



Séparation de sources convolutives

Moussa Akil

► To cite this version:

Moussa Akil. Séparation de sources convolutives. Traitement du signal et de l'image [eess.SP]. Université Joseph-Fourier - Grenoble I, 2008. Français. NNT: . tel-00414263

HAL Id: tel-00414263

<https://theses.hal.science/tel-00414263>

Submitted on 8 Sep 2009

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.

UNIVERSITÉ JOSEPH FOURIER - GRENOBLE 1

N° attribué par la bibliothèque

THESE

pour obtenir le grade de

DOCTEUR DE L'UNIVERSITÉ JOSEPH FOURIER

Spécialité : «Signal, Image, Parole et Télécoms»

préparée au Laboratoire Grenoble Image Parole Signal et Automatique : GIPSA-lab

dans le cadre de l'École Doctorale «Électronique, Électrotechnique,
Automatique, Télécommunications et Signal»

présentée et soutenue publiquement

par

Moussa AKIL

le 22 Avril 2008

Titre :

SÉPARATION DE SOURCES CONVOLUTES

Directrice de thèse : Christine SERVIÈRE

JURY

Mr	J.-L. LACOUME,	Président,	Professeur, INPG de Grenoble
Mr	R.-B. RANDALL,	Rapporteur,	Professeur, UNSW Sydney, Australie
Mr	F. GUILLET,	Rapporteur,	Professeur, Université de Saint-Etienne
Mr	J. ANTONI,	Examineur,	MdC HDR, UTC de Compiègne
Mme	C. SERVIÈRE,	Directrice de thèse,	CR HDR, CNRS de Grenoble

Pour remonter à la source, il faut nager à contre-courant

Stanislaw Jerzy Lec

Remerciements

Je tiens à adresser mes remerciements à M. Jean-Marc CHASSERY, Directeur du laboratoire GIPSA (Grenoble, Image, Parole, Signal, Automatique), pour m'avoir accueilli au sein de son laboratoire.

Cette thèse n'aurait vu le jour sans la confiance, la patience et la générosité de ma directrice de recherche, Madame Christine SERVIÈRE, que je veux vivement remercier. Je voudrais aussi la remercier pour la confiance, le temps et la patience qu'elle m'a accordé tout au long de ces années. De plus, les conseils qu'elle m'a divulgués tout au long de la rédaction, ont toujours été clairs et succincts, me facilitant grandement la tâche et me permettant d'aboutir à la production de cette thèse.

Je tiens à exprimer ma gratitude à M. Jean-Louis LACOUME pour avoir accepté de présider le Jury de ma thèse et pour tous les conseils qu'il m'a donnés afin d'améliorer ce mémoire.

C'est un grand honneur que m'ont fait Messieurs Bob RANDALL, professeur à l'Université New South Wales (UNSW), Sydney, Australie et François GUILLET, professeur à l'Université de Saint Etienne, en s'intéressant à mes travaux de recherche et en acceptant de les juger. Je les remercie pour leur travail de rapporteur, pour le temps qu'ils y ont consacré, et pour les conseils qu'ils ont su me donner afin d'améliorer le document final.

Merci à Monsieur Jérôme Antoni, Maître de conférences à l'Université de Technologie de Compiègne, qui m'a fait l'honneur de participer au Jury de soutenance en tant qu'examinateur, je lui en remercie profondément.

Un profond merci à Benoit et Grégoire, avec qui j'ai eu le plaisir de partager de très bons moments de détente tout au long de ces années. Je remercie aussi toutes les membres du Laboratoire pour leur sympathie, leur aide et leur soutien, alors à vous tous et toutes MERCI !

Ces remerciements s'adressent aussi à l'ensemble des personnes que j'ai croisées lors de

ces très bonnes années passées au GIPSA-lab.

Merci aussi à tous mes amis de Grenoble et d'ailleurs.

Je remercie profondément toute ma famille. Ces études auxquelles je mets un terme, c'est avant tout à eux que je les dois, à eux qui, depuis toujours et en particulier tout au long de ces dernières années ont su partager avec patience et amour mes instants de joie et me soutenir dans mes moments de découragement.

Existe-t-il un mot suffisamment intérieur et avec assez de souffle pour dire à ma douce Myriam, tout ce que je ressens au moment de mettre ce point final ?
Ce que je te dois est immense ... Je t'aime pour toujours.

Table des matières

Introduction	13
1 État de l'art	19
1.1 Généralités sur la séparation de sources	21
1.1.1 Introduction	21
1.1.2 Modèles et problèmes de la séparation de sources	22
1.2 Méthodes de séparation de sources	29
1.2.1 Indépendance statistique	30
1.2.2 Méthodes à l'ordre 2 : Décorrélation	35
1.2.3 Méthodes aux ordres supérieurs : Fonctions de contraste	38
1.3 Quelques méthodes de séparation en mélange convolutif	41
1.3.1 Approche temporelle	42
1.3.2 Approche fréquentielle	48
1.4 Conclusion	49
2 Séparation de sources : Information mutuelle et filtre de Wiener	51
2.1 Critère d'indépendance	52
2.1.1 Introduction	52
2.1.2 Mesure de l'indépendance statistique	53
2.1.3 Mesure de la divergence entre les densités de probabilité	54
2.1.4 Mesure de l'information mutuelle	54
2.1.5 Mesure de l'indépendance dans le cas de mélange convolutif	55
2.1.6 Minimisation de l'information mutuelle	57
2.2 Filtre de Wiener	58
2.2.1 Définition	59
2.2.2 Modélisation du filtre de Wiener	59
2.3 Méthode de séparation proposée	60

2.3.1	Modélisation du problème	61
2.3.2	Calcul des contributions des sources sur un capteur	63
2.4	Algorithme de séparation	71
2.4.1	Calcul des gradients	71
2.4.2	Simulations et résultats obtenus	73
2.5	Conclusion	81
3	Statistiques d'ordre supérieur : Fonctions de contraste	83
3.1	Introduction sur les statistiques d'ordre 4	85
3.1.1	Définition	86
3.1.2	Algorithme de diagonalisation conjointe JADE	87
3.1.3	Implantation et simulations	91
3.2	Fonctions de contraste	104
3.2.1	Définitions et rappels	104
3.2.2	Technique d'annulation des cumulants croisés	107
3.2.3	Fonction de contraste utilisant les contributions des sources sur les capteurs	108
3.2.4	Implantation et simulations	111
3.2.5	Comparaison avec les résultats de la méthode développée dans le chapitre 2	118
3.3	Conclusion	119
	Conclusions et perspectives	121
	Annexes	127
A	Estimation des fonctions scores	129
A.1	Définitions	129
A.1.1	Fonction score	129
A.1.2	Fonction score marginale	129
A.1.3	Fonction score conjointe	130
A.1.4	Fonction score différentielle	130
A.2	Estimation de la densité de probabilité	130
B	Algorithme de diagonalisation conjointe JADE	133
	Bibliographie	134

Préambule

Le travail présenté dans ce manuscript a été réalisé dans le laboratoire **G**renoble **I**mage **P**arole **S**ignal **A**utomatique (GIPSA-lab) sous la direction de Madame Christine SERVIÈRE.

Le sujet abordé tout au long de ce mémoire consiste à l'étude et au développement de méthodes en séparation de sources dans le cadre des mélanges convolutifs.

Introduction

La science est une chose merveilleuse... tant qu'il ne faut pas en vivre!

Albert Einstein

LA SÉPARATION DE SOURCES est un domaine de recherche important en traitement du signal et analyse de données. L'Analyse en Composantes Indépendantes (ACI) est l'une des méthodes les plus développées qui a été introduite dans les années 80 dans les travaux de recherche de C. Jutten et J. Hérault. L'ACI peut être considérée comme une extension de l'Analyse en Composantes principales (ACP). Ainsi, elle consiste à remplacer la recherche des composantes décorrélées, par une recherche de composantes indépendantes.

Le terme *séparation de sources* consiste à extraire des signaux provenant de systèmes physiques différents à partir seulement des mélanges observés de ces signaux. En général, les *sources* sont séparées sous l'hypothèse de signaux statistiquement indépendants et de mélanges linéaires. Un problème de séparation de sources peut être résolu par l'ACI. Il peut aussi faire appel à d'autres approches comme l'analyse temps-fréquence ou le filtrage.

Cette problématique de séparation de sources trouve des retombées directes dans différents domaines d'applications comme la séparation des signaux de parole, les systèmes de communication, les systèmes électromécaniques,...,Etc.

Contexte et problématique

Dans cette thèse, nous nous intéressons au problème de **séparation de sources en mélange convolutif**. C'est un problème qui consiste à retrouver les sources $(y_1, \dots, y_M)$ à partir des observations $(x_1, \dots, x_N)$ d'un mélange (mélange inconnu A) de sources $(s_1, \dots, s_M)$. Ce problème est devenu un thème de recherche important en raison de son intérêt théorique et des nombreuses applications possibles.

$$x(t) = \mathcal{A} * s(t) \tag{1}$$

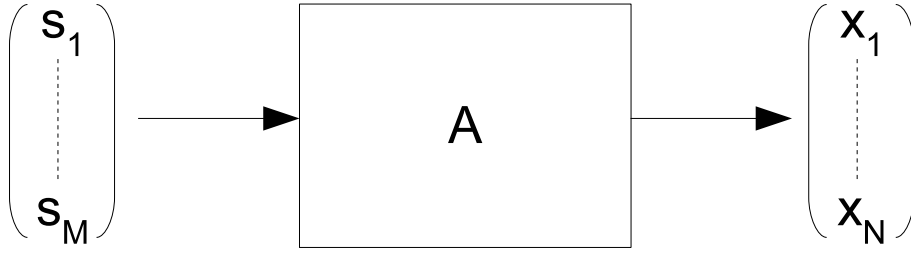


FIG. 1 – Modèle général de séparation de sources

où $*$ représente le produit de convolution.

Dans ce mémoire, nous abordons principalement le problème de séparation de sources dans le cadre des mélanges convolutifs. Nous proposons une méthodologie de séparation basée sur le calcul des **contributions des sources sur les capteurs**. Ces contributions seront obtenues en minimisant l'erreur quadratique entre les sources estimées et les capteurs à l'aide d'un filtrage de Wiener. La méthodologie proposée simplifie la mesure de certains critères d'indépendance développés dans des travaux antérieurs et est primordiale dans un contexte de diagnostic où on cherche à surveiller toutes les contributions des sources sur les capteurs.

L'objectif principal à atteindre est d'identifier le système inconnu A . En d'autre terme, le problème revient à trouver un système noté A^{-1} , l'inverse de A dans le but de retrouver les sources.

Dans le cadre de cette thèse, seul le modèle de mélange convolutif est traité. Ainsi, les M sources $(s_1, \dots, s_M)$ sont mélangées par une matrice de filtre A (système inconnu). En sortie du filtre, nous récupérons N observations au niveau des capteurs. En effet, il s'agit de trouver un filtre matriciel B qui inverse le filtre A afin de reconstituer les sources $s(t)$. L'idée principale dans ce manuscrit est de calculer la contribution de chaque source s_i sur chaque capteur x_i . **La contribution d'une source i** représente le signal que l'on observe au niveau des capteurs si aucune source j n'interfère avec la source i ($i \neq j$).

Plan du mémoire

Cette thèse fait l'étude de deux approches de séparation de sources. La première technique utilise un critère d'indépendance basé sur la minimisation de l'information mutuelle. Et la deuxième exploite les statistiques d'ordre supérieur en utilisant le concept des fonctions de contraste. Tout au long de cette thèse, nous allons utiliser le filtrage de Wiener. Cette procédure va nous permettre de calculer les contributions des sources sur les capteurs. Le calcul de ces contributions est basé sur la projection des sources estimées sur chaque capteur afin de minimiser une erreur quadratique entre les sorties estimées et les observations. Dans le cadre des mélanges convolutifs, la mesure d'indépendance doit être effectuée pour tous les retards k . Ainsi, le calcul des contributions des sources permet de simplifier le critère d'indépendance utilisé et améliorer la qualité de la séparation dans les méthodes développées.

Le présent mémoire est composé de trois chapitres :

- **Chapitre 1 :**

Dans le premier chapitre, nous présentons un état de l'art sur la séparation de sources. Nous allons expliquer le concept de la séparation de sources, les différents types de mélanges, les critères d'indépendance développés,...,Etc. Ensuite, nous présentons une synthèse générale sur les différentes techniques et approches développées dans le cadre des mélanges convolutifs.

- **Chapitre 2 :**

Dans le deuxième chapitre, nous présentons le modèle du filtre de Wiener. Les sorties de ce filtre constituent les contributions des sources sur les capteurs. Ensuite, nous développons un algorithme de séparation basé sur l'utilisation de l'information mutuelle comme critère d'indépendance. Le fait d'insérer les contributions des sources sur les capteurs dans la procédure de séparation va nous amener

à une simplification du critère de séparation et à des résultats numériques avantageux.

- **Chapitre 3 :**

Dans le chapitre 3, nous exploitons les statistiques d'ordre 4. Dans un premier temps, nous présentons une approche basée sur la diagonalisation conjointe des matrices de cumulants (algorithme JADE). La particularité de cette approche est que les matrices de cumulants à diagonaliser sont calculées à partir des sorties estimées et des contributions des sources sur les capteurs calculées par un filtrage de Wiener. Cette approche nécessite une phase de pré-blanchiment des observations au préalable.

Dans la deuxième partie de ce chapitre, nous généralisons une fonction de contraste utilisant comme paramètres les contributions des sources sur les capteurs. L'approche proposée utilise la technique d'annulation des cumulants croisés. Ces cumulants sont calculés à partir des contributions des sources sur les capteurs. Les résultats prouvent l'efficacité du contraste proposé dans le cadre des mélanges convolutifs.

- Ce mémoire se termine par une conclusion générale sur l'ensemble des chapitres ainsi que différentes perspectives qui s'ouvrent à nos recherches.

Chapitre 1

État de l'art

Sommaire

1.1	Généralités sur la séparation de sources	21
1.1.1	Introduction	21
1.1.2	Modèles et problèmes de la séparation de sources	22
1.1.2.1	Modèles	22
1.1.2.2	Indéterminations	24
1.1.2.3	Hypothèses	25
1.1.2.4	Mélange instantané	25
1.1.2.5	Mélange convolutif	26
1.2	Méthodes de séparation de sources	29
1.2.1	Indépendance statistique	30
1.2.1.1	Information mutuelle	31
1.2.1.2	Mesure du maximum de vraisemblance	32
1.2.1.3	Moments et cumulants	33
1.2.2	Méthodes à l'ordre 2 : Décorrélation	35
1.2.2.1	Blanchiment	35
1.2.2.2	Estimation de la matrice de séparation	37
1.2.3	Méthodes aux ordres supérieurs : Fonctions de contraste	38
1.2.3.1	Fonctions de contraste	38
1.2.3.2	Exemples de contrastes	41
1.3	Quelques méthodes de séparation en mélange convolutif . . .	41
1.3.1	Approche temporelle	42
1.3.1.1	Méthode adaptative (<i>travaux de C. Jutten et H.L. Nguyen Thi</i>)	42

1.3.1.2	Annulation des bispectres croisés (<i>travaux de D. Yellin et E. Weinstein</i>)	43
1.3.1.3	Les fonctions de contraste	43
1.3.1.4	Méthode séquentielle (<i>Approche par déflation</i>)	47
1.3.2	Approche fréquentielle	48
1.4	Conclusion	49

C'est par la séparation qu'on évalue la force des liens

Gérard Gévry

LBUT DE CE CHAPITRE est de présenter un état de l'art sur le problème de la séparation de sources en traitement du signal. Nous commençons par une introduction générale sur la séparation de sources et une présentation de la problématique liée à ce domaine en traitement du signal.

Ensuite, nous distinguons les différents types de mélange utilisés en séparation de sources. Après un résumé sur chaque type de mélange, nous présentons les caractéristiques appropriées, et, nous nous intéressons aux différentes méthodes développées pour chaque type de mélange.

Après avoir présenté les bases de la séparation de sources et les modèles utilisés, nous posons la problématique de ce travail qui consiste à élaborer des méthodes pour la séparation de sources dans le cas de mélange convolutif. En effet, nous montrons certains points bloquants et des hypothèses qui doivent être prise en compte pour augmenter l'efficacité des méthodes de séparation vis à vis d'une approche totalement aveugle, soit sur la propagation des sources, soit sur leurs propriétés statistiques.

1.1 Généralités sur la séparation de sources

1.1.1 Introduction

Depuis une vingtaine d'années, la séparation de sources est devenue un des problèmes importants en traitement du signal [COM94, AC97]. Le problème de la séparation de sources a tout d'abord été développé par Héault et Jutten. Ils ont décrit le problème biologique qui a initié la séparation aveugle de sources [HJA85]. Ensuite, Comon, en 1994, a formalisé le lien entre l'analyse en composantes indépendantes (ICA) et la séparation aveugle de sources (BSS). Depuis les années 90, diverses méthodes ont été développées pour résoudre le problème de séparation aveugle de source dans le cadre de mélange linéaire

et non linéaire et qui ont mené à différents algorithmes. Diverses applications se sont développées grâce à la mise en oeuvre de techniques de séparation, témoignant de l'intérêt scientifique et industriel de ce thème de recherche. Parmi ces applications, on peut citer le domaine des communications [AMR95, CCM93, SGO97], la surveillance des systèmes [DA96, DPV97], et les applications biomédicales [LMV95, KHV⁺97].

La technique de séparation de sources consiste à retrouver plusieurs composantes linéaires (appelées sources) à partir de combinaisons linéaires de ces sources reçues sur plusieurs capteurs (appelées observations). On retrouve cette problématique dans des domaines variés tels que les télécommunications, le diagnostic des machines tournantes, le domaine médical, ou la finance...etc.

1.1.2 Modèles et problèmes de la séparation de sources

Le problème de la séparation de sources consiste à développer des méthodes capables d'extraire les M sources inconnues observées à travers N mélanges inconnus. Les N mélanges sont obtenus à partir d'un réseau de N capteurs. Dans la suite de ce manuscrit, on supposera autant de sources que de capteurs. En effet, si le nombre de sources M est inférieur au nombre de capteurs N , il est possible de se ramener au cas précédent par projection sur le sous-espace signal.

1.1.2.1 Modèles

Dans un contexte général, le problème de la séparation aveugle de sources se formule de la façon suivante : On considère M sources aléatoires notées $s_i(t)$, tels que $1 \leq i \leq M$ et $S(t) = [s_1(t), \dots, s_M(t)]^T$ est appelé le vecteur des sources. Soit N le nombre de capteurs utilisés, on considère $x_j(t)$ représentant les N observations tels que $1 \leq j \leq N$ et $X(t) = [x_1(t), \dots, x_N(t)]^T$ est appelé le vecteur d'observations. Dans le cas général, les signaux $x_j(t)$ observés au niveau des capteurs sont une fonction non-linéaire de toutes les sources $s_i(t)$, on note cette fonction A .

$$X(t) = A[S(t)] \quad (1.1)$$

La séparation de sources consiste à estimer un système ou une fonction, noté B et les sources estimées notées $Y(t) = [y_1(t), \dots, y_M(t)]^T$

$$Y(t) = B[X(t)] \quad (1.2)$$

Donc, le but est de déterminer, à partir des observations $X(t)$, la fonction B qui permet de calculer les sources estimées $Y(t)$ qui correspondent aux sources réelles $S(t)$. En général, le problème de la séparation de sources peut être divisé en deux parties (cf. Fig. 1.1) :

- **Modélisation du problème** : le mélange qui définit le lien entre les sources réelles et les observations.
- **Traitement du problème** : la séparation qui génère les sources estimées à partir des observations afin de résoudre le problème.

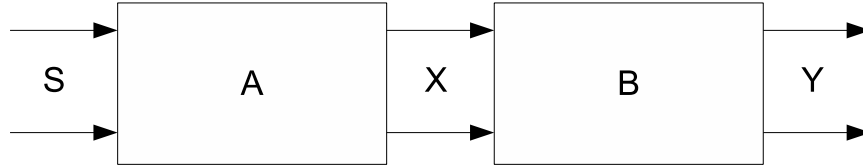


FIG. 1.1 – Structure générale du problème de séparation de sources

La solution idéale consiste à prendre $B = A^{-1}$, mais en séparation de sources, seules les observations sont supposées connues, ce qui contribue à certaines indéterminations qui seront citées dans le paragraphe suivant. Pour séparer, il est nécessaire d'avoir des hypothèses supplémentaires sur le modèle utilisé, particulièrement sur la nature du mélange. Le modèle doit correspondre au phénomène physique liant les sources aux observations. Donc, le choix de modèle est nécessaire parce qu'il permet de choisir la structure de séparation adaptée au problème traité. Les modèles de mélange les plus utilisés sont les suivants :

- **Mélange instantané** : l'observation à un instant donné t est une combinaison linéaire des sources au même instant t . Ainsi, A se réduit à une matrice de mélange

scalaire.

- **Mélange convolutif** : l'observation à l'instant donné t est une combinaison linéaire des valeurs des sources aux instants précédents. La fonction de transfert entre les sources et les observations peut être modélisée par une matrice de filtres linéaires.
- **Mélange post non-linéaire** : les observations sont considérées comme des fonctions non linéaires instantanées des sorties d'un mélange instantané, ou parfois convolutif.

1.1.2.2 Indéterminations

La seule information connue sur les sources est leur indépendance statistique. Comme les sources et le mélange sont inconnus, leur estimation ne peut pas être unique. Donc, en considérant une matrice diagonale inversible D inversible et une matrice de permutation P , on peut écrire :

$$X(t) = A[S(t)] = AD^{-1}D[S(t)] = A_1[S_1(t)] \quad (1.3)$$

en supposant que $A_1 = AD^{-1}$ et $S_1(t) = DS(t)$

On peut aussi écrire :

$$X(t) = A[S(t)] = AP^{-1}P[S(t)] = A_2[S_2(t)] \quad (1.4)$$

en supposant que $A_2 = AP^{-1}$ et $S_2(t) = PS(t)$

En plus, si les composantes du vecteur $S(t)$ sont indépendantes, il en est de même pour $S_1(t)$ et $S_2(t)$. Donc, le problème de séparation de sources possède deux indéterminations :

- **Indétermination d'échelle** : les sources sont estimées à un facteur près, la puissance des sources ne peut pas être retrouvée.
- **Indétermination sur l'ordre des sources** : les sources ne peuvent être retrouvées qu'à une permutation près.

Tenant compte de ces indéterminations, le problème de la séparation de sources est résolu quand la fonction de transfert $C = BA$ obtenue est de la forme suivante :

$$C = PD \quad (1.5)$$

où P est une matrice de permutation et D une matrice diagonale.

1.1.2.3 Hypothèses

Dans cette section, on présente quelques hypothèses sur le mélange et sur les sources pour résoudre le problème de séparation de sources.

- **H1** : Dans la plupart des techniques de séparation de sources, les sources $s_i(t)$ sont supposées statistiquement indépendantes.
- **H2** : Cette hypothèse concerne le type de mélange utilisé. La matrice de séparation B est supposée inversible. La plupart des méthodes traite le cas des mélanges linéaires.
- **H3** : Toutes les sources sont non gaussiennes sauf au plus une.
- **H4** : Les sources $s_i(t)$ sont des sources à moyennes nulles.

1.1.2.4 Mélange instantané

En général, le cas le plus simple de la séparation de sources est d'envisager le mélange linéaire instantané. Chaque observation $x_j(t)$ à l'instant t est une combinaison linéaire des sources $s_i(t)$ au même instant t . La matrice du mélange A ne dépend de l'observation j et de la source i et pas des instants t . On peut écrire :

$$x_j(t) = \sum_{i=1}^N a_{ji}(t) s_i(t) \quad (1.6)$$

où $a_{ij} \in \mathbb{R}$ sont les coefficients du mélange, pour $i \in \{1, \dots, M\}$ et $j \in \{1, \dots, N\}$.

$$x(t) = As(t) \quad (1.7)$$

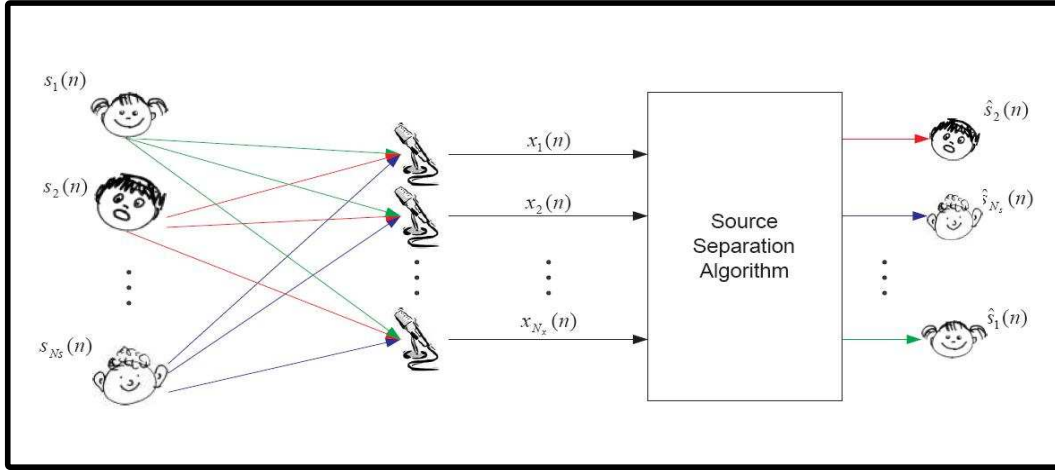


FIG. 1.2 – Exemple d'un mélange instantané de sources

Si la matrice A n'est pas singulière, il existe une matrice de scalaires B qui permet d'estimer les sources $y(t)$ à partir des observations $x(t)$

$$y(t) = Bx(t) \quad (1.8)$$

Donc, la solution idéale est de prendre une matrice $B = A^{-1}$. Mais, en séparation de sources, seules les observations sont connues.

1.1.2.5 Mélange convolutif

Dans la plupart des applications, le mélange de sources est un mélange convolutif. Dans ce cas, le problème consiste à trouver un filtre matriciel, opérant sur la suite des observations, tel que les différentes sources estimées en sortie du filtre séparateur soient aussi indépendantes que possible. Les premiers résultats ont été présentés par Jutten [JTD⁺91] dans le cas de deux sources et deux capteurs modélisés par deux filtres à réponse impulsionnelle finie (RIF). Les coefficients du filtre ont été estimés par une généralisation du critère [JH88b], qui était utilisé dans le cas du mélange instantané. On peut modéliser un mélange convolutif de la façon suivante :

$$x_j(t) = \sum_{i=1}^N a_{ji}(t) * s_i(t) = \sum_{i=1}^N \sum_{k=0}^M a_{ji}(t-k) s_i(t) \quad (1.9)$$

avec :

- $a_{ji}(t)$ représente la réponse impulsionnelle de la source i sur le capteur j ,
- $*$ est l'opérateur de convolution,
- M est l'ordre du filtre de mélange dont les coefficients sont a_{ji} .

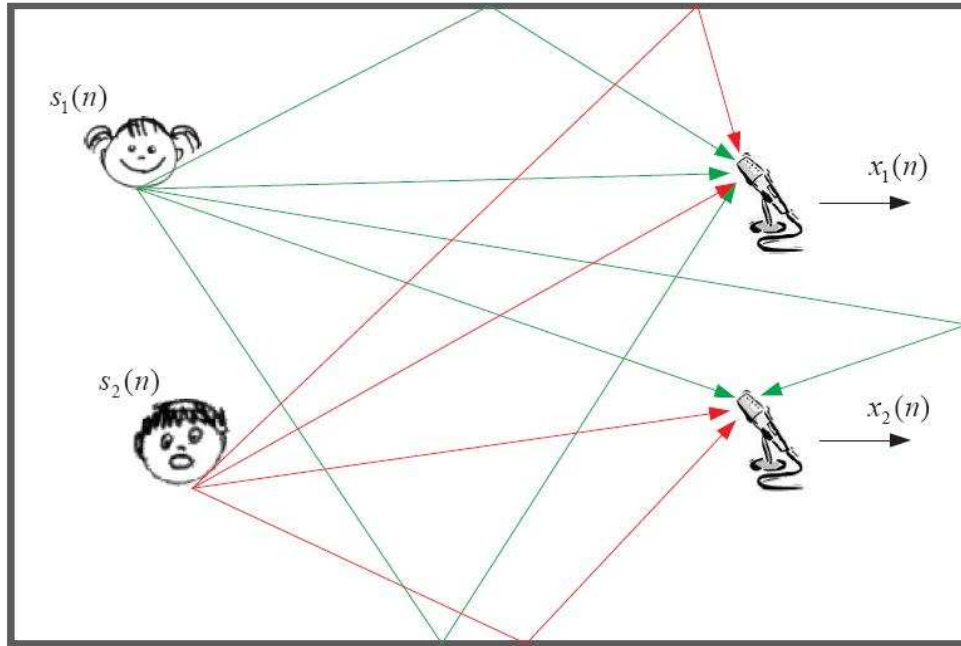


FIG. 1.3 – Exemple d'un mélange convolutif de deux sources et deux capteurs

On peut aussi écrire l'équation 1.9 sous forme matricielle de la façon suivante :

$$\underline{X} = \underline{A} \cdot \underline{S} \quad (1.10)$$

où

- $\underline{A}$ représente la matrice des filtres dont les coefficients sont a_{ji} ,

$$\underline{A} = \begin{pmatrix} \underline{a}_{11}^T & \cdots & \underline{a}_{1N}^T \\ \vdots & \ddots & \vdots \\ \underline{a}_{N1}^T & \cdots & \underline{a}_{NN}^T \end{pmatrix} \quad (1.11)$$

- $\underline{S}$ représente le vecteur des sources

$$\underline{S} = \begin{pmatrix} s_1 \\ \vdots \\ s_N \end{pmatrix} \quad (1.12)$$

- $\underline{X}$ représente le vecteur des observations

$$\underline{X} = \begin{pmatrix} x_1 \\ \vdots \\ x_N \end{pmatrix} \quad (1.13)$$

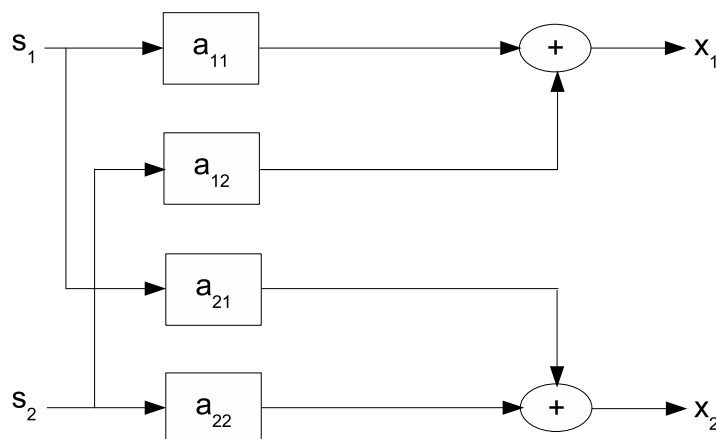


FIG. 1.4 – Principe du mélange avec deux sources et deux capteurs

Comme dans le cas des mélanges instantanés, on peut considérer des signaux stationnaires, i.i.d ou non i.i.d. En général, les approches qui ont été développées dans le cas de mélanges convolutifs peuvent être regroupées en deux catégories différentes :

- Approches globales : Ce sont des méthodes inspirées des algorithmes dans le cas instantané, souvent basées sur le blanchiment des observations et ensuite, le calcul des coefficients du filtre de séparation aux indéterminations près. En général, ces méthodes utilisent les statistiques d'ordre deux ou les statistiques d'ordre supérieur.
- Approches par déflation : Ces méthodes sont basées sur l'extraction et la soustraction des sources du mélange une par une.

