



HAL
open science

Contribution à la séparation de sources cyclo-stationnaires : application aux signaux de télécommunications, mécaniques et biomécaniques

Amine Brahmi

► **To cite this version:**

Amine Brahmi. Contribution à la séparation de sources cyclo-stationnaires : application aux signaux de télécommunications, mécaniques et biomécaniques. Traitement du signal et de l'image [eess.SP]. Université de Lyon; Université Sidi Mohamed ben Abdellah (Fès, Maroc), 2017. Français. NNT : 2017LYSES051 . tel-02110470

HAL Id: tel-02110470

<https://theses.hal.science/tel-02110470v1>

Submitted on 25 Apr 2019

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.



N° d'ordre NNT : 2017LYSES051

THESE de DOCTORAT DE L'UNIVERSITE DE LYON
opérée au sein de
l'Université Jean Monnet Saint-Etienne

Ecole Doctorale N° 488
Sciences Ingénierie Santé

Spécialité / discipline de doctorat :
Traitement de Signal

Soutenue publiquement le 30/11/2017, par :
Amine Brahmi

**CONTRIBUTION À LA SÉPARATION DE SOURCES CYCLO-
STATIONNAIRES : APPLICATION AUX SIGNAUX DE
TÉLÉCOMMUNICATIONS, MÉCANIQUES
& BIOMÉCANIQUES**

Devant le jury composé de :

Mme. EL AMRANI EL IDRISSI, Najiba	PES	Université Sidi Mohamed Ben Abdellah de Fès	Présidente
M. BELOUHRANI, Adel	PES	Ecole Nationale Polytechnique d'Alger	Rapporteur
M. MOREAU, Eric	PU	Université de Toulon	Rapporteur
M. ABDI, Farid	PES	Université Sidi Mohamed Ben Abdellah de Fès	Rapporteur
M. GUILLET, François	PU	Université Jean Monnet	Directeur de thèse
M. LAHBABI, M'hammed	PES	Université Sidi Mohamed Ben Abdellah de Fès	Directeur de thèse
M. CORBIER, Christophe	MCF HDR	Université Jean Monnet	Co-directeur de thèse
M. GHENNIQUI, Hicham	HDR	Université Sidi Mohamed Ben Abdellah de Fès	Co-directeur de thèse

À mes chers parents : Abdessalm & Khadija, pour tous leurs sacrifices, leur amour, leur soutien et leurs prières tout au long de mes études. ♡

À mon cher petit frère Youssef pour son soutien et son encouragement. ♡

Remerciements

CETTE thèse est le fruit d'un travail de réflexion à la fois laborieux et passionnant qui, sans le soutien et dévouement de certaines personnes, n'aurait pu être menée à son terme. Il est donc opportun en préambule de ce manuscrit d'exprimer ma gratitude à toute personne ayant contribué d'une manière ou d'une autre à la réalisation de ce travail.

Qu'il me soit permis en premier lieu d'adresser mes vifs remerciements à mes deux Directeurs de thèse, Monsieur **François GUILLET**, Professeur à l'Université Jean Monnet et Directeur adjoint du LASPI, ainsi que Monsieur **M'hammed LAHBABI**, Professeur à l'Université Sidi Mohamed Ben Abdellah et Directeur du LSSC, pour l'ample intérêt qu'ils ont accordé à mes travaux de recherches. Je leur suis très reconnaissant pour leur précieux conseils, leur aide incessante et leur confiance toujours renouvelée.

Tous mes sincères remerciements vont à mon co-encadrant de thèse, Monsieur **Hicham GHENNIQUI**, Professeur à l'Université Sidi Mohamed Ben Abdellah et Directeur adjoint du LSSC, pour m'avoir fait découvrir le domaine de la séparation de sources, pour l'assistance compétente qu'il m'a apportée grâce à son expérience et ses idées scientifiques ingénieuses. Il m'a sans cesse aidé dans les moments cruciaux à résoudre les problèmes auxquels j'ai été confrontés. Je lui témoigne ma très haute gratitude pour son grand cœur et sa patience.

Que mes chaleureux remerciements aillent à mon co-encadrant de thèse, Monsieur **Christophe CORBIER**, Professeur à l'Université Jean Monnet pour son appui scientifique et morale tout au long de ces trois dernières années. Il a toujours été là pour me soutenir et me conseiller dans les moments où les choses n'allaient pas très bien. Je le remercie également pour m'avoir fait découvrir le métier d'enseignement et pour son apport incontestable à la correction et l'amélioration de ce manuscrit.

Tous mes profonds remerciements vont à Monsieur **Mohamed El BA-**

DAOUI, Professeur à l'Université Jean Monnet et Directeur du LASPI, pour le grand intérêt qu'il a montré à mes travaux, ses remarques et analyses pertinentes lors des différents séminaires organisés au sein du LASPI. Veuillez trouver ici, le témoignage de mon admiration et de mon respect.

Ce travail n'aurait pas été possible sans le soutien financier du programme de bourses d'échanges et de coopération entre l'Europe et l'Afrique du Nord : **Erasmus Mundus - Al Idrisi II**, qui m'a permis, de me consacrer sereinement à ficeler cette thèse. Tous mes remerciements vont à Monsieur **Thomas GUILLOBEZ**, chargé des relations internationales à l'Université Jean Monnet pour sa disponibilité, ses promptes réponses relatives au programme et surtout pour sa bonne humeur.

Monsieur **Nabih NEJJAR**, Directeur de l'Institut Universitaire Technologique (IUT) de Roanne fût le premier à m'offrir la possibilité de dispenser mes premiers enseignements en tant qu'enseignant vacataire, puis tout récemment en tant qu'enseignant titulaire au sein de son établissement. Merci infiniment.

Mes remerciements vont à Monsieur **Salah KENNOUF**, Directeur du Centre Universitaire Roannais (CUR) qui a eu la gentillesse de m'accorder l'occasion d'intervenir en tant qu'enseignant vacataire dans sa structure.

Je remercie du fond du cœur tous les membres du LASPI et IUT pour leur constante sympathie, pour les encouragements qu'ils m'ont apportés tout au long de ces années et les moments agréables passés ensemble. J'en oublie sûrement encore et je m'en excuse. Encore mille mercis à tous pour m'avoir conduit à ce jour remarquable.

Roanne, le 10 Septembre 2017.

Avant-propos

LE travail présenté dans ce manuscrit a été réalisé conjointement dans les laboratoires : LASPI (Laboratoire d'Analyse de Signaux et de Processus Industriels), de l'Université Jean Monnet, Saint-Etienne, sous la direction de Monsieur François GUILLET, codirigée par Monsieur Christophe CORBIER et LSSC (Laboratoire des Signaux, Systèmes et Composants), de l'Université Sidi Moahmed Ben Abdellah, Fès, sous la direction de Monsieur M'hammed LAHBABI et codirigée par Monsieur Hicham GHENNIQUI.

Ce travail rentre dans le cadre d'une convention de co-tutelle franco-marocaine de thèse entre les deux universités financé par le programme de bourses d'échanges et de coopération entre l'Europe et l'Afrique du Nord : Erasmus Mundus - Al Idrisi II.

Le sujet de cette thèse consiste à développer de méthodes d'identification de mélange et de séparation aveugle de sources en exploitant les propriétés de cyclo-stationnarité. Le choix de cette considération s'explique par la nature cyclo-stationnaire des applications étudiées. En effet, les signaux de télécommunications, les signaux vibratoires issue de machines tournantes et les signaux biomécaniques présentent ces propriétés.

Liste des acronymes

De manière à rendre la lecture plus facile, nous avons conservé la version anglaise pour la plupart des acronymes au lieu de les traduire systématiquement en français.

ICA	Independent Component Analysis
PCA	Principal Component Analysis
SCA	Sparse Component Analysis
NMF	Non negative Matrix Factorization
DCM/T	Décomposition Conjointes de Matrices/Tenseurs
<i>c.à.d.</i>	<i>c'est-à-dire</i>
DC	Diagonalisation Conjointe
BDC	Bloc-Diagonalisation Conjointe
DVP	Décomposition en Valeurs Propres
SVD	Singular Value Decomposition
SAS	Séparation Aveugle de Sources
MIMO	Multiple-Input Multiple-Output
SIMO	Single-Input Multiple-Output
MU-MIMO	Multi User MIMO
FIR	Finite Impulse Response
ECG	ElectroCardioGramme
FECG	ElectroCardioGramme Fœtal
EEG	ElectroEncéphaloGramme
MEG	MagnétoEncéphaloGramme
EMG	ElectroMyoGramme
AMUSE	Algorithm for Multiple Unknown Signals Extraction
SOBI	Second Order Blind Identification
JADE	Joint Approximate Diagonalization of Eigen-matrices
BFGS	Broyden-Fletcher-Goldfarb-Shanno
AM	Amplitude Modulation
COM1	COntラスト Maximization 1
COM2	COntラスト Maximization 2
PAJOD	PArtial JOint Diagonalization
TDSEP	Temporal Decorrelation Source SEParation

RSOBI	Robust SOBI
RTFS	Représentation Temps-Fréquence-Spatiale
BGML	Block-Gaussian Maximum Likelihood
SEONS	Second-Order Non-stationary source Separation
CPD	Canonical Polyadic Decomposition
PARAFAC	PARAllel FACtorization
FOBI	Fourth Order Blind Identification
FOOBI	Fourth Order Only Blind Identification
ICAR	Independent Component Analysis using Redundancie)
BIRTH	Blind Identification of mixtures of sources using Redundancies in the daTa Hexacovariance matrix
BIOME	Blind Identification of Overcomplete Mixtures of sourcEs
FOBIUM	Fourth Order Blind Identification of Underdetermined Mixtures of sources
DEFA	DEFlation Approach for blind source separation
SCORE	Spectral Self-Coherence Restoral
CYCLOBLISS	CYCLOstationary BLInd Source Separation
CYMOD	CYclic Multiple Output Deconvolution
DSC	Densité Spectral Cyclique
DSCi	Densité Spectral Cyclique intégrée
FRS	Force de Reaction au Sol

Liste des notations

Afin de discerner au mieux la nature des grandeurs rencontrées tout au long de ce manuscrit, les matrices seront notées par des lettres majuscules en caractères gras, les vecteurs par des lettres minuscules en caractères gras, tandis que les scalaires seront représentés par des lettres minuscules en caractères simples. Pour le reste des notations, nous avons opté pour les conventions suivantes :

\mathbb{C}	ensemble des complexes
\mathbb{R}	ensemble des réels
\mathbb{Z}	ensemble des entiers relatifs
\Re	partie réelle d'un nombre complexe
\Im	partie imaginaire d'un nombre complexe
$\arg\{.\}$	argument d'un nombre complexe
$ \cdot $	module d'un nombre complexe
$\ \cdot\ _F$	norme de Frobenius ou de Hilbert-Schmidt
n	nombre de sources
m	nombre de capteurs
$\mathbf{s}(t)$	vecteur des sources de dimension $n \times 1$
$\hat{\mathbf{s}}(t)$	vecteur des sources estimées de dimension $n \times 1$
$\mathbf{r}(t)$	vecteur des bruits perturbateurs additifs de dimension $m \times 1$
$\mathbf{x}(t)$	vecteur des observations de dimension $m \times 1$
$\mathbf{x}_w(t)$	vecteur des observations blanchies de dimension $m \times 1$
ρ	ordre du filtre
ρ'	nombre d'observations retardées
$\mathbf{S}_i(t)$	vecteur contenant la i -ème source et ses versions retardées de dimension $\nu \times 1$ ($\nu = \rho + \rho'$)
$\mathbf{X}_j(t)$	vecteur contenant la j -ème observation et ses versions retardées de dimension $\rho' \times 1$
$\mathbf{R}_j(t)$	vecteur contenant le j -ème bruit et ses versions retardées de dimension $\rho' \times 1$
$\mathbf{S}(t)$	vecteur contenant les sources et ses versions retardées de dimension $N \times 1$ ($N = n\nu$)

