et des sources i.i.d. PAM4. Les sources ont été extraites en utilisant la méthode de maximisations successives d'un contraste avec référence décrite au paragraphe précédent. Le passage à l'extraction MIMO a été mené par des étapes de déflation successives décrites au paragraphe 3.2.1 et en prenant en compte la remarque du paragraphe 5.3.4.

Les résultats provenant de 100 réalisations de Monte-Carlo sont regroupés dans les tableaux 5.5, 5.6 et 5.7. Nous pouvons constater que la méthode offre des résultats satisfaisants.

Numéro source	EQM moyenne	% de réalisations avec une EQM :		
		$< 5.10^{-2}$	$< 10^{-2}$	$< 5.10^{-3}$
1	0.0006	100	100	100
2	0.0569	96	95	87
3	0.0251	96	91	30

TAB. 5.5 – EQM moyenne sur chacune des 3 sources extraites. résultats pour 5000 échantillons et, pour chaque extraction, 15 itérations de maximisation d'un contraste avec référence.

Numéro source	EQM moyenne	% de réalisations avec une EQM :		
		$< 5.10^{-2}$	$< 10^{-2}$	$< 5.10^{-3}$
1	0.0003	100	100	100
2	0.0145	98	98	97
3	0.0251	98	98	93

TAB. 5.6 – EQM moyenne sur chacune des 3 sources extraites. résultats pour 10000 échantillons et, pour chaque extraction, 10 itérations de maximisation d'un contraste avec référence.

Numéro source	EQM moyenne	% de réalisations avec une EQM :		
		$< 5.10^{-2}$	$< 10^{-2}$	$< 5.10^{-3}$
1	3.0512e-5	100	100	100
2	0.0095	99	99	99
3	0.0053	99	99	99

TAB. 5.7 – EQM moyenne sur chacune des 3 sources extraites. résultats pour 100000 échantillons et, pour chaque extraction, 5 itérations de maximisation d'un contraste avec référence.

5.5 Conclusion

Partant de la notion de signal de référence, nous avons introduit une classe de contrastes MISO utilisables *a priori* dans un cas semi-aveugle. La validité de ces critères de contraste a été établie aussi bien dans le cas de sources i.i.d. que dans le cas non i.i.d. Compte tenu des conditions de validité de ces contrastes, les connaissances nécessaires à l'obtention d'un signal de référence se sont révélées cependant relativement faibles. Les simulations ont ainsi mis en évidence la possibilité d'utiliser ces contrastes dans un cadre aveugle. Enfin, l'avantage majeur des critères introduits réside dans leur dépendance quadratique vis à vis des paramètres. Il en résulte la possibilité de les optimiser par une méthode rapide et efficace. Ainsi, ces contrastes MISO offrent une méthode performante d'extraction d'une source. Combinés avec des étapes de déflations, nos simulations laissent augurer d'excellentes perspectives pour la séparation de mélanges MIMO.

Chapitre 6

Conclusion

6.1 Résumé des travaux réalisés

Les travaux que nous avons menés ont concerné la séparation de sources dans le cadre des mélanges convolutifs. Parmi les principales difficultés que soulève cette problématique nous pouvons citer :

- la recherche de critères de séparation et leur étude, en particulier en ce qui concerne l'existence et les propriétés des maxima locaux ;
- la recherche de procédures d'optimisation efficaces, qui permettent le cas échéant de s'affranchir des maxima locaux indésirables ;
- l'extension au cas non i.i.d. des méthodes existantes, généralement proposées dans le cadre i.i.d. ;
- la recherche de méthodes de séparation conjointes offrant une alternative aux méthodes séquentielles, notamment à la déflation.

Ce travail s'est d'abord concentré sur les méthodes conjointes et l'étude des contrastes qui leur sont associés. Par le biais d'une approche fréquentielle, une famille de contrastes temporels, valables pour des sources non i.i.d. a été mise en évidence. L'optimisation de ces critères est toutefois délicate en raison de l'existence de maxima locaux non séparants. D'autre part, l'étape nécessaire de préblanchiment peut soulever des difficultés supplémentaires.

Les méthodes séquentielles présentent pour leur part des avantages non négligeables. D'abord, l'avantage décisif de ces approches réside en la possibilité de choisir des contrastes de telle sorte que tout maximum local soit séparant. Ensuite, les contrastes MISO auxquels font appel ces approches demeurent valables dans le cas non i.i.d. ce qui permet d'élargir l'horizon des applications potentielles. Enfin, l'absence de préblanchiment facilite une mise en œuvre concrète.

L'approche séquentielle que nous avons proposée apparaît comme une alternative à l'approche dite par déflation. Comme pour l'ensemble des méthodes séquentielles, on peut observer un phénomène gênant d'accumulation des erreurs au fur et à mesure que les sources sont extraites. La possibilité de mise en œuvre d'une post-optimisation locale et sans contraintes a ainsi été justifiée. Ce post-traitement permet de rétablir une symétrie dans le traitement des sources et ainsi de limiter le problème de dégradation des résultats au cours des extractions.

Deux applications des méthodes séquentielles ont pu être envisagées en faisant appel à des connaissances extérieures au domaine intrinsèque de la séparation de sources. Des résultats sur la démodulation aveugle de signaux CPM ont permis, en liaison avec les

nôtres, d'obtenir la séparation et la détection de plusieurs signaux CPM. D'autre part, la séparation de sources a été généralisée aux mélanges convolutifs multidimensionnels et aux images en particulier, par le recours à des résultats sur les matrices polynomiales à plusieurs indéterminées.

L'optimisation des contrastes utilisés dans les approches MISO se réalise en général par une méthode itérative, qui conduit à une importante charge de calcul. La possibilité d'utiliser une connaissance *a priori* supplémentaire a mené à la notion de contrastes avec référence. Ces derniers combinent à eux seuls la plupart des atouts souhaitables : la facilité d'optimisation permet surtout d'obtenir aisément et en un nombre fini d'opérations le maximum global. De ce fait, le problème de l'existence de maxima locaux ne se pose pas. De plus, ces critères restent valables dans le cas de sources non i.i.d. Même si de tels critères ont été envisagés initialement dans un cadre semi-aveugle, les connaissances requises pour leur utilisation sont en réalité faibles et leur étude a fait également apparaître leur grand intérêt dans le cadre aveugle.

6.2 Perspectives

La séparation de sources dans les cas convolutifs demeure un problème ouvert et attrayant ; un certain nombre de pistes méritent encore d'être explorées :

- Méthodes conjointes : la difficulté majeure semble actuellement celle du préblanchiment. La recherche de solutions efficaces à ce problème est une condition préalable pour envisager des applications pragmatiques de ces méthodes. Les difficultés de maximisation des critères de contraste, en partie dues à l'existence de maxima locaux font également partie des préoccupations importantes, même si l'on peut envisager de compenser ce problème par le recours à des méthodes numériques plus lourdes.
- Approches hiérarchiques : l'étude des propriétés locales des contrastes MISO s'est révélée fructueuse. Il pourrait être intéressant de mieux exploiter ces propriétés locales qui ont souvent été négligées au profit d'études globales.
- Application aux CPM : les résultats présentés mériteraient d'être complétés par l'étude du cas où la période d'échantillonnage n'est plus égale à la période symbole. Dans ce cas, le problème devient celui de la séparation de sources cyclo-stationnaires.
- Application aux images : la séparation de mélanges d'images devrait s'inspirer des méthodes de traitement d'image et introduire des connaissances *a priori*. Ainsi, l'utilisation d'un modèle sur les images ou sur le filtre de mélange pourrait vraisemblablement améliorer les résultats.
- Contrastes quadratiques : les résultats sont extrêmement encourageants. Il convient de les approfondir et de se pencher plus attentivement sur le cas de sources non i.i.d. Par ailleurs, toutes les extensions proposées dans le cadre des approches hiérarchiques méritent en parallèle un examen dans le cas des contrastes quadratiques : étude du cas non stationnaire, étude locale, ...