On peut bien évidemment imaginer d'autres modèles de mélanges, mais les deux mélanges présentés permettent de couvrir et de refléter une large généralité des différentes situations réelles. Dans ce travail, nous nous restreindrons essentiellement aux mélanges de type convolutif.

On distingue deux approches différentes pour la séparation de sources en mélange convolutifs : les approches temporelles et les approches fréquentielles. En général, l'approche temporelle est basée sur la minimisation d'une fonction de coût qui va permettre l'estimation des coefficients des filtres de séparation. Quant à l'approche fréquentielle, la solution revient à décomposer le mélange convolutif étudié en plusieurs mélanges instantanés.

1.2 Méthodes de séparation de sources

Dans cette section, nous allons présenter les différentes techniques utilisées pour la séparation de mélanges linéaires selon l'ordre des statistiques utilisées. Les premières méthodes utilisées ont été basées sur les statistiques d'ordre supérieur (ordre 4). Nous rappelons que la seule hypothèse sur les sources est leur indépendance statistique. Les techniques que nous allons présenter exploitent cette hypothèse. Nous présenterons aussi quelques approches dites **fonctions de contraste**.

L'utilisation des statistiques d'ordre dépend du type des signaux utilisés. Par exemple, dans le cas des signaux colorés, éventuellement des signaux gaussiens, nous pouvons nous limiter aux statistiques d'ordre 2. Alors que dans le cas des signaux blancs et non gaussiens, il sera nécessaire d'utiliser des statistiques d'ordre supérieur à 2 et des hypothèses à priori sur les critères d'indépendance utilisés.

Dans le contexte de l'analyse en composantes indépendantes en séparation de sources, il est nécessaire de pouvoir évaluer la dépendance statistique des variables aléatoires avant de définir les critères de séparation. Nous donnons d'abord quelques rappels sur la définition de l'indépendance statistique des variables aléatoires avant de présenter certains critères d'indépendance qui seront décrits dans la suite de ce manuscrit.

En général, l'indépendance statistique des composantes du vecteur de sorties $y(t)$ n'im-

plique pas nécessairement la séparation des sources.

1.2.1 Indépendance statistique

Soit un vecteur aléatoire $X = (X_1, \dots, X_n)$, les composantes du vecteur X sont mutuellement indépendantes si et seulement si :

$$p_X(x) = \prod_{i=1}^n p_{X_i}(x_i) \quad (1.14)$$

L'idée de l'analyse en composantes indépendantes est d'estimer la matrice de séparation à partir des observations $x_i, i = \{1, \dots, N\}$, et mesurer l'indépendance statistique entre les sorties $y_i, i = \{1, \dots, N\}$ estimées. Dans la pratique, les méthodes itératives sont basées sur la maximisation ou la minimisation de fonctions de cout comme l'information mutuelle, l'entropie, les statistiques d'ordre supérieur (cumulants), les moments non linéaires (algorithme de Héroult-Jutten)...etc, comme un critère d'indépendance.

En général, l'indépendance statistique des composantes du vecteur de sorties $y(t)$ n'implique pas nécessairement la séparation des sources. Dans [COM94], P. Comon a montré que l'indépendance statistique des signaux est une condition suffisante pour pouvoir résoudre un problème de séparation aveugle des sources. L'Analyse en Composantes Indépendantes (ACI) est une des méthodes qui permettent d'estimer des sources statistiquement indépendantes si la seule connaissance à priori sur les sources est leur indépendance statistique. Cette méthode a été introduite en 1988 par C. Jutten dans [JH88a], ensuite, elle a été développée par P. Comon en 1994 [COM94]. C'est une extension de la méthode d'Analyse en Composantes Principales (ACP) qui cherche des composantes décorréliées.

Nous allons voir que la caractérisation de l'indépendance statistique à l'ordre 2 n'est pas toujours suffisante pour la séparation d'un mélange linéaire de sources. Par exemple, dans le cas des sources non gaussiennes, il est possible d'exprimer l'indépendance statistique au delà de l'ordre 2. En utilisant les résultats de G. Darmon [DAR53] et de V.-P. Skitovich [SKI54], P. Comon a montré qu'il suffit d'obtenir des sources estimées statistiquement indépendantes si au plus une des sources est de loi gaussienne (hypothèse **H3**). Sous cette

condition de non gaussiannité, la procédure de séparation consiste à trouver une matrice de séparation B qui rend les sorties estimées statistiquement indépendantes. Pour cela, il faut pouvoir mesurer cette indépendance des estimées (Eq. 1.14).

Théorème (Darmois-Skitovich) :

Soient $s_1, \dots, s_N$, N variables aléatoires indépendantes,

$$y_1 = a_1 s_1 + a_2 s_2 + \dots + a_N s_N$$

$$y_2 = b_1 s_1 + b_2 s_2 + \dots + b_N s_N$$

Si y_1 et y_2 sont deux variables indépendantes, alors $a_k b_k \neq 0$ implique que les sources s_i sont gaussiennes.

1.2.1.1 Information mutuelle

La divergence de Kullback-Leibler (KL) permet de mesurer l'écart entre les distributions des densités de probabilité. L'équation de la divergence de KL entre deux densités de probabilité est comme suit :

$$K(f(u)|g(u)) = \int_{-\infty}^{+\infty} f(u) \log \frac{f(u)}{g(u)} du \quad (1.15)$$

où f et g sont deux densités de probabilité multidimensionnelles.

La divergence de Kullback-Leibler entre deux vecteurs aléatoires est positive ou nulle. Cette divergence s'annule si et seulement si les deux vecteurs aléatoires ont la même distribution de probabilité.

En effet, l'information mutuelle entre N variables aléatoires $x_1, x_2, \dots, x_N$ est définie par la divergence de Kullback-Leibler entre $p_x(X)$ et $\prod_{i=1}^N p_{x_i}(x_i)$, L'équation de l'information mutuelle s'écrit de la façon suivante :

$$I(x) = KL \left(p_x(X) \parallel \prod_{i=1}^N p_{x_i}(x_i) \right) = \int_x p_x(X) \ln \frac{p_x(X)}{\prod_{i=1}^N p_{x_i}(x_i)} dx \quad (1.16)$$

où $X = (x_1, x_2, \dots, x_N)^T$, p_x est la densité de probabilité conjointe de X , p_{x_i} est la densité de probabilité de x_i et $\prod_{i=1}^N p_{x_i}(x_i)$ est le produit des densités de probabilité marginales.

L'information mutuelle peut s'écrire aussi comme suit :

$$I(x) = \sum_{i=1}^N H(x_i) - H(x) \quad (1.17)$$

où $H(.) = -E \{ \ln p_{(.)}(.) \}$ est l'entropie de Shannon.

Selon les propriétés de la divergence de Kullback-Leibler [KUL59, CT91], $I(x)$ est toujours positive et égale à zéro si et seulement si $p_x(X) = \prod_i p_{x_i}(x_i)$. Dans ce cas, $x_1, x_2, \dots, x_N$ sont indépendantes. Par conséquent, on peut utiliser l'information mutuelle comme un critère d'indépendance pour la séparation d'un mélange de signaux, et la minimisation de la fonction $I(x)$ conduit à la séparation des sources.

Le problème de séparation de sources est ramené à l'annulation de l'information mutuelle des sources estimées $I(\hat{s}_1, \hat{s}_2, \dots, \hat{s}_N)$. Diverses méthodes et algorithmes ont été développés en utilisant l'information mutuelle pour résoudre un problème de séparation de sources ou l'analyse en composantes indépendantes (ACI). Pour plus de détails sur ces méthodes, voir [ALM01, CAR98, COM94, PHA99].

1.2.1.2 Mesure du maximum de vraisemblance

L'utilisation du maximum de vraisemblance pour la mesure de l'indépendance entre les sources a été introduit par M. Gaeta et J.-L. Lacoume en 1990 [GL90]. Ensuite, D.-T. Pham a utilisé cette approche de maximum de vraisemblance pour la séparation aveugle des sources [PG97].

On considère N sources $s_i(t)$ ayant pour distribution de probabilité $f_{s_i}(.)$ avec $(i = 1, \dots, N)$. Ces sources sont mélangées à l'aide d'une matrice de mélange notée $\mathcal{A}$ telle que les observations sont $x(t) = \mathcal{A}s(t)$. Le logarithme de la vraisemblance s'écrit comme suit :

$$L_T = T \left\{ \sum_{i=1}^N E \log [f_{s_i}(e_i \mathcal{A}^{-1} x(t))] - \log |\det \mathcal{A}^{-1}| \right\} \quad (1.18)$$

- e_i représente la i ème colonne de la matrice identité I .
- E est une moyenne temporelle.

L'estimation du maximum de vraisemblance nécessite la connaissance des densités de probabilités des sources $s(t)$. Ce n'est pas le cas dans le cas de séparation aveugle de sources. Ainsi, on utilise la notion des *fonctions score* pour l'estimation des densités de probabilités (voir Annexe ??). Dans la pratique, on remplace la densité des sources $s_i(t)$ par la densité des sources estimées $y_i(t)$ telle que $y(t) = \hat{\mathcal{A}}^{-1}x(t)$. Ensuite, on cherche à maximiser l'équation (1.18) :

$$E \left[\psi_{y_i}(e_i \hat{\mathcal{A}}^{-1}x(t)) e_j \hat{\mathcal{A}}^{-1}x(t) \right] = 0 \quad i \neq j = 1, \dots, N \quad (1.19)$$

où $\psi_{y_i}(\cdot)$ représente une fonction score :

$$\psi_{y_i} = -\frac{f'_{y_i}}{f_{y_i}}$$

1.2.1.3 Moments et cumulants

Les propriétés des statistiques d'ordre deux ne sont pas suffisantes pour la caractérisation de l'indépendance de signaux non gaussiens. Les statistiques d'ordre trois s'annulent pour des densités de probabilité symétriques, donc, il est nécessaire de faire appel à des approximations à l'ordre 4. En général, les plus utilisées sont les kurtosis, les cumulants d'ordre quatre normalisés [LAC97]. L'utilisation des statistiques d'ordre supérieur à deux pose un problème dans le cas des signaux gaussiens car les cumulants d'ordre supérieur à 2 d'un signal gaussien sont tous nuls. Donc la séparation des sources gaussiennes est impossible sauf dans le cas où on a une seule source gaussienne dans le mélange.

Avant de présenter le principe des méthodes utilisant les statistiques d'ordre 4 comme critères d'indépendance, nous introduisons quelques définitions de ces statistiques.

Définition 1 : Première fonction caractéristique :

Soit $X = (X_1, \dots, X_n)^T$ un vecteur aléatoire ayant une densité de probabilité f_X . La première fonction caractéristique de X , notée $\Phi_X(u)$, la valeur moyenne de la fonction $\exp(ju^T X)$. Cette première fonction caractéristique est le complexe conjugué de la

transformée de Fourier de f_X . Il s'agit d'une fonction complexe de la variable $u \in \mathbb{R}^n$, définie par :

$$\Phi_X(u) = \int f_X(x) \exp(ju^T X) dx \quad (1.20)$$

Définition 2 : Seconde fonction caractéristique :

Soit $X = (X_1, \dots, X_n)^T$ un vecteur aléatoire ayant une densité de probabilité f_X . La seconde fonction caractéristique de X , notée $\psi_X(u)$, est définie par le logarithme népérien de la première fonction caractéristique :

$$\psi_X(u) = \ln(\Phi_X(u)) \quad (1.21)$$

Définition 3 : Cumulant :

On peut définir les cumulants à partir de la seconde fonction caractéristique. Le cumulant d'ordre k de X est donné par :

$$\text{Cum}(X_1, \dots, X_k) = (-j)^k \left. \frac{\partial^k \psi_X(u)}{\partial x_1 \dots \partial x_k} \right|_{u=0} \quad (1.22)$$

Le cumulant de k variables aléatoires $(X_1, X_2, \dots, X_k)$, noté $\text{Cum}(X_1, X_2, \dots, X_k)$ est défini par :

$$\text{Cum}_p^{q,X} = \text{Cum}(\underbrace{X_{i_1}(\tau_1), \dots, X_{i_p}(\tau_p)}_p, \underbrace{X_{i_{p+1}}^*(\tau_{p+1}), \dots, X_{i_k}^*(\tau_k)}_{q=k-p}) \quad (1.23)$$

Le premier algorithme de séparation exploitant les statistiques au delà de l'ordre 2 est présenté dans [HJ86, HJA85], il est basé sur la minimisation du critère :

$$E(f(u_i)g(u_j)) \quad (1.24)$$

avec f et g deux fonctions impaires.

On cite également l'algorithme JADE basé sur la minimisation du critère suivant :

$$\sum_{ijkl \neq iikl} \text{Cum}(u_i, u_j, u_k, u_l)^2 \quad (1.25)$$

1.2.2 Méthodes à l'ordre 2 : Décorrélation

Dans un premier temps, nous présentons les méthodes de séparation de sources exploitant les statistiques d'ordre 2. On rappelle brièvement les statistiques d'ordre 2 utilisées par ces méthodes.

L'intercorrélation entre deux signaux aléatoires stationnaires centrés $x(t)$ et $y(t)$ est :

$$\Gamma_{xy}(\tau) = E \{x(t)y(t - \tau)\} \quad (1.26)$$

Cette quantité, pour des différentes valeurs de τ , définit les statistiques d'ordre 2 des signaux x et y , en particulier, pour $\tau = 0$, $\Gamma_x(0) = Var(x)$, $\Gamma_y(0) = Var(y)$ et $\Gamma_{xy}(\tau) = Cov(x, y)$.

Les méthodes à l'ordre 2 n'exploitent pas complètement l'indépendance statistique des sources mais se limitent à leur décorrélation. Elles exploitent une seconde hypothèse sur les sources qui ne doivent pas être blanches $\Gamma_x(\tau) = E \{x(t)x(t - \tau)\} \neq 0$ pour $\tau \neq 0$. Dans la suite, nous allons détailler le principe de ces méthodes qui s'effectue en deux étapes, tout d'abord, l'étape de blanchiment, ensuite l'estimation d'une matrice orthogonale [BMCM97].

1.2.2.1 Blanchiment

Cette première étape consiste à calculer une matrice W appelée matrice de blanchiment. Cette matrice est exprimée à partir de la décomposition en valeurs propres de la matrice de covariance des observations $X(t)$. W , appliquée aux observations $x(t)$, fournit les signaux décorrélés $v = Wx$

$$\begin{aligned} \Gamma_{XX}(0) &= E \{X(t)X^T(t)\} \\ &= AE \{S(t)S^T(t)\} A^T \\ &= A\Gamma_{SS}(0)A^T \\ &= AA^T \end{aligned} \quad (1.27)$$

La matrice $\Gamma_{XX}(0)$ est symétrique et non négative. On peut ainsi écrire la décomposition en valeurs propres :

$$\Gamma_{XX}(0) = UDU^T \quad (1.28)$$

où :

- U est une matrice orthogonale composée des vecteurs propres de Γ_{XX} .
- D est une matrice diagonale composée des valeurs propres de Γ_{XX} .

La matrice $W = D^{-\frac{1}{2}}U^T$ est une matrice de blanchiment car les signaux $V(t) = WX(t)$ sont de covariance unitaire. En effet, on a :

$$\begin{aligned} \Gamma_{VV}(0) &= W\Gamma_{XX}(0)W^T \\ &= D^{-\frac{1}{2}}U^TUDU^TUD^{-\frac{1}{2}} \\ &= I_d \end{aligned} \quad (1.29)$$

Après le blanchiment par W , on obtient des signaux $v(t)$ décorrélés et normés, mais la séparation n'est pas achevée. Donc, on considère la décomposition en valeur singulière (SVD) de la matrice A :

$$A = UD^{\frac{1}{2}}Q^T \quad (1.30)$$

où D est une matrice diagonale composée des valeurs propres de A , U et Q des matrices orthogonales.

Donc, il reste encore un mélange qui correspond à une matrice orthonormale :

$$\begin{aligned} V &= WAS \\ &= D^{-\frac{1}{2}}U^TUD^{\frac{1}{2}}Q^TS \\ &= Q^TS \end{aligned} \quad (1.31)$$

Donc, V est un mélange des sources S par une matrice orthogonale Q^T . Il reste à estimer cette matrice Q^T pour séparer le mélange.

Cette première étape qui est le blanchiment des observations, est une opération préalable, utilisée dans diverses méthodes de séparation [COM94, CS93] basées sur les statistiques d'ordre 2, afin de simplifier le problème. Cette opération de blanchiment correspond à une Analyse en Composantes Principales (ACP) des observations et les composantes principales sont les observations blanchies.

1.2.2.2 Estimation de la matrice de séparation

L'estimation de la matrice de séparation notée Q permet de reconstruire l'estimation des sources s . Cette méthode présentée par de nombreux auteurs [BAMCM97, FET88, MS94, TLSH91] est possible si les sources ont des fonctions de corrélation différentes. La méthode présentée utilise le calcul de la covariance des observations blanchies :

$$\begin{aligned}
 \Gamma_{VV}(\tau) &= E \{v(t)v(t-\tau)^T\} \\
 &= WAE \{s(t)s(t-\tau)^T\} A^T W^T \\
 &= Q^T \Gamma_{ss}(\tau) Q
 \end{aligned} \tag{1.32}$$

$\Gamma_{ss}(\tau)$ est une matrice diagonale car les sources sont statistiquement indépendantes et centrées. Si tous les termes diagonaux de cette matrice sont différents il est possible de retrouver Q et $\Gamma_{ss}(\tau)$ en faisant une diagonalisation de la matrice $\Gamma_{zz}(\tau)$. Par contre, si deux des termes diagonaux sont identiques, deux colonnes de Q ne pourront pas être estimées. Elles seront choisies orthogonales et normées dans l'espace vectoriel de dimension 2. Mais, il y a une infinité de choix possibles, dont un seul correspond à la vraie matrice Q .

Le choix du retard doit se faire de sorte que les termes diagonaux de $\Gamma_{uu}(\tau)$ soient tous différents. Ainsi, une condition nécessaire de succès pour cette méthode est l'existence d'un retard $\tau \neq 0$ tel que les fonctions d'autocorrélation $\Gamma_{u_i u_i}$ des sources en τ soient non nulles et distinctes, c'est-à-dire $\forall i, j, i \neq j$,

$$\Gamma_{u_i u_i}(\tau) \neq \Gamma_{u_j u_j}(\tau) \tag{1.33}$$

Dans la pratique, on travaille avec plusieurs retards $\tau_i \neq 0$ et on diagonalise conjointement les matrices de covariance $\Gamma_{vv}(\tau_i)$ (pour plus de détails, voir l'algorithme SOBI proposé par A. Belouchrani en 1997 [BAMCM97]). La matrice $\hat{Q}$ résultat de la diagonalisation conjointe est une estimation de la matrice Q

$$\hat{A}^{-1} = \hat{Q}W \quad (1.34)$$

C'est un problème d'optimisation qui consiste à minimiser la somme quadratique des éléments hors diagonaux des matrices $\hat{Q}\Gamma_{vv}(\tau_i)\hat{Q}^T$ [CS96].

Le problème de la séparation des sources est achevé en calculant la matrice de séparation B telle que (cf. Fig. 1.5) :

$$B = \hat{Q}W \quad (1.35)$$

avec $\hat{Q}$ l'estimée de la matrice Q .

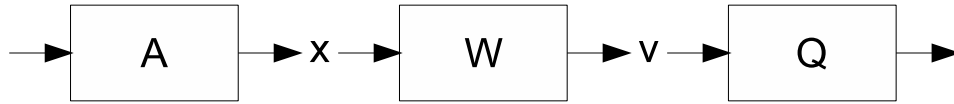


FIG. 1.5 – Structure de séparation à l'ordre 2

Les résultats utilisant cette méthode sont représentés dans la figure 1.6. Ils représentent la distribution de probabilité du mélange des sources, des observations blanchies et des sources obtenues. On remarque qu'après l'étape de blanchiment des observations, il ne reste plus qu'à effectuer une rotation de la distribution de probabilité des observations blanchies afin de retrouver les sources originales.

1.2.3 Méthodes aux ordres supérieurs : Fonctions de contraste

1.2.3.1 Fonctions de contraste

Les statistiques d'ordre supérieur (SOS) sont des outils qui permettent l'évaluation du critère d'indépendance statistique de variables aléatoires non gaussiennes. Les cumulants d'ordre supérieur à deux d'une distribution gaussienne sont nuls. Pour résoudre un problème de séparation aveugle de sources, les cumulants permettent de construire

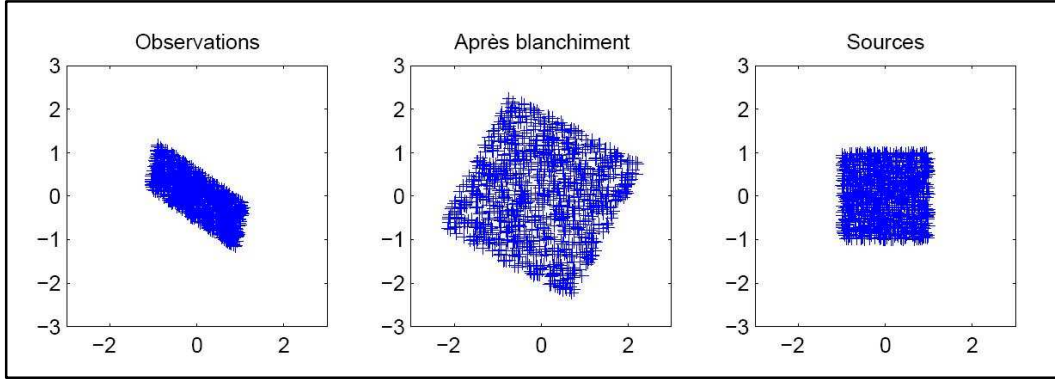


FIG. 1.6 – Exemple de résultats d’une méthode de séparation de deux sources basée sur les statistiques d’ordre 2 [JAL06]

un critère d’indépendance partielle [MTM99, PM01], et on peut les considérer comme des outils d’approximation d’autres critères d’indépendance telle que la néguentropie [COM94, CAR99].

Il est difficile d’utiliser un critère d’indépendance à partir des densités de probabilité comme on a vu dans le cas de l’utilisation de l’information mutuelle, d’où l’utilisation des fonctions qui caractérisent l’indépendance statistique, appelées **fonctions de contraste**. Ces fonctions sont maximales si et seulement si la condition d’indépendance statistique est vérifiée. Dans cette partie, nous présentons les fonctions de contrastes, qui sont des critères d’optimisation pour la séparation de sources et basées sur les statistiques d’ordre 4.

En statistique, la notion de fonctions de contraste permettait le calcul des estimateurs [EGU83, PFA73] appelés estimateurs de minimum de contraste. Ensuite, en traitement du signal, [DON81a] a défini la notion des fonctions de contraste pour le problème de la déconvolution et [COM94] pour l’Analyse en Composantes Indépendantes (ACI).

Pour résoudre le problème de séparation de sources, certaines mesures d’indépendance utilisent les fonctions de contraste qui ont été initiées par P. Comon en 1994 [COM94] et qui ont été généralisées par E. Moreau en 1999 [MTM99].

P. Comon a introduit la fonction suivante :

$$\Psi(y) = \sum_{k=1}^N |(cum(y_k(n), y_k^*(n), y_k(n), y_k^*(n)))|^2 \quad (1.36)$$

Lemme 1 : *La fonction ψ est maximisée avec un maximum global si et seulement si les composantes du vecteur y sont mutuellement indépendantes.*

Les fonctions vérifiant ce lemme permettent de définir des algorithmes de séparation de sources. Ces fonctions sont appelées **fonctions de contrastes**.

Définition : *Fonction de contraste :*

Un contraste est une fonction $\mathcal{I}(\cdot)$ sur un ensemble $\mathcal{H}$ satisfaisant les propriétés suivantes :

- **Propriété 1 :** Le contraste ne peut pas changer sur l'ensemble des solutions possibles.

$$\forall y \in \mathcal{H}, \forall G \in \mathcal{G},$$

$$I(Gy) = I(y) \quad (1.37)$$

- **Propriété 2 :** Si la séparation de sources est achevée, alors tout filtre n'appartenant pas à $\mathcal{G}$ minimise le contraste.

$$\forall a \in \mathcal{A}, \forall S \in \mathcal{S},$$

$$I(Sa) \leq I(a) \quad (1.38)$$

- **Propriété 3 :** Le maximum de la fonction de contraste est atteint pour des filtres liés les uns aux autres par des filtres triviaux.

$$\forall a \in \mathcal{A}, \forall S \in \mathcal{S},$$

$$I(Sa) = I(a) \Rightarrow S \in \mathcal{P}. \quad (1.39)$$

La méthode permettant de résoudre le problème de séparation de sources est donc d'estimer la matrice de séparation B telle $y(n) = Bx(n)$ maximise la fonction ψ . Une des méthodes permettant de maximiser la fonction ψ est l'algorithme JADE [CS93] basé sur la diagonalisation conjointe des matrices de cumulants.

Plusieurs contrastes ont été proposés par Comon [COM94], en se plaçant après blanchiment spatial (ACP). l'information mutuelle définie dans la section 1.2.1.1 est considérée comme un contraste sur l'ensemble des variables aléatoires ayant au plus une composante gaussienne. Mais cette dernière est difficile à calculer directement. On peut chercher une approximation avec un développement en série d'Edgeworth ou de Gram-Charlier de la densité de probabilité [GL90].

1.2.3.2 Exemples de contrastes

L'utilisation de l'information mutuelle pour résoudre un problème de séparation de sources, nécessite la connaissance ou l'estimation des densités de probabilité des vecteurs considérés, ce qui n'est pas un problème simple. Donc, il est nécessaire d'avoir une meilleure approximation du contraste $-I(x)$ où $I(x)$ est l'information mutuelle définie précédemment.

Une autre fonction de contraste a été proposée par Comon aussi [COM96] :

$$\mathcal{J}_1(y) = \sum_{i=1}^N |C[y_i]|^2 \quad (1.40)$$

La maximisation de $\mathcal{J}_1^2(y)$ est équivalent à la minimisation de la somme des carrés de tous les cumulants croisés. Le contraste le plus utilisé est : $\sum |C(y_i)|^2$

D'autres types de fonctions de contrastes ont été proposés par Cardoso [CS93] et qui comportent que des cumulants croisés. Ce type de contraste s'écrit comme suit :

$$\mathcal{J}_2(y) = - \sum_{i_1, i_2, i_3}^N |Cum[y_{i_1}, y_{i_1}, y_{i_2}, y_{i_3}]|^2 \quad (1.41)$$

L'intérêt de ce contraste réside dans son optimisation car il revient à une diagonalisation conjointe d'un ensemble de matrices. Ce contraste a ensuite été généralisé pour des cumulants d'ordre quelconque supérieur ou égal à trois [MOR99].

1.3 Quelques méthodes de séparation en mélange convolutif

Après avoir présenté un état de l'art sur le principe du problème de la séparation de sources, nous allons décrire dans cette section, quelques approches de séparation dans le cas de mélange convolutif de sources. Ce type de mélange intervient dans beaucoup de domaines d'applications. Certaines approches utilisent les statistiques d'ordre supérieur 2 pour la conception du critère d'indépendance. D'autres algorithmes se restreignent aux statistiques d'ordre 2. En général, dans la plupart des approches développées, la seule information disponible est celle des observations enregistrées au niveau des capteurs.

1.3.1 Approche temporelle

1.3.1.1 Méthode adaptative (*travaux de C. Jutten et H.L. Nguyen Thi*)

Les premiers travaux pour résoudre un problème de séparation de sources en mélange convolutif ont été introduit par C. Jutten au début des années 90 [JTD⁺91, TC95, TJ93, THI93, TJC91]. Dans ces travaux, C. Jutten et H.L. Nguyen Thi ont généralisé leur approche temporelle développée pour le cas de mélange instantané des sources.

Le principe est basé sur le fait de chaque sources s_i est obtenue en fonction des contributions des autres sources sur les capteurs :

$$y_i(t) = x_i(t) - \sum_{i=1}^p \sum_{k=0}^M C_{ij}(k) y_j(t-k) \quad (1.42)$$

où :

- $x(t)$ représente les observations,
- $y(t)$ représente les sorties estimées,
- $C_{ii} = 0$ et $i \neq j$.

Comparant au cas de mélange instantané, on peut distinguer plusieurs façons d'ajustement des paramètres C_{ij} selon le critère de mesure d'indépendance entre les sources :

- Le critère d'indépendance est l'annulation des cumulants croisés d'ordre impair :

$$E[f(s_i)g(s_j)]$$

On utilise :

$$C_{ij}(t+1, k) = C_{ij}(t, k) + \mu f(y_i(t))g(y_j(t-k))$$

- Le critère d'indépendance est l'annulation de l'intercumulant symétrique :

$$C1 = cum(s_i(t), s_i(t), s_j(t-k), s_j(t-k))$$

On utilise :

$$C_{ij}(t+1, k) = C_{ij}(t, k) + \mu \text{signe} \left(\frac{\partial C1}{\partial C_{ij}(t, k)} \right) C1$$

- Le critère d'indépendance est l'annulation de l'intercumulant dissymétrique :

$$C2 = cum(s_i(t), s_j(t-k), s_j(t-k), s_j(t-k))$$

On utilise :

$$C_{ij}(t+1, k) = C_{ij}(t, k) + \mu \text{signe} \left(\frac{\partial C2}{\partial C_{ij}(t, k)} \right) C2$$

1.3.1.2 Annulation des bispectres croisés (*travaux de D. Yellin et E. Weinstein*)

Dans cette approche, l'estimation des filtres de séparation est basée sur l'annulation des bispectres croisés des sorties $y(t)$ [YW93, YW94].

Pour décrire cette approche, nous considérons le cas classique de deux sources et deux capteurs. On note $H(z)$ la matrice des filtres de mélange et $G(z)$ la matrice des filtres de séparation telles que :

$$H(z) = \begin{bmatrix} H_{11}(z) & H_{12}(z) \\ H_{21}(z) & H_{22}(z) \end{bmatrix}$$

et

$$G(z) = \begin{bmatrix} G_{11}(z) & G_{12}(z) \\ G_{21}(z) & G_{22}(z) \end{bmatrix}$$

Les auteurs proposent l'équation de décorrélation suivante :

$$G_{12}(e^{i\omega}) = \frac{P_{y_2 y_1}(\omega) - (G_{12}(e^{i\omega}))^* P_{y_2 y_2}(\omega)}{P_{y_1 y_1}(\omega) - (G_{12}(e^{i\omega}))^* P_{y_1 y_2}(\omega)} \quad (1.43)$$

où $P_{y_i y_j}(\omega)$ représente la densité interspectrale de puissance entre y_i et y_j .

1.3.1.3 Les fonctions de contraste

L'utilisation des critères d'indépendance à base des densités de probabilité des sources n'est pas toujours valable. Ainsi, on retrouve d'autres approches qui consistent à l'utilisation des fonctions appelées **fonctions de contraste** permettant maximiser

l'indépendance entre les signaux. Ces fonctions sont maximales si et seulement si les signaux traités sont mutuellement indépendants.

L'utilisation des fonctions de contraste dans le cadre des mélanges convolutifs a été introduit en 1996 [COM96, MP97]. Les auteurs ont démontré la possibilité de généraliser des fonctions de contrastes développées dans le cadre des mélanges instantanés pour qu'elles soient appliquées aux mélanges convolutifs. Dans un premier temps, nous rappelons la définition d'une fonction de contraste (mélange convolutif) introduite par P. Comon.