$\hat{\mathbf{S}}(t)$	vecteur contenant les sources estimées et ses versions retardées de dimension $N \times 1$
$\mathbf{X}(t)$	vecteur contenant les observations et ses versions retardées de dimension $M \times 1$ ($M = m\rho'$)
$\mathbf{X}_w(t)$	vecteur contenant les observations blanchies et ses versions retardées de dimension $M \times 1$
$\mathbf{R}(t)$	vecteur contenant les bruits et ses versions retardées de dimension $M \times 1$
$\mathcal{A}\{\cdot\}$	application définie de \mathbb{C}^n vers \mathbb{C}^m
$a_{ij}(t)$	réponse impulsionnelle entre la j -ème source et le i -ème capteur
\mathbf{A}	matrice d'un système de mélange instantané de scalaires a_{ij}
$\mathbf{A}(t)$	matrice d'un système de mélange convolutif
$\hat{\mathbf{A}}$	matrice estimée de la matrice de mélange \mathbf{A}
\mathbf{W}	matrice de blanchiment de dimension $N \times M$
\mathbf{B}	matrice de séparation de dimension $N \times M$
$\hat{\mathbf{B}}$	matrice de séparation estimée de dimension $N \times M$
\mathbf{G}_{ij}	matrice de structure Toeplitz de dimension $\rho' \times \nu$ contenant les coefficients $a_{ij}(\ell)$; $\ell = 1, \dots, \rho$
\mathbf{G}	matrice bloc-Toeplitz de dimension $M \times N$
Δ	détermination d'un mélange
$\langle \cdot \rangle$	moyenne
$(\cdot)^{-1}$	opérateur inverseur
$(\cdot)^T$	opérateur transposé
$(\cdot)^H$	opérateur transconjugué
$(\cdot)^*$	opérateur conjugué complexe
$*$	produit de convolution
\otimes	produit de Kronecker
\odot	produit de Hadamard
$\mathbf{E}\{\cdot\}$	espérance mathématique
$\mathbf{C}_x(t, \tau)$	matrice de corrélation d'un vecteur $\mathbf{x}(t)$ de signaux aléatoires
\mathbf{C}_x^ξ	fonction d'autocorrélation cyclique
$\delta(\cdot)$	distribution de Dirac
$\text{tr}(\mathbf{M})$	trace de la matrice \mathbf{M}
$\text{rank}\{\mathbf{M}\}$	rang de la matrice \mathbf{M}
$\det\{\mathbf{M}\}$	déterminant de la matrice \mathbf{M}
$\mathbf{M}^\#$	pseudo-inverse de la matrice \mathbf{M}
$\text{diag}(\mathbf{m})$	matrice diagonale carrée contenant les éléments du vecteur \mathbf{m}

$\text{Diag}\{\cdot\}$	opérateur qui extrait la matrice diagonale de son argument
$\text{OffDiag}\{\cdot\}$	opérateur qui extrait la matrice zéro diagonale de son argument
$\text{Bdiag}_{(n_j)}\{\cdot\}$	opérateur qui extrait la matrice bloc-diagonale de son argument. Les r matrices bloc sont de dimension $n_j \times n_j$
$\text{OffBdiag}_{(n_j)}\{\cdot\}$	opérateur qui extrait la matrice zéro bloc-diagonale de son argument. Les r matrices bloc sont de dimension $n_j \times n_j$
\mathbf{I}_m	matrice identité de dimension $m \times m$
$\mathbf{1}_n$	matrice ne contenant que des 1 de dimension $n \times n$
$s^{(h)}(t)$	transformée d'Hilbert de $s(t)$
$s^{(a)}(t)$	représentation analytique de $s(t)$
$\text{Cum}\{\cdot\}$	opérateur cumulatif
$\text{vec}(\cdot)$	opérateur de vectorisation permet de concaténer les colonnes d'une matrice en un vecteur colonne
$D_{\text{KL}}\{\cdot\}$	divergence de Kullback-Leibler
\mathbf{T}_{Diag}	matrice de transformation définie par $\mathbf{T}_{\text{Diag}} = \text{diag}\{\text{vec}\{\text{Bdiag}_{(n)}\{\mathbf{1}_N\}\}\}$
\mathbf{T}_{Off}	matrice de transformation définie par $\mathbf{T}_{\text{Off}} = \mathbf{I}_{N^2} - \mathbf{T}_{\text{Diag}}$

Table des matières

REMERCIEMENTS	iii
AVANT-PROPOS	v
LISTE DES ACRONYMES	vi
LISTE DES NOTATIONS	viii
LISTE DES FIGURES	xiv
INTRODUCTION GÉNÉRALE	1
1 INTRODUCTION À LA SÉPARATION DE SOURCES ET À LA CYCLO-STATIONNARITÉ	1
1.1 Quelques applications de la SAS	2
1.1.1 Problème de la soirée cocktail	2
1.1.2 Systèmes de communications MIMO	3
1.1.3 Diagnostic mécanique	4
1.1.4 Génie biomédical	4
1.2 Modélisation du problème	6
1.3 Modèles de mélanges	7
1.3.1 Mélanges linéaires	7
1.3.2 Mélanges non-linéaires	9
1.4 Détermination des mélanges	10
1.5 Hypothèses classiques de SAS	10
1.5.1 Hypothèses sur les sources	10
1.5.2 Hypothèses sur le système de mélange	11
1.5.3 Hypothèses sur le bruit	11
1.6 Ambiguïtés de séparation	12
1.6.1 Ambiguïté de permutation	12
1.6.2 Ambiguïté d'échelle	13
1.6.3 Ambiguïté totale	14
1.7 Blanchiment à l'ordre deux	15

1.8	Critère d'évaluation des performances	17
1.9	Outils de la cyclo-stationnarité	17
1.9.1	Stationnarité au sens strict	18
1.9.2	Stationnarité au sens large	18
1.9.3	Cyclo-stationnarité au sens strict	18
1.9.4	Cyclo-stationnarité à l'ordre n	19
1.9.5	Cyclo-stationnarité au sens large	20
1.9.6	Poly-cyclo-stationnarité	20
1.9.7	Passage au domaine de Fourier	20
1.9.8	Exemples de processus cyclo-stationnaires	21
	Conclusion	26
2	TOUR D'HORIZON DES TECHNIQUES DE SÉPARATION DE SOURCES	27
2.1	SAS par analyse en composantes indépendantes	28
2.1.1	Contraste et mesure de l'indépendance	28
2.1.2	Lien avec la bloc diagonalisation conjointe matricielle	31
2.1.3	Présentation de quelques méthodes de SAS	35
2.2	SAS par analyse en composantes parcimonieuses	44
2.2.1	Représentations parcimonieuses	45
2.2.2	Principe des méthodes parcimonieuses	46
2.3	SAS par factorisation en matrices non-négatives	47
2.4	Discussions et conclusion	49
3	SÉPARATION DE SOURCES CYCLO-STATIONNAIRES : APPLICATION AUX SIGNAUX DE TÉLÉCOMMUNICATIONS	51
3.1	Séparation de mélanges convolutifs de sources cyclo-stationnaires basée sur la bloc-diagonalisation conjointe matricielle	52
3.1.1	Formulation du problème et hypothèses	52
3.1.2	Préliminaires théoriques	53
3.1.3	Procédure de sélection automatique des points retards-fréquences cycliques	56
3.1.4	Bloc-diagonalisation conjointe des matrices basée sur l'algorithme BFGS	59
3.1.5	Simulations et résultats numériques	63
3.2	Identification de la matrice de mélange dans le cas instantané	67
3.2.1	Modèle des sources et propriété importante	67
3.2.2	Approche proposée combinant la DVP et une méthode de regroupement hiérarchique	71

3.2.3	Simulations et résultats numériques	73
3.3	Conclusion	76
4	SÉPARATION DE SOURCES CYCLO-STATIONNAIRES : APPLI- CATION AUX SIGNAUX MÉCANIQUES ET BIOMÉCANIQUES	77
4.1	Séparation de signaux vibratoires mécaniques, application au cas de roulements	78
4.1.1	Généralités sur les roulements	78
4.1.2	Etude de la cyclo-stationnarité d'un roulement endommagé	81
4.1.3	Séparation de roulements endommagés réels	84
4.2	Séparation des contributions cyclo-stationnaires aux ordres 1 et 2, application aux signaux biomécaniques	92
4.2.1	Introduction	92
4.2.2	Formulation du problème	95
4.2.3	Préliminaires théoriques	95
4.2.4	Méthode proposée	97
4.2.5	Application sur des signaux FRS synthétiques	99
4.2.6	Application sur des signaux FRS réels	99
4.3	Conclusion	102
	CONCLUSION GÉNÉRALE	105
A	ANNEXES	108
A.1	Matrice de corrélation des observations	109
A.1.1	Matrice de corrélation de mélanges de sources aléatoires . .	109
A.1.2	Matrice de corrélation de mélanges de sources déterministes	109
A.2	Relation entre la densité spectrale cyclique intégrée et l'ana- lyse d'enveloppe	111
A.3	Expression des coefficients du polynôme d'ordre 4 impliqué dans le calcul du pas d'adaptation	112
A.3.1	Quelques propriétés et définitions importantes	112
A.3.2	Coefficients du polynôme d'ordre 4	113
A.4	Rappel sur la transformée de Hilbert	114
A.4.1	Définition	114
A.4.2	Propriétés	114
	LISTE DES PUBLICATIONS ASSOCIÉE À LA THÈSE	115
	BIBLIOGRAPHIE	117

Liste des figures

1.1	Schéma illustratif du problème de la soirée cocktail.	3
1.2	Système de transmission MIMO.	4
1.3	Exemple industriel où la SAS est appliquée.	5
1.4	Extraction de l'ECG du fœtus à partir de l'ECG de la mère. . .	5
1.5	Schéma général de la relation sources-observations en présence de bruit.	6
1.6	Schéma fonctionnel des méthodes basées sur le blanchiment.	16
1.7	Différentes représentations d'un signal.	22
1.8	Signal simulé $x(t)$ avec une fréquence $f_0 = 60\text{Hz}$ et une phase nulle.	23
1.9	Spectre d'amplitude du signal simulé $x(t)$	23
1.10	Amplitude de la fonction d'auto-corrélation cyclique du signal $x(t)$ comme une fonction de τ et α	24
1.11	Fonction d'auto-corrélation cyclique du signal $x(t)$ pour un retard nul.	25
1.12	Fonction d'auto-corrélation cyclique du signal $\sum_{k \in \mathbb{Z}} a(k)w(t - kT_s)$	26
2.1	Principe de séparation par analyse en composantes indépendantes.	28
2.2	Le problème de bloc diagonalisation conjointe en format tensoriel.	32
2.3	(a) Signal cosinus bruité dans le domaine temporel. (b) Une transformée en cosinus discrète révèle le comportement parcimonieux du signal.	45
2.4	Principe de séparation par analyse en composantes parcimonieuses.	47
3.1	Coutour de la matrice de corrélation cyclique des sources $C_s^\xi(\tau)$	64
3.2	Contour de la matrice de corrélation cyclique des sources $C_s^\xi(\tau)$	66
3.3	Évolution de l'indice de performances en fonction du RSB avec la matrice A_2	67

3.4	Évolution de l'indice de performances : (a) en fonction du RSB, (b) en fonction du nombre des matrices sélectionnées avec la matrice A_1	68
3.5	Évolution de l'indice de performances : (a) en fonction du RSB, (b) fonction du nombre des matrices sélectionnées. . . .	69
3.6	Évolution de l'indice de performances en fonction de la taille des signaux.	70
3.7	Contour de la matrice de corrélation cyclique des sources analytiques $C_{s^{(a)}}^{\xi}(\tau)$	75
3.8	Evolution de l'indice de performances en fonction du RSB. . .	75
4.1	Éclaté d'un roulement à billes.	79
4.2	Défaut de bague interne.	82
4.3	Signal vibratoire acquis d'un accéléromètre placé sur le palier comportant le roulement défaillant.	82
4.4	Spectre d'amplitude du signal de roulement.	83
4.5	Densité spectrale cyclique du roulement.	84
4.6	Densité spectrale cyclique intégrée du roulement.	85
4.7	Roulement à rotules sur rouleaux de type <i>22208-EK-C3-SKF</i> . .	86
4.8	Banc expérimental de roulement.	86
4.9	Position des capteurs accélérométriques.	87
4.10	Signaux vibratoires des observations : (a) observation 1, (b) observation 2.	88
4.11	Estimation des signaux sources : (a) source 1, (b) source 2. .	88
4.12	Densité spectrale cyclique intégrée : (a) observation 1, (b) observation 2.	89
4.13	Principe de la méthode proposée.	90
4.14	Densité spectrale cyclique intégrée : (a) source estimée 1, (b) source estimée 2.	91
4.15	Coureur sur le tapis roulant et exemple de signal mesuré . . .	93
4.16	Évolution de la FRS et cinématique des pieds pendant la course (a) avec l'appui talon (b) et l'appui en plante de pied.	94
4.17	Organigramme de la méthode proposée pour la séparation des composantes de la FRS.	98
4.18	(c) Signal $x(t)$ modélisant la FRS composé (a) d'une sinusoïde et (b) d'une modulation PAM.	100
4.19	Résultats de la séparation par l'OMP avec l'erreur quadratique moyenne : (a) partie active, (b) partie passive.	100
4.20	Résultats de la séparation par le LASSO avec l'erreur quadratique moyenne : (a) partie active, (b) partie passive.	101

4.21 Localisation des capteurs.	101
4.22 Signal FRS réel.	103
4.23 Résultats de la séparation des composantes de la FRS : (a) composante active, (b) composante passive.	103
4.24 Évolution de l'énergie du signal résiduel selon le nombre d'ité- rations.	104

Introduction générale

Cadre général et objectifs

EN mécanique, télécommunication et divers autres domaines, nous avons affaire lors de l'acquisition d'une grandeur physique sur un ensemble de capteurs à des mélanges auxquels agissent plusieurs contributions désirables ou indésirables. Ces mélanges nécessitent un pré-traitement afin d'en extraire l'information utile, en d'autres termes, les signaux d'intérêt, communément appelés sources. La séparation de sources consiste à estimer n sources à partir de m observations. Ces observations sont des mélanges de ces sources et proviennent de capteurs (accéléromètres, microphones, antennes...). On ne dispose a priori d'aucune information concernant la façon dont se sont mélangées les sources à travers le milieu de propagation. Les méthodes de séparation se basent généralement sur des hypothèses relatives à la nature des sources (stationnarité / non stationnarité / cyclo-stationnarité), l'indépendance statistique, parcimonie, etc. . .