Annexe A

Inversibilité des systèmes MIMO et matrices polynomiales

Le but de cette annexe est d'expliquer comment les résultats d'inversibilité des systèmes MIMO dans le cas monodimensionnel s'étendent au cas multidimensionnel. Ces résultats justifient que l'on ait pu étendre les méthodes de séparation de sources au cas des images. Nous nous ramènerons à l'étude des matrices polynomiales à propos desquelles cette annexe se propose de rassembler les résultats utiles dans notre cadre. Ces résultats ont parfois été établis indépendamment et il n'est pas toujours évident d'en faire la synthèse. Les résultats essentiels dans le cas monodimensionnel seront brièvement rappelés. Le cas multidimensionnel nous obligera ensuite à aborder le cas des matrices polynomiales à plusieurs indéterminées.

A.1 Introduction et notations

A.1.1 Systèmes MIMO et matrices polynomiales

En séparation de sources, nous considérons un système de mélange MIMO à Q sorties et N entrées. Si ce système est supposé RIF et causal, sa transformée en z s'écrit dans le cas monodimensionnel :

$$\mathbf{M}[z] = \sum_{k=0}^{L-1} \mathbf{M}(k)z^{-k} \quad (\text{A.1})$$

Nous nous intéressons aux conditions sous lesquelles le système ci-dessus admet un inverse à gauche, c'est-à-dire les conditions d'existence d'une matrice polynomiale $\mathbf{W}[z]$ telle que

$$\mathbf{W}[z]\mathbf{M}[z] = \mathbf{Id}_N. \quad (\text{A.2})$$

Cette condition d'inversion est cependant plus forte que celle requise par la séparation de sources dans la mesure où toute séparation ne peut être réalisée qu'à un filtre près ou, dans le meilleur des cas à un retard et une amplitude près. Il est par conséquent suffisant de chercher $\mathbf{W}[z]$ polynomiale telle que :

$$\mathbf{W}[z]\mathbf{M}[z] = \begin{pmatrix} z^{-d_1} & 0 & \dots & 0 \\ 0 & z^{-d_2} & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \dots & 0 & z^{-d_N} \end{pmatrix} \quad (\text{A.3})$$

où d_1, \dots, d_N sont des entiers (généralement positifs si les systèmes considérés sont causaux). En multipliant l'équation ci-dessus à gauche par la matrice $\text{Diag}(z^{d_1}, \dots, z^{d_N})$, la condition ci-dessus apparaît comme équivalente à la condition (A.2) si l'on considère que les matrices ont leurs coefficients choisis dans l'anneau des *polynômes de Laurent*. Bien qu'il soit possible de toujours se ramener au cas des polynômes classiques, ceci constitue un détour non nécessaire (dans [61] les auteurs sont ainsi amenés à affaiblir la notion d'inversibilité) et il est bien plus naturel pour la séparation de sources de considérer les polynômes de Laurent.

Le cas des systèmes multidimensionnels nous intéressera plus particulièrement ; cette dimension sera notée p et il suffira de considérer $p = 2$ dans le cas particulier des images. Nous considérerons ainsi les matrices polynomiales à p indéterminées z_1, \dots, z_p . Nous noterons en caractère gras le p -uplet $\mathbf{z} := (z_1, \dots, z_p)$, ainsi que tous les p -uplets que nous serons amenés à considérer. Ainsi, à un système MIMO, RIF et causal est associée la matrice :

$$\mathbf{M}[z_1, \dots, z_p] = \sum_{k_1=0}^{L_1-1} \dots \sum_{k_p=0}^{L_p-1} \mathbf{M}(k_1, \dots, k_p) z^{-k_1} \dots z^{-k_p} \quad (\text{A.4})$$

ce que nous noterons également :

$$\mathbf{M}[\mathbf{z}] = \sum_{\mathbf{k}=0}^{\mathbf{L}-1} \mathbf{M}(\mathbf{k}) \mathbf{z}^{-\mathbf{k}} \quad (\text{A.5})$$

La problématique peut être exprimée de façon identique au cas unidimensionnel : sous quelle condition existe-t-il une matrice polynomiale à plusieurs indéterminées (polynômes de Laurent) telle que $\mathbf{W}[\mathbf{z}]\mathbf{M}[\mathbf{z}] = \mathbf{Id}$?

A.1.2 Notations

Dans cette annexe, \mathbb{K} désigne un corps quelconque. \mathbb{R} et \mathbb{C} nous intéresseront plus particulièrement, mais dans le cas multidimensionnel, il peut être utile de prendre pour \mathbb{K} un corps de fractions rationnelles. Nous notons alors :

- $\mathbb{K}[z]$ l'anneau des polynômes classiques à coefficients dans \mathbb{K} et en l'indéterminée z .
- $\mathbb{K}[z, z^{-1}]$ l'anneau des polynômes de Laurent à coefficients dans \mathbb{K} et en l'indéterminée z .
- $\mathbb{K}[\mathbf{z}]$ l'anneau des polynômes classiques à coefficients dans \mathbb{K} et en les indéterminées z_1, \dots, z_p .
- $\mathbb{K}[\mathbf{z}, \mathbf{z}^{-1}]$ l'anneau des polynômes de Laurent à coefficients dans \mathbb{K} et en les indéterminées z_1, \dots, z_p .

A.2 Systèmes MIMO monodimensionnels

Dans cette partie sont rappelés les résultats d'inversibilité qui ont été utilisés à propos des filtres MIMO à réponse impulsionnelle finie. On pourra consulter [18, 50] pour une introduction aux matrices polynomiales ; d'autres éléments sont disponibles dans [36]. Les résultats d'inversibilité que nous rappelons sont également cités et utilisés dans [39] et [42]. Notons également qu'une synthèse a été réalisée dans [61] ; mais cette dernière reste dans le cadre des polynômes classiques et requiert de ce fait l'introduction non indispensable de conditions dites de «réversibilité». Enfin, les résultats donnés peuvent être déduits du cas général exposé dans le paragraphe A.3.

Matrices polynomiales à une indéterminée On appelle matrice polynomiale une matrice dont les éléments sont dans l'anneau des polynômes ($\mathbb{C}[z]$ ou $\mathbb{C}[z, z^{-1}]$ selon le cas).

Définition 8 Une matrice polynomiale carrée $\mathbf{M}[z] \in \mathbb{C}[z]^{N \times N}$ (ou $\mathbb{C}[z, z^{-1}]$) est unimodulaire lorsqu'elle admet un inverse qui est une matrice polynomiale (c'est-à-dire dans $\mathbb{C}[z]^{N \times N}$ ou $\mathbb{C}[z, z^{-1}]$ respectivement).

Le cas d'une matrice ayant plus de lignes que de colonnes nous intéresse également :

Définition 9 Soit $\mathbf{M}[z] \in \mathbb{C}[z]^{Q \times N}$ (resp. $\mathbf{M}[z] \in \mathbb{C}[z, z^{-1}]^{Q \times N}$) une matrice polynomiale. $\mathbf{M}[z]$ est dite «coprime» lorsque ses déterminants mineurs de taille maximale sont premiers entre eux.