Définition : Une fonction de contraste ϕ est une fonction de l'ensemble $\mathcal{S}$ des vecteurs aléatoires s dans $\mathfrak{R}$ satisfaisant les propriétés suivantes :

- **P1** : $\forall s \in \mathcal{S}$, $\phi(s)$ ne dépend que de la loi de probabilité de s .
- **P2** : $\forall s \in \mathcal{S}$, $\phi(Ps) = \phi(s)$, où P est une permutation.
- **P3** : $\forall s \in \mathcal{S}$, $\phi(Ds) = \phi(s)$, où D est une matrice régulière constante.
- **P4** : Si les variables aléatoires s sont à composantes mutuellement indépendantes, alors $\phi([A(z)]s) \leq \phi(s)$, où $A(z)$ est une matrice inversible. $\phi([A(z)]s) = \phi(s)$ si et seulement si $A(z)$ est de la forme $PD(z)$ où P est une matrice de permutation et $D(z)$ est une matrice de retards.

En général, les méthodes utilisant des fonctions de contraste sont souvent basées sur des statistiques aux ordres supérieurs. En utilisant des fonctions de contraste en séparation de sources, on cherche à estimer la matrice de séparation notée B telle que :

$$B = \text{Max}(\phi(Ay)) \quad (1.44)$$

où

- A est la matrice du mélange.
- y représente le vecteur des sorties estimées.

Le calcul d'un critère de séparation basé sur les cumulants nécessite un temps de calcul assez lourd. Ainsi, nous exploitons des approches qui ont été développées dans [MPTM99,

TMM00] permettant de faire face à cette difficulté.

Dans la plupart des cas, les contrastes en convolutif sont vus comme des extensions des contraste en instantané. L'expression d'un contraste pour le cas convolutif d'un mélange de signaux non i.i.d s'écrit comme suit [TMM00] :

$$\mathcal{I} = \sum_{i \neq j} \sum_{l_1, l_2, l_3} |cum \{y_i(n), y_i^*(n + l_1), y_j(n + l_2), y_j^*(n + l_3)\}|^2 \quad (1.45)$$

où $l_i, i = 1, 2, 3$ représente le vecteur des retards. D'une manière plus simple, l'expression de ce contraste s'écrit aussi :

$$\mathcal{I} = \sum_i \sum_{l_1, l_2, l_3} |cum \{y_i(n), y_i^*(n + l_1), y_i(n + l_2), y_i^*(n + l_3)\}|^2 \quad (1.46)$$

Ce contraste temporel est basé sur les cumulants d'ordre 4 et valable dans le cas général non i.i.d. Ce même contraste a été utilisé d'une façon complètement temporelle dans [LI02]. Cependant, en raison des retards l_i dont leur sommation qui est infinie, l'utilisation de ce contraste devient très difficile.

Afin de simplifier l'utilisation du contraste de l'équation (1.46), il faut prendre en compte des propriétés des sources utilisées qui vont faire disparaître tous les termes qui s'annulent lorsque la séparation est effectuée.

Par exemple, si les sources $s_i(n)$ sont supposées i.i.d. Dans ce cas, on peut se limiter à $l_1 = l_2 = l_3 = 0$ dans l'équation (1.46), ainsi, on obtient à nouveau le contraste proposé par P. Comon [COM96] qui s'écrit comme suit :

$$\mathcal{I} = \sum_i |cum \{y_i(n), y_i^*(n), y_i(n), y_i^*(n)\}|^2 \quad (1.47)$$

En utilisant des contrastes fréquentiels de la même forme, nous ne rencontrons pas ce problème de sommation infinie sur les retards l_i mais on a une triple sommation sur les canaux de fréquences. Cette approche spectrale a été développée dans [CPP05]. Les auteurs expriment les tri-spectres des observations en fonction du tri-spectre des sources pour conclure que l'égalité entre les énergies de ces grandeurs considérées est atteinte si et seulement si la séparation est effectuée. Par contre, dans le domaine temporel, l'estimation

de ces contrastes exigent la prise en compte de certaines propriétés sur la nature des sources.

Fonctions de contraste avec référence

Dans ce paragraphe, nous allons présenter une approche utilisant les fonctions de contrastes dite **contraste avec référence** [CMP04] utilisant des cumulants croisés de type $cum(y(n), y(n), r(n), r(n))$ entre une estimation $r(n)$ d'une source et un signal de référence $r(n)$. Ce dernier peut être choisi comme une version filtrée quasiment quelconque des sources. L'intérêt majeure de cette approche est de conduire à des critères quadratiques dont la maximisation est facile à réaliser. La figure (1.7) montre la problématique de l'approche étudiée.

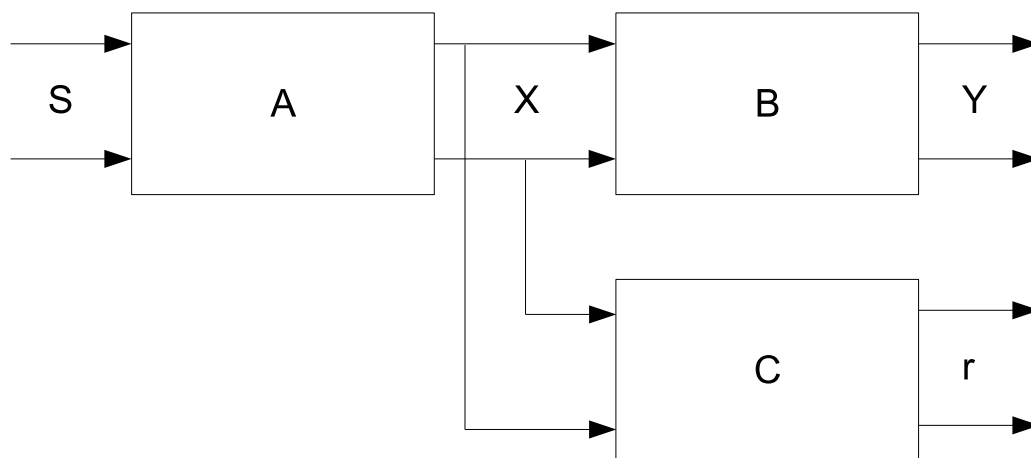


FIG. 1.7 – Modèle de système de séparation avec référence

- A et B représentent respectivement les matrices de mélange et de la séparation.
- C représente la matrice de séparation de *référence* avec $r(n)$ le vecteur des signaux *référence*.

Dans le cadre des mélanges instantanés, les auteurs dans [AMA03a, AMA03b, AMA04],

proposent le contraste avec référence qui s'écrit comme suit :

$$\mathcal{I} \{y(n), r(n)\} = \sum_{i,j=1}^N \left| cum \left(y_i(n), y_i^*(n), \underbrace{r_j(n), r_j^*(n), \dots, r_j^*(n)}_{R-2fois} \right) \right|^2 \quad (1.48)$$

Ce contraste utilise une approche conjointe de séparation. En parallèle, M. Castella and *al.* proposent dans [CMP04] un contraste utilisant l'approche par déflation qui est resté dans un premier temps valable pour des mélanges instantanés et des sources i.i.d. ayant des cumulants de même signe. Ensuite, ces mêmes auteurs ont généralisé des contrastes avec référence pour les mélanges convolutifs en utilisant toujours une approche séquentielle. Ces contrastes ont été validés sur des sources i.i.d et non i.i.d.

Les problèmes qui se posent souvent en utilisant des contrastes avec référence, sont les connaissances nécessaires à l'obtention d'un signal de référence qui sont relativement faibles. Dans [RCM06], les auteurs proposent de prendre un capteur blanchi puis ils appliquent une procédure itérative où une estimation provisoire de la source sert comme signal de référence. La méthode converge finalement vers les sources correctes. En revanche, l'avantage majeur de ce type de critères réside dans leur dépendance quadratique vis à vis des paramètres.

1.3.1.4 Méthode séquentielle (*Approche par déflation*)

Pour résoudre un problème de séparation de sources, les méthodes séquentielles permettent d'extraire les sources les unes après les autres, contrairement aux méthodes citées précédemment permettant de retrouver les sources simultanément. Ce type de méthodes a été introduit par P. Loubaton en 1993 [LR93] pour séparer un mélange convolutif de sources. Ensuite, cette approche a été développée par N. Delfosse en 1995 dans le cas de mélange instantané [DL95] et par J.-K. Tugnait en 1997 dans le cas de mélange convolutif [TUG97, TUG99].

On note $s(n)$ les sources originales et $x(n)$ les observations telles que $x(n) = [A(z)]s(n)$. On a considéré le cas de N sources et N capteurs. Le principe de cette approche repose sur les deux étapes suivantes :

- Extraction d'une sources du mélange
- Soustraction de la contribution de cette source du mélange

Extraction d'une source du mélange

Cette étape consiste à estimer un filtre noté $B(z)$ de façon à que le signal estimé $y(n) = [B(z)]x(n)$ ne dépende que d'une source et une seule. Afin de retrouver le filtre $B(z)$, on utilise souvent des fonctions appelées *fonctions de contraste*. Le maximum de ces fonctions est atteint si et seulement si le signal estimé dépend d'une seule source.

Soustraction de la source extraite au mélange

Cette étape consiste à soustraire la contribution de la source extraite du mélange pour obtenir ainsi, un nouveau mélange de $N - 1$ sources. Pour effectuer cette étape, on cherche à estimer un filtre $t(z)$ qui minimise le critère suivant :

$$E (||x(n) - [t(z)]y(n)||^2) \quad (1.49)$$

On considère $A_i(z)$ ($i = 1, \dots, N$) les colonnes de la matrice $A(z)$. Ainsi, l'équation (1.49) s'écrit comme suit :

$$E \left(\left\| ([A_1(z) - t(z)c(z)])s_1(n) + \sum_{i=2}^N [A_i(z)]s_i(n) \right\|^2 \right) \quad (1.50)$$

Ce critère est minimisé si $t(z)c(z) = A_1(z)$. Dans [DEL95, SIM99, JAL06], nous trouvons plus de détails sur le calcul des différentes étapes de l'approche par déflation.

Nous citons aussi des méthodes appelées *approches hiérarchiques* [TFLT01]. Dans ce type de méthodes, la séparation consiste à la maximisation successive de plusieurs critères dans lesquels on introduit à chaque étape un terme de pénalisation. Cette pénalisation représente la décorrélation spatiale des signaux estimées.

1.3.2 Approche fréquentielle

L'utilisation d'une approche fréquentielle pour résoudre un problème de séparation de sources permet de décomposer le mélange convolutif en plusieurs mélanges

instantanés. Cette approche consiste à effectuer une séparation instantanée par bandes de fréquence. La difficulté de l'utilisation d'une telle méthode réside dans l'indétermination de permutation entre les différentes séparations instantanées. La méthode fréquentielle a été introduite par C. Servière et *al.* dans les années 90 [CAP95, CSL94, CSL95a, CSL95b, DC03, DC01, DSC03]. Plusieurs méthodes aussi ont été proposées pour résoudre ce problème majeur de permutation en posant des hypothèses supplémentaires sur les sources [SP04] et sur les matrices de mélange [PS00, PSB03]. D'autres travaux aussi ont été développés récemment et qui utilisent une approche temps-fréquence pour la séparation aveugle des sources [PD05, RBR01, YR04] dans le cadre des mélanges convolutifs.

Récemment, les approches fréquentielles en séparation de sources utilisent des statistiques d'ordre supérieur des transformés de Fourier des observations. D'autres approches développées aussi dernièrement exploitent les multi-spectres des observations [CP01].

1.4 Conclusion

Dans ce premier chapitre, nous avons présenté d'une manière assez générale le principe de la séparation de sources. Nous avons présenté les problèmes et les modèles utilisés en séparation des sources. Ensuite, nous avons cité brièvement les hypothèses sur les sources et sur les mélanges, les types de mélanges utilisés, particulièrement, le mélange convolutif qui va susciter notre intérêt dans la suite de cette thèse.

Enfin, nous avons présenté une synthèse générale sur les différents méthodes et approches utilisées pour résoudre le problème de séparation de sources.

Nous avons aussi mentionné les principaux outils nécessaires à la séparation de sources aveugles dans le cas de mélange linéaire convolutif en citant des méthodes utilisant la décorrélation et les statistiques d'ordre supérieur.

Nous avons présenté quelques méthodes de séparation sources dans le cadre des mélanges convolutifs. Dans certains cas, une approximation ou l'utilisation des statistiques

de second ordre permet la simplification du mélange traité et de le ramener à un mélange instantané. D'autres approches utilisent des critères dits *fonctions de contraste*. Souvent, ces approches nécessitent un temps de convergence assez important et peuvent conduire dans certains cas à des minima locaux et donc, à des solutions non-optimales. D'autre part, l'utilisation d'une étape de préblanchiment peut soulever des difficultés supplémentaires.

Dans la suite de ce manuscrit, on présentera quelques idées théoriques utilisant l'information mutuelle comme critère d'indépendance ou les statistiques d'ordre 4 afin de généraliser un contraste. Pour augmenter l'efficacité des méthodes développées, nous avons utilisé le filtre de Wiener pour le calcul de la contribution de chaque source sur chaque capteur. L'intérêt de l'utilisation du filtre de Wiener dans la phase de séparation permet de minimiser l'erreur quadratique entre les sources estimées et les observations. Les simulations informatiques montrent la convergence de ces algorithmes dans plusieurs cas différents.

Chapitre 2

Séparation de sources : Information mutuelle et filtre de Wiener

Sommaire

2.1	Critère d'indépendance	52
2.1.1	Introduction	52
2.1.2	Mesure de l'indépendance statistique	53
2.1.3	Mesure de la divergence entre les densités de probabilité	54
2.1.4	Mesure de l'information mutuelle	54
2.1.5	Mesure de l'indépendance dans le cas de mélange convolutif	55
2.1.6	Minimisation de l'information mutuelle	57
2.2	Filtre de Wiener	58
2.2.1	Définition	59
2.2.2	Modélisation du filtre de Wiener	59
2.3	Méthode de séparation proposée	60
2.3.1	Modélisation du problème	61
2.3.2	Calcul des contributions des sources sur un capteur	63
2.3.2.1	Premier cas	64
2.3.2.2	Deuxième cas	65
2.4	Algorithme de séparation	71
2.4.1	Calcul des gradients	71
2.4.2	Simulations et résultats obtenus	73
2.5	Conclusion	81

Un physicien moderne étudie la physique quantique les lundis, mercredis et vendredis et médite sur la théorie de la relativité gravitationnelle les mardis, jeudis et samedis. Le dimanche, il prie... pour que quelqu'un trouve la corrélation entre les deux.

Norbert Wiener

CE CHAPITRE présente les outils utilisés dans cette thèse. Ces outils servent d'une part à comprendre la problématique générale de la séparation de sources et d'autre part à présenter les méthodes développées. Dans cette thèse, nous cherchons à calculer toutes les contributions des sources sur tous les capteurs, et pas les sources elles-mêmes. Ces contributions permettront d'effectuer un diagnostic du système étudié. Dans un premier temps, nous allons présenter un bref descriptif de l'information mutuelle qui sera utilisée comme critère d'indépendance en décrivant l'intérêt de la méthode développée, basée sur l'utilisation du filtre de Wiener. Ce filtre permettra de calculer les contributions de chaque source sur chaque capteur. L'algorithme proposé, destiné à la séparation d'un mélange convolutif de sources, est basé sur la minimisation de l'information mutuelle et en même temps, sur la projection de chaque source estimée sur chaque capteur en utilisant le filtre de Wiener. L'intérêt d'insérer les contributions des sources sur les capteurs permet de simplifier la minimisation du critère d'indépendance utilisé. Nous montrerons par un calcul théorique l'efficacité de la méthode proposée. Ensuite, quelques simulations seront finalement présentées dans ce chapitre afin d'évaluer la performance de la méthode et l'algorithme proposés.

2.1 Critère d'indépendance

2.1.1 Introduction

Dans cette partie, nous allons présenter une méthode de séparation de sources convolutives, utilisant l'information mutuelle comme critère d'indépendance. Dans un premier temps, nous présenterons la mesure de l'indépendance statistique dans le cas de mélange

convolutif ainsi que les propriétés de l'information mutuelle. Nous allons montrer que l'utilisation de ce critère d'indépendance pour les mélanges convolutifs est assez complexe. Pour simplifier la minimisation de l'information mutuelle dans ce cas, nous allons démontrer dans la deuxième partie que le fait de projeter les sources estimées sur chaque capteur permettra de simplifier la minimisation du critère d'indépendance et par la suite, d'avoir une meilleure séparation des sources.

La méthode qui sera décrite dans la suite de ce chapitre, utilise l'approche de gradient afin de minimiser une fonction de coût (dans notre cas, cette fonction est l'information mutuelle). L'algorithme de séparation calculera la contribution de chaque source sur chaque capteur en minimisant l'erreur quadratique entre les sources estimées et les observations. Ces contributions calculées par le filtre de Wiener permettront de simplifier le critère d'indépendance à minimiser, d'estimer au mieux la matrice de séparation et par la suite, les sources recherchées.

2.1.2 Mesure de l'indépendance statistique

Pour résoudre un problème de séparation de sources, on cherche à maximiser l'indépendance statistique entre les différentes composantes d'un vecteur (les composantes de ce vecteur sont les sources estimées).

On définit l'indépendance comme suit :

Définition : *La mesure d'indépendance est une fonction des variables aléatoires x telle que le maximum de cette fonction est atteint si et seulement si les composantes x_i du vecteur x sont indépendantes.*

Propriété : *D'après la définition précédente, l'opposé de la mesure de divergence entre la densité de probabilité conjointe $p_x(x)$ et le produit des densités de probabilité marginales $\prod_{i=1}^N p_{x_i}(x_i)$ est considéré comme une mesure d'indépendance selon l'équation*

suivante :

$$-D \left(p_x(x) \parallel \prod_{i=1}^N p_{x_i}(x_i) \right) \leq 0 \quad (2.1)$$

En conséquence, pour que la séparation des sources soit effectuée ($X = AS$ et $Y = BX$), l'indépendance statistique est maximale quand $B = A^{-1}$.

2.1.3 Mesure de la divergence entre les densités de probabilité

Définition : *Divergence de Kullback Leibler KL :*

soient p et q deux densités de probabilités la divergence KL entre p et q peut s'écrire de la façon suivante :

$$KL(p||q) = \int p(x) \log \frac{p(x)}{q(x)} dx \quad (2.2)$$

Propriétés :

- $KL(p||q) \geq 0$ et égale à 0 si et seulement si $p = q$
- $KL(p||q) = KL(\varphi(p)||\varphi(q))$ pour toute transformation linéaire inversible φ appliquée sur p et q

2.1.4 Mesure de l'information mutuelle

Comme nous l'avons introduit dans le chapitre 1, l'information mutuelle est définie par la divergence de Kullback Leibler entre la densité de probabilité conjointe $p_x(x)$ et le produit de densités de probabilité marginales $\prod_{i=1}^N p_{x_i}(x_i)$ de N variables aléatoires $x_1, x_2, \dots, x_N$:

$$I(x) = KL \left\{ p_x(x) \parallel \prod_{i=1}^N p_{x_i}(x_i) \right\} = \int p_x(x) \ln \frac{p_x(x)}{\prod_{i=1}^N p_{x_i}(x_i)} dx \quad (2.3)$$

avec $x = (x_1, x_2, \dots, x_N)^T$ un vecteur de N variables aléatoires. Cette équation de l'information mutuelle est toujours positive et égale à zéro si et seulement si toutes les variables aléatoires x_i sont indépendantes [CT91]. Les propriétés de la divergence de Kullback Leibler sont plus détaillées dans [KUL59, GRA90] en dehors du contexte de la séparation de sources.

L'information mutuelle est liée à l'entropie de Shannon par l'équation suivante :

$$I(x) = \sum_{i=1}^N H(x_i) - H(x) \quad (2.4)$$

où :

- $H(x_i)$ est l'entropie de Shannon marginale
- $H(x)$ est l'entropie de Shannon conjointe

Propriétés de l'information mutuelle :

- $I(x) \geq 0$ et égale à 0 si et seulement si les composantes du vecteur x sont mutuellement indépendantes
- L'information mutuelle est asymétrique : $KL(p_x, \prod_{i=1}^N p_{x_i}) \neq KL(\prod_{i=1}^N p_{x_i}, p_x)$
- $I(x) = I(\varphi(x))$ pour toute transformation linéaire inversible φ

Pour résoudre un problème de séparation de sources, nous avons utilisé l'information mutuelle comme un critère pour la mesure de l'indépendance entre les différentes sources estimées. Plusieurs méthodes ont été développées dans [ALM01, CAR98, PHA99, TJ97, TJ99] basées sur l'utilisation de l'information mutuelle. La maximisation de ce critère d'indépendance entre les sources permet d'estimer la matrice de mélange et par la suite l'obtention des sources originales.

2.1.5 Mesure de l'indépendance dans le cas de mélange convolutif

D'après les propriétés de l'information mutuelle, la mesure de l'indépendance entre les sources estimées permet de résoudre un problème de séparation de sources. Cette mesure d'indépendance statistique diffère selon le type de mélange obtenu (instantané, convolutif,...,etc). Dans ce paragraphe, nous allons présenter l'utilisation de l'information mutuelle pour séparer un mélange de sources convolutives.

On considère le modèle général constitué de deux sources et deux capteurs, soient $s_1(n)$ et

$s_2(n)$ les sources et $x_1(n)$ et $x_2(n)$ les observations mesurées au niveau de chaque capteur.

On peut écrire :

$$x_j(n) = \sum_{i=1}^2 a_{ji} * s_i(n) \quad (2.5)$$

où a_{ij} représente la réponse impulsionnelle de la source i ($i = 1, 2$) sur le capteur j ($j = 1, 2$), $*$ signifie que le type de mélange utilisé est convolutif. On cherche à estimer un filtre non causal b_{ij} entre la source i et le capteur j tel que les sources sont estimées à un filtre linéaire après. L'équation des sources estimées $y(n)$ s'écrit comme suit :

$$\begin{aligned} y_j(n) &= \sum_{i=1}^2 b_{ji} * x_i(n) \\ &= \sum_{i=1}^2 \sum_{k=-L}^L b_{ji}(k) x_i(n-k) \end{aligned} \quad (2.6)$$

Dans certains cas spécifiques, la séparation des sources est effectuée quand les sources estimées $y_1(n)$ et $y_2(n)$ deviennent indépendantes [COM96]. Par exemple, dans le cas des sources i.i.d normalisées, la somme des cumulants d'ordre quatre calculées au niveau des sources estimées est une fonction de contraste [COM94].

Dans le cas général d'un mélange convolutif, il faut tenir compte des temps de retard m qui doivent être introduits dans le calcul de la fonction de contraste. La séparation d'un mélange convolutif de sources est réalisée quand les différentes composantes du vecteur $y(n)$ des sources estimées sont indépendantes au sens des variables stochastiques [BZ02].

En d'autres termes, $y_1(n)$ et $y_2(n-m)$ doivent être indépendantes pour tout retard m .

Par exemple, l'utilisation de l'information mutuelle comme critère d'indépendance nécessite la minimisation du critère J qui s'écrit comme suit :

$$J = \sum_{m=-\infty}^{+\infty} I(y_1(n), y_2(n-m)) \quad (2.7)$$

où $m \in$ au support de dépendance des capteurs.

Ce critère d'indépendance est très complexe et nécessite un temps de calcul important.

Une méthode de minimisation de ce critère a été proposée par [BZ02], utilisant l'algorithme

du gradient. La mesure de l'indépendance a été faite pour un retard m choisi aléatoirement à chaque itération en utilisant toujours $I(y_1(n), y_2(n - m))$. En conséquence, après un certain nombre d'itérations, on suppose qu'on a pris en compte tous les retards m . Ce qui rend la convergence de l'algorithme plus rapide.

Ce critère d'indépendance (équation 2.7) doit être normalement minimisé pour tout retard m , I représente l'information mutuelle qui sera toujours positive et égale à zéro si et seulement si toutes les composantes du vecteur $y(n)$ sont statistiquement indépendantes.

$$I(y) = \int_R p_y(y) \ln \left(\frac{p_y(y)}{\prod_{i=1}^M p_{y_i}(y_i)} \right) \quad (2.8)$$

Les retards m sont considérés à priori entre $[-K, \dots, K]$, ce choix dépend du rang des filtres utilisés dans le mélange convolutif. Donc, la minimisation du critère est assez lourde en terme de temps de calcul.

Dans la suite de ce chapitre, nous allons présenter brièvement l'algorithme de la descente de gradient qui a été utilisé dans le développement de la méthode proposée, ensuite, nous allons développer l'aspect théorique de l'utilisation du filtre de Wiener afin de simplifier la mesure de l'indépendance pour tous les retards m . Nous allons montrer que le fait de projeter les sources estimées sur les capteurs permettront d'augmenter l'efficacité de la méthode utilisée et par la suite la séparation de sources.

2.1.6 Minimisation de l'information mutuelle

Après avoir présenté les différentes propriétés de l'information mutuelle I , cette quantité sera utilisée comme un critère d'indépendance pour résoudre un problème de séparation de sources. Le but est de minimiser l'information mutuelle des sources estimées. Pour répondre à cet objectif, nous allons utiliser l'algorithme de la descente de gradient. Selon l'équation (2.8), on remarque qu'il nécessite le calcul des densités de probabilités marginales $p_{y_i}(y_i)$ et conjointes $p_y(y)$. Pour résoudre ce problème et en utilisant l'algorithme de gradient, la matrice de séparation B doit être estimée de façon à ce que

l'information mutuelle des sorties $y = Bx$ soit minimisée.

$$B \leftarrow B - \mu \frac{\partial I(y)}{\partial B_{ij}} \quad (2.9)$$

D'après l'équation (2.39), l'utilisation de l'algorithme de la descente de gradient nécessite l'estimation de $\frac{\partial I(y)}{\partial B_{ij}}$. Si on considère $y = Bx$, la densité de probabilité de y s'écrit comme suit :

$$p_y(y) = \frac{p_x(x)}{|\det B|} \quad (2.10)$$

et le logarithme népérien de l'équation (2.10) est :

$$\ln p_y(y) = \ln p_x(x) - \ln |\det B| \quad (2.11)$$

En utilisant les équations de l'entropie de Shannon $H(\cdot) = -E \{ \ln p_{(\cdot)}(\cdot) \}$ et de l'information mutuelle $I(y) = \sum_i H(y_i) - H(y)$, on peut écrire :

$$I(y) = \sum_i H(y_i) - H(x) - \ln |\det B| \quad (2.12)$$

Le calcul du terme $\frac{\partial I(y)}{\partial B}$ nécessite l'estimation des densités de probabilité de y . Cette estimation utilise le calcul des fonctions scores. Le concept de ces fonctions scores sera détaillé dans l'Annexe A, ainsi que les propriétés et les méthodes d'estimation de ces fonctions.

2.2 Filtre de Wiener

Dans cette section, nous présentons un état de l'art sur le filtrage de Wiener. Nous décrivons brièvement la modélisation d'un filtre de Wiener. Ensuite, nous focalisons sur le descriptif de la méthode que nous proposons pour réaliser une séparation efficace d'un mélange convolutif de sources. Nous présentons en détail les principes de ce traitement basé sur le calcul de la contribution de chaque source sur chaque capteur afin de minimiser l'erreur quadratique entre les deux signaux que sont la source et le capteur.

2.2.1 Définition

La technique du filtrage de Wiener est un problème d'estimation où on dispose d'une connaissance a priori sur le signal à estimer. Typiquement, on utilise cette technique pour la soustraction de bruit qui permet d'estimer un signal noyé dans le bruit.

Un filtre de Wiener est conçu pour minimiser l'erreur quadratique moyenne entre sa sortie et une sortie désirée. Il est le filtre optimum au sens du critère de l'erreur quadratique moyenne et nous verrons que les coefficients du filtre sont liés à la fonction d'autocorrélation du signal d'entrée et à l'intercorrélacion entre les signaux d'entrée et de sortie désirée.

2.2.2 Modélisation du filtre de Wiener

Nous présentons un modèle de filtrage de Wiener utilisé pour optimiser la séparation d'un signal et d'un bruit de mesure. On considère les deux signaux stationnaires $z(t)$ et $y(t)$.

où :

- $\Gamma_{zz}(\tau) = E[z(t)z(t - \tau)]$: Fonction d'autocorrélation du signal $z(t)$.
- $\Gamma_{zy}(\tau) = E[z(t)y(t - \tau)]$: Fonction d'intercorrélacion entre $z(t)$ et $y(t)$.
- γ_{zz} : Densité spectrale de puissance de $z(t)$.
- γ_{zy} : Densité interspectrale de puissance de $z(t)$ et $y(t)$.

On cherche à estimer un filtre $W(z)$ dont la sortie, notée $\hat{z}(t)$ minimise l'erreur quadratique entre le signal $z(t)$ et sa prédiction $\hat{z}(t)$ à partir du signal filtré (cf fig. 2.1).

Donc, le problème qui se pose est comment retrouver $z(t)$ à partir de $y(t)$. Une solution consiste à filtrer $y(t)$ de telle sorte que la sortie $\hat{z}(t)$ soit la plus proche possible de $z(t)$. On peut mesurer la qualité de l'estimation par le signal d'erreur $\epsilon(t)$, défini par :

$$\epsilon(t) = z(t) - \hat{z}(t) \quad (2.13)$$

Plus le signal d'erreur $\epsilon(t)$ sera faible, plus l'estimation sera bonne. On cherche donc un filtre qui minimisera l'erreur. Il est pratique de chercher à minimiser $\epsilon^2(t)$ car c'est

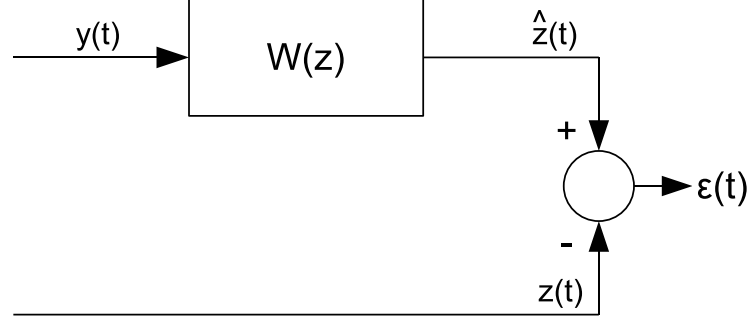


FIG. 2.1 – Shéma du filtre de Wiener

une fonction quadratique facilement dérivable. En revanche, étant donné que les signaux traités sont souvent aléatoires, la fonction coût qui sera à minimiser est l'erreur quadratique moyenne (MSE : Mean Square Error) définie par :

$$\xi(t) = E(\epsilon^2(t)) \quad (2.14)$$

Le filtre optimal de Wiener correspond au filtre qui minimisera la MSE (équation 2.22).

2.3 Méthode de séparation proposée

Dans cette section, nous développons une méthode de calcul pour la séparation d'un mélange convolutif de sources dans le domaine temporel. Cette méthode est basée sur la minimisation de l'information mutuelle et utilisant le filtre de Wiener. En général, dans le cas de mélange convolutif, le critère d'indépendance doit être minimisé pour les différents temps de retard.

La méthode proposée dans la suite de ce manuscrit permet la simplification du critère d'indépendance. Nous démontrons que le test d'indépendance entre les contributions de toutes les sources sur le même capteur au même instant t conduit à une séparation efficace des sources. Ces contributions seront obtenues par la projection des sources sur chaque capteur en utilisant le filtre de Wiener. Dans cette méthode, nous montrons que les contributions des sources calculées par un filtrage de Wiener rendent le critère d'indépendance utilisé plus simple et plus facile à minimiser en terme de temps de calcul.

2.3.1 Modélisation du problème

On considère un modèle de mélange convolutif de sources constitué de p sources et p capteurs. Nous commençons par définir les notations des signaux mis en jeu. Chaque source est un signal noté $s_i(n)$ et $S(n) = [s_1(n), s_2(n), \dots, s_p(n)]$ est le vecteur des sources. Nous considérons que le nombre de capteurs utilisés est égal au nombre de sources formant le mélange. Il y a donc p observations notées $x_i(n)$ ($i = 1, \dots, p$) et $X(n) = [x_1(n), x_2(n), \dots, x_p(n)]$ est appelé le vecteur d'observations.