La séparation aveugle de sources (SAS) est un des thèmes majeurs et ardu du traitement du signal. De plus, le développement de systèmes d'acquisition multi-voies, multi-capteurs, a renforcé l'intérêt de ces méthodes. La SAS a connu un essor considérable depuis les années 80 et continue à être un thème de recherche extrêmement actif, en témoigne la littérature très abondante sur le sujet. Ce problème a été initialement modélisé par les premiers acteurs de cette discipline par un produit mathématique simple. Ceci correspond au cas nommé instantané de la séparation de sources. Plus tard, cette modélisation s'est montrée inconvenable à toutes les situations pratiques, notamment lorsque les sources étaient éloignées des capteurs, ou si elles étaient à larges bandes. Ainsi, des modélisations plus réalistes ont été suggérées, supposant que le canal de transmission soit caractérisé par une application dépendante du temps et réalise un produit de convolution, afin d'engendrer les mélanges dits convolutifs.

Un grand nombre d’algorithmes pour la séparation de sources stationnaires et non stationnaires sont présentés dans la littérature, mais on retrouve cependant moins de travaux pour les sources cyclo-stationnaires que nous proposons de privilégier dans ce manuscrit. Ces problèmes de séparation de sources cyclo-stationnaires sont pourtant présents dans beaucoup de domaines tels que la mécanique où le cycle de fonctionnement des machines tournantes est à l’origine des propriétés de cyclo-stationnarité des signaux. C’est aussi le cas dans le domaine biomécanique où les signaux de pression générés lors de la marche humaine, ont un caractère cyclo-stationnaire. Pareillement, dans les systèmes de télécommunications les signaux présentent aussi des propriétés cyclo-stationnaires. Ceci peut être facilement vérifié par un simple examen de leurs fonctions auto-corrélations qui révèlent alors l’existence de périodicités, propriété qui caractérise un signal cyclo-stationnaire.

Le sujet de cette thèse a pour but de revisiter les méthodes de séparation de sources dans un contexte cyclo-stationnaire utilisant des approches basées sur les décompositions matricielles. Les outils développés seront validés sur des signaux synthétiques en télécommunications et des données expérimentales en mécanique et en biomécanique. Pour les signaux en mécanique, les méthodes proposées peuvent être utilisées pour caractériser un défaut potentiel créé par un organe défectueux dans une machine tournante à partir d’un signal recueilli par un accéléromètre. Pour les signaux en télécommunications, nous cherchons à identifier les systèmes MIMO FIR pilotés par des signaux de télécommunications de fréquences cycliques inconnues. Quant aux signaux biomécaniques, l’objectif est de séparer les composantes de la force de réaction au sol (contribution musculaire ou osseuse) lors de la course à pieds chez les sportifs de haut niveau. Cette séparation permettra aux médecins de proposer un suivi des activités physiques dans le but d’améliorer des performances des sportifs et de détecter les premiers signes de la fatigue musculaire ou des maladies neurologiques chez les personnes âgées.

Plan du manuscrit

Le document est organisé de la façon suivante :

- Le **premier chapitre** sera consacré aux généralités liées à la séparation de sources : introduction à travers des applications de la SAS, modélisation mathématique du problème de la SAS, nature des mé-

langes, déterminations, hypothèses considérées pour la résolution de problème d'intérêt, indéterminations inhérentes au problème de la SAS, opération de blanchiment présente dans certaines méthodes. Ce chapitre sera terminé par la présentation des descripteurs de la cyclo-stationnarité et quelques exemples de signaux cyclo-stationnaires.

- Dans le **deuxième chapitre**, nous ferons un tour d'horizon des méthodes de séparation classées selon trois grandes familles soit l'analyse en composantes indépendantes et décompositions conjointes de matrices/tenseurs, l'analyse en composantes parcimonieuses et la factorisation non-négative des matrices, et nous tenterons de répondre aux questions suivantes : quelles statistiques utilisent elles ? Dans quel domaine opèrent-elles ? Requièrent elles une étape de blanchiment de données ou non ? Quels schémas d'optimisation utilisent-elles ? Qu'appelle t-on hypothèse de cyclo-stationnarité ? Parcimonie ? Quel ajout peut apporter la contrainte de non négativité des matrices dans le cadre de la SAS ?

- Dans le **troisième chapitre**, nous présenterons deux nouvelles solutions pour la séparation de mélange linéaire et sur-déterminé de sources cyclo-stationnaires. La première solution traitera le cas de signaux partageant une ou plusieurs fréquences cycliques inconnues, basée sur la bloc-diagonalisation conjointe non unitaire d'un ensemble de matrices construit à partir des statistiques cycliques d'ordre deux par détection automatique. La diagonalisation utilise l'algorithme BFGS. Cette solution a donné lieu à une publication dans la revue *Mathematical Problems in Engineering* [42], un chapitre d'ouvrage dans *Lecture Notes in Computer Science* (Springer) [45] et 2 articles de conférences internationales avec comité de lecture et actes [46], [41].
La deuxième solution traitera le cas d'un mélange instantané de sources cyclo-stationnaires de fréquences cycliques inconnues et différentes. Le but est d'identifier la matrice de mélange à travers un processus itératif combinant la détection des matrices de rang un, la décomposition en éléments propres du produit de matrices sélectionnées et une méthode de regroupement hiérarchique. Les deux solutions sont appliquées aux signaux de télécommunications. Cette solution fait l'objet de 2 articles de conférences internationales avec comité de lecture et actes [43], [44].

-
- L'objectif du **quatrième chapitre** sera d'appliquer la SAS à des données issues de signaux mécaniques et biomécaniques. Nous testerons d'abord un des algorithmes proposés dans le troisième chapitre sur des signaux mécaniques découlant d'un banc d'essai de roulements défaillants. Ensuite, nous proposerons une méthode qui s'appuie sur l'analyse en composantes parcimonieuses pour séparer les composantes de la force de réaction au sol lors de la course à pieds chez les sportifs de haut niveau. En effet, ces composantes présentent des propriétés cyclo-stationnaires à l'ordre 1 et 2. Cette méthode a donné lieu à un article de conférence internationale avec comité de lecture et actes [111].

 - Une conclusion sur l'ensemble des résultats obtenus pour chaque voie d'étude, quelques perspectives de recherche qui se dégagent de ce travail, des annexes ainsi que des références bibliographiques mettront fin à ce manuscrit.

Introduction à la séparation de sources et à la cyclo-stationnarité



Sommaire

1.1	Quelques applications de la SAS	2
1.1.1	Problème de la soirée cocktail	2
1.1.2	Systèmes de communications MIMO	3
1.1.3	Diagnostic mécanique	4
1.1.4	Génie biomédical	4
1.2	Modélisation du problème	6
1.3	Modèles de mélanges	7
1.3.1	Mélanges linéaires	7
1.3.2	Mélanges non-linéaires	9
1.4	Détermination des mélanges	10
1.5	Hypothèses classiques de SAS	10
1.5.1	Hypothèses sur les sources	10
1.5.2	Hypothèses sur le système de mélange	11
1.5.3	Hypothèses sur le bruit	11
1.6	Ambiguïtés de séparation	12
1.6.1	Ambiguïté de permutation	12
1.6.2	Ambiguïté d'échelle	13
1.6.3	Ambiguïté totale	14
1.7	Blanchiment à l'ordre deux	15
1.8	Critère d'évaluation des performances	17
1.9	Outils de la cyclo-stationnarité	17
1.9.1	Stationnarité au sens strict	18

1.9.2	Stationnarité au sens large	18
1.9.3	Cyclo-stationnarité au sens strict	18
1.9.4	Cyclo-stationnarité à l'ordre n	19
1.9.5	Cyclo-stationnarité au sens large	20
1.9.6	Poly-cyclo-stationnarité	20
1.9.7	Passage au domaine de Fourier	20
1.9.8	Exemples de processus cyclo-stationnaires	21
	Conclusion	26

DANS ce chapitre, nous abordons les aspects théoriques sur lesquels la SAS est fondée. Nous présentons quelques applications réelles de la SAS, la modélisation mathématique, les modèles de mélanges et leur détermination, les hypothèses classiques faites souvent sur les sources, bruits et le système de mélange. Les indéterminations inhérentes au problème de la SAS, seront examinés en détail, y compris l'opération de blanchiment à l'ordre deux présente dans certaines méthodes de SAS. Ce chapitre inclura également quelques outils basés sur la cyclo-stationnarité.

1.1 Quelques applications de la SAS

Depuis sa naissance, la SAS a fait l'objet de nombreux travaux et développements. En effet les champs d'applications sont très larges, nous ciblerons dans cette section quelques applications faisant appel aux techniques de SAS.

1.1.1 Problème de la soirée cocktail

L'illustration classique de la SAS est le problème de la soirée cocktail comme le montre la Fig.(1.1). Lors d'une telle soirée, on dispose m microphones dans une salle dense, où n personnes discutent par groupes de tailles diverses. Chaque microphone enregistre la superposition des discours des personnes à ses alentours et le problème consiste à retrouver la voix de chaque personne. Cet exemple chez les humains peut-être également rencontré chez les animaux [128, 4, 15]. Les animaux doivent souvent identifier des compagnons, des descendants ou des adversaires dans des environnements encombrés, contenant de nombreux animaux qui vocalisent à la fois. Les grenouilles, par exemple, doivent utiliser des vocalisations conspécifiques pour localiser les compagnons de nuit, quand il y a peu de repères visuels, au milieu de centaines d'autres grenouilles.

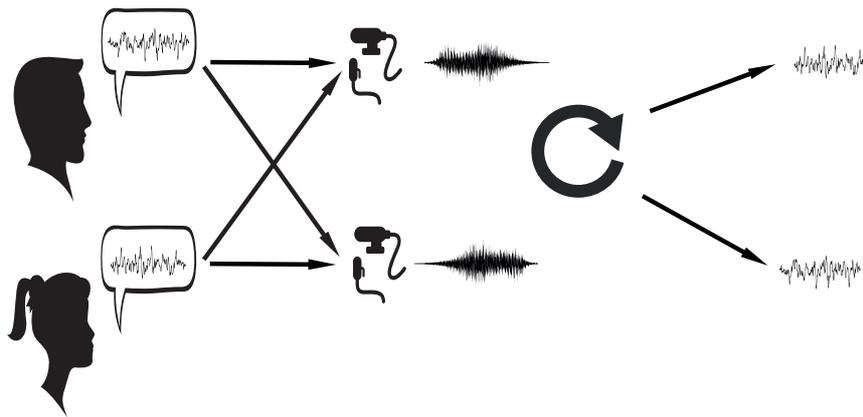


FIGURE 1.1 – Schéma illustratif du problème de la soirée cocktail.

Dans une soirée cocktail, il y a deux défis distincts pour un auditeur. Le premier est le problème de la ségrégation sonore. Les sons dans une scène auditive s'associent pour générer un signal qui pénètre les oreilles. Généralement, l'humain est intéressé particulièrement par une source sonore donnée (un partenaire potentiel, par exemple). Le système auditif doit dériver les propriétés des sons individuels du mélange entrant les oreilles. Le deuxième défi consiste à diriger l'attention sur la source sonore d'intérêt tout en ignorant les autres et à changer l'attention entre les sources en suivant intermittamment deux conversations. La plupart de nos processus cognitifs sont mono tâches, nous sélectionnons naturellement une source sonore particulière sur laquelle nous nous concentrons.