Le résultat ci-dessous permet de caractériser les matrices «coprime» :

Propriété 4 Soit $\mathbf{M}[z] \in \mathbb{C}[z]^{Q \times N}$ (resp. $\mathbf{M}[z] \in \mathbb{C}[z, z^{-1}]^{Q \times N}$) une matrice polynomiale avec $Q \geq N$. $\mathbf{M}[z]$ est «coprime» si et seulement si l'une des conditions équivalentes ci-dessous est vérifiée :

1. $\mathbf{M}[z]$ est irréductible, c'est-à-dire $\text{rang}(\mathbf{M}[z]) = N, \forall z \in \mathbb{C}$ (resp. $\forall z \in \mathbb{C}^*$).
2. $\mathbf{M}[z]$ admet un inverse à gauche, c'est-à-dire il existe $\mathbf{W}[z] \in \mathbb{C}[z]^{Q \times N}$ (resp. $\mathbb{C}[z, z^{-1}]^{Q \times N}$) tel que $\mathbf{W}[z]\mathbf{M}[z] = \text{Id}$.

Inversibilité d'un système MIMO-FIR monodimensionnel Nous pouvons déduire de la propriété 4 ci-dessus qu'un filtre RIF-MIMO à N entrées et Q sorties de transformée en z $\mathbf{M}[z] \in \mathbb{C}[z, z^{-1}]^{Q \times N}$ admet un inverse FIR si et seulement si il satisfait la condition :

$$Q \geq N \text{ et } : \forall z \in \mathbb{C}^*, \text{rang}(\mathbf{M}[z]) = N \quad (\text{A.6})$$

A.3 Matrices à plusieurs indéterminées

A.3.1 Notions de primalité

Dans le cas de polynômes à plusieurs indéterminées, la notion de polynômes premiers entre eux est plus délicate à définir : la difficulté vient en effet en partie du fait que l'anneau des polynômes à plusieurs indéterminées n'est plus un anneau principal (voir [28] pour les propriétés élémentaires des polynômes à plusieurs variables). Nous n'irons pas plus loin dans ces considérations algébriques et nous contenterons d'introduire, dans le cas matriciel, les quatre notions suivantes [35] :

Définition 10 Soit $\mathbf{M}[\mathbf{z}] \in \mathbb{K}[\mathbf{z}, \mathbf{z}^{-1}]^{Q \times N}$. On note $\mathcal{I}_{\mathbf{M}}$ l'idéal engendré par les déterminants mineurs de taille maximale de $\mathbf{M}[\mathbf{z}]$. Par définition, $\mathbf{M}[\mathbf{z}]$ est alors dite :

- «right zero prime» (rZP) lorsque $Q \geq N$ et $\mathcal{I}_{\mathbf{M}} = \mathbb{K}[\mathbf{z}, \mathbf{z}^{-1}]$. En particulier, si $Q = N$, une matrice rZP est appelée matrice unimodulaire.
- «right variety prime» (rVP) lorsque $Q \geq N$ et $\mathcal{I}_{\mathbf{M}}$ contient des polynômes de Laurent (non nuls) de $\mathbb{K}[z_i, z_i^{-1}]$ pour tout $i \in \{1, \dots, p\}$.
- «right minor prime» (rMP) lorsque $Q \geq N$ et les déterminants mineurs de $\mathbf{M}[\mathbf{z}]$ n'ont pas de facteurs communs non triviaux (c'est-à-dire qui ne soient pas des éléments inversibles de l'anneau).

– «right factor prime» (rFP) lorsque $Q \geq N$ et dans toute décomposition $\mathbf{M}[\mathbf{z}] = \widetilde{\mathbf{M}}[\mathbf{z}]\mathbf{T}[\mathbf{z}]$ avec $\mathbf{T}[\mathbf{z}]$ carrée, alors $\det \mathbf{T}[\mathbf{z}] \in \mathbb{K}$ (c'est-à-dire $\mathbf{T}[\mathbf{z}]$ est unimodulaire). Les mêmes définitions s'appliquent mutatis mutandis au cas des matrices dont les éléments sont des polynômes classiques de $\mathbb{K}[\mathbf{z}]$.

Les notions de rZP, rVP et rMP apparaissent naturellement dans l'étude des polynômes multivariés, tandis que la notion de rFP n'est utilisée que dans le cas des matrices. On montre aisément les implications suivantes :

$$\text{rZP} \Rightarrow \text{rVP} \Rightarrow \text{rMP} \Rightarrow \text{rFP} \quad (\text{A.7})$$

Dans le cas monodimensionnel, ces quatre définitions sont équivalentes. Dans le cas bidimensionnel, il y a équivalence entre les notions rVP, rMP et rFP. En revanche, dans le cas général multidimensionnel où $p > 2$, aucune réciproque des implications ci-dessus n'est vérifiée. Notons pour terminer que nous pourrions considérer les mêmes notions que celles de la définition 10 mais à gauche. Nous avons privilégié les notions à droite qui sont celles qui nous servent directement pour notre application.

A.3.2 Quelques propriétés

Les notions précédentes permettent de donner une condition nécessaire et suffisante d'existence d'un inverse polynomial. D'après [35], la propriété suivante est vérifiée :

Propriété 5 Soit $\mathbf{M}[\mathbf{z}] \in \mathbb{K}[\mathbf{z}, \mathbf{z}^{-1}]^{Q \times N}$. Les proposition suivantes sont équivalentes :

1. $\mathbf{M}[\mathbf{z}]$ est rZP.
2. $\mathbf{M}[\mathbf{z}]$ peut être bordée de manière à obtenir une matrice unimodulaire de taille $Q \times Q$.
3. Il existe une matrice $Q \times Q$ unimodulaire $\mathbf{U}[\mathbf{z}]$ telle que :

$$\mathbf{U}[\mathbf{z}]\mathbf{M}[\mathbf{z}] = \begin{bmatrix} \mathbf{Id}_N \\ \mathbf{0} \end{bmatrix}. \quad (\text{A.8})$$

4. $\mathbf{M}[\mathbf{z}]$ admet un inverse à gauche polynomial, c'est-à-dire il existe une matrice polynomiale $\mathbf{W}[\mathbf{z}] \in \mathbb{K}[\mathbf{z}, \mathbf{z}^{-1}]^{N \times Q}$ telle que :

$$\mathbf{W}[\mathbf{z}]\mathbf{M}[\mathbf{z}] = \mathbf{Id}_N \quad (\text{A.9})$$

(relation dite de Bezout).

5. pour tout vecteur $\mathbf{u}[\mathbf{z}] \in \mathbb{K}[\mathbf{z}, \mathbf{z}^{-1}]$ qui est rZP, $\mathbf{M}[\mathbf{z}]\mathbf{u}[\mathbf{z}]$ est également rZP.

Compte tenu de ces équivalences, il apparaît qu'une matrice admet un inverse polynomial si et seulement si elle est rZP. La notion de matrice rZP est donc celle qui nous concerne pour l'inversibilité des filtres MIMO RIF multivariés.

A.3.3 Structures algébriques et variétés

Dans cette partie, \mathbb{C} est le corps de départ. Le caractère rZP d'une matrice a été défini d'un point de vue algébrique. De même que dans le cas monodimensionnel, une matrice était coprime si et seulement si elle était irréductible, nous souhaitons ici donner une condition équivalente en terme de variété dans le cas où \mathbb{C} est le corps de départ. Les théorèmes et définitions donnés ici sans démonstration se trouvent dans [28]. Nous donnons en plus les quelques éléments qui permettent d'appliquer les résultats nécessaires au cas des polynômes de Laurent.