Dans le cas de mélange convolutif linéaire, les observations sont des mélanges linéaires des signaux issus des sources par filtrage, $L \in]-\infty, +\infty[$. Le modèle de mélange est :

$$\begin{aligned} x_i(n) &= \sum_{j=1}^p a_{ij}(n) * s_j(n) \\ &= \sum_{j=1}^p \sum_{k=-L}^L a_{ij}(n) s_j(n-k) \end{aligned} \quad (2.15)$$

où $a_{ij}(t)$ est la réponse impulsionnelle de la source j sur le capteur i et $'*$ est le produit de convolution. L'inverse du filtre de mélange n'est pas forcément causal. En revanche, le but de la séparation de sources est d'estimer un filtre de séparation éventuellement non-causal avec une réponse impulsionnelle notée f_{ji} entre la source i et le capteur j .

Le modèle de mélange s'écrit comme suit :

$$x_i(n) = \sum_{j=1}^p A_{ij}(z) * s_j(n) \quad (2.16)$$

où $A_{ij}(z) = \sum_k a_{ij}(k) z^{-k}$ est la fonction de transfert de la source j sur le capteur i et $A_{ij}(z) * s_j(n)$ est le résultat de l'opération de filtrage :

$$A_{ij}(z) * s_j(n) = \sum_{k=-L}^L a_{ij}(k) s_j(n-k) \quad (2.17)$$

Cette équation (2.17) peut aussi être écrite sous forme matricielle de la façon suivante :

$$x(n) = [A(z)] s(n) = \sum_k A(k) s(n-k) \quad (2.18)$$

On note $Y(t) = [y_1(t), \dots, y_p(t)]$ le vecteur des sources estimées. Ces sources seront estimées à un filtrage près. On note $y(n)$ à temps discret :

$$\begin{aligned} y_i(n) &= \sum_{j=1}^p b_{ij} * x_j(n) \\ &= \sum_{j=1}^p \sum_{k=-L}^L b_{ij}(k) x_j(n-k) \end{aligned} \quad (2.19)$$

On appelle sources les p contributions sur le 1^{er} capteur. $x_1(n)$ s'écrit comme suit :

$$x_1(n) = s_1(n) + s_2(n) + \dots + s_p(n) \quad (2.20)$$

Le filtre global du mélange et de la séparation sera noté $g_{ij} = b_{ij} * a_{ij}$ (cf Fig. 2.2)

$$y_j(n) = \sum_{i=1}^p \sum_{k=-L}^L g_{ij}(k) s_i(n-k) \quad (2.21)$$

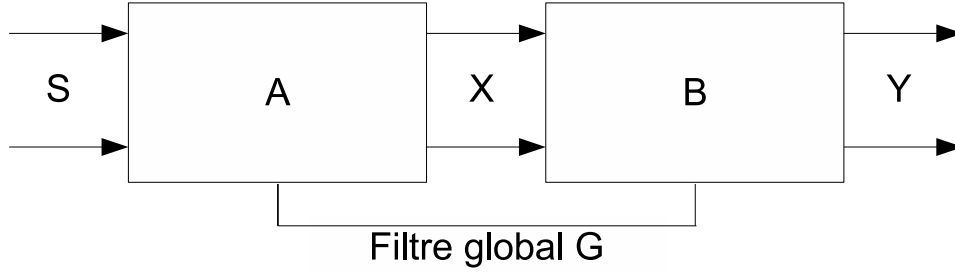


FIG. 2.2 – Système de mélange et de séparation

Si on considère que $y_j(n)$ est un filtrage linéaire d'une source et d'une seule, donc, la contribution de cette source sur le premier capteur est calculée par un filtre de Wiener non causal dont la fonction de transfert est notée $W_j(z)$ avec $j = 1, \dots, p$. Les coefficients de la réponse impulsionnelle du filtre seront calculées de sorte que l'erreur quadratique moyenne entre $x_1(n)$ et $y_j(n)$ soit minimisée. Cette erreur quadratique s'écrit comme suit :

$$E \{ |x_1(n) - w_{1j} * y_j(n)|^2 \} \quad (2.22)$$

Les p contributions des p sources sur le 1^{er} capteur s'écrivent de la façon suivante :

$$\begin{aligned} z_j(n) &= w_{1j} * y_j(n) \\ &= \sum_{k=-L}^L w_{1j}(k) y_j(n-k) \end{aligned} \quad (2.23)$$

En raison de simplification des équations, nous présentons les coefficients du filtre de Wiener utilisé sous forme fréquentielle qui sera tout simplement une transformée de Fourier de la version temporelle. Ces coefficients $w_j(k)$ sont le résultat d'une fonction constituée par la densité interspectrale entre la source $y_j(n)$ et l'observation $x_1(n)$ (le 1^{er} capteur) divisée par la densité spectrale de la même source $y_j(n)$. Une densité spectrale ou interspectrale sera représentée par γ .

Les coefficients du filtre de Wiener permettant de calculer les contributions des p sources sur le 1^{er} capteur s'écrivent comme suit :

$$\begin{aligned} W_1(f) &= \frac{\gamma_{Y_1 X_1}(f)}{\gamma_{Y_1}(f)} \\ W_2(f) &= \frac{\gamma_{Y_2 X_1}(f)}{\gamma_{Y_2}(f)} \\ &\vdots \\ W_p(f) &= \frac{\gamma_{Y_p X_1}(f)}{\gamma_{Y_p}(f)} \end{aligned} \quad (2.24)$$

2.3.2 Calcul des contributions des sources sur un capteur

Nous avons déjà présenté le critère d'indépendance basé sur la minimisation de l'information mutuelle (éq. 2.7). Dans le cadre de notre travail, la mesure de l'indépendance entre les p sources estimées $y_1(n), y_2(n), \dots$, et $y_p(n)$ est une condition suffisante pour résoudre un problème de séparation de sources.

Dans le cas d'un mélange convolutif de p sources, la séparation est effectuée quand les composantes du vecteur $y(n)$ sont indépendantes. En d'autre terme, c'est l'indépendance entre les $y_1(n), y_2(n-1), y_3(n-2), \dots, y_p(n-m_{p-1})$ pour tous les retards $m_1, m_2, \dots, m_{p-1}$.

Le problème se ramène à minimiser le critère J suivant :

$$J = \sum_{m_1} \sum_{m_2} \dots \sum_{m_{p-1}} I(y_1(n), y_2(n - m_1), y_3(n - m_2), \dots, y_p(n - m_{p-1})) \quad (2.25)$$

où I représente l'information mutuelle qui a été définie précédemment par l'équation (2.3). Notre méthode est basée sur l'étude de la séparabilité des contributions $z_1(n), z_2(n), \dots$, et $z_p(n)$ par rapport aux sources estimées $y_1(n), y_2(n), \dots$, et $y_p(n)$. Nous rappelons que les contributions $z_1(n), z_2(n), \dots, z_p(n)$ sont le résultat de la projection des sources estimées sur les capteurs en utilisant le filtre de Wiener (éq. 2.24).

Nous allons montrer que le fait d'insérer les contributions des sources sur les capteurs dans l'algorithme de séparation ne nécessite pas un test d'indépendance pour tout instant $(n-m)$ avec $m = 1, 2, \dots, p-1$. Pour que la séparation soit effectuée, $I(y_1(n), y_2(n), \dots, y_p(n)) = 0$ est une condition nécessaire mais qui n'est pas suffisante. Par conséquence, nous pouvons distinguer deux cas différents.

2.3.2.1 Premier cas

Le premier cas que nous pouvons distinguer est assez simple. Nous utilisons toujours la minimisation de l'information mutuelle comme critère d'indépendance. Cet algorithme de séparation sera plus détaillé dans la suite de ce manuscrit.

Dans le premier cas, si chaque source estimée $y_j(n)$ avec $j = 1, 2, \dots, p$ dépend d'une source et une seule, donc, toutes les sources estimées $(y_1(n), y_2(n), \dots, y_p(n))$ sont également indépendantes dans le sens des processus stochastiques, cela veut dire que le test d'indépendance a été vérifié pour tous les retards m (éq. 2.25). De même, pour les contributions calculées par le filtre de Wiener, $(z_1(n), z_2(n), \dots, z_p(n))$ sont aussi indépendantes. Dans ce cas, l'information mutuelle est égale zéro.

- $I(y_1(n), y_2(n), \dots, y_p(n)) = 0$.
- $I(z_1(n), z_2(n), \dots, z_p(n)) = 0$.

L'information mutuelle sera minimisée pour toutes les sources estimées et pour toutes les contributions des sources sur les capteurs. Dans ce cas, la séparation est effectuée.

2.3.2.2 Deuxième cas

Dans le deuxième cas, les sources estimées $y_j(n)$ peuvent être indépendantes ($I(y_1(n), y_2(n), \dots, y_p(n)) = 0$) au retard 0, mais la séparation n'est pas effectuée. Par exemple, dans le cas de sources i.i.d, les p sorties estimées (éq. 2.26) sont indépendantes à l'instant 0, mais, chaque sortie contient toujours un mélange des sources.

$$\begin{aligned}
 y_1(n) &= s_1(n) + s_2(n) + \dots + s_p(n) \\
 y_2(n) &= s_1(n-1) + s_2(n-1) + \dots + s_p(n-1) \\
 &\vdots \\
 y_p(n) &= s_1(n-(p-1)) + s_2(n-(p-1)) + \dots + s_p(n-(p-1))
 \end{aligned} \tag{2.26}$$

Dans cet exemple, nous remarquons que chaque sortie estimée est toujours composée d'un mélange de sources mais pour des retards différents. Donc, les sorties $y_j(n)$ sont indépendantes juste au retard 0. Dans ce cas, les composantes du vecteur $y(n)$ sont indépendantes, mais ce n'est pas une condition suffisante pour que les composantes $z_j(n) = W_i(z)y_j(z)$ après un filtrage de Wiener soient aussi indépendantes. Le rôle du filtre de Wiener est de maximiser la corrélation entre la contribution $z_1(n)$ et l'observation $x_1(n)$, respectivement $z_2(n)$ et $x_1(n)$, $\dots$, $z_p(n)$ et $x_1(n)$. Nous allons démontrer par un calcul théorique que les contributions $z_1(n), z_2(n), \dots, z_p(n)$ ne peuvent pas être indépendantes dans le cas étudié et par conséquent, $I(y_1(n), y_2(n), \dots, y_p(n))$ n'est pas égal à zéro. En conclusion, le problème de séparation des sources n'est pas résolu dans ce deuxième cas.

En conséquence, la seule solution pour mener une séparation efficace des sources est de minimiser à la fois $I(y_1(n), y_2(n), \dots, y_p(n))$ et $I(z_1(n), z_2(n), \dots, z_p(n))$. Il ne reste plus qu'à décrire le deuxième cas que nous venons de citer. Nous supposons que les sorties estimées $y_1(n), y_2(n), \dots, y_p(n)$ sont toujours constituées d'un mélange de sources (exemple des sources i.i.d cité ci-dessus) et que $I(y_1(n), y_2(n), \dots, y_p(n)) = 0$. Nous calculons les p contributions de ces sources sur le 1^{er} capteur en utilisant le filtre de Wiener décrit dans

la section 2.3.1. Ces contributions $z_j(n)$ sont constituées d'un mélange à la fois de sources.

$$\begin{aligned}
 Z_1(f) &= W_1(f)Y_1(f) \\
 Z_2(f) &= W_2(f)Y_2(f) \\
 &\vdots \\
 Z_p(f) &= W_p(f)Y_p(f)
 \end{aligned} \tag{2.27}$$

où $Z_i(f)$, ($i = 1, \dots, p$) est la transformée de Fourier discrète de $z_i(n)$.

Dans la section (2.3.1), on a déjà noté G le filtre global qui sera la convolution entre la matrice des filtres séparation B et la matrice des filtres de mélange A . les sorties estimées $y_i(n)$ s'écrivent de la façon suivante :

$$\begin{aligned}
 Y_1(f) &= G_{11}(f)S_1(f) + G_{12}(f)S_2(f) + \dots + G_{1p}(f)S_p(f) \\
 Y_2(f) &= G_{21}(f)S_1(f) + G_{22}(f)S_2(f) + \dots + G_{2p}(f)S_p(f) \\
 &\vdots \\
 Y_p(f) &= G_{p1}(f)S_1(f) + G_{p2}(f)S_2(f) + \dots + G_{pp}(f)S_p(f)
 \end{aligned} \tag{2.28}$$

Les contributions des sources sont calculées sur le premier capteur (observation $x_1(n)$).

$$X_1(f) = S_1(f) + S_2(f) + \dots + S_p(f) \tag{2.29}$$

En combinant les deux équations (2.24) et (2.28), la transformée de Fourier discrète des coefficients du filtre de Wiener s'écrit comme suit :

$$\begin{aligned}
 W_1(f) &= \frac{\gamma_{Y_1X_1}(f)}{\gamma_{Y_1}(f)} = \frac{\bar{G}_{11}(f)\gamma_{S_1}(f) + \bar{G}_{12}(f)\gamma_{S_2}(f) + \dots + \bar{G}_{1p}(f)\gamma_{S_p}(f)}{\gamma_{Y_1}(f)} \\
 W_2(f) &= \frac{\gamma_{Y_2X_1}(f)}{\gamma_{Y_2}(f)} = \frac{\bar{G}_{21}(f)\gamma_{S_1}(f) + \bar{G}_{22}(f)\gamma_{S_2}(f) + \dots + \bar{G}_{2p}(f)\gamma_{S_p}(f)}{\gamma_{Y_2}(f)} \\
 &\vdots \\
 W_p(f) &= \frac{\gamma_{Y_pX_1}(f)}{\gamma_{Y_p}(f)} = \frac{\bar{G}_{p1}(f)\gamma_{S_1}(f) + \bar{G}_{p2}(f)\gamma_{S_2}(f) + \dots + \bar{G}_{pp}(f)\gamma_{S_p}(f)}{\gamma_{Y_p}(f)}
 \end{aligned} \tag{2.30}$$

En remplaçant les coefficients de Wiener dans l'équation (2.27), on obtient les contributions

des p sources sur le premier capteur.

$$\begin{aligned}
Z_1(f) &= \frac{|G_{11}(f)|^2 \gamma_{S1}(f) + G_{11}(f) \bar{G}_{12}(f) \gamma_{S2}(f) + \dots + G_{11}(f) \bar{G}_{1p}(f) \gamma_{Sp}(f)}{\gamma_{Y1}(f)} S_1(f) \\
&\quad + \frac{G_{12}(f) \bar{G}_{11}(f) \gamma_{S1}(f) + |G_{12}(f)|^2 \gamma_{S2}(f) + \dots + G_{12}(f) \bar{G}_{1p}(f) \gamma_{Sp}(f)}{\gamma_{Y1}(f)} S_2(f) \\
&\quad \vdots \\
&\quad + \frac{G_{1p}(f) \bar{G}_{11}(f) \gamma_{S1}(f) + G_{1p}(f) \bar{G}_{12}(f) \gamma_{S2}(f) + \dots + |G_{1p}(f)|^2 \gamma_{Sp}(f)}{\gamma_{Y1}(f)} S_p(f) \\
Z_2(f) &= \frac{|G_{21}(f)|^2 \gamma_{S1}(f) + G_{21}(f) \bar{G}_{22}(f) \gamma_{S2}(f) + \dots + G_{21}(f) \bar{G}_{2p}(f) \gamma_{Sp}(f)}{\gamma_{Y1}(f)} S_1(f) \\
&\quad + \frac{G_{22}(f) \bar{G}_{21}(f) \gamma_{S1}(f) + |G_{22}(f)|^2 \gamma_{S2}(f) + \dots + G_{22}(f) \bar{G}_{2p}(f) \gamma_{Sp}(f)}{\gamma_{Y1}(f)} S_2(f) \\
&\quad \vdots \\
&\quad + \frac{G_{2p}(f) \bar{G}_{21}(f) \gamma_{S1}(f) + G_{2p}(f) \bar{G}_{22}(f) \gamma_{S2}(f) + \dots + |G_{2p}(f)|^2 \gamma_{Sp}(f)}{\gamma_{Y1}(f)} S_p(f) \\
&\quad \vdots \\
&\quad \vdots \\
Z_p(f) &= \frac{|G_{p1}(f)|^2 \gamma_{S1}(f) + G_{p1}(f) \bar{G}_{p2}(f) \gamma_{S2}(f) + \dots + G_{p1}(f) \bar{G}_{pp}(f) \gamma_{Sp}(f)}{\gamma_{Y1}(f)} S_1(f) \\
&\quad + \frac{G_{p2}(f) \bar{G}_{p1}(f) \gamma_{S1}(f) + |G_{p2}(f)|^2 \gamma_{S2}(f) + \dots + G_{p2}(f) \bar{G}_{pp}(f) \gamma_{Sp}(f)}{\gamma_{Y1}(f)} S_2(f) \\
&\quad \vdots \\
&\quad + \frac{G_{pp}(f) \bar{G}_{p1}(f) \gamma_{S1}(f) + G_{pp}(f) \bar{G}_{p2}(f) \gamma_{S2}(f) + \dots + |G_{pp}(f)|^2 \gamma_{Sp}(f)}{\gamma_{Y1}(f)} S_p(f)
\end{aligned} \tag{2.31}$$

On considère $u_{ij}(k)$ la fonction de transfert liant les sources $s_i(n)$ aux contributions $z_j(n)$ calculées par le filtre de Wiener.

$$u_{ij}(k) = [w_j * g_{ij}](k) \quad (2.32)$$

où $*$ représente une convolution linéaire. L'équation liant ces contributions aux sources $s_i(n)$ s'écrit comme suit :

$$z_j(n) = \sum_{i=1}^p \sum_{k=-L}^L u_{ij}(k) s_i(n-k) \quad (2.33)$$

D'après le calcul des contributions $z_j(n)$ (éq. 2.31), Nous pouvons déduire les expressions des coefficients $u_{ij}(k)$. Nous montrons d'après le calcul des coefficients que les contributions $z_1(n), z_2(n), \dots, z_p(n)$ ne sont pas indépendantes (le terme $I(z_1(n), z_2(n), \dots, z_p(n))$ est différent de zéro).

Il suffit de montrer que les coefficients $u_{11}(k), u_{12}(k), \dots$, et $u_{1p}(k)$ ne valent pas zéro pour le même temps de retard k . De même pour les coefficients $(u_{21}(k), u_{22}(k), \dots, u_{2p}(k)), (u_{31}(k), u_{32}(k), \dots, u_{3p}(k)), \dots$, et $(u_{p1}(k), u_{p2}(k), \dots, u_{pp}(k))$. Ci-dessus, l'expression des coefficients $u_{11}(k), u_{12}(k), \dots, u_{1p}(k)$ pour l'instant $k = 0$.

$$\begin{aligned} u_{11}(0) &= \sum_f^{N-1} \frac{|G_{11}(f)|^2 \gamma_{S1}(f) + G_{11}(f) \bar{G}_{12}(f) \gamma_{S2}(f) + \dots + G_{11}(f) \bar{G}_{1p}(f) \gamma_{Sp}(f)}{\gamma_{Y1}(f)} \\ u_{12}(0) &= \sum_f^{N-1} \frac{G_{12}(f) \bar{G}_{11}(f) \gamma_{S1}(f) + |G_{12}(f)|^2 \gamma_{S2}(f) + \dots + G_{12}(f) \bar{G}_{1p}(f) \gamma_{Sp}(f)}{\gamma_{Y1}(f)} \\ &\vdots \\ u_{1p}(0) &= \sum_f^{N-1} \frac{G_{1p}(f) \bar{G}_{11}(f) \gamma_{S1}(f) + G_{1p}(f) \bar{G}_{12}(f) \gamma_{S2}(f) + \dots + |G_{1p}(f)|^2 \gamma_{Sp}(f)}{\gamma_{Y1}(f)} \end{aligned} \quad (2.34)$$

Nous considérons le premier de terme $u_{11}(0)$. Pour montrer que ce terme est toujours supérieur ou égal à zéro, on cherche son expression au carré.

$$\begin{aligned}
|u_{11}(0)|^2 &= (r(f))^2 \\
&+ \left| \sum_f^{N-1} \frac{G_{11}(f)\bar{G}_{12}(f)\gamma_{S2}(f) + \dots + G_{11}(f)\bar{G}_{1p}(f)\gamma_{Sp}(f)}{\gamma_{Y1}(f)} \right|^2 \\
&+ r(f) \left(\sum_f^{N-1} \frac{\bar{G}_{11}(f)G_{12}(f)\gamma_{S2}(f) + \dots + \bar{G}_{11}(f)G_{1p}(f)\gamma_{Sp}(f)}{\gamma_{Y1}(f)} \right) \\
&+ r(f) \left(\sum_f^{N-1} \frac{G_{11}(f)\bar{G}_{12}(f)\gamma_{S2}(f) + \dots + G_{11}(f)\bar{G}_{1p}(f)\gamma_{Sp}(f)}{\gamma_{Y1}(f)} \right)
\end{aligned} \tag{2.35}$$

En posant $r(f) = \sum_f^{N-1} |G_{11}(f)|^2 \frac{\gamma_{S1}(f)}{\gamma_{Y1}(f)}$. Si la somme des deux derniers termes de l'équation précédente $r(f) \left(\sum_f^{N-1} \frac{\bar{G}_{11}(f)G_{12}(f)\gamma_{S2}(f) + \dots + \bar{G}_{11}(f)G_{1p}(f)\gamma_{Sp}(f)}{\gamma_{Y1}(f)} \right)$

et

$$r(f) \left(\sum_f^{N-1} \frac{G_{11}(f)\bar{G}_{12}(f)\gamma_{S2}(f) + \dots + G_{11}(f)\bar{G}_{1p}(f)\gamma_{Sp}(f)}{\gamma_{Y1}(f)} \right), \text{ (éq. 2.34)}$$

sont positifs ou nuls, alors, $u_{11}(0)$ est différent de zéro. En revanche, s'ils sont négatifs, $(u_{11}(0))^2$ est toujours supérieur ou égal à une valeur positive (éq. 2.37).

$$\begin{aligned}
|u_{11}(0)|^2 &\geq (r(f))^2 \\
&+ \left| \sum_f^{N-1} \frac{G_{11}(f)\bar{G}_{12}(f)\gamma_{S2}(f) + \dots + G_{11}(f)\bar{G}_{1p}(f)\gamma_{Sp}(f)}{\gamma_{Y1}(f)} \right|^2 \\
&- r(f) \left(\left| \sum_f^{N-1} \frac{\bar{G}_{11}(f)G_{12}(f)\gamma_{S2}(f) + \dots + \bar{G}_{11}(f)G_{1p}(f)\gamma_{Sp}(f)}{\gamma_{Y1}(f)} \right| \right) \\
&- r(f) \left(\left| \sum_f^{N-1} \frac{G_{11}(f)\bar{G}_{12}(f)\gamma_{S2}(f) + \dots + G_{11}(f)\bar{G}_{1p}(f)\gamma_{Sp}(f)}{\gamma_{Y1}(f)} \right| \right)
\end{aligned} \tag{2.36}$$

$$|u_{11}(0)|^2 \geq \left| r(f) - \sum_f^{N-1} \frac{G_{11}(f)\bar{G}_{12}(f)\gamma_{S2}(f) + \dots + G_{11}(f)\bar{G}_{1p}(f)\gamma_{Sp}(f)}{\gamma_{Y1}(f)} \right|^2 \tag{2.37}$$

Nous pouvons déduire que le coefficient $|u_{11}(0)|^2$ est toujours minoré par une valeur positive. Le même calcul peut être appliqué sur tous les coefficients $u_{ij}(0)$ avec $i = 1, \dots, p$ et $j = 1, \dots, p$.

D'après le calcul développé ci-dessus, nous pouvons conclure que le fait de projeter toutes les sources estimées sur les capteurs (dans notre cas, c'était le premier capteur) en utilisant un filtrage de Wiener, permet de montrer que l'information mutuelle calculée sur les contributions $z_j(n)$ est différente de zéro ($I(z_1(n), z_2(n), \dots, z_p(n)) \neq 0$). Donc, il reste le cas où les sources $(y_j(n))_{j=1, \dots, p}$ dépendent d'une source et d'une seule, en d'autre terme, le cas où toutes les sources $(y_j(n))$ sont indépendantes ($I(z_1(n), z_2(n), \dots, z_p(n)) = 0$), le problème de la séparation des sources sera réalisé.

En conséquence, nous pouvons conclure que la minimisation de l'information mutuelle des sources $y_j(n)$ et des contributions $z_j(n)$ sur les capteurs mène à une séparation efficace des signaux issus d'un mélange convolutif de sources. Cette méthode simplifie le calcul du critère d'indépendance à minimiser (2.7) et est basé sur l'estimation de l'information mutuelle au niveau des sources estimées.

2.4 Algorithme de séparation

Dans cette partie, nous appliquons l'approche de gradient pour la séparation d'un mélange convolutif de sources. L'utilisation de l'algorithme du gradient permet la minimisation du critère d'indépendance cité ci-dessus. On a noté B la matrice des filtres de séparation. Le vecteur des sources estimées s'écrit comme suit :

$$y(n) = B_0x(n) + B_1x(n-1) + \dots + B_px(n-p) \quad (2.38)$$

On appelle p le degré des filtres de séparation. Pour le calcul des composantes du vecteur $y(n)$, on cherche à estimer B_k avec $k = 1, \dots, p$. L'estimation des coefficients de B nécessite le calcul du gradient de l'information mutuelle utilisée comme critère d'indépendance. A chaque itération, on estime B_k et par la suite, les sources estimées $y(n)$. Une fois, les sources $y(n)$ calculées, on applique le filtre de Wiener pour calculer les contributions $z_j(n)$ de ces sources sur le capteurs $x_j(n)$.

2.4.1 Calcul des gradients

Comme nous l'avons montré dans la méthode citée dans la section (2.3.2), la séparation des sources convolutives ne nécessite pas la minimisation du critère d'indépendance $I(y_1(n), y_1(n), \dots, y_p(n))$ pour tous les retards m . L'insertion des contributions des sources sur le premier capteur dans l'algorithme de séparation simplifie notre critère d'indépendance et conduit à une séparation efficace des sources recherchées.

L'approche consiste à minimiser l'information mutuelle $I(y_1(n), y_1(n), \dots, y_p(n))$ par rapport à la matrice de séparation B . Cette information mutuelle est une mesure directe de l'indépendance des composantes du vecteur $y(n)$.

On considère le cas de deux sources et deux capteurs. On cherche à calculer le gradient de la fonction $I(y_1(n), y_2(n))$ par rapport à la matrice de séparation. On pourra minimiser directement l'information mutuelle à condition d'estimer les densités de probabilité conjointe et marginale de y . On appelle B la matrice de séparation : $y = B * x$, donc :

$$H(y) = H(x) + \log |\det B| \quad (2.39)$$

Donc, l'information mutuelle $I(y)$ s'écrit comme suit :

$$I(y) = \sum_{i=1}^N H(y_i) - H(x) - \log |det B| \quad (2.40)$$

L'expression du gradient de l'information mutuelle $\frac{\partial I(y)}{\partial B}$ calculé par rapport à la matrice de séparation B est :

$$\frac{\partial I(y)}{\partial B} = \sum_{i=1}^N \frac{\partial H(y_i)}{\partial B} - \frac{\partial}{\partial B} \log |det B| \quad (2.41)$$

En calculant le gradient de la somme par rapport à un coefficient b_{ij} de la matrice de séparation B , on déduit que la i ème source à estimer de la ligne seulement de la i ème ligne de la matrice B .

$$\begin{aligned} \frac{\partial H(y_i)}{\partial b_{ij}(t)} &= -E \left[\frac{\partial \log p_{y_i}(y_i(t))}{\partial b_{ij}(t)} \right] \\ &= E \left[\beta_{y_i}(y_i(t)) \frac{\partial y_i(t)}{\partial b_{ij}(t)} \right] \\ &= E [\beta_{y_i}(y_i(t)) x(t)] \end{aligned} \quad (2.42)$$

où β_{y_i} est la **fonction score** de la source y_i qui est définie dans l'annexe A.

Finalement, le gradient de $I(y)$ vaut :

$$\frac{\partial I(y)}{\partial B} = E [\beta_y(y(t)) x(t)^T] - \frac{\partial}{\partial B} \log |det B| \quad (2.43)$$

L'expression du critère de l'information mutuelle et de son gradient nécessite l'estimation des densités de probabilités marginales p_{y_i} et des fonctions scores β_{y_i} .

Dans le cas de deux sources et deux capteurs, on cherche à estimer les sources notées $y_1(n)$ et $y_2(n)$. Le gradient de $I(y_1(n), y_2(n))$ s'écrit comme suit :

$$\frac{\partial I(y_1(n), y_2(n))}{\partial B} = E [\beta_y(y(t)) x(t)^T] \quad (2.44)$$

Pour chaque couple de sources estimées $(y_1(n), y_2(n))$, on cherche à calculer leurs contributions sur chaque capteur x_i . Ces contributions sont obtenues par la projection de chaque source estimée $y_i(n)$ sur chaque capteur. Le calcul de ces contributions est obtenu en utilisant un filtrage de Wiener (voir section 2.3.1). L'avantage de l'utilisation du filtre de

Wiener pour le calcul des contributions de chaque source sur chaque capteur simplifie le critère d'indépendance utilisé, et ne nécessite pas le test de l'indépendance pour chaque retard m .

2.4.2 Simulations et résultats obtenus

Dans cette section, nous présentons les résultats obtenus utilisant la méthode de séparation basée sur la minimisation de l'information mutuelle et utilisant le filtrage de Wiener pour le calcul des contributions sources sur les capteurs. Pour montrer la performance de la méthode utilisée, nous allons comparer la minimisation de l'information mutuelle calculées à partir des sources estimées $y(n)$ et à partir des contributions $z(n)$ calculées au niveau des capteurs.

Nous considérons le cas d'un mélange convolutif de deux sources et deux capteurs. Dans un premier temps, nous utilisons l'algorithme suivant :

Initialisation : $y(n) = x(n)$

Itérations : Répéter jusqu'à la convergence

- *Estimation des fonctions scores entre $y_1(n)$ et $y_2(n)$: $\beta(y_1(n), y_2(n))$*
- $y \leftarrow y(n) - \mu\beta(y_1(n), y_2(n))$
- *Calcul des coefficients du filtre de Wiener $W_i(z)$*
- *Calcul des contributions des sources sur le premier capteur : $z_j(n) = W_j(n)y_j(n)$*
- *Remplacer $y(n)$ par $z(n)$*

La matrice de mélange A utilisée est la suivante :

$$A(z) = \begin{bmatrix} 1 + 0.2z^{-1} + 0.1z^{-2} & 0.5 + 0.3z^{-1} + 0.1z^{-2} \\ 0.5 + 0.3z^{-1} + 0.1z^{-2} & 1 + 0.2z^{-1} + 0.1z^{-2} \end{bmatrix}$$

L'algorithme de séparation présenté ci-dessus, est basé sur la minimisation de l'information mutuelle des sorties estimées. Afin de montrer l'efficacité de la méthode proposée, nous calculons le gradient de l'information mutuelle des contributions des sources sur les capteurs. Par exemple, dans le cas considéré de deux sources et deux capteurs, nous calculons quatre contributions qui correspondent aux deux sources du mélange. Ainsi, nous proposons l'algorithme suivant :

- Initialisation : $y(n) = x(n)$
 Itérations : Répéter jusqu'à la convergence
- *Estimation des fonctions scores entre $\beta(z_{11}(n), z_{12}(n))$ et $\beta(z_{21}(n), z_{22}(n))$*
 - $y \leftarrow y(n) - \mu [\beta(z_{21}(n), z_{22}(n)) + \beta(z_{11}(n), z_{12}(n))]$
 - *Calcul des coefficients du filtre de Wiener $W_i(z)$*
 - *Calcul des contributions des sources sur le premier capteur : $z_j(n) = W_j(n)y_j(n)$*
 - *Remplacer $y(n)$ par $[z_{11}(n), z_{12}(n)]$*

La figure (2.6) représente l'information mutuelle en fonction du nombre d'itérations. La courbe tracée avec des croix représente l'information mutuelle $I(z_1(n), z_2(n))$ entre les deux contributions $z_1(n)$ et $z_2(n)$. Ces deux contributions sont le résultat de la projection des sources sur le premier capteur en utilisant un filtrage de Wiener. Le calcul des fonctions scores notées $\beta(\cdot)$ utilise la méthode développée par D.-T. Pham dans [PHA02] utilisant des splines.