1.1.2 Systèmes de communications MIMO

Les systèmes de communications MIMO tels que les normes IEEE 802.11n (Wi-Fi), IEEE 802.11ac (Wi-Fi), HSPA + (3G), WiMAX (4G) et Long Term Evolution (4G) peuvent se modéliser exactement sous la forme du problème de SAS (voir Fig.(1.2)). Le cas des canaux multi-trajets où les signaux se propagent par plusieurs chemins et interagissent avec plusieurs objets physiques avant d'atteindre les antennes réceptrices correspond au mélange convolutif tandis que le cas simple trajet correspond au mélange instantané. Le processus $S(t)$ représente le vecteur des signaux transmis par des antennes émettrices après codage source et codage canal, les bruits $R(t)$ symbolisent toute perturbation additive reçue au niveau des antennes réceptrices. Les algorithmes de SAS sont des outils attrayants pour éliminer les interférences entre différents utilisateurs dans le cas des systèmes MU-MIMO (Multi User). Ils permettant de partager le débit radio et d'émettre des flux de données vers plusieurs utilisateurs, ou éliminer les interférences entre symboles pro-

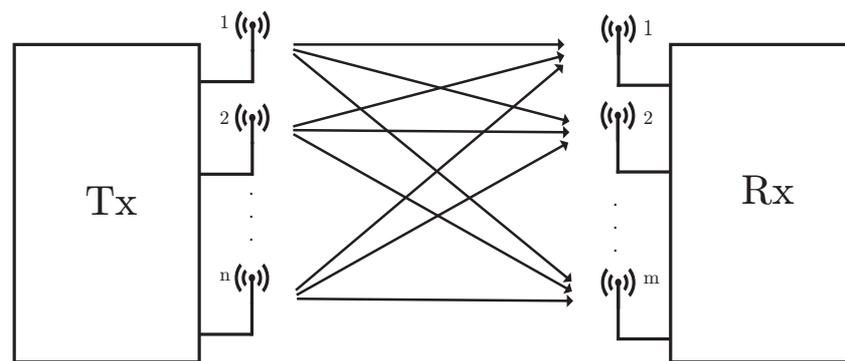


FIGURE 1.2 – Système de transmission MIMO.

venant des multi-trajets de propagation [169, 70, 194, 17].

1.1.3 Diagnostic mécanique

Dans le cas de diagnostic mécanique, nous nous intéressons à surveiller des systèmes souvent complexes, constitués de divers éléments, engrenages roulements, etc. . . Ces éléments peuvent être proches des uns des autres. Ainsi, le signal recueilli par un capteur placé à proximité est constitué de nombreuses composantes vibratoires, qui sont les réponses des divers éléments en mouvement constituant plusieurs sources. L'identification d'un défaut potentiel crée par un organe défectueux est quasiment impossible, lorsque le signal est couplé avec du bruit. La séparation ou l'extraction de sources peut donc être pertinente pour le diagnostic mécanique. Le but étant de détecter et de localiser le plus précocement possible les modes défaillants du système afin de prévenir un dysfonctionnement général. La SAS vise souvent dans ce cas à décomposer les observations selon leur nature à savoir périodique ou aléatoire (voir Fig.1.3). Les composantes périodiques sont liées au déséquilibre, défaut d'alignement et d'excentricité d'une machine. Tandis que les composantes aléatoires se réfèrent à des phénomènes transitoires impulsifs anormaux comme les défauts de roulements [38][13][32].

1.1.4 Génie biomédical

La SAS a été mise en œuvre pour répondre à des besoins cliniques en terme de diagnostic médical sur divers signaux électro-physiologiques acquis à l'aide d'un dispositif à électrode. Parmi ceux-ci, nous citons principalement :

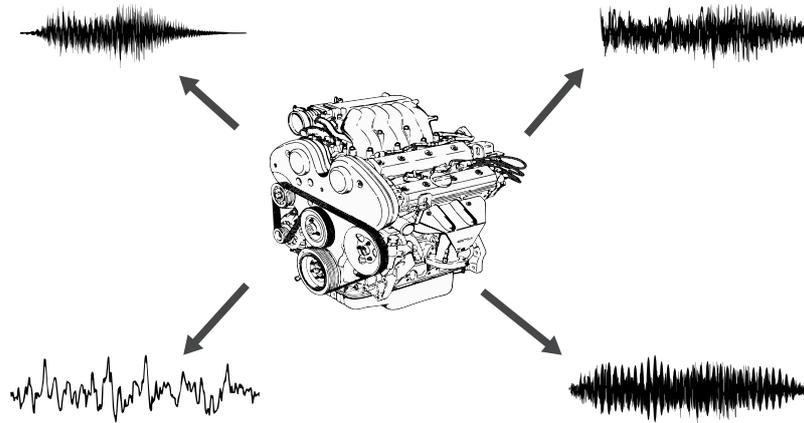


FIGURE 1.3 – Exemple industriel où la SAS est appliquée.

- l'ElectroCardioGramme (ECG) : un examen pratiqué dans le cadre du dépistage ou du diagnostic de différentes cardiopathies,
- l'ElectroEncéphaloGramme (EEG) : un examen permettant de mesurer l'activité électrique du cerveau,
- le MagnétoEncéphaloGramme (MEG) : une technique de mesure des champs magnétiques induits par l'activité électrique des neurones du cerveau,
- l'ElectroMyoGramme (EMG) : un examen qui permet enregistrer l'activité électrique spontanée d'un muscle ou d'un nerf.

Parmi les utilisations les plus notoires de la SAS dans le génie biomédical, nous trouvons l'extraction de l'ECG du fœtus à partir d'un mélange de signaux ECG maternel comme le montre la Fig.(1.4). Cela permet notamment en période intra-uterine fœtus de rassembler le maximum d'informations susceptibles de révéler d'éventuelles malformations ou problèmes cardiaques [83, 198, 200, 123, 143].

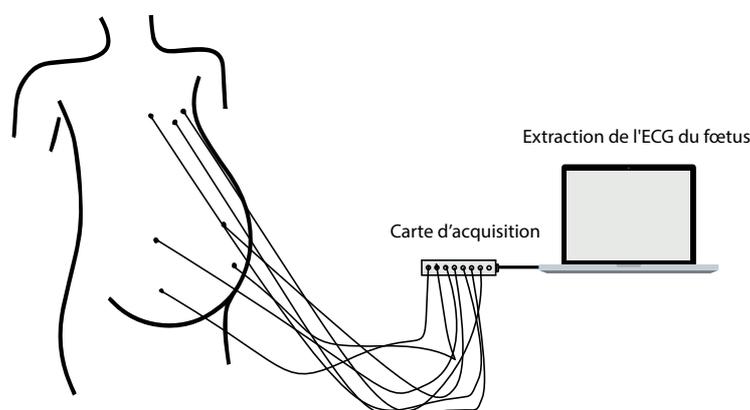


FIGURE 1.4 – Extraction de l'ECG du fœtus à partir de l'ECG de la mère.

1.2 Modélisation du problème

Indépendamment du champ d'application, la modélisation du problème de la SAS se fait de la manière suivante : des signaux émis par n émetteurs appelés sources, se propageant dans un milieu caractérisé par une matrice de mélange de dimension $m \times n$ et sont reçus par m capteurs. A la réception, nous avons accès seulement aux signaux reçus appelés observations. Nous ne disposons a priori d'aucune information concernant la façon dont les sources se sont mélangées à travers le milieu de propagation. L'objectif est de restituer les signaux émis en utilisant uniquement les observations. Mathématiquement, nous modélisons le problème de la manière suivante,

$$\mathbf{x}(t) = \mathcal{A}\{\mathbf{s}(t)\} + \mathbf{r}(t), \quad (1.1)$$

où

- $\mathbf{s}(t) = [s_1(t), s_2(t), s_3(t), \dots, s_n(t)]^T \in \mathbb{C}^n$ représente le vecteur des sources émises.
- $\mathbf{x}(t) = [x_1(t), x_2(t), x_3(t), \dots, x_m(t)]^T \in \mathbb{C}^m$ est le vecteur des observations.
- $\mathcal{A}\{\cdot\}$ est une application définie de \mathbb{C}^n vers \mathbb{C}^m .
- $\mathbf{r}(t) = [r_1(t), r_2(t), r_3(t), \dots, r_m(t)]^T \in \mathbb{C}^m$ est le vecteur des bruits perturbateurs additifs.
- n désigne le nombre de sources émises et m celui de capteurs.

La Fig.(1.5) symbolise cette relation.

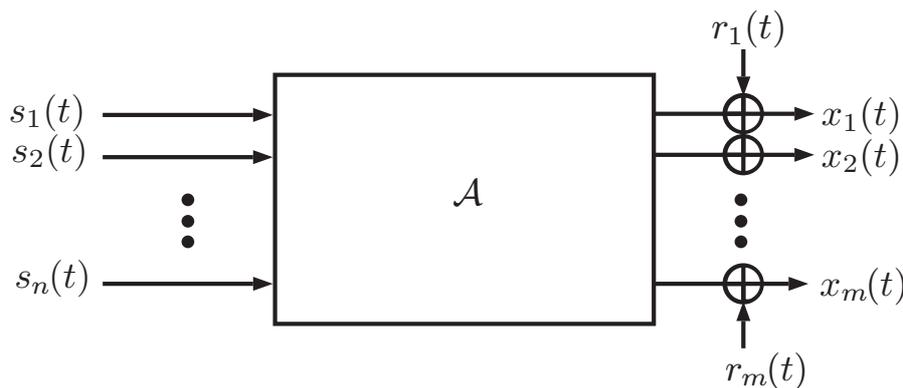


FIGURE 1.5 – Schéma général de la relation sources-observations en présence de bruit.

1.3 Modèles de mélanges

Selon la nature de l'opérateur \mathcal{A} , nous distinguons deux grandes catégories de mélanges à savoir : les mélanges linéaires et les mélanges non-linéaires.

1.3.1 Mélanges linéaires

Nous parlons d'un mélange linéaire lorsque les observations forment des combinaisons linéaires des sources. Il se décline en deux sous-catégories : le mélange convolutif et le mélange instantané.

Mélange convolutif

Le mélange *convolutif* est un système de filtres dont les réponses impulsionnelles permettent de modéliser les propagations le long de trajets multiples entre sources et observations. Ces filtres de mélange comportent des termes de mémoire, *c.à.d.* les signaux captés dépendent à la fois des sources et de leurs versions retardées. L'application $\mathcal{A}\{\cdot\}$ dans ce cas représente un filtre linéaire et invariant dans le temps. Ce modèle se traduit par l'équation suivante,

$$\forall i = 1, \dots, m, \quad x_i(t) = \sum_{j=1}^n \sum_{k=0}^{\rho} a_{ij}(k) s_j(t - k) + r_i(t), \quad (1.2)$$

ou encore sous forme matricielle

$$\begin{aligned} \mathbf{x}(t) &= \sum_{k=0}^{\rho} \mathbf{A}(k) \mathbf{s}(t - k) + \mathbf{r}(t), \\ &= \{\mathbf{A} * \mathbf{s}\}(t) + \mathbf{r}(t), \end{aligned} \quad (1.3)$$

où $*$ est le produit de convolution et $\mathbf{A}(t) = \sum_{i=1}^m \sum_{j=1}^n a_{ij}(t)$ représente la matrice des réponses impulsionnelles finies à temps discret des filtres d'ordre ρ , associés à chaque relation entre une source et une observation. En dehors de la linéarité et l'invariance temporelle, nous supposons souvent que ce système est stable et inversible.