1.3.3-a) Variété et idéal, Nullstellensatz de Hilbert

Définition 11 Soit $f \in \mathbb{C}[\mathbf{z}]$ un polynôme. On appelle variété du polynôme f l'ensemble de ses racines :

$$\mathcal{V}(f) := \{\mathbf{z} \in \mathbb{C}^p \mid f(\mathbf{z}) = 0\} \quad (\text{A.10})$$

Si f_1, \dots, f_r sont des polynômes de $\mathbb{C}[\mathbf{z}]$, la variété qui leur est associée est l'intersection des variétés associées à chacun d'eux :

$$\mathcal{V}(f_1, \dots, f_r) := \{\mathbf{z} \in \mathbb{C}^p \mid f_1(\mathbf{z}) = \dots = f_r(\mathbf{z}) = 0\} \quad (\text{A.11})$$

La propriété suivante résulte du théorème de Hilbert sur les anneaux noethériens¹ et permet de ramener les propriétés sur un idéal à des propriétés sur un ensemble fini de polynômes :

Propriété 6 Tout idéal de $\mathbb{C}[\mathbf{z}]$ est engendré par un nombre fini d'éléments (i.e. $\mathbb{C}[\mathbf{z}]$ est noethérien).

Dès lors, le résultat ci-dessus permet de vérifier aisément la cohérence de la définition suivante de la variété associée à un idéal :

Définition 12 Soit \mathcal{I} un idéal et f_1, \dots, f_r une famille génératrice finie ($r \in \mathbb{N}$) de \mathcal{I} . La variété $\mathcal{V}(\mathcal{I})$ associée à l'idéal \mathcal{I} est la variété associée à la famille f_1, \dots, f_r :

$$\mathcal{V}(\mathcal{I}) := \mathcal{V}(f_1, \dots, f_r) \quad (\text{A.12})$$

Le théorème suivant constitue le premier lien fort entre propriétés géométriques et algébriques [28] :

Théorème 3 (Nullstellensatz (forme faible)) Soit \mathcal{I} un idéal de $\mathbb{C}[\mathbf{z}]$. Si $\mathcal{V}(\mathcal{I}) = \emptyset$, alors $\mathcal{I} = \mathbb{C}[\mathbf{z}]$ (et réciproquement.)

Par application de la forme faible du Nullstellensatz, il est immédiat qu'une matrice $\mathbf{M}[\mathbf{z}] \in \mathbb{C}[\mathbf{z}]^{Q \times N}$ (avec toujours $Q \geq N$) est rZP si et seulement si $\mathbf{M}[\mathbf{z}]$ est de rang maximal quel que soit $\mathbf{z} \in \mathbb{C}^p$. En effet, $\mathbf{M}[\mathbf{z}]$ est rZP si et seulement si l'idéal $\mathcal{I}_{\mathbf{M}}$ engendré par ses mineurs de taille maximale est $\mathbb{C}[\mathbf{z}]$, c'est-à-dire si et seulement si $\mathcal{V}(\mathcal{I}_{\mathbf{M}}) = \emptyset$, ce qui signifie que l'ensemble des points annulant simultanément tous ses mineurs de taille maximale est vide.

Pour démontrer la propriété analogue dans le cas des polynômes de Laurent nous aurons besoin de l'énoncé suivant [28] :

Théorème 4 (Nullstellensatz de Hilbert) Soit $f \in \mathbb{C}[\mathbf{z}]$ et \mathcal{I} un idéal de $\mathbb{C}[\mathbf{z}]$. Si pour tout p -uplet $(a_1, \dots, a_p) \in \mathcal{V}(\mathcal{I})$ on a $f(a_1, \dots, a_p) = 0$, alors il existe $r \in \mathbb{N}^*$ tel que $f^r \in \mathcal{I}$ (et réciproquement).

¹Un anneau est dit noethérien lorsque tout idéal de l'anneau est engendré par un nombre fini d'éléments.

1.3.3-b) Variété de Laurent

Nous adaptons ici au cas des polynômes de Laurent quelques propriétés en général présentées dans le cadre des polynômes classiques².

Définition 13 Soit $f \in \mathbb{C}[\mathbf{z}, \mathbf{z}^{-1}]$ un polynôme de Laurent. On appelle variété de Laurent du polynôme f l'ensemble de ses racines :

$$\mathcal{V}_L(f) := \{\mathbf{z} \in \mathbb{C}^{*p} \mid f(\mathbf{z}) = 0\} \quad (\text{A.13})$$

On peut montrer que $\mathbb{C}[\mathbf{z}, \mathbf{z}^{-1}]$ est noethérien, d'où la possibilité de donner la définition suivante :

Définition 14 Si \mathcal{I} est une collection de polynômes (en particulier si \mathcal{I} est un idéal), sa variété de Laurent $\mathcal{V}_L(\mathcal{I})$ est l'intersection des variétés de Laurent de ses éléments.

Nous pouvons alors démontrer que :

Propriété 7 Soit $\mathbf{M}[\mathbf{z}] \in \mathbb{C}[\mathbf{z}, \mathbf{z}^{-1}]^{Q \times N}$. $\mathbf{M}[\mathbf{z}]$ est rZP si et seulement si $Q \geq N$ et $\mathbf{M}[\mathbf{z}]$ est de rang plein pour tout $\mathbf{z} \in \mathbb{C}^{*p}$ (c'est-à-dire $\mathcal{V}_L(\mathcal{I}_{\mathbf{M}}) = \emptyset$).

Preuve: Si $\mathbf{M}[\mathbf{z}]$ est rZP, alors l'idéal engendré par ses mineurs de taille maximale est $\mathcal{I}_{\mathbf{M}} = \mathbb{C}[\mathbf{z}, \mathbf{z}^{-1}]$ et la variété associée à cet idéal est l'ensemble vide : $\mathcal{V}_L(\mathcal{I}_{\mathbf{M}}) = \emptyset$.

Réciproquement, supposons $\mathcal{V}_L(\mathcal{I}_{\mathbf{M}}) = \emptyset$, ce qui est équivalent à dire que les mineurs de taille maximale de \mathbf{M} , notés m_1, \dots, m_l n'admettent pas de zéros communs. Clairement, il existe un polynôme $\mathbf{z}^k := z_1^k \dots z_p^k$ ($k \in \mathbb{N}$) tel que :

$$\bar{m}_i := \mathbf{z}^k m_i \in \mathbb{C}[\mathbf{z}], \forall i \quad (\text{A.14})$$

De plus, il est possible de choisir k tel que les zéros communs des polynômes classiques $\bar{m}_1, \dots, \bar{m}_l$ soient exactement les points de leur variété suivante :

$$\mathcal{V}(\bar{m}_1, \dots, \bar{m}_l) = \{\mathbf{z} = (z_1, \dots, z_p) \in \mathbb{C}^p \mid \prod_{i=1}^p z_i = 0\} \quad (\text{A.15})$$

En utilisant la Nullstellensatz de Hilbert, il vient qu'il existe $r \in \mathbb{N}$ tel que le polynôme $z_1^r \dots z_p^r$ appartienne à l'idéal associé à $\bar{m}_1, \dots, \bar{m}_l$; et par conséquent il existe des polynômes x_1, \dots, x_l de $\mathbb{C}[\mathbf{z}]$ tels que :

$$z_1^r \dots z_p^r = \sum_{i=1}^l x_i \bar{m}_i. \quad (\text{A.16})$$

Il s'ensuit :

$$1 = \sum_{i=1}^l \bar{x}_i m_i. \quad (\text{A.17})$$

où l'on a défini les polynômes de Laurent $\bar{x}_i := \mathbf{z}^{-r+k} x_i$. Par suite, le polynôme unité appartient à $\mathcal{I}_{\mathbf{M}}$ et ainsi $\mathcal{I}_{\mathbf{M}} = \mathbb{C}[\mathbf{z}, \mathbf{z}^{-1}]$. ■

Ainsi, les propriétés 5 et 7 permettent de justifier la propriété 2 du chapitre 4.