Sur la même figure, la deuxième courbe représente l'information mutuelle en rajoutant les contributions $z_{21}(n)$ et $z_{22}(n)$ sur le deuxième capteur. En comparant les deux courbes sur la figure (2.6), nous remarquons que quelque soit le nombre d'itérations, l'algorithme de séparation a une convergence plus rapide en insérant les contributions des sources sur le deuxième capteur dans la procédure de séparation. Ceci peut s'expliquer car on augmente le nombre de contraintes $I(z_{11}(n), z_{12}(n)) = 0$ et $I(z_{21}(n), z_{22}(n)) = 0$.

Ce résultat valide notre approche qui consiste à insérer les contributions des sources sur les capteurs dans la phase de séparation. Ces contributions permettent de minimiser l'erreur quadratique entre les sources estimées et les observations sur les capteurs. Ainsi, la figure (2.7) représente cette erreur quadratique. De la même façon que pour l'information mutuelle représentée sur la figure (2.6), nous retrouvons une meilleure minimisation de l'erreur quadratique en rajoutant les contributions des sources calculées au niveau des deux capteurs.

Les sources utilisées dans cette simulation sont composées de 1500 échantillons. Chaque source est la somme d'un bruit uniforme indépendant et identiquement distribué et d'une sinusoïde à paramètres déterministes. Les résultats sont moyennés sur 50 réalisations. Après 30 itérations, la valeur de l'information mutuelle vaut $2,6 \cdot 10^{-8}$. Les meilleurs résultats sont obtenus en estimant les fonctions scores par la méthode des splines [PHA02]. La valeur minimale du critère utilisé (information mutuelle) dépend des formes des densités de probabilités (ddp) et de la corrélation de chaque source. Par exemple, dans le cas de deux sources uniformes i.i.d, la valeur minimale est $2,7 \cdot 10^{-15}$. Ceci peut s'expliquer par le fait que l'estimation des fonctions scores n'est pas exacte et dépend de la forme des densités de probabilités. De plus, le critère d'indépendance est plus simple à minimiser car il contient moins de termes.

Nous avons comparé la méthode proposée avec celle proposée par M. Babaie-Zadeh [BZ02]. Dans ses travaux de recherche, l'auteur minimise le critère cité dans l'équation

(2.7), c'est à dire l'information mutuelle $I(y_1(n), y_2(n - m))$ en choisissant un retard m aléatoirement choisi à chaque itération.

Dans le cas de deux sources aléatoires uniformes, la méthode [BZ02] converge lentement (1000 à 1500 itérations). Les sources sont composées de 500 échantillons et les performances ont été moyennées sur 100 réalisations. Les résultats obtenus par M. Babaie-Zadeh sont représentés par les figures (2.4) et (2.5) en utilisant la même matrice de mélange $A(z)$. Les paramètres choisis sont $\mu = 0.3$ et $m \in [-6, 6]$. Dans ce cas, le meilleur taux d'interférence obtenu est 25 dB. Nous avons réalisé la même expérience et nous avons obtenu une convergence semblable aux figures (2.6) et (2.7). Le taux d'interférence (équation 2.45) à convergence (30 itérations) est de 36 dB.

Le taux d'interférence est mesuré par l'équation suivante :

$$T.I = 10 \log_{10} \frac{E \{y_i^2\}}{E \{y_i^2 | s_i=0\}} \quad (2.45)$$

Ces résultats valident l'intérêt de la méthode proposée (simplification du critère d'indépendance et convergence rapide). Les performances sont aussi meilleures pour un nombre d'itérations faible et donc un coût de calcul moindre.

En revanche, la qualité de séparation est moins bonne quand nous utilisons des filtres de mélange ayant des longues réponses impulsionnelles (quelques dizaines de coefficients). En outre, nous avons rencontré certains problèmes d'estimation des densités de probabilité qui sont susceptibles d'influencer l'efficacité de la méthode utilisée. Cette méthode nécessite l'optimisation de certains paramètres en fonction du nombre des coefficients du filtre de mélange et du calcul des densités de probabilités des sources, ce qui peut entraîner une non convergence de l'algorithme en raison de la possibilité d'existence des maximas locaux.

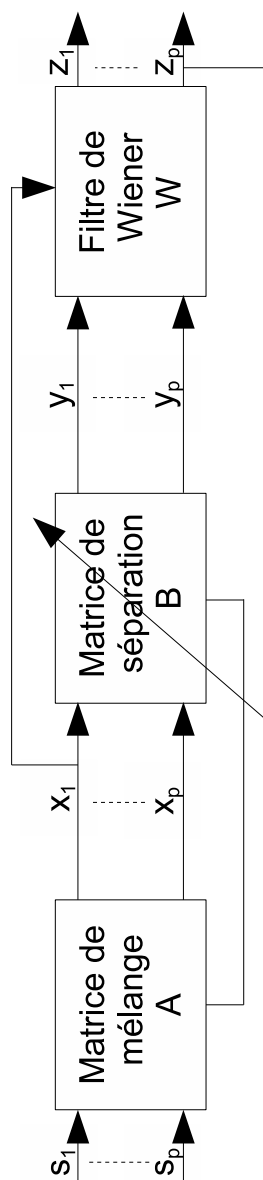


FIG. 2.3 – Modèle de séparation

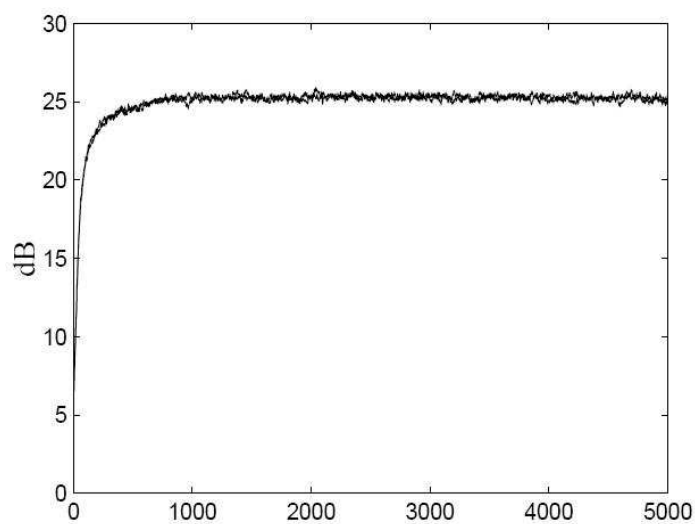


FIG. 2.4 – Approche du gradient : Taux d'interférence en fonction du nombre d'itérations

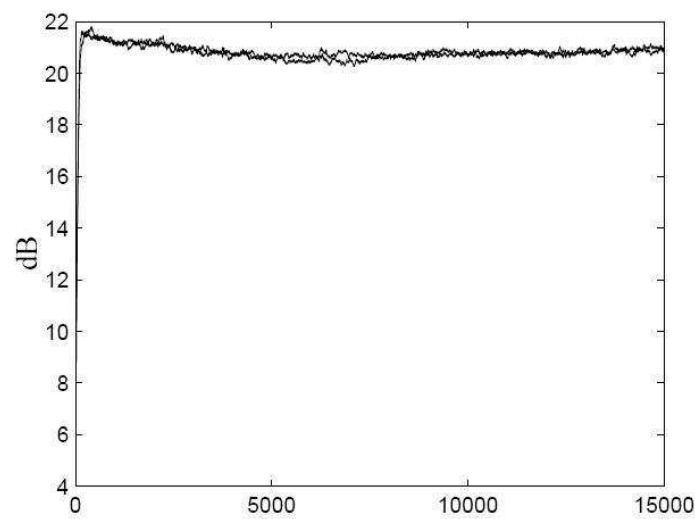


FIG. 2.5 – Approche de projection : Taux d'interférence en fonction du nombre d'itérations.

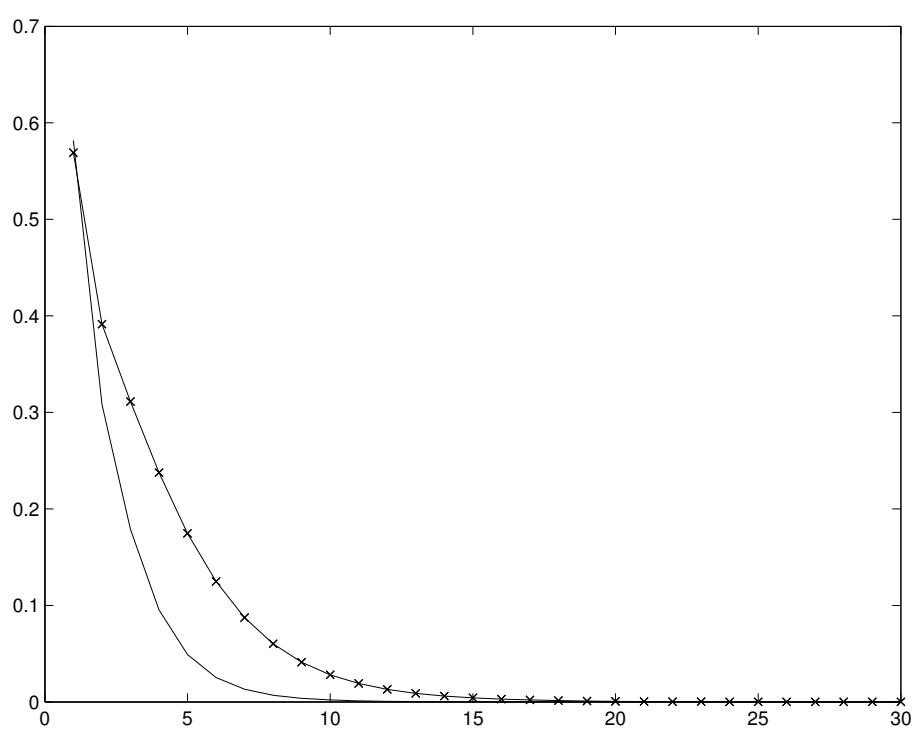


FIG. 2.6 – Tracé de l'information mutuelle en fonction du nombre d'itérations.

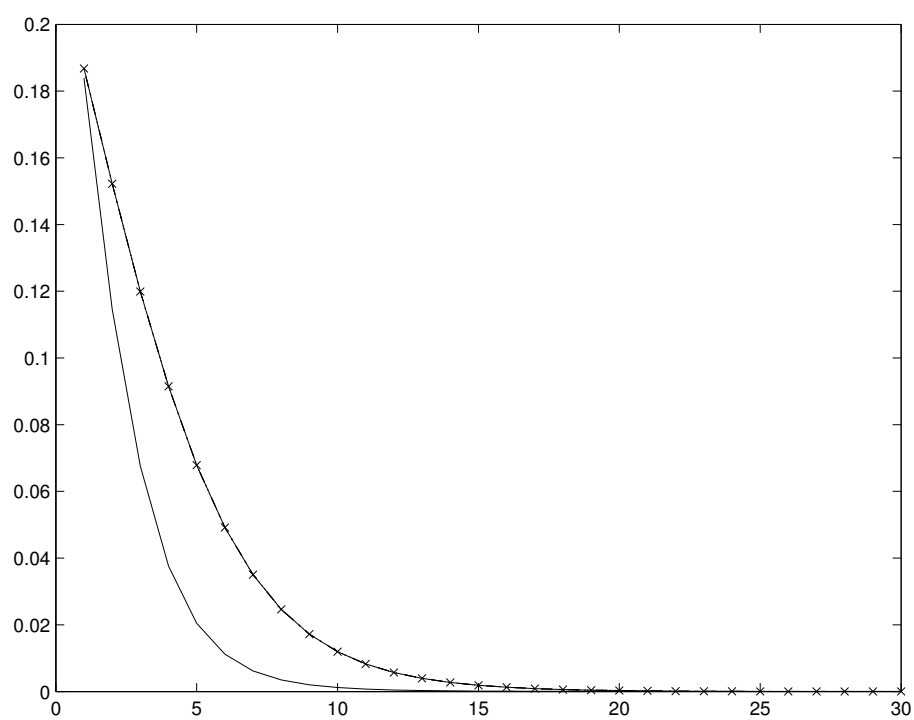


FIG. 2.7 – Erreur quadratique entre la contribution de la première source sur le premier capteur et son estimée en fonction du nombre d'itérations.

2.5 Conclusion

Dans ce chapitre, nous avons développé une méthode de séparation basée sur la minimisation d'information mutuelle. L'utilisation de ce critère d'indépendance pour résoudre un problème de séparation de sources nécessite l'estimation des densités de probabilité des sources (fonctions scores). Dans le cas d'un mélange convolutif de sources, la mesure de l'indépendance statistique doit être réalisée pour les différents retards m . Dans la méthode que nous avons présenté dans ce chapitre, le critère de mesure de l'indépendance entre les contributions des sources sur les capteurs est plus simple. Nous avons démontré théoriquement et par des simulations numériques que le fait de calculer les contributions des sources sur les capteurs permet d'augmenter l'efficacité de l'algorithme de séparation. Pour le calcul de ces contributions, nous avons utilisé le filtrage de Wiener basé sur la minimisation de l'erreur quadratique.

Par contre, l'estimation des densités de probabilité nécessite un choix des paramètres à régler parfois délicat. De plus, bien que la méthode fournisse de très bons résultats, elle est moins performante pour des filtres plus longs que quelques coefficients. Nous nous intéressons donc dans le prochain chapitre à l'utilisation des fonctions de contraste basées sur les statistiques d'ordre 4. Ceci nous emmène donc à proposer une méthode de séparation basée sur le calcul des cumulants d'ordre 4 et utilisant les contributions des sources sur les capteurs. Cette méthode sera développée dans le chapitre 3.

Chapitre 3

Statistiques d'ordre supérieur : Fonctions de contraste

Sommaire

3.1	Introduction sur les statistiques d'ordre 4	85
3.1.1	Définition	86
3.1.2	Algorithme de diagonalisation conjointe JADE	87
3.1.2.1	Modélisation du problème	88
3.1.2.2	Critère de séparation	88
3.1.3	Implantation et simulations	91
3.1.3.1	Mélange de deux signaux i.i.d à distribution uniforme	92
3.1.3.2	Mélange d'un signal sinusoïdal et d'un signal i.i.d à distribution uniforme	97
3.1.3.3	Performances de l'algorithme en fonction de la longueur L du filtre de mélange	100
3.2	Fonctions de contraste	104
3.2.1	Définitions et rappels	104
3.2.1.1	Déconvolution aveugle en séparation de sources	104
3.2.1.2	Fonctions de contraste en mélange instantané	105
3.2.1.3	Fonctions de contraste en mélange convolutif	106
3.2.2	Technique d'annulation des cumulants croisés	107
3.2.3	Fonction de contraste utilisant les contributions des sources sur les capteurs	108
3.2.4	Implantation et simulations	111

3.2.4.1	Performances de l'algorithme en fonction de la longueur L du filtre de mélange	114
3.2.5	Comparaison avec les résultats de la méthode développée dans le chapitre 2	118
3.3	Conclusion	119

Ce que nous appelons le hasard n'est que notre incapacité à comprendre un degré d'ordre supérieur.

Jean Guitton

DANS CE CHAPITRE, nous présentons les principales méthodes de séparation de sources basées sur les statistiques d'ordre 4. Dans le cas des signaux gaussiens, ces méthodes ne sont pas applicables vu que les statistiques d'ordre supérieur 2 n'apportent aucune information supplémentaire. Nous commençons par un bref descriptif sur les méthodes utilisant les notions des statistiques d'ordre 4. Ensuite, nous développons une méthode de séparation basée sur la diagonalisation conjointe d'une matrice de cumulants. Dans cette méthode, nous remplaçons la mesure de l'indépendance (information mutuelle dans le chapitre 2) par une mesure de diagonalisation d'un ensemble de matrices. Dans la deuxième partie, nous présentons une méthode basée sur les statistiques d'ordre 4 et le filtrage de Wiener. Cette technique utilisera la minimisation de deux critères basée sur l'annulation des cumulants croisés entre les contributions des sources estimées sur chaque capteur.

3.1 Introduction sur les statistiques d'ordre 4

Les statistiques d'ordre quatre sont suffisantes pour séparer les sources ([CAR89, COM89, SOU93, SC91]). Nous allons voir dans ce chapitre la suffisance de l'ordre 4 par rapport aux statistiques d'ordre 2, pour des mélanges de deux sources. L'utilisation des statistiques d'ordre 2 fait apparaître un problème lié aux signaux gaussiens. Ainsi, les cumulants d'ordres supérieurs d'un signal gaussien sont tous nuls deux à deux. Alors, la séparation de sources gaussiennes est impossible, sans hypothèse supplémentaire, sauf dans le cas où l'on a une source et une seule qui est gaussienne.

Parmi les premiers algorithmes utilisant les statistiques d'ordre quatre, on peut citer l'algorithme proposé par Jutten *et al.* [JTD⁺91] appliqué sur un mélange convolutif de

deux sources et deux capteurs, modélisé par deux filtres à réponse impulsionnelle finie. la procédure de séparation est une matrice de filtres dont les coefficients sont estimés par une généralisation du critère [JH88b] qui a été utilisé dans le cas de mélange instantané. La séparation est effectuée en annulant les moments croisés des sources estimées prises à plusieurs instants.

$$E(f(x_i(n))g(x_j(n_k))) \quad (3.1)$$

Cet algorithme peut être amélioré en optimisant le choix des fonctions f et g utilisées pour la séparation.

3.1.1 Définition

Dans cette section, nous présentons une méthode de séparation des sources basé sur l'algorithme JADE (**J**oint **A**pproximate **D**iagonalization of **E**igen matrices) proposé par Cardoso en 1993 [CS93]. L'algorithme JADE est un algorithme robuste à base de techniques statistiques pour l'application de séparation aveugle de sources. Cette méthode utilise un algorithme jacobien pour optimiser une fonction de contraste. La fonction qui a été utilisée dans l'algorithme JADE s'appuie sur les cumulants d'ordre 4. Dans cet algorithme, Cardoso cherche une matrice de séparation B en diagonalisant conjointement un ensemble de matrices de cumulants notées $F(M_i)$ en maximisant un critère de la forme suivante :

$$J_{JADE} = \sum_i |diag(BF(M_i)B^T)| \quad (3.2)$$

L'algorithme JADE se résume par les étapes suivantes :

- Initialisation : estimer la matrice de blanchiment, puis transformer les données (observations) par cette matrice.
- Calcul des statistiques : estimer les matrices de cumulants.
- Optimisation d'un contraste orthogonal : trouver une matrice de rotation telle que les matrices de cumulants calculées auparavant deviennent les plus diagonales possibles.
- Séparation : estimer les matrices de séparation par rotation des matrices de blanchiment.

L'algorithme JADE est très attractif (Annexe B.) car il est simple à implémenter et ne nécessite pas un choix de paramètres à optimiser. Les avantages de cette approche se résument par l'efficacité et la rapidité de convergence de cet algorithme.

Nous avons présenté brièvement l'algorithme JADE. Cette présentation fera l'objet de la méthode que nous développerons dans la suite de ce chapitre.

3.1.2 Algorithme de diagonalisation conjointe JADE

L'algorithme JADE est la première approche pour résoudre un problème de séparation de sources par diagonalisation conjointe d'un ensemble de matrices. Les matrices sont représentées par les cumulants d'ordre 4 des vecteurs de sortie. La méthode de diagonalisation conjointe est définie par le terme suivant :

$$D(H, \mathcal{M}) = \sum_{m=1}^M \left(\sum_{n=1}^N |M_{i,i}^H(m)|^2 \right) \quad (3.3)$$

où

$$M_{i,i}^H(m) = \sum_{n_1, n_2} H_{i, n_1} H_{i, n_1} M_{n_1, n_2}(m)$$

Le symbole $\mathcal{M}$ représente l'ensemble des matrices de dimension (N, N) . Une matrice diagonalisant conjointement l'ensemble $\mathcal{M}$ est définie par l'argument du maximum de $D(H, \mathcal{M})$ sur l'ensemble des matrices unitaires.

Si on compare l'algorithme JADE par diagonalisation conjointe à la fonction de contraste définie par Comon $\left(\mathcal{J}(x) = \sum_{i_1, i_2, i_3=1}^N (\text{cum}[y_{i_1}, y_{i_1}, y_{i_2}, y_{i_3}])^2 \right)$, on remarque qu'il existe un lien entre les deux critères. En effet, si on définit :

$$C(i_3, i_4) = C_{i_1, i_2}(i_3 i_4)$$

où

$$C_{i_1, i_2}(i_3 i_4) = \text{cum}(x_{i_1}, x_{i_3}, x_{i_3}, x_{i_4})$$

et en notant que C_4^2 l'ensemble de $C_{i_1, i_2}(i_3 i_4)$ pour i_3 et i_4 allant de 1 à N . Donc, si H est une matrice unitaire, on peut déduire l'équivalence suivante :

$$D(H, C_4^2) = \mathcal{J}(Hx) \quad (3.4)$$

3.1.2.1 Modélisation du problème

Dans cette section, nous développons une méthode de séparation des sources basé sur l'algorithme de JADE et utilisant les contributions des sources sur les capteurs calculées par le filtrage de Wiener [ASR07b]

Nous reprenons les notations utilisées dans le chapitre 2 pour la modélisation du problème. Nous considérons le cas d'un mélange convolutif constitué par p sources et p capteurs. L'équation du mélange se résume comme suit :

$$\begin{aligned} x_i(n) &= \sum_{j=1}^p a_{ij}(n) * s_j(n) \\ &= \sum_{j=1}^p \sum_{k=-L}^L a_{ij}(n) s_j(n-k) \end{aligned} \quad (3.5)$$

Nous avons noté $y(n)$ le vecteur des sources estimées :

$$\begin{aligned} y_i(n) &= \sum_{j=1}^p b_{ij} * x_j(n) \\ &= \sum_{j=1}^p \sum_{k=-L}^L b_{ij}(n) x_j(n-k) \end{aligned} \quad (3.6)$$

Les contributions des sources sur le premier capteur sont calculées par un filtrage de Wiener :

$$\begin{aligned} z_j(n) &= w_{ij} * y_i(n) \\ &= \sum_{k=-L}^L w_{ij}(k) y_i(n-k) \end{aligned} \quad (3.7)$$

3.1.2.2 Critère de séparation

Dans l'approche proposée, le critère de séparation utilise les statistiques d'ordre 2 et les statistiques d'ordre 4 à la fois calculées au niveau des observations $x(n)$. Nous avons noté B la matrice de séparation. Cette matrice B sera décomposé en deux matrices. Nous appelons V une matrice de blanchiment telle que $[V(z)]x(n)$ est un blanchiment spatial des observations et H la matrice de séparation.

On appelle $e(n)$ le vecteur des observations blanchies :

$$e(n) = [V(z)] x(n) \quad (3.8)$$

$y(n)$ est le vecteur des sources estimées :

$$\begin{aligned} y(n) &= [H(z)] e(n) \\ &= [H(z)] [V(z)] x(n) \\ &= [B(z)] x(n) \end{aligned} \quad (3.9)$$

Dans un premier temps, nous exploitons les statistiques d'ordre 2 afin de modéliser la matrice de blanchiement V .

Théorème :

Toute matrice $H(z)$ peut être décomposée en $R(z)U(z)$ telle que $U(z)$ est une matrice para-unitaire et $R(z)$ est une matrice triangulaire à phase minimale. Une matrice est dite para-unitaire si :

$$U(z)U^H(z) = I \quad (3.10)$$

Où I est une matrice identité. La matrice de corrélation calculée au niveau des observations $x(n)$ est exprimée par l'équation suivante :

$$R_x(k) = E \{x(n)x^T(n-k)\} \quad (3.11)$$

La décomposition en valeurs propres de la matrice $R_x(k)$ permet de calculer la matrice de blanchiement V telle que $e(z) = [V(z)]x(n)$ seront appelées les observations blanchies.

$$VAA^HV^H = I \quad (3.12)$$

On définit $U(z) = Q(z)A(z)$. En utilisant l'équation 3.12, on déduit que $U(z)$ est une matrice unitaire.

L'équation de la matrice de covariance des observations blanchies $e(n)$ s'écrit comme suit :

$$\begin{aligned}
 R_{ee}(z) &= E \{ e e^H \} \\
 &= V(z) R_{xx}(z) V^H(z) \\
 &= V(z) A(z) R_{ss}(z) A^H(z) V^H(z) \\
 &= V(z) A(z) A^H(z) V^H(z) \\
 &= U(z) U^H(z) \\
 &= I
 \end{aligned} \tag{3.13}$$

Après l'étape de blanchiement des observations, nous allons utiliser les statistiques d'ordre 4 pour l'estimation de la matrice de séparation H .

On définit la matrice de cumulants notée C_4^{yz} . Cette matrice s'écrit comme suit :

$$C_4^{yz} = \begin{bmatrix} \text{cum}(y_1^* z_1^* y_1 z_1) & \text{cum}(y_2^* z_1^* y_2 z_1) \\ \text{cum}(y_1^* z_2^* y_1 z_2) & \text{cum}(y_2^* z_2^* y_2 z_2) \end{bmatrix} \tag{3.14}$$

z_i représentent la projection des sources sur le premier capteur en utilisant un filtrage de Wiener. Ces contributions sont calculées à l'aide de l'équation suivante :

$$z_j(n) = \sum_{k=-L}^L w_{ij}(k) y_i(n-k) \tag{3.15}$$

Dans le chapitre 2, nous avons utilisé l'information mutuelle comme critère d'indépendance pour estimer la matrice de séparation. La méthode développée dans ce chapitre utilise la matrice de cumulants définie dans l'équation (3.14). Le critère d'indépendance est le suivant :

$$\text{Maximiser} : \sum_{i,j=1}^2 |\text{cum}(y_i^* z_j^* y_i z_j)|^2 \tag{3.16}$$

En remplaçant $y(n)$ par $[H(z)]e(n)$ dans l'équation (3.14), on obtient :

$$C_4^{yz} = \begin{bmatrix} h_1 cum(e_1^* z_1^* e_1 z_1) & h_2 cum(e_2^* z_1^* e_2 z_1) \\ h_1 cum(e_1^* z_2^* e_1 z_2) & h_2 cum(e_2^* z_2^* e_2 z_2) \end{bmatrix} \quad (3.17)$$

$H = [h_1, h_2]^T$ est la matrice de séparation où chaque composante h_i représente un vecteur de dimension $(1 \times n)$.

Pour estimer la matrice de séparation, nous allons maximiser la somme des carrés des éléments de la matrice C_4^{yz} (équation 3.17). Le critère à maximiser sera exprimé comme suit :

$$\sum_{i,j=1}^2 |diag(H.cum(e_i^* z_j^* e_i z_j).H^H)|^2 \quad (3.18)$$

où $|diag(.)|$ est la norme du vecteur diagonal de la matrice des cumulants.

Nous avons déjà démontré dans le chapitre 2 que la mesure d'indépendance entre les différentes contributions des sources sur les capteurs simplifie le critère de séparation utilisé et ne nécessite pas le test de l'indépendance pour tous les retards m dans le cas de mélange convolutif de sources. Pour estimer la matrice de séparation notée H , nous utilisons l'algorithme de diagonalisation conjointe développé par Cardoso.

3.1.3 Implantation et simulations

La technique de séparation de sources proposée utilise la maximisation du critère cité dans l'équation (3.18). Nous avons considéré un mélange convolutif composé de deux sources et deux capteurs. Comme nous l'avons cité précédemment, la première étape de la technique nécessite un blanchiment des observations, ensuite une diagonalisation conjointe des matrices de cumulants calculées à partir des sources estimées et des contributions de ces sources sur le premier capteur afin d'estimer la matrice de séparation.

Donc, le second critère utilisé est :

$$E(|x_1(n) - w_j * y_j(n)|^2)$$

La minimisation de ce critère permet d'obtenir la contribution de la source $y_j(n)$ sur le premier capteur. $(x_1(n) - w_j * y_j(n))$ représente l'erreur entre l'observation $x_1(n)$

sur le premier capteur et la source estimée $y_j(n)$. Maintenant, nous allons détailler le développement pratique de la technique utilisée. On note $A(z)$ la matrice des filtres utilisée pour réaliser un mélange convolutif de deux sources. Les fonctions de transfert des filtres constituant cette matrice sont les suivantes :

$$\begin{aligned} A_{11}(z) &= 1 + 0.3z^{-1} + 0.8z^{-2} \\ A_{12}(z) &= 0.9 + 0.2z^{-1} + 0.3z^{-2} \\ A_{21}(z) &= 0.6 + 0.4z^{-1} + 0.5z^{-2} \\ A_{22}(z) &= 1 + 0.3z^{-1} + 0.1z^{-2} \end{aligned} \tag{3.19}$$

3.1.3.1 Mélange de deux signaux i.i.d à distribution uniforme

Dans cette section, nous illustrons les résultats obtenus. On a considéré le cas de deux sources. Les deux sources $s_1(n)$ et $s_2(n)$ sont des signaux i.i.d à distribution uniforme et sont composées de 1000 points (voir figure 3.1). Les deux sources sont normalisées. Ces deux sources sont mélangées par une matrice de filtres $A(z)$ de dimension (2,2) (équation 3.19). Les filtres de mélange sont de longueur égale à 3. Ci-après, un récapitulatif de l'algorithme basé sur la diagonalisation conjointe des matrices de cumulants :

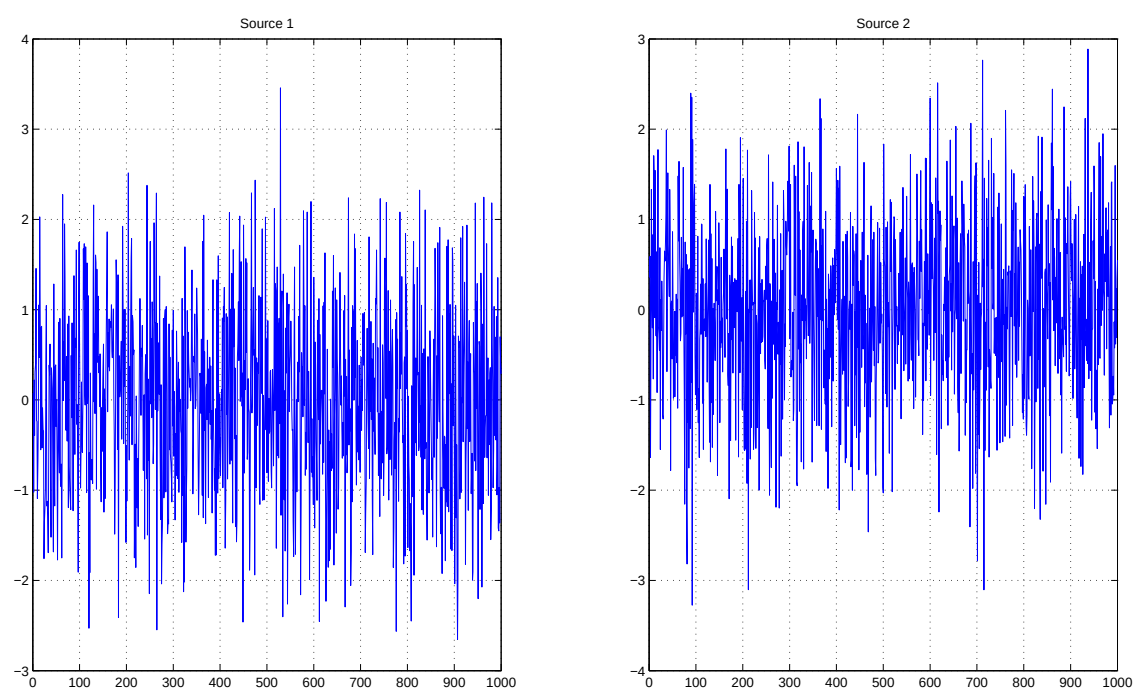


FIG. 3.1 – Schéma des sources utilisées

Initialisation : $y(n) = e(n)$

Itérations : Répéter jusqu'à la convergence

- *Calcul des coefficients du filtre de Wiener*
 $W_{ij}(n)$
- *Calcul des contributions des sources sur le premier capteur : $z_{ij}(n) = W_{ij}(n)y_j(n)$*
- *Calcul de la matrice de séparation notée H en utilisant l'approche de la diagonalisation conjointe selon le critère suivant :*

$$\sum_{i,j=1}^2 |diag(H.cum(e_i^*, z_j^*, e_i, z_j).H^H)|^2$$

Au lieu de travailler directement sur les observations $x_j(n), j = 1, 2$, on a ajouté une étape de préblanchiment ($y(n) = e(n)$ où $e(n)$ est une version blanchie des observations $x(n)$) basée sur l'utilisation des statistiques d'ordre 2. Après cette étape, les observations $x(n)$ seront remplacées par leurs versions normalisées $e_j(n), j = 1, 2$. Le vecteur des sources estimées sera dans ce cas $y(n) = [H(z)]e(n)$ et la matrice de séparation globale notée $B(z)$ vaut :

$$B(z) = H(z)V(z)$$

où $V(z)$ est la matrice de blanchiment appliquée sur les observations.