Mélange instantané

Le mélange *instantané* est un cas particulier du modèle *convolutif* pour lequel chaque filtre composant la matrice \mathbf{A} ne présente pas de mémoire et

se réduit à un simple coefficient, réel ou complexe. Autrement dit :

$$\mathbf{A}(t) = \begin{cases} 0 & \text{si } \forall t \neq 0, \\ \mathbf{A} & \text{sinon.} \end{cases}$$

Ce modèle se traduit par l'équation suivante,

$$\mathbf{x}(t) = \mathbf{A}\mathbf{s}(t) + \mathbf{r}(t) \quad (1.4)$$

$$\begin{bmatrix} x_1(t) \\ x_2(t) \\ \vdots \\ x_m(t) \end{bmatrix} = \begin{bmatrix} a_{1,1} & a_{1,2} & \cdots & a_{1,n} \\ a_{2,1} & a_{2,2} & \cdots & a_{2,n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m,1} & a_{m,2} & \cdots & a_{m,n} \end{bmatrix} \begin{bmatrix} s_1(t) \\ s_2(t) \\ \vdots \\ s_n(t) \end{bmatrix} + \begin{bmatrix} r_1(t) \\ r_2(t) \\ \vdots \\ r_m(t) \end{bmatrix}$$

Transformation d'un modèle convolutif en un modèle instantané

Il existe deux méthodes permettant de reformuler le modèle convolutif en un modèle instantané. La première méthode se fait dans le domaine temporel, tel est le cas pour [35][36][93][94], en considérant trois tenseurs $\mathbf{S}(t)$, $\mathbf{X}(t)$ et $\mathbf{R}(t)$ définis respectivement comme suit :

$$\begin{aligned} \mathbf{S}(t) &= [\mathbf{S}_1(t)^T, \dots, \mathbf{S}_i(t)^T, \dots, \mathbf{S}_n(t)^T]^T, \\ \mathbf{X}(t) &= [\mathbf{X}_1(t)^T, \dots, \mathbf{X}_j(t)^T, \dots, \mathbf{X}_m(t)^T]^T, \\ \mathbf{R}(t) &= [\mathbf{R}_1(t)^T, \dots, \mathbf{R}_j(t)^T, \dots, \mathbf{R}_m(t)^T]^T, \end{aligned}$$

où pour $i = 1, \dots, n$:

$$\mathbf{S}_i(t) = [\mathbf{s}_i(t)^T, \dots, \mathbf{s}_i(t - \nu + 1)^T]^T,$$

et pour $j = 1, \dots, m$:

$$\begin{aligned} \mathbf{X}_j(t) &= [\mathbf{x}_j(t)^T, \dots, \mathbf{x}_j(t - \rho' + 1)^T]^T, \\ \mathbf{R}_j(t) &= [\mathbf{r}_j(t)^T, \dots, \mathbf{r}_j(t - \rho' + 1)^T]^T, \end{aligned}$$

où $[\cdot]^T$ est l'opérateur transposé, ρ' est un entier représentant le nombre de retards et $\nu = \rho + \rho'$. La matrice \mathbf{G} de dimensions $(M \times N)$ avec $M = m \times \rho'$ et $N = n \times \nu$ est une matrice par blocs (ou matrice partitionnée) qui peut être divisée en matrices rectangulaires de dimensions inférieures données par :

$$\mathbf{G} = \begin{bmatrix} \mathbf{G}_{11} & \cdots & \mathbf{G}_{1n} \\ \vdots & \ddots & \vdots \\ \mathbf{G}_{m1} & \cdots & \mathbf{G}_{mn} \end{bmatrix},$$

où

$$\mathbf{G}_{ij} = \begin{bmatrix} a_{ij}(0) & \dots & a_{ij}(\rho) & 0 & \dots & 0 \\ 0 & \ddots & \ddots & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & \ddots & \ddots & 0 \\ 0 & \dots & 0 & a_{ij}(0) & \dots & a_{ij}(\rho) \end{bmatrix},$$

sont des matrices dites de Toeplitz ayant pour dimensions $(\rho' \times \nu)$, dont les coefficients sur une diagonale descendant de gauche à droite sont les mêmes. Dans le cas où $\rho = 0$, $\rho' = 1$ et $\mathbf{G} = \mathbf{A}$, nous retombons alors sur un problème de SAS en contexte linéaire instantané. Finalement, l'équation (1.2) devient :

$$\mathbf{X}(t) = \mathbf{G}\mathbf{S}(t) + \mathbf{R}(t). \quad (1.5)$$

La méthode temporelle nécessite à la fois de connaître approximativement la longueur des filtres caractérisant le système ainsi que des échantillons importants. De plus, la méthode peut s'avérer couteuse en terme de calcul lorsque nous exploitons les statistiques d'ordre supérieur. La deuxième méthode transpose les observations dans le domaine fréquentiel afin de représenter le mélange convolutif par une suite de mélanges instantanés à chaque canal de fréquence. Le lecteur peut se référer aux travaux présentés dans [165] [150] [145] [155]. Cependant, la séparation dans le domaine fréquentiel fait apparaître deux indéterminations majeures. La première est l'ambiguïté de phase de la réponse en fréquence du système, la seconde concerne l'ambiguïté de permutation liée à l'identification de chaque mélange instantané fréquentiel.

1.3.2 Mélanges non-linéaires

En supposant que les non-linéarités du système soient faibles, ou que l'amplitude des signaux reste bornée, le modèle linéaire convolutif ou instantané est un modèle relativement admissible. Cependant, dans de nombreuses situations, il est impossible de remplir ces conditions approximatives. A titre d'exemple, en cas de saturation d'amplificateurs de conditionnement ou utilisation de capteurs en dehors des plages de linéarité. Il est donc nécessaire de considérer le mélange non-linéaire lors de la résolution du problème de la SAS. Les mélanges non-linéaires résultent d'une transformation-non linéaire des sources : $\mathbf{x} = F(\mathbf{s})$ où $F(\cdot)$ est une fonction non-linéaire quelconque. L'extrême complexité de ces mélanges limite son étude à des configurations bien particulières, comme par exemple le cas *post-non-linéaire* dans lequel le mélange est d'abord convolutif avant d'arriver sur les capteurs qui introduisent des effets non linéaires [61][18].

1.4 Détermination des mélanges

Nous définissons la détermination du mélange Δ comme la différence entre le nombre de sources et capteurs. Ainsi, trois type de mélanges se déclinent à partir de cette définition, à savoir,

- si $\Delta = 0$, alors il y a autant de sources que capteurs, nous parlons de mélanges **déterminés**. Dans le cas d'un mélange instantané non bruité, une fois la matrice de mélange estimée $\hat{\mathbf{A}}$, nous pouvons reconstruire les sources par simple inversion matricielle,

$$\hat{\mathbf{s}}(t) = \hat{\mathbf{A}}^{-1}\mathbf{x}(t). \quad (1.6)$$

- si $\Delta < 0$, alors le problème est **sur-déterminé**, *c.à.d.* qu'il y a plus de capteurs que de sources. Ce mélange ne présente aucune difficulté car nous pouvons aisément revenir à un mélange déterminé en ne choisissant que n observations sur les m disponibles.
- sinon le problème est **sous-déterminé**, *c.à.d.* qu'il y a plus de sources que de capteurs. Ce cas est le plus difficile à résoudre, puisque \mathbf{A} n'est pas inversible à gauche¹. Pour reconstruire les sources, les approches parcimonieuses peuvent être utilisées. Dans ce cas, les problèmes d'estimation de la matrice de mélange \mathbf{A} et de la restauration des sources sont deux problèmes bien distincts. Dans le cas du mélange instantané, même lorsque le mélange est parfaitement connu, les sources sont estimées à une indétermination près. Des informations a priori supplémentaires sur les sources comme la parcimonie, sont utiles pour fixer l'indétermination et fournir des solutions appropriées.

1.5 Hypothèses classiques de SAS

La séparation est dite *aveugle* si on est capable de séparer les signaux sans aucune information sur les paramètres du mélange et que les sources ne sont pas observées. Sans hypothèses supplémentaires la SAS apparaît comme un problème insoluble. Nous résumons dans ce paragraphe les hypothèses sur lesquelles la SAS est basée.

1.5.1 Hypothèses sur les sources

\mathcal{H}_1 . La majorité des approches de SAS supposent l'indépendance statistique et la décorrélation des sources. Dans ce cas, nous pouvons utiliser des

1. Si \mathbf{A} est une matrice de taille $(m \times n)$ de rang n alors elle admet une inverse gauche de taille $(m \times n)$ satisfaisant $\mathbf{BA} = \mathbf{I}_n$.

algorithmes de diagonalisation conjointe de matrices, qui sont extrêmement simples et efficaces.

- \mathcal{H}_2 . Généralement, on suppose l'ergodicité du vecteur des sources, ce qui assure que les statistiques peuvent être approchées par l'étude d'une seule réalisation suffisamment longue.
- \mathcal{H}_3 . Les pionniers de la SAS ont fondé leurs approches sur l'hypothèse de la stationnarité des sources. Néanmoins, il s'avère que dans plusieurs applications réelles comme les télécommunications, la mécanique et le génie biomédical, les sources sont très souvent non-stationnaires voir cyclo-stationnaires.
- \mathcal{H}_4 . Dans de nombreuses situations, les sources peuvent être considérées comme à valeurs discrètes ou, de façon plus générale, parcimonieuses dans un certain espace de représentation. Dans un tel espace, la recherche de solutions parcimonieuses conduit à des méthodes de résolution très efficaces, appelées analyse en composantes parcimonieuses, notamment pour la restauration de sources dans le cas de mélanges sous-déterminés.

1.5.2 Hypothèses sur le système de mélange

- \mathcal{H}_5 . La plupart des méthodes de SAS traitent des mélanges linéaires déterminés ou sur-déterminés *i.e* $n \leq m$. Dans le cas où nous traitons un mélange convolutif, il est nécessaire que ρ' soit choisi de telle façon à vérifier l'inégalité suivante :

$$N \leq M \Leftrightarrow n(\rho + \rho') \leq m\rho'.$$

- \mathcal{H}_6 . Pour garantir l'existence de la matrice de séparation dans le cadre des mélanges convolutifs (resp. instantanés), il faut que la matrice \mathbf{G} (resp. \mathbf{A}) soit de rang plein, *c.à.d.* $\text{rank}(\mathbf{G}) = N$ (resp. $\text{rank}(\mathbf{A}) = n$) .

1.5.3 Hypothèses sur le bruit

- \mathcal{H}_7 . Le vecteur $\mathbf{r}(t)$ contenant les bruits est stationnaire au sens strict.
- \mathcal{H}_8 . Les composantes de $\mathbf{r}(t)$ sont, statistiquement indépendantes ou décorrélées des composantes de $\mathbf{s}(t)$.
- \mathcal{H}_9 . Les composantes de $\mathbf{r}(t)$ sont blanches, gaussiennes et mutuellement indépendantes de moyenne nulle et de variance σ_r^2 . Cette hypothèse statistique classique présente un certain nombre d'avantages :

1. les cumulants d'ordre supérieur à 2 d'une variable aléatoire gaussienne sont tous nuls.
2. si un ensemble des variables aléatoires peut être décomposé en deux sous-ensembles indépendants, alors les cumulants croisés de cet ensemble sont tous nuls.
3. Dans ce cas la matrice de corrélation des bruits $\mathbf{C}_r(\tau)$ des bruits $r_i(t)$, $i = 1, \dots, m$ en instantané, et celle $\mathbf{C}_R(\tau)$ du vecteur $\mathbf{R}(t)$ en convolutif, contenant les bruits et leurs versions retardées peuvent s'écrire sous les formes suivantes :

$$\mathbf{C}_r(\tau) = \mathbf{E}\{\mathbf{r}(t)\mathbf{r}^H(t + \tau)\} = \sigma_r^2 \delta(\tau) \mathbf{I}_m \quad (1.7)$$

et,

$$\mathbf{C}_R(\tau) = \begin{bmatrix} \sigma_r^2 \tilde{\mathbf{I}}_{\rho'}(\tau) & 0_{\rho'} & \cdots & 0_{\rho'} \\ 0_{\rho'} & \sigma_r^2 \tilde{\mathbf{I}}_{\rho'}(\tau) & \cdots & 0_{\rho'} \\ \vdots & \vdots & \ddots & \vdots \\ 0_{\rho'} & 0_{\rho'} & \cdots & \sigma_r^2 \tilde{\mathbf{I}}_{\rho'}(\tau) \end{bmatrix} \quad (1.8)$$

où $0_{\rho'}$ est la matrice de dimension $\rho' \times \rho'$ contenant que des zéros, $\delta(\tau)$ désigne l'impulsion de Dirac, \mathbf{I}_m est la matrice identité de dimension $m \times m$, $(.)^H$ est l'opérateur trans-conjuguée et $\tilde{\mathbf{I}}_{\rho'}(\tau)$ est la matrice de dimension $\rho' \times \rho'$ contenant des éléments de valeur un sur la $\tau^{\text{ème}}$ diagonale supérieure si $0 \leq \tau < \rho'$ ou sur la $|\tau|^{\text{ème}}$ sous-diagonale si $-\rho' \leq \tau \leq 0$ et des zéros ailleurs.

Par conséquent, l'hypothèse gaussienne permet aux méthodes de SAS utilisant des cumulants d'ordre supérieur à 2 d'être non perturbées aux bruits gaussiens.