²Remercions à ce titre le professeur Maria Elena Valcher de l'université de Padova pour les éclaircissements qu'elle nous a apportés.

1.3.3-c) Une condition suffisante d'inversibilité pour les systèmes générés aléatoirement

Le caractère générique de l'inversibilité d'une matrice polynomiale a été considéré dans [83]. En particulier, la propriété suivante qui nous est utile y est démontrée :

Propriété 8 *Soit $\mathbf{M}[\mathbf{z}] \in \mathbb{C}[\mathbf{z}]^{Q \times N}$ une matrice polynomiale dont les coefficients sont tirés selon une loi de probabilité à densité continue. Si $\binom{Q}{N} > p$ alors $\mathbf{M}[\mathbf{z}]$ est presque sûrement inversible à gauche (dans $\mathbb{C}[\mathbf{z}]$). Si $\binom{Q}{N} \leq p$, alors $\mathbf{M}[\mathbf{z}]$ est presque sûrement non inversible à gauche (dans $\mathbb{C}[\mathbf{z}]$).*

Table des figures

1.1	Modèle général de la séparation de sources.	11
1.2	Schéma du système considéré pour la séparation de sources.	14
1.3	Schéma du système considéré pour la séparation de sources.	16
2.1	Blanchiment et filtre para-unitaire.	22
2.2	Comparaison des contrastes dans le cas de deux sources i.i.d. PAM4 (moyenne de l'EQM en fonction du nombre d'échantillons). Le filtre de mélange est de longueur 5.	47
2.3	Comparaison des contrastes dans le cas de deux sources multiplicatives du contre-exemple en 2.3.2-b) et pour deux valeurs du paramètre ϵ (EQM moyenne en fonction du nombre d'échantillons). Le filtre de mélange est de longueur 5.	48
2.4	Comparaison des contrastes pour deux sources ARCH(9). Le filtre de mélange est de longueur 5.	49
2.5	Comparaison des contrastes pour deux sources non linéaire obtenues par sous-échantillonnage d'un processus linéaire. Le filtre de mélange est de longueur 3.	49
2.6	EQM moyenne en fonction de la longueur du filtre de mélange pour des signaux du contre-exemple en 2.3.2-b) avec $\epsilon = 0.5$. Le nombre de sources était de 2 et le nombre d'échantillons de 8192.	50
2.7	EQM moyenne en fonction du SNR pour des signaux du contre-exemple en 2.3.2-b) $\epsilon = 0.5$; Le nombre de sources était de 2 ou 3, le nombre d'échantillons de 8192 et la longueur du filtre de mélange de 5.	50
2.8	Représentation du contraste ζ_0 comme fonction des angles de rotation. Les sources sont PAM4 i.i.d. et ont 2048 échantillons.	51
2.9	Erreur d'estimation des contrastes. Les sources sont PAM4 i.i.d. avec 2048 échantillons.	52
3.1	EQM moyenne sur les sources reconstituées. Comparaison sur 100 réalisations de Monte-Carlo (3 sources, 6 capteurs, filtre de longueur 3, 4096 échantillons). Les réalisations ont été classées par ordre croissant de l'EQM pour la méthode hiérarchique.	72
3.2	Séparation de 4 sources de types ARCH(1) : valeur moyenne sur 100 réalisations des critères $\tau_i, i \in \{1, \dots, 4\}$ en fonction du nombre d'échantillons. Le filtre de mélange comportait 8 capteurs, était de longueur 2 et a été tiré aléatoirement selon une loi gaussienne normalisée à chaque réalisation. . . .	74
4.1	Système global à temps discret.	78

4.2	Sources originales (1 ^{ère} ligne) et reconstituées, sans post-optimisation locale (2 ^{ème} ligne) et avec (3 ^{ème} ligne) (Indices de modulations : 0,25 ; 0,75 et 0,4 ; 4 capteurs, filtre de mélange de longueur 4, 1000 échantillons.)	81
4.3	Deux exemples typiques des distributions des valeurs singulières de la matrice de contrainte (obtenus avec un filtre de mélange identique, 10000 échantillons et deux types de sources distincts).	82
4.4	Fonction de répartition du critère $\frac{\sum_{i=1}^3 \tau_i}{3}$ pour un filtre 4×3 fixé. (4000 échantillons)	83
4.5	Fonction de répartition du critère $\frac{\sum_{i=1}^4 \tau_i}{4}$ pour un filtre 5×4 fixé.	84
4.6	Fonction de répartition du critère $\frac{\sum_{i=1}^N \tau_i}{N}$ pour des filtres 4×3 et 5×4 générés aléatoirement. Le nombre d'échantillons a été fixé à 1000.	85
4.7	Exemple typique de résultat de séparation dans le cas de sources n'occupant pas toute la bande : observations au niveau spectral (16384 échantillons).	92
4.8	Valeur du critère τ_1 en fonction du nombre d'échantillons et du pourcentage occupé de la bande.	93
4.9	EQM après séparation de deux sources i.i.d. uniformes 2D.	93
4.10	Exemple de séparation d'une image naturelle et d'un bruit i.i.d. uniforme	95
4.11	Exemple de séparation de deux images naturelles	96
5.1	Modèle du système de séparation avec référence.	98
5.2	EQM moyenne sur les sources reconstituées en fonction du nombre d'échantillons. Les valeurs ont été obtenues avec des filtres de mélange de longueur 3, et à l'aide d'un moyennage sur 1000 réalisations.	113
5.3	EQM sur les sources reconstituées. Comparaison sur 1000 réalisations de Monte-Carlo (3 sources, 4 capteurs, filtres de mélange tirés aléatoirement, 5000 échantillons) du résultat avec le contraste $\mathcal{C}_z\{.\}$ et le contraste classique sans référence $ \kappa_4\{y(n), y(n)\} $. Les réalisations ont été classées par ordre croissant de l'EQM pour le contraste sans référence.	114
5.4	EQM sur les sources reconstituées (3 sources de type PAM4, 4 capteurs, filtres de mélange tirés aléatoirement, 5000 échantillons). L'optimisation du contraste sans référence n'a pas été menée jusqu'à la convergence et le résultat obtenu a servi de référence. Les réalisations ont été classées par ordre croissant de l'EQM pour le contraste sans référence.	115