Ensuite, nous projetons les sources estimées $y(n)$ sur les capteurs en utilisant un filtrage de Wiener. Cette projection permet de retrouver la contribution de chaque source sur chaque capteur (dans notre cas, nous avons calculé la contribution de chaque source sur le premier capteur $x_1(n)$). Cette étape simplifie le test de la mesure d'indépendance entre les différentes sources dans le cas de mélange convolutif.

Nous représentons dans la figure (3.2) la distribution des sources et des observations. Les observations sont obtenu à partir du mélange des deux sources $s_1(n)$ et $s_2(n)$ en utilisant l'équation (3.19) En appliquant la méthode décrite précédemment, nous avons obtenus des

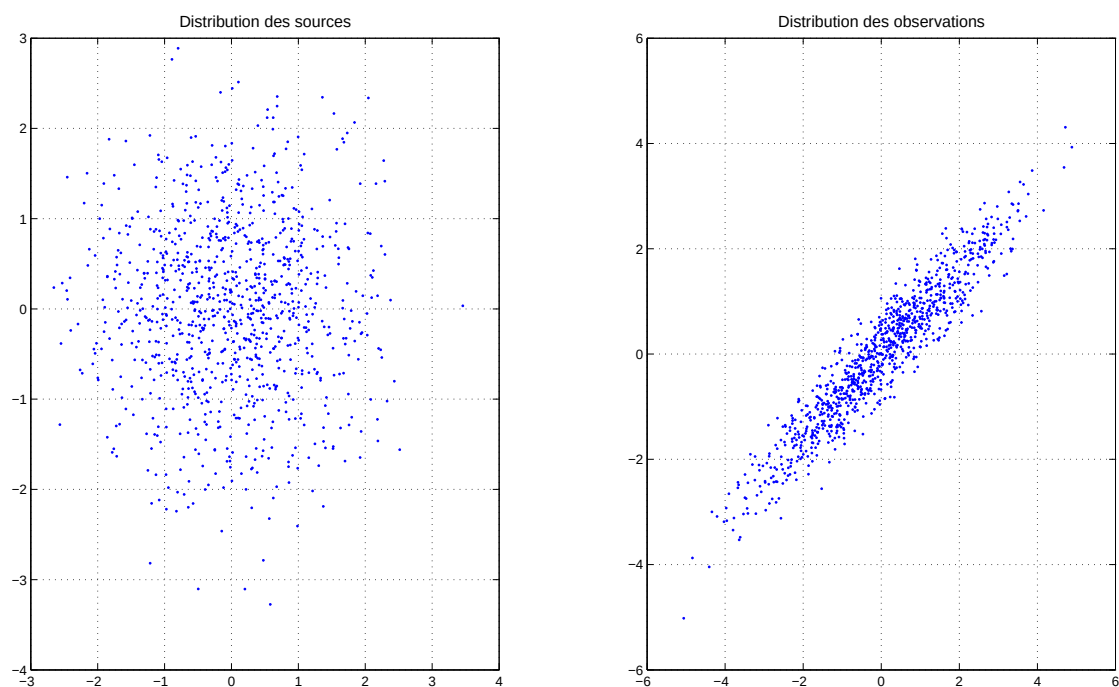


FIG. 3.2 – Distribution des sources et des observations

résultats plutôt positifs. La figure (3.3) représente les sources originales et les estimées.

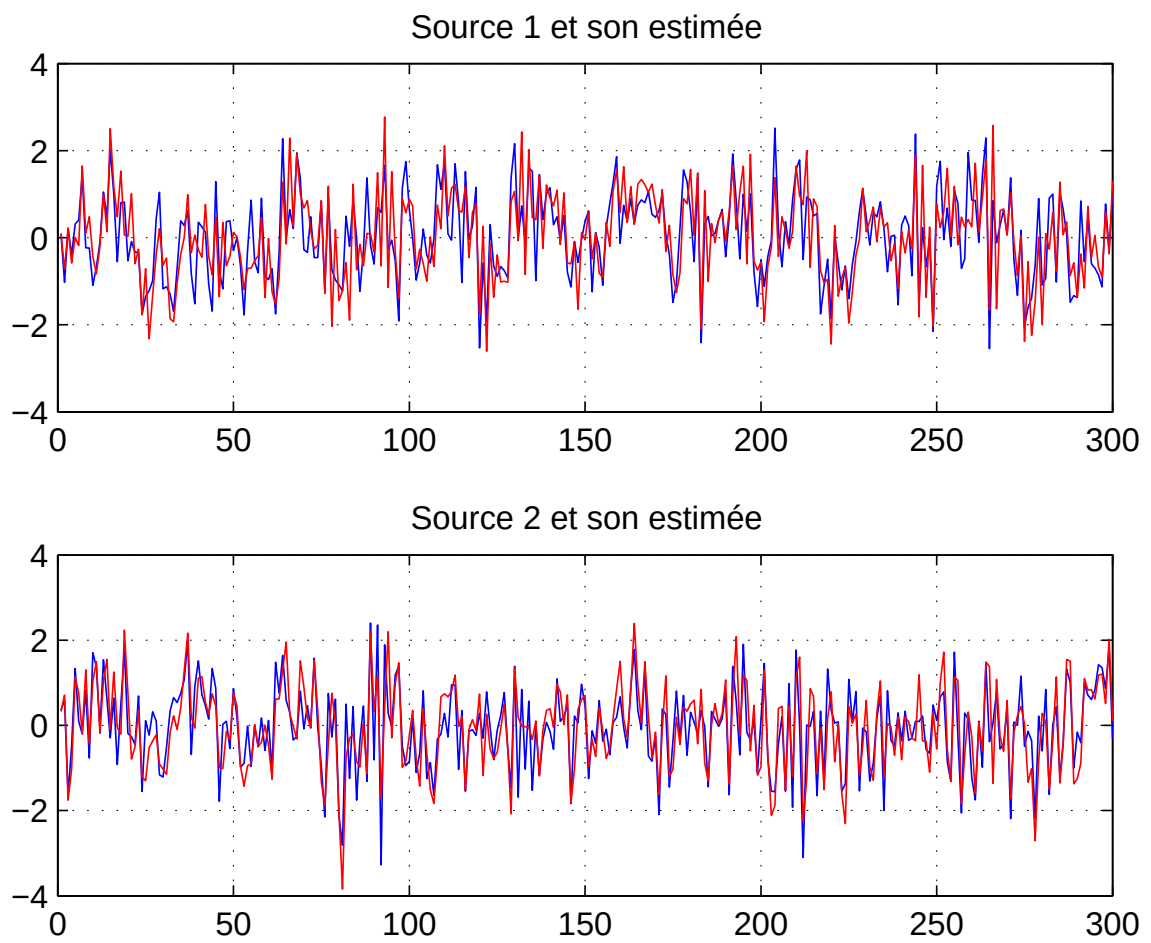


FIG. 3.3 – Combinaisons des sources originales et des estimées

Pour évaluer les performances de la méthode utilisée, on utilise la *fonction de cohérence* proposée par Welch [WEL67]. Cette fonction permet de comparer la source s à son estimée y pour chaque fréquence f . elle vaut 1 quand les signaux sont cohérents et 0 dans le cas contraire.

La cohérence sera calculée pour chaque source estimée $y_i(n)$. De cette manière, elle peut nous renseigner sur la bonne qualité des sources séparées quand elle est proche de la valeur 1. En effet, la cohérence entre deux signaux $s(n)$ et $y(n)$ s'écrit comme suit :

$$C_{sy}(f) = \frac{|P_{sy}(f)|^2}{P_{ss}(f)P_{yy}(f)} \quad (3.20)$$

- P_{ss} et P_{yy} représentent la densité spectrale de puissance de s et de y
- P_{sy} représente la densité spectrale de puissance entre s et y

Nous montrons sur les figures (3.5) et (3.6) la cohérence entre les sources originales et leurs estimées. Nous constatons que l'utilisation des contributions des sources sur les capteurs améliore la qualité de séparation. Ceci illustre l'intérêt du développement de la première partie de ce chapitre.

3.1.3.2 Mélange d'un signal sinusoïdal et d'un signal i.i.d à distribution uniforme

Dans ce paragraphe, nous considérons le cas d'un mélange convolutif d'un signal sinusoïdal et d'un signal i.i.d à distribution uniforme. Les signaux du mélange sont toujours normalisés. Nous avons utilisé la même matrice de mélange de l'équation (3.19).

Nous appliquons le même algorithme de séparation tout en précédant la procédure de séparation par une étape de blanchiment des observations $x(n)$. Nous représentons sur la figure 3.4 les sources originales en bleu et les sources estimées en rouge. Nous remarquons que les performances sont aussi meilleures dans ce cas de mélange, ce qui valide l'intérêt de l'approche proposée. On remarque quelques erreurs d'estimation dans les hautes fréquences de l'estimé du signal sinusoïdal.

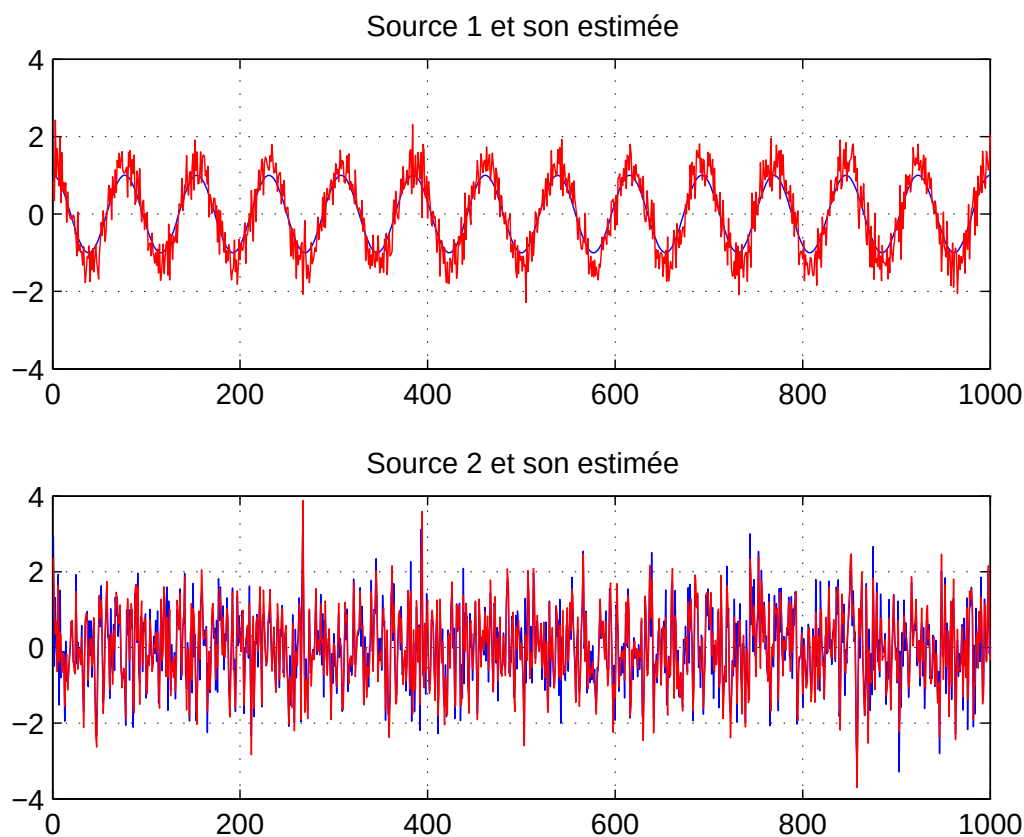


FIG. 3.4 – Combinaisons des sources originales et des estimées

Toutefois, comme on vient de le voir selon les résultats obtenus, l'insertion des contributions des sources sur l'algorithme dans la procédure de séparation apporte effectivement une amélioration des résultats de séparation. Il faut rappeler que l'utilisation de ces contributions permet de simplifier le test de l'indépendance entre les sources dans le cas de mélange convolutif, donc, la charge de calcul est considérablement allégée et on obtient une séparation efficace des sources. Cet allègement se traduit par une accélération remarquable du processus de séparation.

Compte tenu des résultats obtenus, il est intéressant de généraliser une fonction de contraste utilisant les contributions des sources sur les capteurs. Cette approche basée sur la méthode des contrastes sera étudiée et expliquée dans la deuxième partie de ce chapitre.

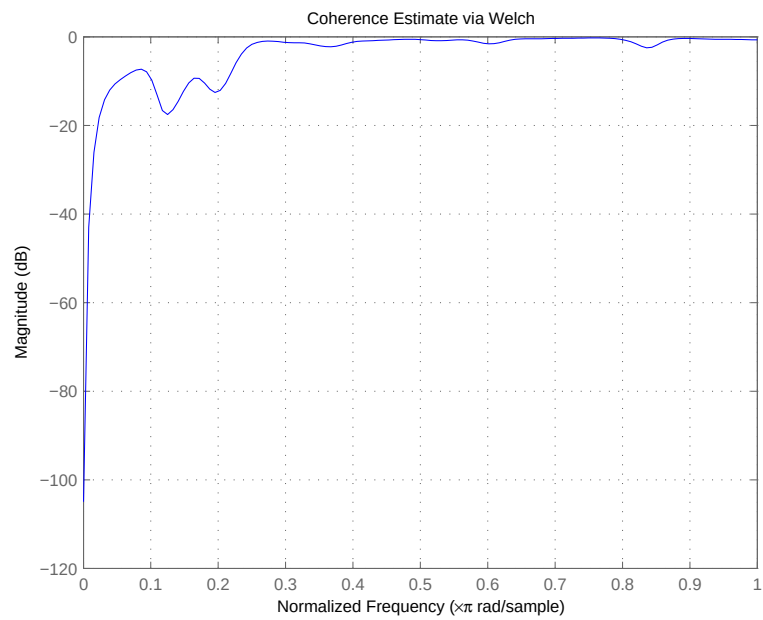


FIG. 3.5 – Mesure de la cohérence entre la source 1 et son estimée

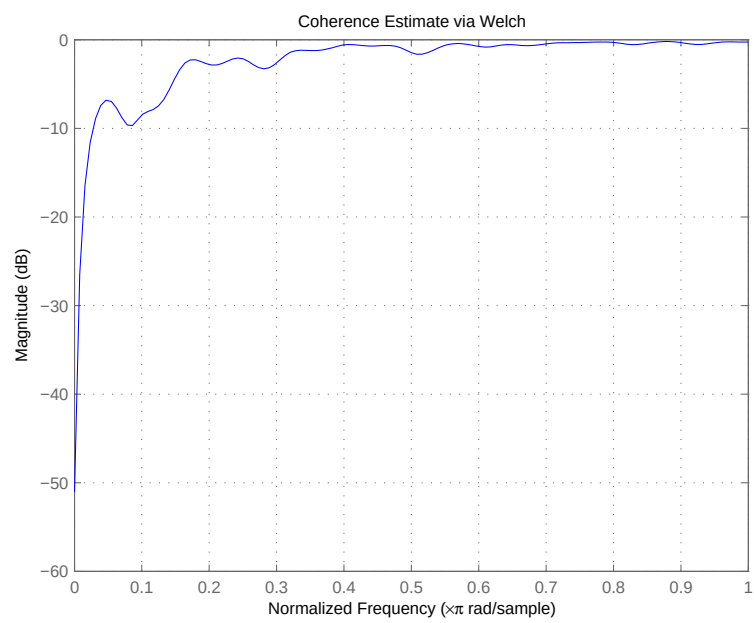


FIG. 3.6 – Mesure de la cohérence entre la source 2 et son estimée

3.1.3.3 Performances de l'algorithme en fonction de la longueur L du filtre de mélange

Dans cette section, nous allons évaluer l'efficacité de l'approche proposée en fonction de la taille du filtre de mélange. Pour effectuer cette manipulation, nous considérons le filtre $f(t)$ (figure 3.7) ayant la fonction de transfert suivante :

$$f(t) = e^{-\lambda t} \quad (3.21)$$

où $\lambda = 1 - \frac{1}{L}$ et L représente une longueur équivalente du filtre $f(t)$. Cette longueur est obtenue en calculant la tangente à l'origine de la fonction $f(t)$.

Nous considérons le cas de deux sources $s_1(t)$ et $s_2(t)$ et deux observations $x_1(t)$ et $x_2(t)$.

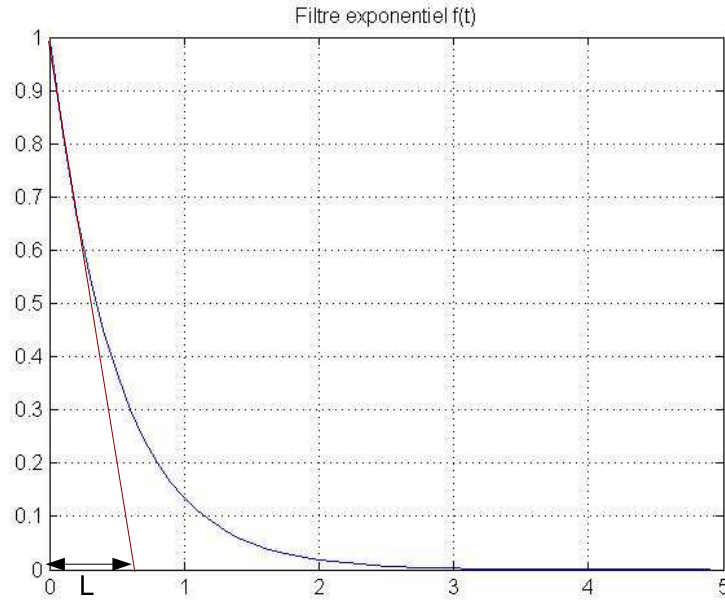


FIG. 3.7 – Tracé du filtre de mélange $f(t)$

Nous utilisons la fonction $f(t)$ comme un filtre de mélange. Ainsi, le problème sera modélisé par l'équation suivante :

$$\begin{aligned} x_1(t) &= f(t) * s_1(t) + s_2(t) \\ x_2(t) &= s_1(t) + f(t) * s_2(t) \end{aligned} \quad (3.22)$$

où $*$ représente le produit de convolution.

Pour montrer les performances de la méthode proposée, nous allons calculer l'erreur quadratique moyenne (EQM, équation 3.23) et le rapport signal sur bruit (RSB, équation 3.24) entre les sources originales $s_i(t)$ et les sources estimées $y_i(t)$.

$$EQM = E \{ (y(t) - s(t))^2 \} \quad (3.23)$$

$$RSB = 20 \log \left| \frac{s(t)}{y(t) - s(t)} \right| \quad (3.24)$$

Nous distinguons deux cas différents en fonction du type des signaux utilisés :

- **Cas de deux sources i.i.d. uniformes :**

Dans ce cas, les sources utilisées sont des sources i.i.d. normalisées. Le tableau 3.1 donne la valeur de l'EQM et du RSB (en dB) en fonction de la longueur du filtre L . Ces valeurs sont représentées sur les figures (3.8) et (3.9).

L	10	20	30	50	100
EQM	$6, 1.10^{-7}$	$4, 2.10^{-6}$	$3, 5.10^{-5}$	$8, 1.10^{-5}$	$9, 2.10^{-4}$
RSB (en dB)	23,5	20,8	19,1	18,5	17,9

TAB. 3.1 – EQM et RSB en fonction de la longueur L du filtre de mélange

- **Cas d'une source sinusoïdale et d'une source i.i.d uniforme :**

Dans ce cas, nous considérons un mélange de sources composées d'un signal sinusoïdal et d'un signal i.i.d. à distribution uniforme. De la même façon, nous calculons l'EQM (figure 3.10) et le RSB (figure 3.11) en fonction de la longueur L (tableau 3.2).

L	10	20	30	50	100
EQM	$3, 8.10^{-7}$	$5, 1.10^{-6}$	$3, 8.10^{-5}$	$0, 8.10^{-5}$	$7, 3.10^{-4}$
RSB (en dB)	22,3	20,2	18,9	18,3	17,1

TAB. 3.2 – EQM et RSB en fonction de la longueur L du filtre de mélange

D'une façon générale, nous avons obtenu des résultats avantageux. Nous avons présenté l'influence de la largeur du filtre de mélange sur la qualité de la séparation tout en obtenant des performances convenables, ce qui affirme la validité de l'approche développée.

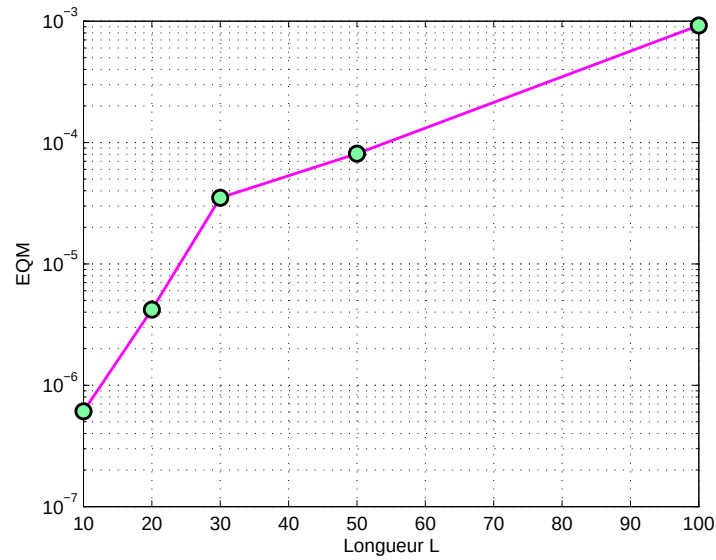


FIG. 3.8 – EQM en fonction de la longueur L du filtre de mélange

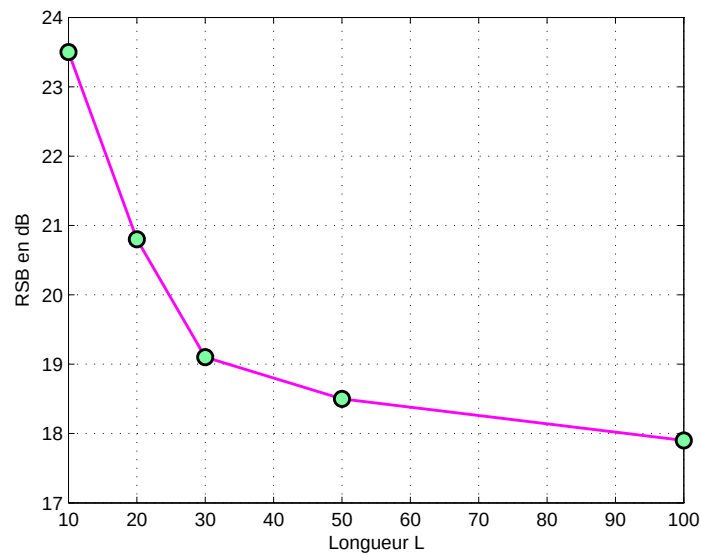


FIG. 3.9 – RSB en fonction de la longueur L du filtre de mélange

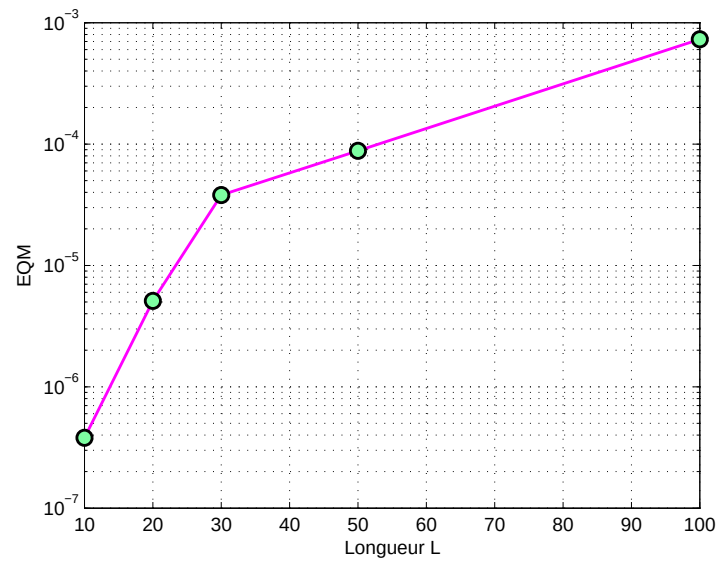


FIG. 3.10 – EQM en fonction de la longueur L du filtre de mélange

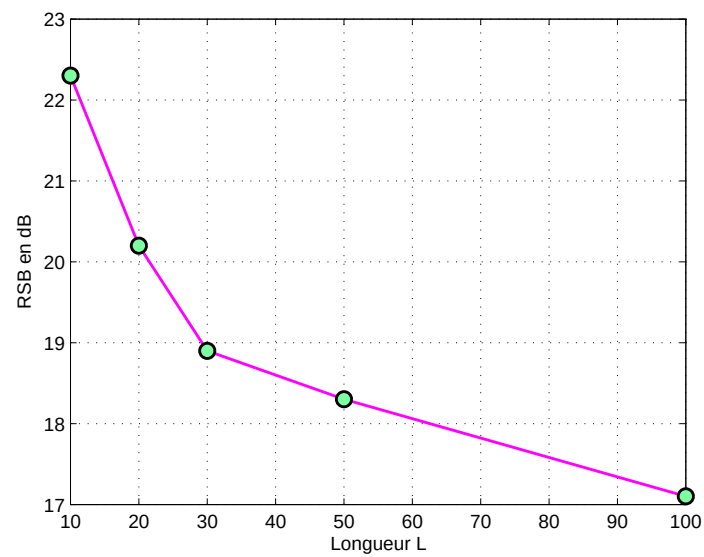


FIG. 3.11 – RSB en fonction de la longueur L du filtre de mélange

3.2 Fonctions de contraste

Dans un problème de séparation de sources, les fonctions de contraste permettent de caractériser l'indépendance statistique entre des variables aléatoires. Ces fonctions sont maximales si et seulement si elles correspondent à des variables aléatoires mutuellement indépendantes.

Dans la suite de ce chapitre, nous allons présenter un rappel sur l'utilisation des fonctions de contraste dans le cadre de la séparation de sources. Dans un premier temps, l'approche des fonctions de contraste a été utilisée dans le cadre de la déconvolution aveugle, ensuite, elle a été utilisée pour la séparation des sources en mélange instantané, puis généralisée pour le cas des mélanges convolutifs de sources. En général, dans la plupart des travaux qui ont été réalisés, les signaux utilisés sont indépendants et identiquement distribués (i.i.d). Ensuite, nous présentons deux critères d'indépendance (fonctions de contraste) basés sur l'annulation des cumulants croisés d'ordre quatre et utilisant le filtre de Wiener pour la projection des sources estimées sur les capteurs afin de retrouver leurs contributions sur ces capteurs. Nous illustrons aussi l'avantage de la méthode proposée par des résultats obtenus sur des signaux simulés.

3.2.1 Définitions et rappels

3.2.1.1 Déconvolution aveugle en séparation de sources

La déconvolution aveugle permet d'estimer le signal $s(n)$ en entrée d'un système dont sa réponse impulsionnelle $f(n)$ inconnue à partir de l'observation du signal en sortie $x(n)$ définie par :

$$x(n) = \sum_k s(n)f(n-k) \quad (3.25)$$

En général, le signal $s(n)$ est i.i.d et non gaussien. Le problème consiste à trouver un filtre $g(n)$ représentant l'inverse du filtre $f(n)$. Les premiers travaux ont été initiés par Donoho en 1981 [DON81b]. L'idée est de construire des fonctions statistiques appelées **fonctions de contraste**, qui sont maximales si et seulement si elles sont appliquées à des variables

aléatoires indépendantes. Cette idée a le même principe du théorème de la limite centrale.

Donoho définit plusieurs fonctions de contraste. Parmi ces fonctions, on peut citer :

- Le contraste à base de cumulant d'ordre m :

$$|\mathcal{C}(s)| = \frac{|C_m(s)|}{|E(s^2)|}$$

où $C_m(s)$ est le cumulant d'ordre m du signal s et $\mathcal{C}(s)$ représente le contraste sur l'ensemble des variables aléatoires de cumulant d'ordre m .

- L'entropie de Shannon notée $E(s)$:

$$-E(s) = \int \ln p_s(x) p_s(x) dx$$

où $p_s(x)$ la densité de probabilité de s .

3.2.1.2 Fonctions de contraste en mélange instantané

Si on considère un mélange instantané de M sources et M capteurs :

$$x(n) = As(n)$$

où $s(n) = (s_1(n), s_2(n), \dots, s_M(n))^T$ représente le vecteur des sources, A est la matrice du mélange supposée inversible et $x(n) = (x_1(n), x_2(n), \dots, x_M(n))^T$ le vecteur des observations au niveau des capteurs. L'objectif nécessite une matrice B appelée matrice de séparation telle que :

$$y(n) = Bx(n)$$

vaut $s(n)$ à des indéterminations près. Pour réaliser la séparation, Comon a proposé de maximiser une fonction de contraste dont sa définition est comme suit :

Fonction de contraste : Si on suppose $\mathcal{S}$ un ensemble de variables aléatoires dont les composantes sont non gaussiennes. On appelle Φ une fonction de contraste si elle vérifie les propriétés suivantes :

- Pour tout $s \in \mathcal{S}$, $\Phi(s)$ ne dépend que la probabilité de s .
- Pour tout $s \in \mathcal{S}$, et pour toute matrice A pour laquelle $As \in \mathcal{S}$, on a :

$$\Phi(Hs) \leq \Phi(s)$$

Parmi les fonctions de contraste qui ont été proposées, on peut citer les suivantes :

$$\Phi(s) = \sum_{k=1}^M |c_4(s_k)|$$

où s représente un vecteur de variables aléatoires.

$$-I(s) = - \int p_s(x) \ln \frac{p_s(x)}{\prod_{k=1, \dots, M} p_{s_k}(x_k)}$$

où $I(s)$ désigne l'information mutuelle qui n'est que la divergence de Kullback-Leibler entre p_s et $\prod_{k=1, \dots, M} p_{s_k}$ avec $p_s(x)$ désigne la densité de probabilité de s au point x et p_{s_k} représente la densité de probabilité de la composante s_k de s au point x_k .