1.6 Ambiguïtés de séparation

1.6.1 Ambiguïté de permutation

Dans le cas d'une séparation aveugle où les sources et la matrice de mélange sont inconnus, les observations $\mathbf{X}(t)$ peuvent être générées à partir d'une infinité de sources $\mathbf{S}(t)$. En effet, l'ordre des signaux n'est pas garanti après la séparation étant donné que toute permutation appliquée sur les sources et sur les lignes de la matrice de mélange \mathbf{G} correspondantes donne le même vecteur $\mathbf{X}(t)$. Considérons par exemple un modèle de mélange linéaire instantané où nous permutons par exemple deux sources (*e.g.* d'indice

1 et 2) dans la matrice de mélange :

$$\forall i = 1, \dots, m, \quad \mathbf{X}_i(t) = \sum_{j=1}^n \mathbf{G}_{ij} \mathbf{S}_j(t) + \mathbf{R}_i(t). \quad (1.9)$$

En utilisant la propriété de commutativité de la multiplication et l'addition, nous pouvons écrire :

$$\begin{aligned} \mathbf{X}_i(t) &= \mathbf{G}_{i1} \mathbf{S}_1(t) + \mathbf{G}_{i2} \mathbf{S}_2(t) + \dots + \mathbf{G}_{in} \mathbf{S}_n(t) + \mathbf{R}_i(t) \\ &= \mathbf{G}_{i2} \mathbf{S}_2(t) + \mathbf{G}_{i1} \mathbf{S}_1(t) + \dots + \mathbf{G}_{in} \mathbf{S}_n(t) + \mathbf{R}_i(t), \end{aligned} \quad (1.10)$$

Il s'ensuit que :

$$\begin{aligned} \mathbf{X}(t) &= \mathbf{G}\mathbf{S}(t) + \mathbf{R}(t) \\ \begin{bmatrix} \mathbf{X}_1(t) \\ \mathbf{X}_2(t) \\ \vdots \\ \mathbf{X}_m(t) \end{bmatrix} &= \begin{bmatrix} \mathbf{G}_{1,2} & \mathbf{G}_{1,1} & \cdots & \mathbf{G}_{1,n} \\ \mathbf{G}_{2,2} & \mathbf{G}_{2,1} & \cdots & \mathbf{G}_{2,n} \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{G}_{m,2} & \mathbf{G}_{m,1} & \cdots & \mathbf{G}_{m,n} \end{bmatrix} \begin{bmatrix} \mathbf{S}_2(t) \\ \mathbf{S}_1(t) \\ \vdots \\ \mathbf{S}_n(t) \end{bmatrix} + \begin{bmatrix} \mathbf{R}_1(t) \\ \mathbf{R}_2(t) \\ \vdots \\ \mathbf{R}_m(t) \end{bmatrix} \end{aligned}$$

Cette permutation est modélisée par une matrice de permutation \mathbf{P} qui contient seulement un élément égal à 1 par ligne et par colonne. En notant la nouvelle matrice de mélange $\tilde{\mathbf{G}}$ et le nouveau vecteur des sources $\tilde{\mathbf{S}}(t)$, la matrice de mélange s'écrira :

$$\begin{aligned} \mathbf{X}(t) - \mathbf{R}(t) &= \tilde{\mathbf{G}}\tilde{\mathbf{S}}(t) \\ &= \mathbf{G}\tilde{\mathbf{P}}\tilde{\mathbf{S}}(t) \\ &= \mathbf{G}\mathbf{P}\mathbf{P}^{-1}\mathbf{S}(t) \\ &= \mathbf{G}\mathbf{I}\mathbf{S}(t) \\ &= \mathbf{G}\mathbf{S}(t). \end{aligned} \quad (1.11)$$

1.6.2 Ambiguïté d'échelle

En plus de l'ambiguïté de permutation, l'échange d'un facteur d'échelle entre un signal source et le vecteur colonne de la matrice de mélange \mathbf{A} correspondant n'affecte en rien les observations. Ceci est illustré à travers ce qui suit : $\forall i = 1, \dots, m, \forall \tau = 1, \dots, (\rho + \rho') + 1, \forall \lambda_{j \setminus j=1, \dots, n}^{(\tau)} \in \mathbb{C}^*$,

$$x_i(t) = \sum_{j=1}^n \sum_{k=0}^{\rho} \lambda_j^{(k)} a_{ij}(k) \frac{s_j(t-k)}{\lambda_j^{(k)}} + r_i(t). \quad (1.12)$$

En factorisant ces scalaires dans une matrice, nous obtenons une matrice bloc-diagonale inversible notée \mathbf{D} . Si $\dot{\mathbf{G}}$ est la nouvelle matrice de mélange et $\dot{\mathbf{S}}(t)$ est le nouveau vecteur des sources alors,

$$\begin{aligned} \mathbf{X}(t) - \mathbf{R}(t) &= \dot{\mathbf{G}}\dot{\mathbf{S}}(t) \\ &= \mathbf{GD}\dot{\mathbf{S}}(t) \\ &= \mathbf{GDD}^{-1}\mathbf{S}(t) \\ &= \mathbf{GIS}(t) \\ &= \mathbf{GS}(t), \end{aligned} \tag{1.13}$$

où

$$\dot{\mathbf{S}}(t) = \begin{bmatrix} \frac{s_1(t)}{\lambda_1^{(0)}} \\ \vdots \\ \frac{s_1(t - (\rho + \rho') + 1)}{\lambda_1^{((\rho + \rho') + 1)}} \\ \frac{s_n(t - (\rho + \rho') + 1)}{\lambda_n^{((\rho + \rho') + 1)}} \end{bmatrix}, \quad \dot{\mathbf{G}} = \begin{bmatrix} \dot{\mathbf{G}}_{11} & \dots & \dot{\mathbf{G}}_{1n} \\ \vdots & \ddots & \vdots \\ \dot{\mathbf{G}}_{1m} & \dots & \dot{\mathbf{G}}_{mn} \end{bmatrix},$$

avec

$$\dot{\mathbf{G}}_{ij} = \begin{bmatrix} \lambda_j^{(0)} a_{ij}(0) & \dots & \lambda_j^{(\rho)} a_{ij}(\rho) & 0 & \dots & 0 \\ 0 & \ddots & \ddots & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & \ddots & \ddots & 0 \\ 0 & \dots & 0 & \lambda_j^{(0)} a_{ij}(0) & \dots & \lambda_j^{(\rho)} a_{ij}(\rho) \end{bmatrix},$$

Nous constatons dans le meilleur des cas que les sources sont restituées à une permutation \mathbf{P} près et à un gain près en modèle convolutif (resp. une amplitude près en instantané, la matrice \mathbf{D} étant dans ce cas là une matrice diagonale inversible).

1.6.3 Ambiguïté totale

Dans le cas où nous associons les ambiguïtés, la solution donnée par la SAS s'écrit comme suit :

$$\begin{aligned} \hat{\mathbf{S}}(t) &= \mathbf{BS}(t) \\ &= \mathbf{BGS}(t) \\ &= \mathbf{PDS}(t). \end{aligned} \tag{1.14}$$

Cela donne :

$$\begin{aligned} \mathbf{BG} &= \mathbf{PD} \\ \Rightarrow \mathbf{B} &= \mathbf{PDG}^\#, \end{aligned} \tag{1.15}$$

où $(.)^\#$ est l'opérateur pseudo-inverse de Moore-Penrose. Ces ambiguïtés sont dues à la méconnaissance des sources et la matrice de mélange. Elles n'impactent pas l'indépendance des sources recherchées par le biais de l'optimisation de certains critères. Notons également que dans le cas des signaux complexes, une autre ambiguïté relative à la phase s'ajoute.

1.7 Blanchiment à l'ordre deux

Le blanchiment constitue une étape importante pour les méthodes de SAS basées sur l'hypothèse d'indépendance statistique, car il permet d'obtenir un vecteur de composantes indépendantes en décorrélant² les observations $\mathbf{X}(t)$ tout en imposant en plus l'unitarité de leur puissance. Cette opération est réalisée en appliquant une transformation \mathbf{W} (voir Fig.(1.6)) dite matrice de blanchiment au vecteur d'observations $\mathbf{X}(t)$ telle que :

$$\mathbf{X}_w(t) = \mathbf{W}\mathbf{X}(t), \quad (1.16)$$

\mathbf{X} est spatialement blanc³, *c.à.d.* que sa matrice de corrélation est de la forme suivante,

$$\begin{aligned} \mathbf{I}_N &= \mathbf{E}\{\mathbf{X}_w(t)\mathbf{X}_w^H(t)\} \\ &= \mathbf{W}\mathbf{C}_\mathbf{X}(t, \tau)\mathbf{W}^H \\ &= \mathbf{W}\mathbf{G}\mathbf{G}^H\mathbf{W}^H \\ &= \mathbf{U}\mathbf{U}^H, \end{aligned} \quad (1.17)$$

où $\mathbf{U} = \mathbf{W}\mathbf{G}$ est une matrice unitaire de dimension $(N \times N)$. Le filtrage des observations par la matrice \mathbf{W} permet de réduire la dimension de la matrice de mélange \mathbf{W} en passant de $(M \times N)$ à $(N \times N)$. Notons que la matrice de blanchiment n'est pas unique car la multiplication de la matrice de blanchiment par la gauche et d'une matrice unitaire ne change pas le résultat. Si \mathbf{W} est une matrice de blanchiment, $\mathbf{K}\mathbf{W}$ est aussi une matrice de blanchiment lorsque \mathbf{K} est une matrice unitaire.

Nous présentons dans l'Annexe (A.1) la définition de la matrice de corrélation des observations selon la nature de ces signaux. Ces définitions nous seront indispensable pour calculer la matrice de blanchiment \mathbb{W} .

La matrice de blanchiment est obtenue par le biais de la décomposition en valeurs singulières (SVD) des matrices $\mathbf{C}_\mathbf{X}(t, \tau)$. Il est à noter que dans certains algorithmes de séparation ce bloc est inexistant. Considérons le modèle

2. Généralement la décorrélation n'implique pas l'indépendance statistique sauf s'il s'agit d'un processus gaussien.

3. Un vecteur est dit blanc si ses composantes sont décorrélées et de puissance unitaire.

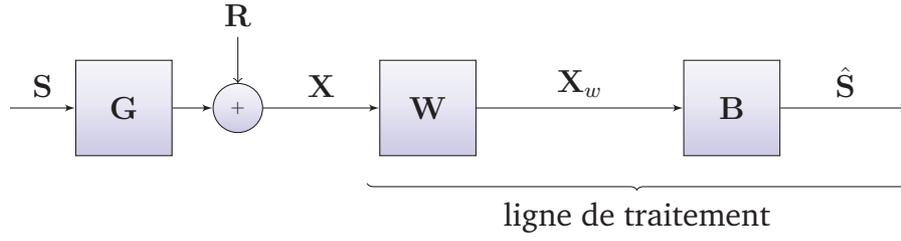


FIGURE 1.6 – Schéma fonctionnel des méthodes basées sur le blanchiment.

linéaire convolutif sur-déterminé, sous les hypothèses \mathcal{H}_6 et \mathcal{H}_9 et les sources ont une puissance unité $\mathbf{C}_S(t, 0) = \mathbf{I}_M$. La matrice de corrélation des observations s'écrit comme suit :

$$\begin{aligned} \mathbf{C}_X(t, 0) &= \mathbf{E}\{\mathbf{X}(t)\mathbf{X}(t)^H\} \\ &= \mathbf{A}\mathbf{C}_S(t, 0)\mathbf{A}^H + \mathbf{C}_R(t, 0) \\ &= \mathbf{A}\mathbf{A}^H + \sigma_r^2\mathbf{I}_M. \end{aligned} \quad (1.18)$$

La SVD de la matrice $\mathbf{C}_X(t, 0)$ s'écrit sous la forme suivante,

$$\mathbf{C}_X(t, 0) = \mathbf{\Gamma}\mathbf{\Sigma}\mathbf{\Gamma}^H = \begin{bmatrix} \mathbf{\Gamma}_S & \mathbf{\Gamma}_R \end{bmatrix} \begin{bmatrix} \mathbf{\Sigma}_S & 0 \\ 0 & \mathbf{\Sigma}_R \end{bmatrix} \begin{bmatrix} \mathbf{\Gamma}_S^H \\ \mathbf{\Gamma}_R^H \end{bmatrix} \quad (1.19)$$

$$= \mathbf{\Gamma}_S\mathbf{\Sigma}_S\mathbf{\Gamma}_S^H + \mathbf{\Gamma}_R\mathbf{\Sigma}_R\mathbf{\Gamma}_R^H, \quad (1.20)$$

où $\mathbf{\Sigma}_S$ est une matrice diagonale de dimension $N \times N$ contenant les N plus grandes valeurs propres de $\mathbf{C}_X(t, 0)$, rangées dans l'ordre décroissant et donnée par $\mathbf{\Sigma}_S = \text{diag}(\lambda_1, \dots, \lambda_N)$, avec $\mathbf{\Gamma}_S = \text{diag}(u_1, \dots, u_N)$ matrice des vecteurs singuliers correspondants engendrant le sous-espace signal. La matrice $\mathbf{\Sigma}_R = \text{diag}(\lambda_{N+1}, \dots, \lambda_M)$ contient les $(M - N)$ valeurs singulières restantes et $\mathbf{\Gamma}_R = \text{diag}(u_{N+1}, \dots, u_M)$ contient les $(M - N)$ vecteurs singuliers correspondants, ils engendrent le sous-espace bruit.