Liste des tableaux

2.1	EQM moyenne et pour des intervalles de confiance à 90% en fonction du nombre d'échantillons.	47
2.2	Pourcentage de succès de la séparation pour deux sources ARCH(5) mélangées par un filtre de longueur 3 (La séparation a été considérée comme réussie lorsque l'EQM moyenne sur les deux sources était inférieure à 1e-2).	52
3.1	Comparaison des EQM moyennes sur 1000 réalisations données par différentes méthodes de séparation séquentielles. Cas de 3 sources i.i.d. à distribution uniforme, 6 capteurs, et d'un filtre de mélange de longueur 3.	71
4.1	Valeurs moyennes des critères $\tau_i, i \in \{1, \dots, 4\}$ pour un filtre 5×4 fixé, 4000 échantillons, 100 réalisations de Monte-Carlo.	84
4.2	Valeurs médiane des critères $\tau_i, i \in \{1, \dots, 4\}$ pour un filtre 5×4 fixé, 4000 échantillons, 100 réalisations de Monte-Carlo.	86
4.3	Valeur moyenne des critères $\tau_i, i \in \{1, \dots, 3\}$ pour un filtre de mélange 4×3 fixé et pour différentes valeurs du SNR. Le nombre de réalisations de Monte-Carlo est de 100.	86
4.4	Estimation du BER pour un filtre de mélange 4×3 fixé et pour différents SNR. Le nombre de réalisations de Monte-Carlo est de 1000.	86
4.5	Valeur du critère τ_1 en fonction du nombre d'échantillons et du pourcentage occupé de la bande.	92
5.1	Valeur moyenne du critère τ_1 sur 1000 réalisations. Le mélange était tiré aléatoirement, à coefficients complexes avec 3 sources et 4 capteurs.	115
5.2	Optimisation itérée d'un contraste avec référence tirée au hasard : résultat pour 3 sources i.i.d. PAM4 et 5000 échantillons.	116
5.3	Optimisation itérée d'un contraste avec référence tirée au hasard : résultat pour 3 sources i.i.d. PAM4 et 10000 échantillons.	116
5.4	Optimisation itérée d'un contraste avec référence tirée au hasard : résultat pour 3 sources i.i.d. PAM4 et 100000 échantillons.	116
5.5	EQM moyenne sur chacune des 3 sources extraites. résultats pour 5000 échantillons et, pour chaque extraction, 15 itérations de maximisation d'un contraste avec référence.	117
5.6	EQM moyenne sur chacune des 3 sources extraites. résultats pour 10000 échantillons et, pour chaque extraction, 10 itérations de maximisation d'un contraste avec référence.	117

- 5.7 EQM moyenne sur chacune des 3 sources extraites. résultats pour 100000 échantillons et, pour chaque extraction, 5 itérations de maximisation d'un contraste avec référence. 118

Bibliographie

- [1] <http://www.shef.ac.uk/~gaipp/ga-toolbox/>.
- [2] Karim Abed-Meraim, Philippe Loubaton, and Eric Moulines. A subspace algorithm for certain blind identification problems. *IEEE Trans. on Information Theory*, 43(2) :499–511, March 1997.
- [3] Abdellah Adib, Eric Moreau, and Driss Aboutajdine. Referenced contrasts a new approach to the problem of blind source separation. In *Seventh International Symposium on Signal Processing and its Applications (ISSPA'2003)*, Paris, July 2003.
- [4] Abdellah Adib, Eric Moreau, and Driss Aboutajdine. Source separation contrasts using a reference signal. *IEEE Signal Processing Letters*, 11(3) :312–315, March 2004.
- [5] Massoud Babaie-Zadeh, Christian Jutten, and Kambiz Nayebi. A minimization-projection (MP) approach for blind separating convolutive mixtures. In *Proc. IEEE Int. Conf. on Acoustics, Speech and Signal Processing (ICASSP)*, pages 533–536, Montreal, Canada, May 2004.
- [6] Anthony J. Bell and Terrence J. Sejnowski. An information-maximisation approach to blind separation and blind deconvolution. *Neural computation*, 7(6) :1129–1159, November 1995.
- [7] A. Belouchrani and Andrzej Cichocki. Robust whitening procedure in blind separation context. *Electronic Letters*, 36(24) :2050–2051, 2000.
- [8] Adel Belouchrani, Karim Abed-Meraim, Jean-François Cardoso, and Eric Moulines. A blind source separation technique using second-order statistics. *IEEE Trans. Signal Processing*, 45(2) :434–444, February 1997.
- [9] Pascal Bianchi. *Démodulation aveugle de modulations non linéaires à phase continue*. PhD thesis, Université de Marne-la-Vallée, 2003.
- [10] Pascal Bianchi, Philippe Loubaton, and François Sirven. On the blind equalization of continuous phase modulation using a constant modulus criterion. In *Proc. SPAWC*, Roma, Italy, 2003.
- [11] Tim Bollerslev, Ray Y. Chou, Narayanan Jayaraman, and Kenneth F. Kroner. Les modèles ARCH en finance : un point sur la théorie et les résultats empiriques. *Annales d'économie et de statistique*, (24), 1991.
- [12] Peter J. Brockwell and Richard A. Davis. *Time Series : Theory and Methods*. Springer, 2nd edition, 1991.
- [13] Véronique Capdevielle, Christine Servière, and Jean-Louis Lacoume. Blind separation of wide-band sources in the frequency domain. In *Proc. IEEE Int. Conf. on Acoustics, Speech and Signal Processing (ICASSP)*, pages 2080–2083, Detroit, 1995.

- [14] Jean-François Cardoso. Blind signal separation : statistical principles. *Proceedings of the IEEE. Special issue on blind identification and estimation*, 9(10) :2009–2026, October 1998.
- [15] Jean-François Cardoso and Antoine Souchoumiac. Blind beamforming for non gaussian signals. In *IEEE- Proceedings-F*, volume 140, pages 362–370, 1993.
- [16] Jean-François Cardoso. Independent component analysis of the cosmic microwave background. In *Proc. of ICA'03*, pages 1111–1116, Nara, Japan, 2003.
- [17] Marc Castella, Eric Moreau, and Jean-Christophe Pesquet. A quadratic MISO contrast function for blind equalization. In *Proc. IEEE Int. Conf. on Acoustics, Speech and Signal Processing (ICASSP)*, pages 681–684, Montréal, Canada, May 2004.
- [18] François Chaplais. Théorie algébrique des systèmes linéaires algébriques stationnaires multivariables. (cours de DEA de l'Université de Paris-Dauphine) http://cas.ensmp.fr/~chaplais/FTP/Cours_Systemes/.
- [19] Binning Chen and Athina P. Petropulu. Multiple-input-multiple-output blind system identification based on cross-polyspectra. In *Proc. IEEE Int. Conf. on Acoustics, Speech and Signal Processing (ICASSP)*, pages 584–587, 2000.
- [20] Binning Chen and Athina P. Petropulu. Frequency domain blind MIMO system identification based on second- and higher order statistics. *IEEE Trans. Signal Processing*, 49(8) :1677–1688, August 2001.
- [21] H. Cirpan and M.K. Tsatsanis. Blind receiver for non linearly modulated signals in multipath. In *Proc. SPAWC*, pages 149–152, 1997.
- [22] Pierre Comon. Analyse en composantes indépendantes et identification aveugle. *Traitement du Signal*, 7(5) :435–450, Décembre 1990.
- [23] Pierre Comon. Independent component analysis, a new concept. *Signal Processing*, 36(3) :287–314, April 1994.
- [24] Pierre Comon. Contrasts for multichannel blind deconvolution. *IEEE Signal Processing Letters*, 3(7) :209–211, July 1996.
- [25] Pierre Comon and Jérôme Lebrun. Contrast functions for deterministic blind source separation. In *Proc. SPAWC*, Roma, Italy, June 2003.
- [26] Pierre Comon and Eric Moreau. Blind MIMO equalization and joint-diagonalization criteria. In *Proc. IEEE Int. Conf. on Acoustics, Speech and Signal Processing (ICASSP)*, Salt Lake City, USA, May 2001.
- [27] Pierre Comon and Ludwig Rota. Blind separation of independent sources from convolutive mixtures. *IEICE Trans. on Fundamentals of Electronics, Communications, and Computer Sciences*, E86-A(3) :542–549, March 2003.
- [28] David Cox, John Little, and Donal O'Shea. *Ideals, Varieties and Algorithms : An Introduction to Computational Algebraic Geometry and Commutative Algebra*. Springer, 2nd edition, 1996.
- [29] Adriana Dapena and Luis Castedo. Blind source separation in the frequency domain : A novel solution to the amplitude and the permutation indeterminacy. In *Proc. IWANN*, pages 603–310, Granada, Spain, June 2001.
- [30] Adriana Dapena and Luis Castedo. A novel frequency domain approach for separating convolutive mixtures of temporally-white signals. *Digital Signal Processing*, 13(2) :301–316, April 2003.