3.2.1.3 Fonctions de contraste en mélange convolutif

Récemment, la notion des fonctions de contraste a été utilisée pour la séparation de mélange convolutif. Les premiers travaux ont été initié par Comon en 1996 [COM96]. C'est en quelque sorte une généralisation des fonctions de contraste utilisées en mélange instantané pour qu'elles soient valides dans le cas de mélange convolutif. La formulation d'un mélange convolutif de sources s'écrit comme suit :

$$x(n) = [A(z)]s(n)$$

où $A(z) = \sum_k A_k z^{-k}$ est la fonction de transfert du filtre de mélange. L'objectif consiste à estimer à partir des observations $x(n)$ une version filtrée de chaque source.

Ayant une version filtrée de chaque source, on peut ainsi utiliser le filtre de Wiener qui minimise une erreur quadratique entre une source estimée et une observation sur un capteur. En d'autre terme, cette manipulation va nous permettre de calculer la contribution de chaque source sur chaque capteur. Cette méthode sera présentée plus en détails dans la suite de ce chapitre.

La démarche est de retrouver un filtre noté $B(z)$ qui, appliqué aux observations $x(n)$, permet d'obtenir des signaux $y(n)$ mutuellement indépendants. Les fonctions de contraste qui ont été présentées précédemment dans le cas de mélange instantané permettent de tester l'indépendance des sorties estimées $y_i(n)$ et $y_j(n)$ pour $i \neq j$. Donc, on peut déduire

que ce type de fonctions de contraste ne peut pas être appliqué dans le cas d'un mélange convolutif. Il est possible d'utiliser des contrastes à base de cumulants d'ordre quatre à condition d'effectuer un blanchiment au préalable aux observations $x(n)$ (voir ce que nous avons présenté dans la première partie de ce chapitre sur l'utilisation de la diagonalisation conjointe et le filtre de Wiener pour résoudre un problème de sources convolutives).

3.2.2 Technique d'annulation des cumulants croisés

Dans ce chapitre, nous décrivons les méthodes de séparation de sources basées sur les statistiques d'ordre supérieur. Dans ce paragraphe, nous allons présenter une approche utilisant les cumulants d'ordre quatre qui est plus simple que les méthodes présentées auparavant.

Cette méthode consiste à estimer le filtre de séparation $B(z)$ tel que les cumulants croisés d'ordre quatre entre les $y_i(n)$, composantes de $y(n)$ valent zéro.

On considère le cas de deux sources et deux capteurs. Les filtres de séparations contenus dans la matrice $A(z)$ sont des filtres à réponse impulsionnelle finie. Dans [TJC92], il a été proposé d'annuler les cumulants cités ci-dessous :

$$\text{cum}(y_i(n), y_i(n), y_j(n-k), y_j(n-k)) \quad (3.26)$$

ou

$$\text{cum}(y_i(n), y_i(n), y_i(n), y_j(n-k)) \quad (3.27)$$

et

$$\text{cum}(y_j(n), y_j(n), y_j(n), y_i(n-k)) \quad (3.28)$$

Dans la méthode que nous allons présenter dans la suite, nous utilisons cette technique d'annulation des cumulants croisés. Ces cumulants seront exprimés à partir du calcul des contributions des sources sur les capteurs en utilisant la technique de Wiener.

De même, dans le cas des signaux complexes. Si on prend par exemple le cumulant suivant :

$$\text{cum}(y_i(n), y_i(n), y_i(n), y_j(n-k)) \quad (3.29)$$

Les différentes combinaisons possibles pour ce type de signaux s'écrivent comme suit :

$$\text{cum}(y_i(n), y_i(n), y_i(n), y_j(n-k))$$

$$\text{cum}(y_i^*(n), y_i(n), y_i(n), y_j(n-k))$$

$$\text{cum}(y_i^*(n), y_i^*(n), y_i(n), y_j(n-k))$$

$$\text{cum}(y_i^*(n), y_i^*(n), y_i^*(n), y_j(n-k))$$

3.2.3 Fontion de contraste utilisant les contributions des sources sur les capteurs

Dans ce chapitre, nous avons présenté une méthode utilisant les cumulants d'ordre 4 pour l'estimation de la matrice de séparation. La méthode proposée utilise l'algorithme de diagonalisation conjointe JADE. Dans cette section, nous allons présenter deux critères d'indépendance basés sur l'annulation des cumulants croisés d'ordre 4 entre les différentes contributions $z_{ij}(n)$ des sources sur les capteurs calculées par un filtrage de Wiener entre les sources estimées $y_i(n)$ et les observations $x_j(n)$.

Les deux critères utilisés sont les suivants :

$$C_1 = \min \left(\sum_m \sum_{i,j,k,l \neq iiii} |\text{cum}(z_{mi}^*, z_{mj}^*, z_{mk} z_{ml})| \right) \quad (3.30)$$

$$C_2 = \min \left(\sum_m \sum_{i,j,k,l \neq iiii} |\text{cum}(z_{mi}, z_{mj}^*, x_{mk} x_{ml}^*)| \right) \quad (3.31)$$

La minimisation de ces deux critères d'une façon itérative conduit à la séparation [AS06, FK98]. Nous avons considéré le cas de deux sources et deux capteurs. Nous appelons sources les deux contributions sur le premier capteur.

$$x_1(n) = s_1(n) + s_2(n) \quad (3.32)$$

On note z_{mi} les contributions des sources sur le premier capteur, calculées par le filtre de Wiener.

$$z_{mi} = \sum_{k=-L}^L w_{mi}(k) y_i(n-k) \quad (3.33)$$

La transformée de Fourier discrète (TFD) des coefficients du filtre de Wiener $w_{mi}(k)$ est représentée par la densité insterspectrale notée $\gamma_{Y_j X_1}(f)$ entre l'observation $x_1(n)$ et la source estimée $y_j(n)$ divisée par la densité spectrale $\gamma_{Y_j}(f)$ de la même source $y_j(n)$.

$$\begin{aligned} W_{11}(f) &= \frac{\gamma_{Y_1 X_1}(f)}{\gamma_{Y_1}(f)} \\ W_{12}(f) &= \frac{\gamma_{Y_2 X_1}(f)}{\gamma_{Y_1}(f)} \end{aligned} \quad (3.34)$$

Nous avons appelé $G(z)$ le filtre de global qui est le produit de convolution de la matrice de séparation B par la matrice du mélange A . Les deux contributions des sources sur le premier capteur en fonction de $G(z)$ s'écrivent comme suit :

$$\begin{aligned} Z_{11}(f) &= \frac{|G_{11}(f)|^2 \gamma_{S_1}(f) + G_{11}(f) \bar{G}_{12}(f) \gamma_{S_2}}{\gamma_{Y_1}(f)} S_1(f) \\ &\quad + \frac{\bar{G}_{11}(f) G_{12}(f) \gamma_{S_1}(f) + |G_{12}(f)|^2 \gamma_{S_2}}{\gamma_{Y_1}(f)} S_2(f) \\ Z_{12}(f) &= \frac{|G_{21}(f)|^2 \gamma_{S_1}(f) + G_{21}(f) \bar{G}_{22}(f) \gamma_{S_2}}{\gamma_{Y_2}(f)} S_1(f) \\ &\quad + \frac{\bar{G}_{21}(f) G_{22}(f) \gamma_{S_1}(f) + |G_{22}(f)|^2 \gamma_{S_2}}{\gamma_{Y_2}(f)} S_2(f) \end{aligned} \quad (3.35)$$

L'algorithme proposé est constitué de deux étapes principales. L'étape de calcul des contributions des sources sur le premier capteur et l'étape de séparation. Ces deux phases utilisent les statistiques d'ordre 2 (calcul des corrélations entre les sources $s_1(n)$ et $s_2(n)$) et les statistiques d'ordre 4 (indépendance entre les sources $s_1(n)$ et $s_2(n)$). La convergence de l'algorithme de séparation est représentée par :

$$\begin{aligned} \frac{|G_{11}(f)|^2 \gamma_{S_1}(f) + G_{11}(f) \bar{G}_{12}(f) \gamma_{S_2}}{\gamma_{Y_1}(f)} &\longrightarrow G_{11}(f) \\ \frac{|G_{12}(f)|^2 \gamma_{S_2} + G_{12}(f) \bar{G}_{11}(f) \gamma_{S_1}(f)}{\gamma_{Y_1}(f)} &\longrightarrow G_{12}(f) \end{aligned}$$

$$\begin{aligned} [\bar{G}_{11}(f) - |G_{11}(f)|^2] \gamma_{S1}(f) + [\bar{G}_{12}(f) - |G_{12}(f)|^2] \gamma_{S2}(f) &= 0 \\ [\bar{G}_{21}(f) - |G_{21}(f)|^2] \gamma_{S1}(f) + [\bar{G}_{22}(f) - |G_{22}(f)|^2] \gamma_{S2}(f) &= 0 \end{aligned}$$

Ceci implique que :

$$G_{ij}(f) = 0 \quad \text{ou} \quad G_{ij}(f) = 1$$

L'application du filtre de Wiener à chaque itération permet de maximiser la corrélation entre les sources estimées et les observations au niveau des capteurs.

$$\sum_m \sum_{i,j,k,l \neq iiii} |cum(z_{mi}^*, z_{mj}^*, z_{mk} z_{ml})| = 0 \quad (3.36)$$

On rappelle que $z_{11}(n) = s_1(n)$ et $z_{12}(n) = s_2(n)$. Si on suppose que :

$$x_2(n) = F_{11}(s_1(n)) + F_{12}(s_2(n))$$

Donc, on peut déduire que :

$$z_{21}(n) = F_{11}(s_1(n)) \quad \text{et} \quad z_{22}(n) = F_{12}(s_2(n))$$

L'utilisation des statistiques d'ordre 2 signifie qu'on pourra retrouver deux sources orthogonales quelconques, notées $s'_1(n)$ et $s'_2(n)$ telles que leur somme vaut l'observation sur le premier capteur.

$$x_1(n) = s'_1(n) + s'_2(n)$$

Supposant qu'on retrouve :

$$z_{11}(n) = s'_1(n) = \alpha s_1(n) + \beta s_2(n)$$

$$z_{12}(n) = s'_2(n) = \gamma s_1(n) + \delta s_2(n)$$

Les sources $s'_1(n)$ et $s'_2(n)$ sont toujours considérées comme des signaux orthogonaux. Dans ce cas, le mélange des sources est devenu de type mélange instantané. Le critère

de séparation :

$$\sum_{i,j,k,l \neq iiii} |cum(z_{1i}^*, z_{1j}^*, z_{1k}z_{1l})| = 0$$

si et seulement si :

$$z_{11}(n) = s_1(n) \quad \text{et} \quad z_{12}(n) = s_2(n)$$

3.2.4 Implantation et simulations

Dans ce paragraphe, nous allons donner des détails sur l'implantation pratique de la méthode proposée. Nous avons proposé de minimiser deux critères d'indépendance $C1$ et $C2$ (eq. 3.30 et 3.31). La minimisation de ces critères a été faite d'une façon itérative. A chaque itération, nous mettons à jour le calcul de la matrice des filtres de séparation notée $B(z)$.

Nous avons considéré le cas d'un mélange convolutif de deux sources normalisées et deux capteurs. Chaque source composée de 1500 échantillons est constituée par la somme d'un signal aléatoire à distribution uniforme avec un signal sinusoidal. Les filtres de mélange utilisés sont de longueur 3. (équation (3.37))

$$\begin{aligned} A_{11} &= 1 + 0.2z^{-1} + 0.1z^{-2} \\ A_{12} &= 0.5 + 0.3z^{-1} + 0.1z^{-2} \\ A_{21} &= 0.5 + 0.3z^{-1} + 0.1z^{-2} \\ A_{22} &= 1 + 0.2z^{-1} + 0.1z^{-2} \end{aligned} \tag{3.37}$$

La procédure de séparation consiste à estimer la matrice des filtres notée $B(z)$ en minimisant d'une façon itérative les critères $C1$ et $C2$ tout en appliquant à chaque itération la procédure de Wiener afin de retrouver les contributions des sources estimées sur les capteurs. L'algorithme de séparation se résume par les étapes suivantes :

Initialisation : $y(n) = x(n)$

Itérations : Répéter jusqu'à la convergence

- Estimation des critères $C1$ ou $C2$ en utilisant les contributions des sources sur les capteurs $z_j(n) = [W_i(z)]y_j(n)$:

$$C1 = \min \left(\sum_m \sum_{i,j,k,l \neq iiii} |cum(z_{mi}^*, z_{mj}^*, z_{mk}, z_{ml})| \right)$$

$$C2 = \min \left(\sum_m \sum_{i,j,k,l \neq iiii} |cum(z_{mi}, x_m^*, z_{mk}, x_m^*)| \right)$$

- Calcul de la matrice des filtres :

$$B(z) \leftarrow B(z) - \mu C1$$

- Calcul des sources en sortie : $y_j(n) = [B(z)]x(n)$
- Calcul des nouvelles contributions des sources sur les capteurs : $z_j(n) = [W_i(z)]y_j(n)$
- Remplacer $y(n)$ par les contributions $[z_{11}(n), z_{12}(n)]$

Nous montrons la performance de l'algorithme décrit par les figures (3.12) et (3.13). L'observation des tracés des EQM (**E**rreur **Q**uadratique **M**oyenne) de ces figures est encourageante [ASR07a]. Ces EQMs sont moyennés sur 50 réalisations des sources. Sur la figure (3.12), on représente l'erreur quadratique moyenne en fonction des itérations pour $\mu = 0.03$. Les courbes en noir et en bleu représentent respectivement l'EQM des critères $C1$ et $C2$. La courbe en rouge représente l'EQM du critère $C2$ dans le cas où on applique une étape de préblanchiment sur les observations $x(n)$. Les algorithmes sont comparés pour la même valeur de $\mu = 0.03$. Celle-ci est optimale pour la dernière courbe (figure 3.13) dans le sens où son EQM est minimale pour le nombre d'itérations choisies.

Les mêmes résultats ont été obtenus dans [RCM06] en utilisant l'approche d'une fonction de contraste avec référence. En comparant les différentes courbes de la figure (3.12), on

remarque que l'utilisation du critère $C1$ conduit à l'algorithme de séparation le plus rapide en terme de la vitesse de convergence pour une valeur de μ constante (tableau 3.3). Ceci se traduit par le fait que dans le critère $C1$, nous minimisons plus de cumulants croisés que dans le critère $C2$. Par exemple, dans le cas de deux sources et deux capteurs, le critère $C1$ est minimisé conjointement pour 6 cumulants croisés et 2 cumulants croisés pour le critère $C2$.

Dans la figure (3.13), le terme μ est optimal pour chacun des critères $C1$ et $C2$ en terme

	$C1$	$C2$	$C3$
EQM	2.10^{-7}	3.10^{-7}	5.10^{-7}

TAB. 3.3 – EQM des critères $C1$, $C2$ et $C3$

d'EQM pour comparer leur vitesse de convergence. L'EQM est de 3.10^{-7} et on remarque que $C1$ (courbe en noire) a toujours la vitesse de convergence la plus rapide par rapport à $C2$ (courbe en rouge). La valeur de μ était égale à 0.01

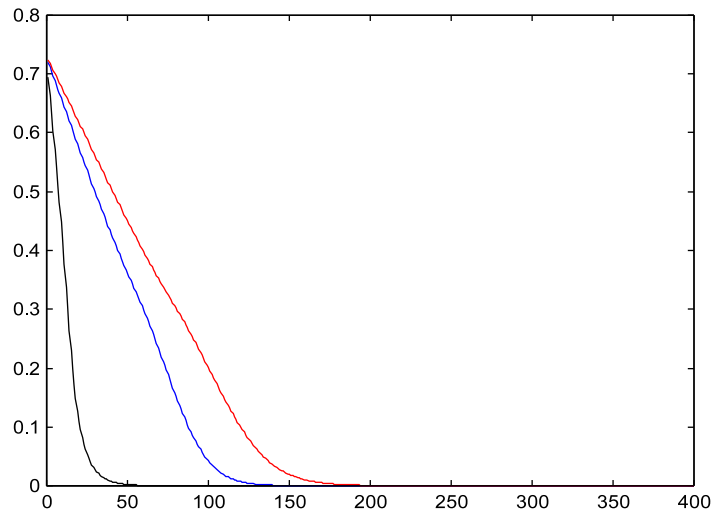


FIG. 3.12 – EQM des critères $C1$, $C2$ et $C3$ en fonction des itérations ($\mu = 0.03$)

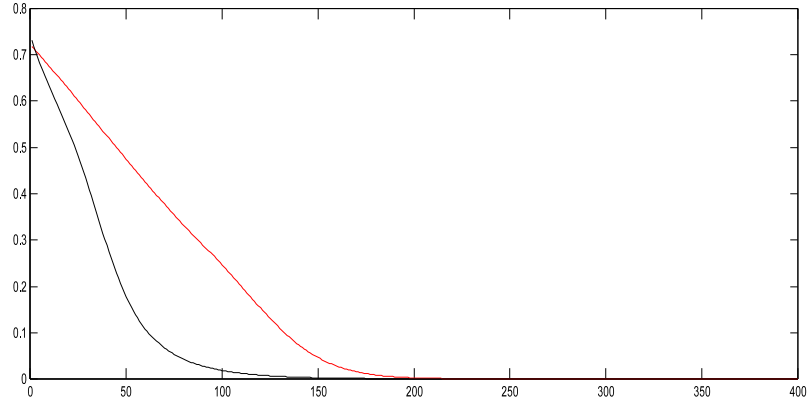


FIG. 3.13 – EQM des critères $C1$ et $C2$ en fonction des itérations ($\mu = 0.01$)

3.2.4.1 Performances de l'algorithme en fonction de la longueur L du filtre de mélange

Pour étudier les performances du critère de séparation proposé, nous proposons de mesurer l'EQM et le RSB entre les sources originales et les sources estimées. Comme ci-avant, nous distinguons deux cas différents en fonction du type des sources.

- **Cas de deux sources i.i.d. uniformes :**

Dans ce paragraphe, nous avons considéré un mélange convolutif de deux sources i.i.d. uniformes. Chaque source est normalisée et composée de 1500 échantillons. Comme dans le paragraphe 3.1.3.3, nous calculons l'EQM et le RSB en fonction de la longueur L du filtre de mélange. Le tableau 3.4 représente les valeurs de l'EQM et du RSB en utilisant le critère $C1$ comme critère de séparation.

Les figures (3.14) et (3.15) tracent l'EQM et le RSB en fonction de la longueur du filtre de mélange. Il apparaît que les performances du critère $C1$ sont légèrement moins bonnes en augmentant le paramètre L du filtre de mélange. En revanche, les résultats restent corrects.

L	10	20	30	50	100
EQM	$3,8 \cdot 10^{-7}$	$2,1 \cdot 10^{-6}$	$3,2 \cdot 10^{-5}$	$7,4 \cdot 10^{-5}$	$8,9 \cdot 10^{-5}$
RSB (en dB)	25,6	23,8	20,4	19,3	18,2

TAB. 3.4 – Critère C1 : EQM et RSB en fonction de la longueur L du filtre de mélange

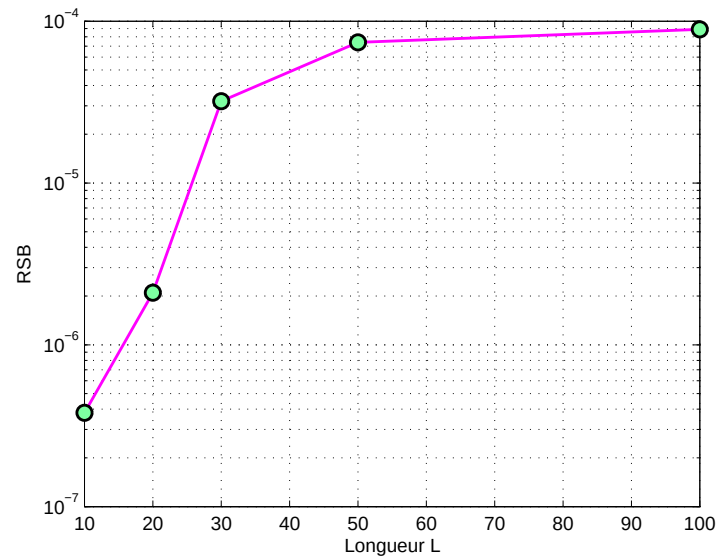


FIG. 3.14 – Critère C1 : EQM en fonction de la longueur L du filtre de mélange

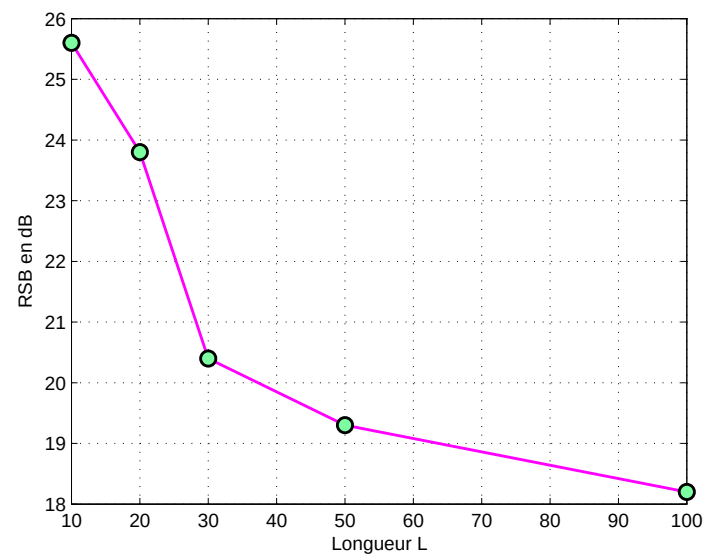


FIG. 3.15 – Critère C1 : RSB en fonction de la longueur L du filtre de mélange

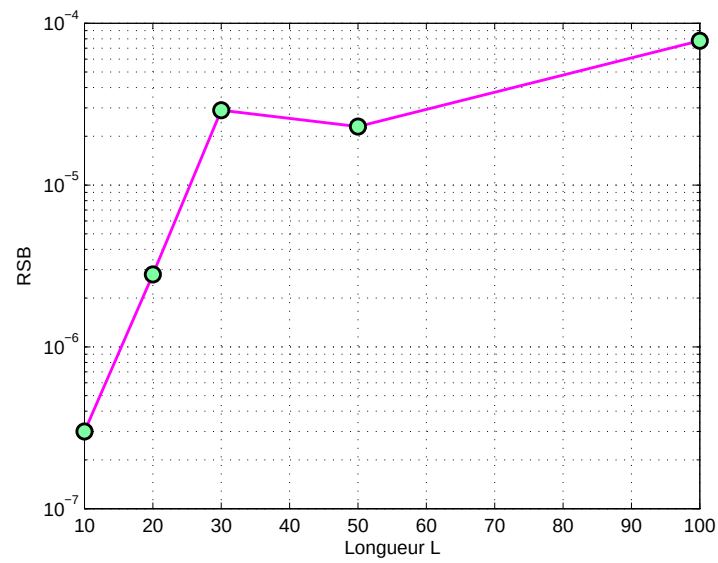
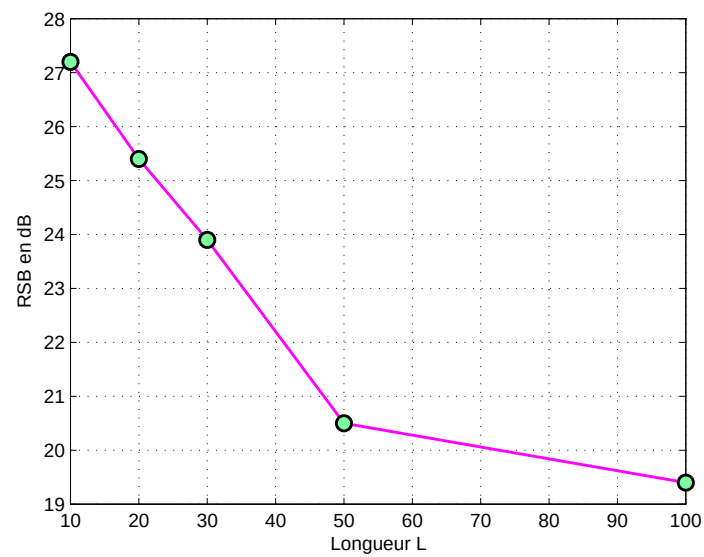
- **Cas de deux sources composées chacune d'une somme d'un signal sinusoïdal et d'une distribution i.i.d. uniforme :**

Dans ce paragraphe, chaque source est la somme d'un bruit uniforme indépendant et identiquement distribué (i.i.d.) et d'une sinusoïde à paramètres déterministes. De la même façon, nous mesurons l'EQM et le RSB des sources en fonction de la longueur du filtre (tableau 3.5).

Les figures (3.16) et (3.17) permettent de voir l'influence du paramètre L . Comparant aux résultats de la méthode utilisant la diagonalisation conjointe (paragraphe 3.1.3.3), on remarque que la longueur L du filtre de mélange influence moins la qualité de séparation. De la même manière, nous pouvons évaluer les performances des critères $C2$ et $C3$ qui donnent des résultats comparables. Par contre, l'avantage de l'utilisation du critère $C1$ réside aussi dans sa vitesse de convergence.

L	10	20	30	50	100
EQM	$3 \cdot 10^{-7}$	$2,8 \cdot 10^{-6}$	$2,9 \cdot 10^{-5}$	$2,3 \cdot 10^{-5}$	$7,8 \cdot 10^{-5}$
RSB (en dB)	27,2	25,4	23,9	20,5	19,4

TAB. 3.5 – Critère $C1$: EQM et RSB en fonction de la longueur L du filtre de mélange

FIG. 3.16 – Critère C1 : EQM en fonction de la longueur L du filtre de mélangeFIG. 3.17 – Critère C1 : RSB en fonction de la longueur L du filtre de mélange

3.2.5 Comparaison avec les résultats de la méthode développée dans le chapitre 2

Dans ce paragraphe, nous allons présenter un bref récapitulatif des résultats obtenus. Pour montrer les performances des méthodes proposées, nous comparons les résultats de la méthode de contraste au résultat de la méthode basée sur la minimisation de l'information mutuelle décrite dans le chapitre 2. Pour réaliser cette comparaison, nous avons utilisé le même type de sources (signal sinusoïdal et signal i.i.d. à distributions uniforme) et le même filtre de mélange $A(z)$. La réponse impulsionnelle de ce filtre compte 3 coefficients. Les sources utilisées sont composées de 1500 échantillons et les résultats ont été moyennés sur 50 réalisations.

Nous avons établi un tableau récapitulatif (tableau 3.6) permettant de comparer les différentes EQMs et le nombre d'itérations nécessaire à la convergence de l'algorithme. Selon le tableau (3.6), on peut conclure que la meilleure méthode est celle basée sur la

	Information mutuelle	$C1$	$C2$	$C3$
EQM	$2,6 \cdot 10^{-8}$	$3 \cdot 10^{-7}$	$8 \cdot 10^{-7}$	$5 \cdot 10^{-7}$
Temps de convergence (nombre d'itérations)	30	150	200	200

TAB. 3.6 – EQM et temps de convergence

minimisation de l'information mutuelle pour des filtres à courte réponse impulsionnelle. $C2$ et $C3$ sont très proches et le blanchiment des capteurs ne s'impose pas.

Dans le chapitre 2, les simulations ont confirmé l'intérêt de la méthode basée sur la minimisation de l'information mutuelle. Nous avons constaté clairement une bonne convergence de l'algorithme pour des filtres de mélange à courte réponse impulsionnelle. Par contre, nous avons rencontré certains problèmes de non convergence dans le cas où on augmente la longueur équivalente du filtre de mélange. Ces mêmes problèmes sont dûs aussi dans certains cas à une mauvaise estimation des densités de probabilité des sources qui nécessite un réglage de paramètres parfois délicat. Alors que dans le chapitre 3, les résultats obtenus montrent la validité de l'algorithme pour des filtres à longue réponse impulsionnelle (voir

les paragraphes (3.1.3.3) et (3.2.4.1)). La première méthode basée sur la diagonalisation conjointe des matrices de cumulants nécessite une étape de blanchiment au préalable. La deuxième méthode est basée sur l'utilisation des fonctions de contraste. Elle présente des résultats très avantageux, surtout en terme de vitesse de convergence (critère $C1$).

3.3 Conclusion

Au cours de la première partie de ce chapitre, nous avons rappelé l'algorithme de diagonalisation conjointe des matrices de cumulants (JADE). Ensuite, nous avons expliqué comment on peut insérer la procédure de Wiener dans cette approche dans le cadre de mélange convolutif de sources. Nous avons procédé par une étape de blanchiment des observations, ensuite le calcul des matrices de cumulants en utilisant comme paramètres les sources estimées et les contributions de ces sources sur le premier capteur. Nous avons fini cette première partie par une illustration de l'approche proposée par des simulations numériques qui montrent la qualité de la séparation.

Dans la deuxième partie de ce chapitre, nous nous sommes concentrés sur des approches appelées *fonctions de contraste* car elles présentent certaines avantages par rapport à d'autres méthodes. Plusieurs contrastes ont été proposés dans la littérature. Nous nous sommes intéressés par la technique d'annulation des cumulants croisés. Ainsi, nous avons introduit la notion des fonctions de contraste en séparation de sources. La performance de ce type d'approche dans un processus de séparation nous a poussé à généraliser une fonction de contraste à base des contributions des sources sur les capteurs obtenues avec un filtrage de Wiener. Les simulations présentées nous a ainsi permis de mettre en évidence l'efficacité des notions de contraste.

Enfin, l'avantage majeur des critères développés dans ce chapitre réside dans la projection des sources sur les capteurs afin de retrouver leurs contributions sur chaque capteur. Il en résulte la possibilité de simplifier les critères d'indépendance dans le cas de mélange convolutif de sources. Ainsi, l'utilisation des fonctions de contraste offre une bonne qualité de séparation des sources.

Conclusions et perspectives

Conclusions

Notre travail de recherche a concerné le problème de la séparation de sources dans le cadre de mélange convolutif. Nous nous sommes intéressés à étudier les méthodes existantes en analyse en composantes indépendantes et à établir de nouvelles approches pour la séparation des sources en mélange convolutif. Certaines méthodes ont fait l'objet d'une démonstration théorique et d'autres ont simplement été généralisées et implantées. L'ensemble de ces méthodes ont été testés sur des simulations numériques.

Le contexte de notre étude est celui de l'extraction des sources où l'on dispose d'un mélange de signaux reçus sur plusieurs capteurs. Le traitement de cette problématique nécessite le bon choix du critère d'indépendance qui sera utilisé pour assurer une séparation efficace des sources en vue d'améliorer les performances par rapport aux méthodes existantes. La difficulté majeure que soulève cette problématique dans le cas d'un mélange convolutif de sources est la mesure d'indépendance pour tous les instants m .

Ce travail s'est d'abord concentré sur les approches utilisant l'information mutuelle comme critère d'indépendance pour la séparation. Plusieurs méthodes utilisant ce critère existent dans la littérature et montrent les bonnes performances de ce critère. Dans un premier temps, nous nous sommes focalisés sur le calcul des contributions des sources sur chaque capteur. L'estimation de ces contributions a été réalisée grâce à un filtrage de Wiener en minimisant l'erreur quadratique entre les sources estimées et les observations. Dans le cadre des mélanges convolutifs, l'obtention des composantes indépendantes nécessite le test d'indépendance à tous les retards m , ce qui rend le critère d'indépendance utilisé très complexe. Nous montrons théoriquement que l'insertion des contributions des sources sur les capteurs dans la procédure de séparation simplifie le test d'indépendance. En d'autre terme, le test d'indépendance à l'instant 0 suffit pour résoudre le problème de séparation des sources.