Dans le cas où $M = N$ la matrice de blanchiment est définie comme,

$$\mathbf{W} = \mathbf{\Sigma}^{-\frac{1}{2}}\mathbf{\Gamma}^H. \quad (1.21)$$

Si $M > N$, la matrice de blanchiment est calculée en utilisant seulement le sous-espace signal comme,

$$\mathbf{W} = \mathbf{\Sigma}_S^{-\frac{1}{2}}\mathbf{\Gamma}_S^H. \quad (1.22)$$

Nous pouvons avoir une estimation de la variance du bruit $\hat{\sigma}_r^2$ en faisant la moyenne des $M - N$ plus petites valeurs propres de $\mathbf{C}_X(t, 0)$,

$$\hat{\sigma}_r^2 = \frac{1}{M - N} \sum_{i=N+1}^M \lambda_i. \quad (1.23)$$

Maintenant que nous avons calculé la matrice de blanchiment, le vecteur des observations de dimension $(N \times 1)$ s'écrit de la façon suivante :

$$\begin{aligned}\mathbf{X}_w(t) &= \mathbf{W}\mathbf{X}(t) \\ &= \mathbf{W}\mathbf{G}\mathbf{S}(t) + \mathbf{W}\mathbf{R}(t) \\ &= \mathbf{U}\mathbf{S}(t) + \mathbf{W}\mathbf{R}(t).\end{aligned}\tag{1.24}$$

La séparation passe donc par l'estimation de la matrice unitaire \mathbf{U} .

1.8 Critère d'évaluation des performances

Afin d'évaluer la qualité de l'estimation de la matrice de mélange, nous pouvons mesurer l'indice de performance proposé dans le cadre des mélanges convolutifs par Ghennioui et al. dans [104], [102], généralisant celui introduit dans [137],[108] pour les mélanges instantanés :

$$I_{\text{perf}} = \frac{1}{\nu(\nu - 1)} \left[\sum_{i=1}^{\nu} \left(\sum_{j=1}^{\nu} \frac{\|(\mathbf{Z})_{i,j}\|_F^2}{\max_{\ell} \|(\mathbf{Z})_{i,\ell}\|_F^2} - 1 \right) + \sum_{j=1}^{\nu} \left(\sum_{i=1}^{\nu} \frac{\|(\mathbf{Z})_{i,j}\|_F^2}{\max_{\ell} \|(\mathbf{Z})_{\ell,j}\|_F^2} - 1 \right) \right],\tag{1.25}$$

où $\nu = \rho + \rho'$, $(\mathbf{Z})_{i,j} \forall i, j \in \{1, \dots, \nu\}$ est la (i, j) -ème matrice bloc (carrée) de $\mathbf{Z} = \hat{\mathbf{B}}\mathbf{G}$. Cet indice mesure l'écart entre $\hat{\mathbf{B}}\mathbf{G}$ et le produit $\mathbf{P}\mathbf{D}$ d'une matrice de permutation par une matrice bloc-diagonale. Une bonne séparation est obtenue lorsque la matrice \mathbf{Z} est égale à une matrice contenant un seul bloc non-nul par bloc-ligne et par bloc-colonne (resp. en instantané, cela signifie qu'un seul terme non-nul est obtenu par ligne et par colonne). Une séparation parfaite correspond à un indice I_{perf} tendant vers $-\infty$ en échelle logarithmique (0 en échelle linéaire). Nous rappelons que lorsque $\rho = 0$, $\rho' = 1$ et $\mathbf{G} = \mathbf{A}$, nous retrouvons alors bien l'indice de performances qui détermine la qualité d'identification de la matrice de mélange en instantané.

1.9 Outils de la cyclo-stationnarité

La stationnarité fixe la propriété d'invariance par rapport au temps t des lois de probabilité qui caractérisent le processus stochastique. Elle est une propriété très forte qui a défini un bon nombre de méthodes de traitement de divers processus réels. Dès lors que cette propriété n'est pas vérifiée, des complications naissent quand nous définissons la méthode d'analyse. Toutefois, certains processus réels non-stationnaires issus de diverses applications tels que les communications numériques ou les vibrations des ma-

chines tournantes, ont des propriétés statistiques qui présentent des périodicités ou quasi-périodicités. Cette propriété n'est rien d'autre que la cyclo-stationnarité dont l'utilisation aujourd'hui est en pleine expansion, permettant entre autres, de généraliser la plupart des résultats obtenus dans le contexte stationnaire au contexte cyclo-stationnaire (pour plus de détails sur la notion de cyclo-stationnarité, nous renvoyons le lecteur aux références suivantes [99, 142]). Dans cette section, nous commençons par rappeler la stationnarité puis nous définirons quelques notions liées à la théorie de cyclo-stationnarité puis nous citerons des exemples de signaux cyclo-stationnaires.

1.9.1 Stationnarité au sens strict

Un processus stochastique $\{\mathbf{X}(t)\}_{t \in \mathbb{R}}$ complexe ou réel est dit **stationnaire au sens strict** si sa fonction de répartition est invariante pour tout changement de l'origine du temps,

$$\mathbf{F}_{\mathbf{X}}(v_1, \dots, v_n; t_1, \dots, t_n) = \mathbf{F}_{\mathbf{X}}(v_1, \dots, v_n; t_1 + \tau, \dots, t_n + \tau) \quad \forall \tau \in \mathbb{R}, \quad (1.26)$$

1.9.2 Stationnarité au sens large

La stationnarité au sens large⁴ est une notion plus pratique qui est définie à l'ordre deux et généralement à partir des deux premiers moments du processus. $\{\mathbf{X}(t)\}_{t \in \mathbb{R}}$ est un processus stochastique **stationnaire au sens large** si,

$$\begin{cases} m_{\mathbf{X}}(t) = m_{\mathbf{X}}, \\ \mathbf{C}_{\mathbf{X}}(t, \tau) = \mathbf{C}_{\mathbf{X}}(\tau). \end{cases} \quad (1.27)$$

où $\mathbf{C}_{\mathbf{X}}(t, \tau)$ est la matrice d'auto-corrélation au retard τ .

1.9.3 Cyclo-stationnarité au sens strict

Soit $\{\mathbf{X}(t)\}_{t \in \mathbb{R}}$ un processus stochastique complexe ou réel. $\{\mathbf{X}(t)\}_{t \in \mathbb{R}}$ est dit **cyclo-stationnaire au sens strict** si sa fonction de répartition est périodique de période T_0 ,

$$\mathbf{F}_{\mathbf{X}}(v_1, \dots, v_n; t_1, \dots, t_n) = \mathbf{F}_{\mathbf{X}}(v_1, \dots, v_n; t_1 + T_0, \dots, t_n + T_0) \quad \forall n \in \mathbb{N}^* \quad (1.28)$$

que l'on peut écrire de manière plus compacte,

$$\mathbf{F}_{\mathbf{X}}(\mathbf{v}; \mathbf{t}) = \mathbf{F}_{\mathbf{X}}(\mathbf{v}; \mathbf{t} + \mathbf{T}_0). \quad (1.29)$$

4. Il est à noter qu'un processus stationnaire au sens strict est stationnaire au sens large mais la réciproque n'est pas nécessairement vraie.

avec $\mathbf{T}_0 = [T_0, \dots, T_0]$, $\mathbf{t} = [t_1, \dots, t_n]$ et $\mathbf{v} = [v_1, \dots, v_0]$.

1.9.4 Cyclo-stationnarité à l'ordre n

$\{\mathbf{X}(t)\}_{t \in \mathbb{R}}$ est dit **cyclo-stationnaire à l'ordre n** si son moment d'ordre n existe et périodique de période \mathbf{T}_0 ,

$$\mu_{\mathbf{X}(n)}(\mathbf{t}) = \mu_{\mathbf{X}(n)}(\mathbf{t} + \mathbf{T}_0). \quad (1.30)$$

Cas particuliers : ordre 1 et 2

La cyclo-stationnarité à l'ordre 1, se traduit par la périodicité de la moyenne,

$$m_{\mathbf{X}}(t) = \mathbf{E}\{\mathbf{X}(t)\} = m_{\mathbf{X}}(t + T_0). \quad (1.31)$$

où $\mathbf{E}\{\cdot\}$ est l'espérance mathématique. $\{\mathbf{X}(t)\}_{t \in \mathbb{R}}$ est dit cyclo-stationnaire à l'ordre 2 si sa fonction d'auto-corrélation est périodique de période $T_0 = \frac{1}{\alpha_0}$,

$$\mathbf{C}_{\mathbf{X}}(t, \tau) = \mathbf{E}\left\{\mathbf{X}\left(t - \frac{\tau}{2}\right)\mathbf{X}^H\left(t + \frac{\tau}{2}\right)\right\} \quad (1.32)$$

$$= \mathbf{C}_{\mathbf{X}}(t + T_0, \tau). \quad (1.33)$$

Dans ce cas, l'auto-corrélation admet un développement en séries de Fourier selon la variable t comme suit,

$$\mathbf{C}_{\mathbf{X}}(t, \tau) = \sum_{\alpha \in \Psi} \mathbf{C}_{\mathbf{X}}^{\alpha}(\tau) e^{j2\pi\alpha t}, \quad (1.34)$$

où α est appelée fréquence cyclique et $\Psi = \{\alpha = \frac{k}{T_0}, k \in \mathbb{Z}\}$ représente l'ensemble des fréquences cycliques (la fréquence fondamentale et ses harmoniques). Notons que la somme dans l'équation (1.34) est effectuée sur les multiples entiers de la fréquence fondamentale $\alpha_0 = \frac{1}{T_0}$ et les coefficients de Fourier $\mathbf{C}_{\mathbf{X}}^{\alpha}$ sont appelés les **fonctions d'auto-corrélations cycliques**

$$\mathbf{C}_{\mathbf{X}}^{\alpha}(\tau) = \lim_{T_0 \rightarrow \infty} \frac{1}{T_0} \int_{-\frac{T_0}{2}}^{\frac{T_0}{2}} \mathbf{C}_{\mathbf{X}}(t, \tau) e^{-j2\pi\alpha t} dt. \quad (1.35)$$

Remarques importantes :

- La fonction d'auto-corrélation cyclique $\mathbf{C}_{\mathbf{X}}^{\alpha}(\tau)$ est continue par rapport à la variable τ et discrète par rapport à la variable α .

- Si la fonction d'auto-corrélation cyclique $C_{\mathbf{X}}^{\alpha}$ est nulle pour toutes les valeurs discrètes de α sauf pour $\alpha = 0$, l'auto-corrélation ne dépend pas de t et le processus est stationnaire (puisque la valeur de fonction d'auto-corrélation cyclique obtenue pour la fréquence cyclique nulle est l'auto-corrélation du processus, *c.à.d.* l'auto-corrélation de sa composante stationnaire).
- Si la fonction d'auto-corrélation cyclique est non nulle pour des valeurs discrètes de α , le processus est cyclo-stationnaire.

1.9.5 Cyclo-stationnarité au sens large

$\{\mathbf{X}(t)\}_{t \in \mathbb{R}}$ est dit **cyclo-stationnaire au sens large**, s'il est cyclo-stationnaire conjointement à l'ordre un et à l'ordre deux.

1.9.6 Poly-cyclo-stationnarité

$\{\mathbf{X}(t)\}_{t \in \mathbb{R}}$ est dit **poly-cyclo-stationnaire** ou **presque cyclo-stationnaire**, s'il existe plusieurs périodes $T_1, \dots, T_k, \dots, T_K$ non multiples les unes par rapport aux autres pour lesquelles $\{\mathbf{X}(t)\}$ est cyclo-stationnaire.