- [31] Adriana Dapena, Christin Servière, and Luis Castedo. Inversion of the sliding fourier transform using only two frequency bins and its application to source separation. *Signal Processing*, 83(2) :453–457, February 2003.
- [32] Nathalie Delfosse and Philippe Loubaton. Adaptive blind separation of independent sources : a deflation approach. *Signal Processing*, 45 :59–83, 1995.
- [33] Konstantinos I. Diamantaras, Athina P. Petropulu, and Binning Chen. Blind two-input-two-output FIR channel identification based on frequency domain second-order statistics. *IEEE Trans. Signal Processing*, 48(2) :534–542, February 2000.
- [34] R. F. Engle. Autoregressive conditional heteroscedasticity with estimates of the variance of united kingdom inflation. *Econometrica*, 50(4) :987–1008, 1982.
- [35] Ettore Fornasini and Maria Elena Valcher. nD polynomial matrices with applications to multidimensional signal analysis. *Multidimensional Systems and Signal Processing*, 8 :387–408, 1997.
- [36] Jr. G. David Forney. Minimal bases of rational vector spaces, with application to multivariable linear systems. *SIAM J. Control*, 13(3) :493–520, May 1975.
- [37] Giorgios B. Giannakis and Robert W. Heath. Blind identification of multichannel FIR blurs and perfect image restoration. *IEEE Trans. on Image Processing*, 9(11) :1877–1896, November 2000.
- [38] Dominique N. Godard. Self-recovering equalization and carrier tracking in two-dimensional data communication systems. *IEEE Trans. on Communications*, 28(11) :1867–1875, November 1980.
- [39] Alexei Gorokhov and Philippe Loubaton. Subspace based techniques for blind separation of convolutive mixtures with temporally correlated sources. *IEEE Trans. Circuits and Systems I*, 44(9) :813–820, September 1997.
- [40] Dominique Guégan. *Séries chronologiques non linéaires à temps discret*. ECONOMICA, 1994.
- [41] Sébastien Houcke, Antoine Chevreuil, and Philippe Loubaton. Estimation of the symbol period : the frequency offset case. In *XI European Signal Processing Conference (EUSIPCO)*, volume 1, pages 635–668, Toulouse, France, September 2002.
- [42] Yingbo Hua, Senjian An, and Yong Xiang. Blind identification of FIR MIMO channels by decorrelating subchannels. *IEEE Trans. Signal Processing*, 51(5) :1143–1155, May 2003.
- [43] Yingbo Hua and Jitendra K. Tugnait. Blind identifiability of FIR-MIMO systems with colored input using second order statistics. *IEEE Signal Processing Letters*, 7(12) :348–350, December 2000.
- [44] Aapo Hyvärinen, Patrik O. Hoyer, and Jarmo Hurri. Extensions of ICA as models of natural images and visual processing. In *Proc. of ICA '03*, pages 963–974, Nara, Japan, 2003.
- [45] Aapo Hyvärinen and Erkki Oja. Independent component analysis : Algorithms and applications. *Neural Networks*, 13(4-5) :411–430, 2000.
- [46] Yujiro Inouye and Takehito Sato. Iterative algorithms based on multistage criteria for multichannel blind deconvolution. *IEEE Trans. Signal Processing*, 47(6) :1759–1764, June 1999.
- [47] Yujiro Inouye and Kazuaki Tanebe. Super-exponential algorithms for multichannel blind deconvolution. *IEEE Trans. Signal Processing*, 48(3) :881–888, March 2000.

- [48] Björn Jellonek, Dieter Boss, and Karl-Dirk Kammeyer. Generalized eigenvector algorithm for blind equalization. *Signal Processing*, 61(3) :237–264, September 1997.
- [49] Christian Jutten and J. Herault. Blind separation of sources, part i : An adaptative algorithm based on neuromimetic architecture. *Signal Processing*, 24(1) :1–10, 1991.
- [50] Thomas Kailath. *Linear Systems*. Prentice Hall Inc., 1980.
- [51] G.K. Kaleh and R. Vallet. Joint parameter estimation and symbol detection for linear or nonlinear unknown channels. *IEEE Trans. on Communications*, 42(7) :2406–2413, July 1994.
- [52] Mitsuru Kawamoto and Yujiro Inouye. Blind deconvolution of MIMO-FIR systems with colored inputs using second-order statistics. *IEICE Trans. on Fundamentals of Electronics, Communications, and Computer Sciences*, E86-A(3) :597–604, March 2003.
- [53] Mitsuru Kawamoto and Yujiro Inouye. Blind separation of multiple convolved colored signals using second-order statistics. In *Proc. of ICA'03*, Nara, Japan, April 2003.
- [54] Mitsuru Kawamoto and Yujiro Inouye. A deflation algorithm for the blind source-factor separation of MIMO-FIR channels driven by colored sources. *IEEE Signal Processing Letters*, 10(11) :343–346, November 2003.
- [55] D. Kundur and D. Hatzinakos. Blind image deconvolution. *IEEE Signal Processing Mag.*, 13 :43–64, May 1996.
- [56] Jean-Louis Lacoume, Pierre-Olivier Amblard, and Pierre Comon. *Statistiques d'Ordre Supérieur pour le Traitement du Signal*. Masson, 1997.
- [57] Jean-Louis Lacoume and P. Ruiz. Separation of independent sources from correlated inputs. *IEEE Trans. Signal Processing*, 40(12) :3074–3078, December 1992.
- [58] Lieven De Lathauwer, Bart De Moor, and Joos Vandewalle. Independent component analysis and (simultaneous) third-order tensor diagonalization. *IEEE Trans. Signal Processing*, 49(10) :2262–2271, October 2001.
- [59] P.A. Laurent. Exact and approximate construction of digital phase modulations by superposition of amplitude modulated pulse. *IEEE Trans. on Communications*, COM-34(2) :150–160, February 1986.
- [60] Ruey-Wen Liu and Yujiro Inouye. Blind equalization of MIMO-FIR channels driven by white but higher order colored source signals. *IEEE Trans. Information Theory*, 48(5) :1206–1214, May 2002.
- [61] Ruey-Wen Liu, Yujiro Inouye, and Hui Luo. A system-theoretic foundation for blind signal separation of MIMO-FIR convolutive mixtures - a review. In *Proc. of ICA'00*, Helsinki, Finland, June 2000.
- [62] Philippe Loubaton and Philippe Regalia. Blind deconvolution of multivariate signals : a deflation approach. In *Proceedings of ICC*, pages 1160–1164, Geneva, Switzerland, May 1993.
- [63] David G. Luenberger. *Linear and nonlinear programming*. Addison-Wesley, 2nd edition, 1984.
- [64] Ali Mansour and Mitsuru Kawamoto. ICA papers classified according to their applications and performances. *IEICE Trans. on Fundamentals of Electronics, Communications, and Computer Sciences*, E86-A(3) :620–633, March 2003.