Deux méthodes ont été proposées dans ce manuscrit. La première repose sur l'utilisation de l'information mutuelle. Nous cherchons à minimiser ce critère d'indépendance en utilisant l'approche de gradient. Dans l'algorithme de séparation, nous remplaçons les sources estimées par les contributions de ces sources dans les observations, ce qui augmente la performance de la convergence de l'algorithme. La deuxième méthode exploite les statistiques d'ordre quatre. Dans un premier temps, nous reprenons l'algorithme de diagonalisation conjointe (JADE) basé sur la diagonalisation de plusieurs matrices de cumulants. Le calcul de ces matrices nécessite la connaissance des contributions des sources sur les capteurs obtenues par un filtrage de Wiener. Ensuite, nous généralisons une fonction de contraste basée sur l'annulation des cumulants croisés, calculés toujours à partir des contributions des sources sur les capteurs. L'avantage de l'utilisation des statistiques d'ordre supérieur par rapport l'information mutuelle réside dans le choix des paramètres à régler. Nous rappelons que l'utilisation de l'information mutuelle nécessite l'estimation des densités de probabilité des sources. Le principal atout des algorithmes proposés est probablement la simplicité du test d'indépendance entre les sources estimées et la facilité d'implantation.

Plusieurs expériences de séparation ont été simulées. Au vue de la performance et de la qualité de séparation, évaluées par la convergence de l'algorithme ou la cohérence entre les sources estimées et les sources originales, les premiers résultats obtenus montrent que les approches proposées fournissent des résultats favorables malgré la complexité et la difficulté du contexte.

Perspectives

Plusieurs approches ont été proposées pour résoudre un problème de séparation d'un mélange instantané de sources. Dans le cadre d'un mélange convolutif, ce problème reste ouvert et attirant. Ainsi, il sera intéressant d'exploiter d'autres informations et hypothèses pour améliorer la performance des méthodes existantes.

Plusieurs perspectives peuvent être envisageables :

- Le problème traité dans ce manuscrit concerne les mélanges convolutifs et des signaux stationnaires. Nous avons validé les approches proposées avec des filtre de mélange à réponse impulsionnelle finie (RIF) contenant une dizaine de coefficients. Mais il serait intéressant de valider ces méthodes sur des signaux provenant des machines tournantes qui présentent des réponses impulsionnelles plus longues et en présence de bruit.
- Une étude peut être menée pour exploiter la propriété de la cyclostationnarité des signaux de moteurs électriques, qui peut probablement jouer un rôle important dans l'obtention de bons résultats.
- Les mélanges traités dans ce manuscrit sont linéaires. Par contre, il existe très peu de travaux sur des mélanges de type non-linéaires. Donc, il serait intéressant d'étudier ce type de mélanges dans le but de concevoir de nouvelles approches de séparabilité des sources ou de nouvelles fonctions de contraste pour tester l'indépendance entre les signaux extraits tout en utilisant des contributions des sources sur le capteurs.

Annexes

Annexe A

Estimation des fonctions scores

Dans le chapitre 2, nous avons utilisé l'information mutuelle comme critère d'indépendance pour la séparation des sources. L'utilisation de ce critère nécessite l'estimation des densités de probabilités marginales f_{x_i} et des fonctions scores marginales ψ_{x_i} . Dans cet annexe, nous allons présentons quelques définitions et méthodes pour l'estimation de ces fonctions.

A.1 Définitions

A.1.1 Fonction score

Soient x une variable aléatoire et $p_x(x)$ est sa densité de probabilité. La fonction score de x s'écrit comme suit :

$$\psi_x(x) = -\frac{d}{dx} \ln p_x(x) = -\frac{p'_x(x)}{p_x(x)}$$

A.1.2 Fonction score marginale

$$\psi_x(x) = (\psi_1(x_1), \psi_2(x_2), \dots, \psi_N(x_N))^T$$

où :

$$\psi_i(x_i) = -\frac{d}{dx_i} \ln p_{x_i}(x_i) = -\frac{p'_{x_i}(x_i)}{p_{x_i}(x_i)}$$

A.1.3 Fonction score conjointe

$$\phi(x) = (\phi_1(x), \phi_2(x), \dots, \phi_N(x))^T$$

où :

$$\phi_i(x) = -\frac{\partial}{\partial x_i} \ln p_x(x) = -\frac{\frac{\partial}{\partial x_i} p_x(x)}{p_x(x)}$$

A.1.4 Fonction score différentielle

$$\beta_x(x) = \psi_x(x) - \phi_x(x)$$

A.2 Estimation de la densité de probabilité

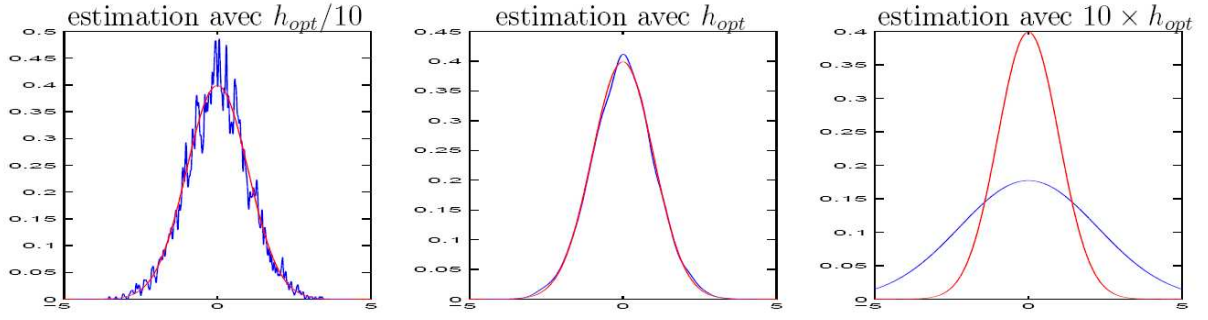
L'estimation de la densité de probabilité d'un processus aléatoire $x = [x(1), x(2), \dots, x(N)]$ en utilisant la méthode des noyaux est décrite par l'équation suivante :

$$f(x) = \frac{1}{Nh} \sum_{i=1}^N K\left(\frac{x - x(i)}{h}\right)$$

- $K(\cdot)$ représente le noyau,
- h représente la largeur du noyau.

$$K(u) = \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{u^2}{2}\right)$$

Le paramètre h permet d'optimiser l'estimation de la densité de probabilité de x . La figure (A.1) montre l'importance du choix de la largeur du noyau. Ce figure représente l'exemple de l'estimation d'une gaussienne avec un noyau gaussien. h_{opt} représente la valeur optimale de h pour une meilleure estimation de la densité de probabilité. La courbe théorique est représentée par la courbe en rouge et la densité de probabilité estimée est en bleu.

FIG. A.1 – Influence du paramètre h pour l'estimation de la densité de probabilité

La méthode décrite dans le chapitre 2 utilise l'estimation des fonctions scores ψ en utilisant un noyau gaussien. Nous avons utilisé l'information mutuelle noté I comme un critère d'indépendance pour la séparation.

En notant B la matrice de séparation recherchée, nous avons calculé le gradient de $I(y)$ par rapport à B . y représente le vecteur des sources estimées.

$$I(y) = \sum_{i=1}^N H(y_i) - H(x) - \log|B|$$

où H est l'entropie de Shannon et x est le vecteur des observations.

L'expression du gradient de l'information mutuelle s'écrit comme suit :

$$\frac{\partial I(y)}{\partial B} = \sum_{i=1}^N \frac{\partial H(y_i)}{\partial B} - \frac{\partial}{\partial B} \log|det B|$$

L'expression de ce gradient en fonction de ψ_{y_i} la fonction score de la source y_i s'exprime finalement comme suit :

$$\frac{\partial I(y)}{\partial B} = E[\psi_y(y(t))x(t)^T] - \frac{\partial}{\partial B} \log|det B|$$

où $\psi_y(y(t)) = [\psi_{y_1}(y_1(t)), \psi_{y_2}(y_2(t)), \dots, \psi_{y_N}(y_N(t))]^T$.

Il existe d'autres méthodes qui sont moins coûteuses en temps de calcul, notamment celle développée par D.-T. Pham [PHA02] utilisant des splines.

Annexe B

Algorithme de diagonalisation conjointe JADE

J.-F. Cardoso et A. Souloumiac ont introduit et développé un algorithme robuste à base de techniques statistiques pour l'application de séparation aveugle des sources. Cette approche emploie un algorithme Jacobien pour optimiser une fonction de contraste. La fonction utilisée par JADE (*Joint Approximate Diagonalization of Eigen-matrices*) s'appuie sur les cumulants d'ordre croissant (le quatrième cumulant). JADE opère en quatre étapes :

- **L'initialisation** : Estimer une matrice de blanchiment, puis transformer les données par cette matrice

$$\bar{x}(k) = Qx(k)$$

où $\bar{x}$ représente les observations blanchies.

- **Former les statistiques** : Estimer un ensemble maximal de matrices cumulants

$$C_{\bar{x}}(I) = \frac{1}{M} \sum_{k=1}^M [\bar{x}^T(k) \bar{x}(k) \bar{x}(k) \bar{x}^T(k)] - 2\hat{R}_{\bar{x}}(0)\hat{R}_{\bar{x}}(0) - \text{tr}(\hat{R}_{\bar{x}}(0))\hat{R}_{\bar{x}}(0) = \hat{U}\Lambda_I\hat{U}^T$$

où $\hat{R}_{\bar{x}}(0) = \frac{1}{M} \sum_{k=1}^M [\bar{x}(k) \bar{x}^T(k)]$ et $\hat{U} = [\hat{u}_1, \hat{u}_2, \dots, \hat{u}_m]$.

$$C_{\bar{x}}(I) = \frac{1}{M} \sum_{k=1}^M [\bar{x}^T(k) E_p \bar{x}(k) \bar{x}(k) \bar{x}^T(k)] - \hat{R}_{\bar{x}}(0) E_p \hat{R}_{\bar{x}}(0) - \text{tr}(E_p \hat{R}_{\bar{x}}(0)) \hat{R}_{\bar{x}}(0) - \hat{R}_{\bar{x}} E_p^T \hat{R}_{\bar{x}}$$

où $E_p = \hat{u}_p \hat{u}_p^T, p = 1, 2, \dots, m$

- **Optimiser un contraste orthogonal** : Trouver une matrice de rotation telle que les matrices de cumulants deviennent les plus diagonales possibles.
- **Séparation** : Estimer les matrices de séparation par rotation des matrices de blanchiment.

Ainsi, la première étape de l'algorithme JADE est la normalisation par blanchiment de la partie signal à l'aide de la matrice de covariance. La deuxième étape consiste à l'estimation des sources [CS93]. L'identification de la matrice de mélange consiste seulement à l'estimation d'une matrice unitaire à l'aide des cumulants des signaux blanchis. Pour JADE, la matrice unitaire qui maximise le critère :

$$\sum_{i,k,l}^M |cum(x_i, x_i^*, x_k, x_l^*)|^2$$

Deux hypothèses sont maintenant nécessaires pour l'estimation de cette matrice, les signaux $s_i(t)$ doivent être indépendants et non gaussiens (nous pouvons séparer au plus un signal gaussien). Ainsi J.-F. Cardoso et A. Souloumiac proposent dans [SOU93] une transformation unitaire des signaux blanchis grâce à une diagonalisation conjointe d'une famille de matrices propres.

L'algorithme JADE est très attractif car il est simple à implémenter et ne requiert pas la gestion de paramètres d'optimisation. Il présente l'avantage par rapport aux autres algorithmes d'ACI d'être beaucoup plus rapide dans la convergence.

Bibliographie

- [AC97] S. AMARI et J.-F. CARDOSO. Blind source separation. semi-parametric statistical approach. *IEEE Trans. on Signal Processing*, 45(11) :2692–2700, November 1997.
- [ALM01] L. B. ALMEIDA. Ica of linear and nonlinear mixtures based on mutual information. *International Joint Conference on Neural Networks*, July 2001.
- [AMA03a] A. ADIB, E. MOREAU et D. ABOUTAJDINE. Blind sources separation by simultaneous generalized referenced contrasts and diagonalisation. In *Proc. of the Fourth Int. Conf. on Independent Component Analysis and Blind Source Separation, ICA2003*, pages 657–661, Nara, Japan, April 2003.
- [AMA03b] A. ADIB, E. MOREAU et D. ABOUTAJDINE. Referenced contrasts a new approach of blind source separation. In *In Seventh International Symposium on Signal Processing and its applications (ISSPA'2003)*, pages 97–100, Paris, France, July 2003.
- [AMA04] A. ADIB, E. MOREAU et D. ABOUTAJDINE. Source separation contrasts using a reference signal. *IEEE Signal Processing Letters*, 11(3) :312–315, March 2004.
- [AMR95] K. ANAND, G. MATHEW et V. REDDY. Blind separation of multiple co-channel bpsk signals arriving at an antenna array. *IEEE Signal Processing Letters*, 2 :176–178, September 1995.
- [AS06] M. AKIL et C. SERVIÈRE. Separability of convolutive mixtures : Application to the separation of combustion noise and piston-slap in diesel engine. In

- 6th international conference on Independent Component Analysis and Blind Source Separation, ICA2006*, pages 319–326, Charleston, USA, March 2006.
- [ASR07a] M. AKIL, C. SERVIÈRE et R.-B. RANDALL. Contrast function based on fourth-order statistics and wiener filtering for separation of convolutive mixtures. In *14th International Congress on Sound and Vibration (ICSV14)*, Cairns, Australia, July 2007.
- [ASR07b] M. AKIL, C. SERVIÈRE et R.B. RANDALL. Separability of convolutive mixtures using high-order statistics and wiener filtering. In *Proceeding of Physics in Signal and Image Processing, PSIP 2007*, Mulhouse, France, February 2007.
- [BAMCM97] A. BELOUCHRANI, K. ABED-MERAIM, J.-F. CARDOSO et E. MOULINES. A blind source separation technique using second-order statistics. *IEEE Trans. on Signal Processing*, 45(2) :434–444, February 1997.
- [BMCM97] A. BELOUCHRANI, K. ABED MERAIM, J.F. CARDOSO et E. MOULINES. A blind source separation technique using second order statistics. *IEEE Transactions on Signal Processing*, 45 :210–220, 1997.
- [BZ02] M. BABAIE-ZADEH. *On blind source separation in convolutive and non-linear mixtures*. PhD thesis, Institut National Polytechnique de Grenoble, Grenoble, France, 2002.
- [CAP95] V. CAPDEVIELLE. *Séparation de sources large bande à l'aide des moments d'ordre supérieur*. PhD thesis, Institut National Polytechnique de Grenoble, Grenoble, France, 1995.
- [CAR89] J.-F. CARDOSO. Source separation using higher order moments. In *Proceedings ICASSP*, pages 2109–2112, Glasgow, Scotland, May 1989.
- [CAR98] J.-F. CARDOSO. Blind signal separation : statistical principles. In *Proceedings IEEE*, volume 9, pages 2009–2025, 1998.
- [CAR99] J.-F. CARDOSO. High-order contrasts for independent component analysis. *Neural Computation*, 11(1) :157–192, 1999.

- [CCM93] E. CHAUMETTE, P. COMON et D. MULLER. Ica-based technique for radiating sources estimation :application to airport surveillance. *IEE Proceedings-F*, 140 :395–401, December 1993.
- [CMP04] M. CASTELLA, E. MOREAU et J.-C PESQUET. A quadratic miso contrast function for blind equalization. In *Proc. of the IEEE. int. conf. on Acoustics, Speech and Signal Processing (ICASSP2004)*, pages 681–684, Montréal, Canada, May 2004.
- [COM89] P. COMON. Separation of sources using higher-order cumulants. In *SPIE vol. 1152 Advanced Algorithms and Architectures for Signal Processing IV*, San Diego, USA, August 8-10 1989.
- [COM94] P. COMON. Independent component analysis, a new concept ? *Signal Processing*, 36 :287–314, 1994.
- [COM96] P. COMON. Contrasts for multichannel blind deconvolution. *IEEE Signal Processing Letters*, 3(7), July 1996.
- [CP01] B. CHEN et A.-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.
- [CPP05] M. CASTELLA, J.-F. PESQUET et A.-P. PETROPULU. A family of frequency- and time-domain contrasts for blind separation of convolutive mixtures of temporally dependent signals. *IEEE Trans. on Signal Processing*, 53(1) :107–120, January 2005.
- [CS93] J.-F. CARDOSO et A. SOULOUMIAC. Blind beamforming for non gaussian signals. *IEE Proceedings-F*, 140(6) :362–370, 1993.
- [CS96] J.-F. CARDOSO et A. SOULOUMIAC. Jacobi angles for simultaneous diagonalization. *SIAM Journal on Matrix Analysis and Applications*, 17(1) :161–164, January 1996.
- [CSL94] V. CAPDEVIELLE, C. SERVIÈRE et J.-L. LACOUME. Application of source separation to wide band signals. In *SSAP*, 1994.

- [CSL95a] V. CAPDEVIELLE, C. SERVIÈRE et J.-L. LACOUME. Blind separation of wide-band source in the frequency domain. In *Proceedings IEEE Int. Conf. on Acoustic, Speech and Signal Porcessing ICASSP*, pages 2080–2083, 1995.
- [CSL95b] V. CAPDEVIELLE, C. SERVIÈRE et J.-L. LACOUME. Blind separation of wide band sources. In *Proceedings IEEE Signal Processing*, pages 66–70, June 1995.
- [CT91] T. M. COVER et J. A. THOMAS. *Elements of Information Theory*. Wiely Series in Telecommunications, 1991.
- [DA96] Y. DEVILLE et L. ANDRY. Application of blind source separation techniques to multi-tag contactless identification system. *IEICE Transactions on fundamentals of electronics, Communications and Computer sciences*, pages 1694–1699, 1996.
- [DAR53] G. DARMOIS. Analyse générale des liaisons stochastiques. *Revue de l'Institut International de Statistique*, 21 :2–8, 1953.
- [DC01] A. DAPENA et L. CASTEDO. Blind source separation in the frequency domain : A novel solution to the amplitude and the permutation indeterminacy. In *Proceedings IWANN*, pages 603–610, Granada, Spain, June 2001.
- [DC03] A. DAPENA et L. CASTEDO. A novel frequency domain approach for separating convolutive mixtures of temporally-white signals. *Digital Signal Processing*, 13(2) :301–316, April 2003.
- [DEL95] N. DELFOSSE. *Séparation aveugle adaptative de mélanges de signaux indépendants*. PhD thesis, Université de Marne la Vallée, Décembre 1995.
- [DL95] N. DELFOSSE et P. LOUBATON. Adaptive blind separation of independent sources : a deflation approach. *IEEE Transactions on Signal Processing*, 45 :59–83, 1995.
- [DON81a] D. DONOHO. On minimum entropy deconvolution. *Applied Time Series Analysis*, pages 565–608, 1981.

- [DON81b] D. DONOHO. On minimum entropy deconvolution. In *Proc. second Applied Time Series Symposium*, pages 565–608, New York, USA, 1981.
- [DPV97] G. D’URSO, P. PRIEUR et C. VINCENT. Blind identification methods applied to edf civil works and power plants monitoring. In *HOS*, Canada, June 1997.
- [DSC03] A. DAPENA, C. SERVIÈRE et L. 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.
- [EGU83] S. EGUSHI. Second order efficiency of minimum contrast estimators in a curved exponential family. *The Annals of Statistics*, pages 793–803, 1983.
- [FET88] L. FETY. *Méthode de traitement d’antenne adaptées aux radiocommunications*. PhD thesis, ENST, 1988.
- [FK98] M. FENG et K.D. KAMMEYER. Blind source separation for communication signals using antenna arrays. In *Proc. ICUPC*, pages 665–669, Florence, Italy, October 1998.
- [GL90] M. GAETA et J.-L. LACOUME. Source separation without a priori knowledge : the maximum likelihood solution. In *EUSIPCO*, volume 2, pages 621–641, Barcelone, September 1990.
- [GRA90] R.M. GRAY. *Entropy and information theory*. Springer-Verlag, New York, 1990.
- [HJ86] J. HERAULT et C. JUTTEN. Space or time adaptive signal processing by neural network models. *AIP Conference Proceedings*, 51 :206–211, August 1986.
- [HJA85] J. HERAULT, C. JUTTEN et B. ANS. Détection de grandeurs primitives dans un message composite par une architecture de calcul neuromimétique en apprentissage non supervisé. In *GRETSI*, pages 1017–1022, Nice, France, Mai 1985.

- [JAL06] P. JALLON. *Séparation de mélanges convolutifs de sources cyclostationnaires*. PhD thesis, Université de Marne la Vallée, Mai 2006.
- [JH88a] C. JUTTEN et J. HERAULT. Independent component analysis versus pca. In *Proc. International Conference EUSIPCO*, pages 643–646, Grenoble, France, 1988.
- [JH88b] C. JUTTEN et J. HERAULT. Une solution neuromimétique du problème de séparation de sources. *Traitement du Signal*, 5 :389–403, 1988.
- [JTD⁺91] C. JUTTEN, L. NGUYEN THI, E. DIJKSTRA, E. VITTOZ et J. CAELEN. Blind separation of sources : an algorithm for separation of convolutive mixtures. In *International Signal Processing Workshop on Higher Order Statistics*, pages 273–276, Chamrousse, France, July 1991.
- [KHV⁺97] J. KARHUNEN, A. HYVARINEN, R. VIGARIO, J. HURRI et E. OJA. A class of neural blind separation to signal and image processing. In *ICASSP*, volume 1, pages 131–134, 1997.
- [KUL59] S. KULLBACK. *Information theory and statistics*. John Wiley and Sons, 1959.
- [LAC97] J.-L. LACOUME, P.O. AMBLARD et P. COMON. *Statistiques d'ordre supérieur pour le traitement du signal*. Masson, 1997.
- [LI02] R.-W. LIU et Y. INOUE. 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.
- [LMV95] L. DE LATHAUWER, B. DE MOOR et J. VANDEWALLE. Fetal electrocardiogram extraction by source subspace separation. In *HOS*, Aiguablava, Spain, June 1995.
- [LR93] P. LOUBATON et P. REGALIA. Blind deconvolution of multivariate signals : a deflation approach. In *Proceedings ICC*, pages 1160–1164, June 1993.

- [MOR99] E. MOREAU. A generalization of joint diagonalization criteria for source separation. *IEEE Trans. Signal Processing*, 1999.
- [MP97] E. MOREAU et J.-C. PESQUET. Generalized contrasts for multichannel blind deconvolution of linear systems. *IEEE Signal Processing Letters*, 4(6) :182–183, June 1997.
- [MPTM99] E. MOREAU, J.-F. PESQUET et N. THIRION-MOREAU. An equivalence between non symmetrical contrasts and cumulant matching for blind signal separation. In *First Int. Workshop Independent Component Analysis and Signal Separation*, Aussois, France, January 1999.
- [MS94] L. MOLGEDEY et H.G. SCHUSTER. Separation of independent signals using timedelayed correlations. *Phys. Reviews Letters*, 72(23) :3634–3637, 1994.
- [MTM99] E. MOREAU et N. THIRION-MOREAU. Nonsymmetrical contrasts for source separation. *IEEE Trans. Signal Processing*, 47 :2241–2252, 1999.
- [PD05] M. PUIGT et Y. DEVILLE. Time-frequency ratio-based blind separation methods for attenuated and time-delayed sources. *Signal Processing*, 19 :1348–1379, 2005.
- [PFA73] J. PFANZAGL. Asymptotic expansions related to minimum contrast estimators. *Annals of Statistics*, pages 993–1026, 1973.
- [PG97] D.-T PHAM et P. GARAT. Blind separation of mixtures of independent sources through a quasi maximum likelihood approach. *IEEE Transactions on Signal Processing*, 45(7) :1712–1725, 1997.
- [PHA99] D. T. PHAM. Mutual information approach to blind separation of stationary sources. In *Proceedings of ICA 99*, volume 9, pages 215–220, Aussois, France, January 1999.
- [PHA02] D.-T. PHAM. Mutual information approach to blind separation of stationary sources. *IEEE Transactions on Information Theory*, 48(7) :1935–1946, July 2002.

- [PM01] J.-C. PESQUET et E. MOREAU. Cumulant-based independence measures for linear mixtures. *IEEE Transaction on Information Theory*, 47(5) :1947–1956, 2001.
- [PS00] L. PARRA et C. SPENCE. Convolutional blind separation of non stationary sources. *IEEE Transactions on Speech and Audio Processing*, 8(3) :320–327, May 2000.
- [PSB03] D.-T. PHAM, C. SERVIÈRE et H. BOUMARAF. Blind separation of convolutional audio mixtures using nonstationarity. In *Proc. Int. Conf. Independent Component Analysis and Blind Source Separation (ICA)*, Nara, Japan, April 2003.
- [RBR01] S. RICKARD, R. BALAN et J. ROSCA. Real-time time-frequency based blind source separation. In *Proc. Int. Conf. Independent Component Analysis and Blind Source Separation ICA*, San Diego, USA, December 2001.
- [RCM06] S. RIHAOUI, M. CASTELLA et E. MOREAU. Contrast functions for blind source separation in a convolutional context. In *6th international conference on Independent Component Analysis and Blind Source Separation, ICA2006*, pages 327–334, Charleston, USA, March 2006.
- [SC91] A. SOULOUMIAC et J.-F. CARDOSO. Comparaison de méthodes de séparation de sources. *GRETSI*, pages 661–664, Septembre 1991.
- [SGO97] A. SWINDLEHURST, M. GORIS et B. OTTERSTEN. Some experiments with array data data collected in actual urban and suburban environments. *IEEE workshop on Signal Processing Advances in Wireless Communications*, 140 :301–304, April 1997.
- [SIM99] C. SIMON. *Séparation aveugle des sources en mélange convolutif*. PhD thesis, Université de Marne la Vallée, Novembre 1999.
- [SKI54] V. P. SKITOVICH. Linear forms of independent random variables and the normal distribution law. *Izvestiya Akademii Nauk SSSR. Seriya Matematicheskaya*, 18 :185–200, 1954.

- [SOU93] A. SOULOUMIAC. *Utilisation des statistiques d'ordre supérieur pour le filtrage et la séparation*. PhD thesis, Ecole Nationale Supérieure des Télécommunications (ENST), Paris, Février 1993.
- [SP04] C. SERVIÈRE et D.-T. PHAM. A novel method for permutation correction in frequency-domain in blind separation of speech mixtures. In *Proc. Int. Conf. Independent Component Analysis and Blind Source Separation (ICA)*, pages 807–815, Granada, Spain, 2004.
- [TC95] H.-L. NGUYEN THI et C. JUTTEN. Blind source separation for convolutive mixtures. *IEEE Trans. Signal Processing*, 45 :209–229, August 1995.
- [TFLT01] A. TOUZNI, I. FIJALKOW, M.-G. LARIMORE et J.-R. TREICHLER. A globally convergent approach for blind mimo adaptive deconvolution. *IEEE Transactions on Signal Processing*, 49(6) :1166–1178, June 2001.
- [THI93] H.-L. NGUYEN THI. *Séparation aveugle de sources large bande dans un mélange convolutif*. PhD thesis, Institut National Polytechnique de Grenoble, Janvier 1993.
- [TJ93] H.-L. NGUYEN THI et C. JUTTEN. Comparaison de quelques algorithmes adaptatifs de séparation aveugle de sources dans un mélange convolutif. In *Proc. 14ème colloque GRETSI*, pages 333–336, Juan les Pins, Septembre 1993.
- [TJ97] A. TALEB et C. JUTTEN. Entropy optimization, application to blind source separation. In *ICANN*, pages 529–534, Lausanne,Switzerland, October 1997.
- [TJ99] A. TALEB et C. JUTTEN. Source separation in post nonlinear mixtures. *IEEE Transactions on Signal Processing*, 47(10) :2807–2820, 1999.
- [TJC91] H.-L. NGUYEN THI, C. JUTTEN et J. CAELEN. Séparation aveugle de parole et de bruit dans un mélange convolutif. In *Proc. 13ème colloque GRETSI*, volume 45, pages 737–740, Juan les Pins, Septembre 1991.

- [TJC92] H.L. NGUYEN THI, C. JUTTEN et J. CAELEN. Speech enhancement :analysis and comparaison of methods in various real situations. In *Proc. International Conference EUSIPCO*, pages 303–306, Bruxelles, Belgium, Aout 1992.
- [TLSH91] L. TONG, R.-W. LIU, V.C. SOON et Y.-F. HUANG. Indeterminacy and identifiability of blind identification. *IEEE Trans. on Circuits and Systems*, 38(5) :499–509, 1991.
- [TMM00] N. THIRION-MOREAU et E. MOREAU. New criteria for blind signal separation. *IEEE Workshop on Statistical Signal and Array Processing*, pages 344–348, August 2000.
- [TUG97] J.-K. TUGNAIT. blind spatio-temporal equalization and impulse response estimation for mimo channels using a godard cost function. *IEEE Transactions on Signal Processing*, 45(1) :268–271, January 1997.
- [TUG99] J.-K. TUGNAIT. Adaptive blind source separation of convolutive mixtures of independent linear signals. *Signal Processing*, 73 :139–152, January 1999.
- [WEL67] P.D. WELCH. The use of fast fourier transform for the estimation of power spectra : A method based on time averaging over short, modified periodograms. *IEEE Trans. Audio Electroacoust.*, 15 :70–73, June 1967.
- [YR04] O. YILMAZ et S. RICKARD. Blind separation of speech mixtures via time-frequency masking. *IEEE Transactions on Signal Processing*, 52(7) :1830–1847, July 2004.
- [YW93] D. YELLIN et E. WEINSTEIN. Mutlichannel separation based on cross-bispectra. In *Proc. Workshop on HOS*, pages 270–274, Lake Taheo, USA, June 1993.
- [YW94] D. YELLIN et E. WEINSTEIN. Criteria for multichannel signal separation. *IEEE Transactions on Signal Processing*, 42(8) :2158–2168, August 1994.

Résumé

La séparation de sources consiste à extraire des signaux appelés sources à partir seulement de mélanges observés de ces signaux appelés observations. De nombreuses approches ont été brièvement présentées dans ce manuscrit. La plupart de ces approches exploite l'hypothèse d'indépendance des sources.

Dans cette thèse, nous avons considéré le cas des mélanges linéaires convolutifs. Nous proposons de calculer les contributions des sources sur les capteurs afin d'optimiser la procédure de séparation. L'estimation des contributions dans les observations est réalisée grâce à un critère quadratique optimisé par un filtre de Wiener. Ensuite, nous étudions deux approches de séparation de sources. La première utilise l'information mutuelle comme critère d'indépendance et la seconde dite fonction de contraste est basée sur les statistiques d'ordre quatre. L'utilisation des contributions des sources sur les capteurs dans la phase de séparation nous permet de proposer deux algorithmes de séparation, qui constituent deux généralisations d'algorithmes classiques.

Mots clés : Séparation de sources, Analyse en Composantes Indépendantes (ACI), Mélange convolutif, Filtre de Wiener, Information mutuelle, Fonction de contraste, Statistiques d'ordre supérieur, Diagonalisation conjointe.

Abstract

Blind Source Separation is a method for recovering a set of unknown source signals from the observations of their mixtures. We present a short review of source separation techniques. Most methods exploit the source independence hypothesis.

In this thesis, we focus on the separability of convolutive mixtures, expressed in time-domain. We propose to insert the contributions of sources on each sensor in the separation procedure. The estimation of these contributions in the observations is realized by the means of a quadratic criterion optimized by a Wiener filtering. We first propose a method based on the minimization of the mutual information and using wiener filtering. Then, we generalize a contrast function based on fourth-order statistics. In the convolutive context, we prove that testing the independence between the contributions of sources on the same sensor lead to separability, without testing the independence between delayed outputs.

Key words : Blind Source Separation (BSS), Independant Component Analysis (ICA), Convolutional mixture, Wiener filter, Mutual information, Contrast function, High order statistics, Joint diagonalisation.