- [65] Rubén Martín-Clemente, José I. Acha, and Carlos G. Puntonet. Eigendecomposition of self-tuned cumulant-matrices for blind source separation. *Signal Processing*, 84(7) :1201–1211, July 2004.
- [66] Amit Mathur, Arvind V. Keerthi, John J. Shynk, and Richard P. Gooch. Convergence properties of the multistage constant modulus array for correlated sources. *IEEE Trans. Signal Processing*, 45(1) :280–286, January 1997.
- [67] Eric Moreau. An any order generalization of JADE for complex source signals. In *Proc. IEEE Int. Conf. on Acoustics, Speech and Signal Processing (ICASSP)*, 2001.
- [68] Eric Moreau and Jean-Christophe Pesquet. Generalized contrasts for multichannel blind deconvolution of linear systems. *IEEE Signal Processing Letters*, 4(6) :182–183, June 1997.
- [69] Eric Moreau, Jean-Christophe Pesquet, and Nadège Thirion-Moreau. An equivalence between non symmetrical contrasts and cumulant matching for blind signal separation. In *First Int. Workshop on Independent Component Analysis and Signal Separation*, Aussois, France, January 1999.
- [70] Hoang Nguyen and B.C. Levy. Blind ML detection of CPM signals via the EMV algorithm. In *Proc. ICASSP*, pages 2457–2460, Orlando, USA, 2002.
- [71] Masashi Ohata and Kiyotoshi Matsuoka. Stability analysis of information-theoretic blind separation algorithms in the case where the sources are nonlinear processes. *IEEE Trans. Signal Processing*, 50(1) :69–77, January 2002.
- [72] Gilles Pagès and Claude Bouzitat. *En passant par hasard... Les probabilités de tous les jours*. Vuibert, 1999.
- [73] L. Parra and C. Spence. Convolutional blind separation of non-stationary sources. *IEEE Trans. Speech and Audio Processing*, 8(3) :320–327, May 2000.
- [74] Jean-Christophe Pesquet, Binning Chen, and Athina P. Petropulu. Frequency-domain contrast functions for separation or convolutional mixtures. In *Proc. IEEE Int. Conf. on Acoustics, Speech and Signal Processing (ICASSP)*, volume 5, pages 2765–2768, Salt Lake City, USA, May 2001.
- [75] Jean-Christophe Pesquet and Eric Moreau. Cumulant-based independence measures for linear mixtures. *IEEE Trans. Information Theory*, 47(5) :1947–1956, July 2001.
- [76] Dinh Tuan Pham. Mutual information approach to blind separation of stationary sources. *IEEE Trans. on Information Theory*, 48(7) :1935–1946, July 2002.
- [77] Dinh-Tuan Pham and Jean-François Cardoso. Blind separation of instantaneous mixtures of nonstationary sources. *IEEE Trans. Signal Processing*, 9(49) :1837–1848, September 2001.
- [78] Dinh-Tuan Pham and Philippe Garat. Blind separation of mixtures of independent sources through a quasi-maximum likelihood approach. *IEEE Trans. Signal Processing*, 45(7) :1712–1725, 1997.
- [79] Dinh-Tuan Pham, Christine Servière, and Hakim Boumaraf. Blind separation of convolutional audio mixtures using nonstationarity. In *Proc. of ICA '03*, pages 975–980, Nara, Japan, 2003.
- [80] Guillaume Planquette. *Etude de certaines propriétés algébriques et spectrales des codes correcteurs d'erreurs*. PhD thesis, Université de Rennes 1, Décembre 1996.
- [81] John G. Proakis. *Digital Communications*. Electrical Engineering Series. McGraw-Hill International Editions, 3rd edition, 1995.

- [82] R. Raheli, A. Polydoros, and Ching-Kae Tzou. Per-survivor processing : a general approach to MLSE in uncertain environments. *IEEE Trans. on Communications*, 43(234) :354–364, Feb/March/April 1995.
- [83] Ravikiran Rajagopal and Lee C. Potter. Multivariate MIMO FIR inverses. *IEEE Trans. on Image Processing*, 12(4) :458–465, April 2003.
- [84] Phillip A. Regalia. On the equivalence between the Godard and Shalvi-Weinstein schemes of blind equalization. *Signal Processing*, 73 :185–190, February 1999.
- [85] Ofir Shalvi and Ehud Weinstein. New criteria for blind deconvolution of nonminimum phase systems (channels). *IEEE Trans. on Information Theory*, 36(2) :312–321, March 1990.
- [86] Albert N. Shiryaev. *Essentials of stochastic finance : Facts, Models, Theory*, volume 3 of *Advanced Series on Statistical science & Applied Probability*. World Scientific, 1999.
- [87] Carine Simon. *Séparation aveugle des sources en mélange convolutif*. PhD thesis, Université de Marne-la-Vallée, November 1999.
- [88] Carine Simon, Philippe Loubaton, and Christian Jutten. Separation of a class of convolutive mixtures : a contrast function approach. *Signal Processing*, (81) :883–887, 2001.
- [89] Anisse Taleb and Christian Jutten. Source separation in post-nonlinear mixtures. *IEEE Trans. Signal Processing*, 47(10) :2807–2820, October 1999.
- [90] Nadège Thirion and Eric Moreau. New criteria for blind signal separation. In *IEEE Workshop on Statistical Signal and Array Processing*, pages 344–348, Pocono Manor, Pennsylvania, USA, August 2000.
- [91] Nadège Thirion-Moreau and Eric Moreau. Generalized criteria for blind multivariate signal equalization. *IEEE Signal Processing Letters*, 9(2) :72–74, 2002.
- [92] Azzédine Touzni, Inbar Fijalkow, Michael G. Larimore, and John R. Treichler. A globally convergent approach for blind MIMO adaptive deconvolution. *IEEE Trans. Signal Processing*, 49(6) :1166–1178, June 2001.
- [93] Jitendra K. Tugnait. Identification and deconvolution of multichannel linear non-gaussian processes using higher order statistics and inverse filter criteria. *IEEE Trans. Signal Processing*, 45(3) :658–672, March 1997.
- [94] Jitendra K. Tugnait. Adaptive blind separation of convolutive mixtures of independent linear signals. *Signal Processing*, 73 :139–152, February 1999.
- [95] Filip Šroubek and Jan Flusser. Multichannel blind iterative image restoration. *IEEE Trans. on Image Processing*, 12(9) :1094–1106, September 2003.
- [96] P. P. Vaidyanathan. *Multirate systems and filter banks*. Prentice Hall, Englewood Cliffs, 1993.
- [97] Alle-Jan van der Veen, Shilpa Talwar, and Arogyaswami Paulraj. Blind estimation of multiple digital signals transmitted over fir channels. *IEEE Signal Processing Letters*, 2(5) :99–102, May 1995.
- [98] Hsiao-Chun Wu and Jose C. Principe. Simultaneous diagonalization in the frequency domain (SDIF) for source separation. In *Proc. of ICA '99*, pages 245–250, Aussois, France, January 1999.
- [99] Daniel Yellin and Ehud Weinstein. Criteria for multichannel signal separation. *IEEE Trans. Signal Processing*, 42(8) :2158–2168, August 1994.