1.9.7 Passage au domaine de Fourier

Dans le domaine de Fourier, les signaux cyclo-stationnaires sont caractérisés par la Densité Spectrale Cyclique (DSC), qui selon le théorème de Wiener-Khinchin-Einstein n'est rien d'autre que la transformée de Fourier de la fonction d'auto-corrélation cyclique définie comme suit :

$$\text{DSC}_{\mathbf{X}}^{\alpha}(f) = \sum_{\tau} C_{\mathbf{X}}^{\alpha}(\tau) e^{-j2\pi f\tau}. \quad (1.36)$$

L'intégration de la DSC sur f donne la densité spectrale cyclique intégrée (DSCi) :

$$\text{DSCi}_{\mathbf{X}}(\alpha) = \int_{\mathbb{R}} \text{DSC}_{\mathbf{X}}^{\alpha}(f) df. \quad (1.37)$$

La DSCi peut être obtenue également en calculant la transformée de Fourier du module au carré du signal étudié^{5 6} :

$$\text{DSCi}_{\mathbf{X}}(\alpha) = \lim_{T \rightarrow \infty} \frac{1}{T} \int_{-\frac{T}{2}}^{\frac{T}{2}} \mathbf{E}[|\mathbf{X}(t)|^2] e^{-j2\pi\alpha t} dt. \quad (1.38)$$

5. Ce résultat est exclusivement valide pour les signaux à moyenne nulle. La preuve est donnée dans l'Annexe (A.2).

6. Selon Randall et Antoni, il y a une équivalence entre la DSCi et l'analyse spectrale de l'enveloppe du signal étudié [156].

Pour l'étude d'un processus donné, quatre domaines d'analyse sont disponibles comme cela est décrit dans la Fig.(1.7). Le choix d'un domaine d'analyse dépend de la nature du processus. Autrement dit, nous utilisons l'analyse de la fonction d'auto-corrélation instantanée (dépend du couple temps - retard (t, τ)) ou encore du spectre de Wigner-Ville (dépend du couple temps - fréquence (t, f)) pour les processus non-stationnaires⁷. Cependant, nous employons l'analyse de la fonction d'auto-corrélation cyclique (dépend du couple fréquence cyclique - retard (α, τ)) ou de la densité spectrale cyclique (dépend du couple fréquence cyclique - fréquence (α, f)) ou la densité spectrale cyclique intégrée (dépend seulement du retard α) pour les processus cyclo-stationnaires. Ces deux dernières représentations sont particulièrement importantes car elles présentent dans le cas d'un processus cyclo-stationnaire des raies dans le domaine de Fourier en les fréquences cycliques et leurs multiples entiers.

1.9.8 Exemples de processus cyclo-stationnaires

Exemple 1 : modulation d'amplitude (AM)

La plupart des signaux modulés sont cyclo-stationnaires car les opérations classiques du traitement de signal comme l'échantillonnage, le multiplexage, la modulation, etc. . . , créent des périodicités dans le signal qui peuvent les rendre cyclo-stationnaires. Certains signaux de télécommunication non périodiques (à l'ordre 1), ont une énergie périodique [98]. Soit $x(t)$ un signal modulé en amplitude tel que,

$$x(t) = a(t) \cos(2\pi f_0 t + \theta), \quad (1.39)$$

où $a(t)$ est signal stationnaire aléatoire avec une phase nulle et une fréquence porteuse $f_0 = 60\text{Hz}$ ayant les propriétés suivantes,

$$\begin{aligned} \langle a(t) \rangle &= 0 \\ \langle a(t)a^H(t - \tau) \rangle &\neq 0 \\ \langle a(t)a^H(t - \tau)e^{-j2\pi\alpha t} \rangle &\neq 0, \quad \forall \alpha \neq 0, \end{aligned}$$

ou encore,

$$\langle a(t)e^{-j2\pi\alpha t} \rangle \neq 0, \quad \forall \alpha \neq 0,$$

7. Dans le cas des processus stationnaires, nous utilisons la fonction d'auto-corrélation qui dépend seulement du retard τ ou la densité spectrale de puissance qui dépend seulement de la fréquence f .

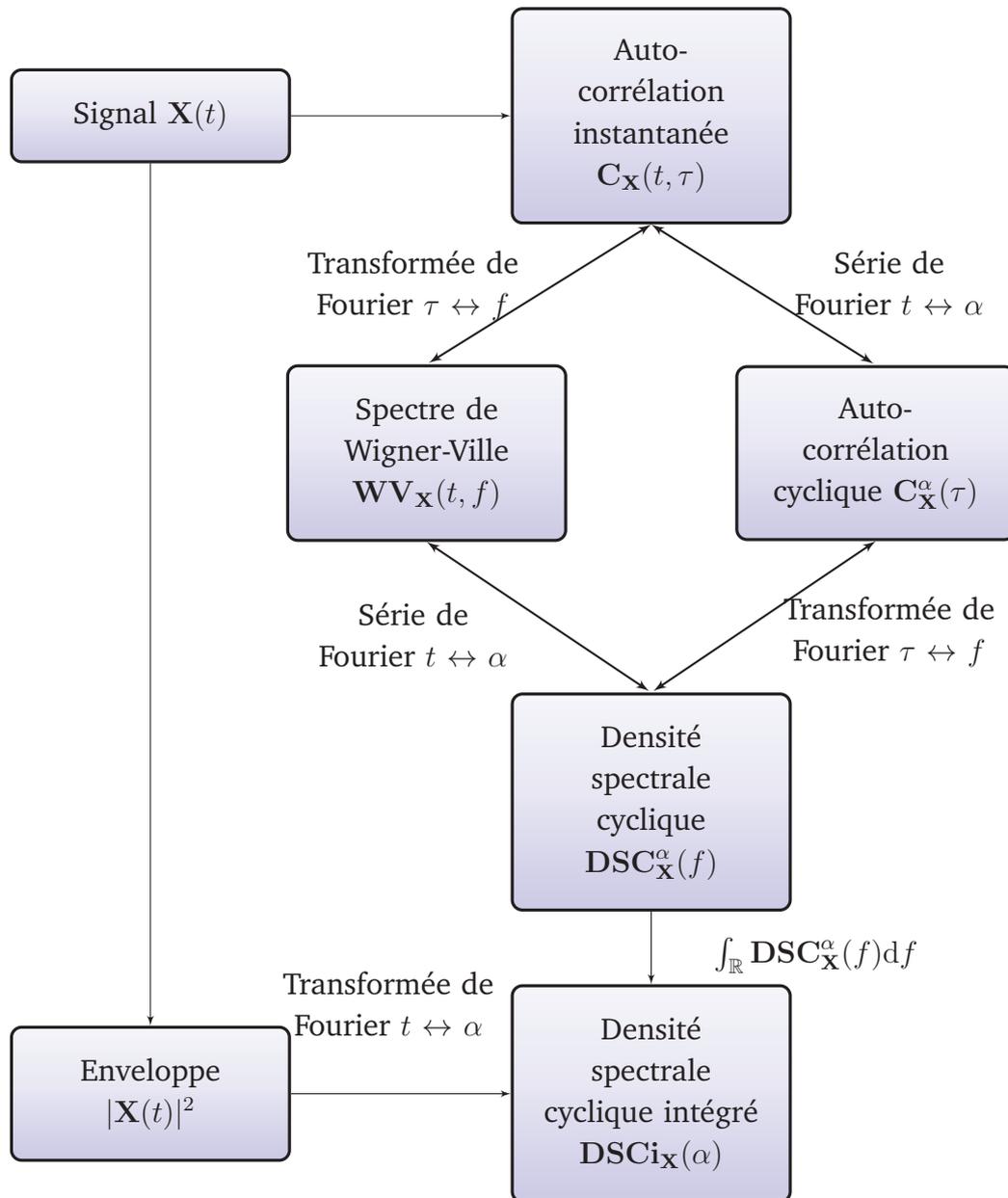


FIGURE 1.7 – Différentes représentations d'un signal.

avec $\langle \cdot \rangle = \lim_{T_0 \rightarrow \infty} \frac{1}{T_0} \int_{-\frac{T_0}{2}}^{\frac{T_0}{2}} (\cdot) dt$. La Fig.(1.8) montre le signal simulé. L'examen du spectre d'amplitude de ce signal ne donne aucune information sur ses propriétés cyclo-stationnaires comme le montre la Fig.(1.9).

Pour étudier la cyclo-stationnarité de signal $x(t)$, nous proposons d'analyser sa fonction d'auto-corrélation cyclique. En appliquant la formule d'Euler sur (1.39), nous obtenons,

$$\mathbf{x}(t) = \frac{1}{2}a(t) [e^{j(2\pi f_0 t + \theta)} + e^{-j(2\pi f_0 t + \theta)}]$$

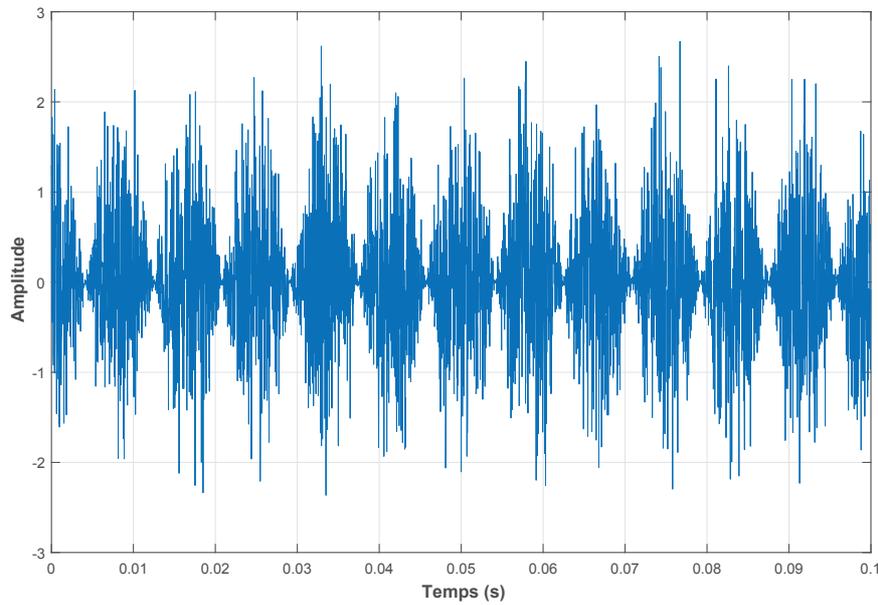


FIGURE 1.8 – Signal simulé $\mathbf{x}(t)$ avec une fréquence $f_0 = 60\text{Hz}$ et une phase nulle.

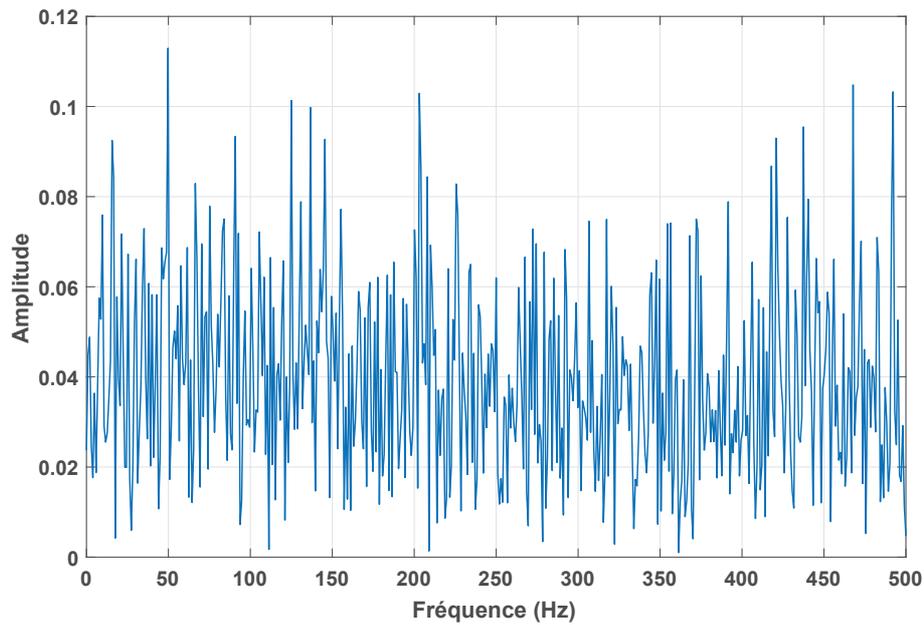


FIGURE 1.9 – Spectre d'amplitude du signal simulé $\mathbf{x}(t)$.

Calculant maintenant le produit, $\mathbf{x}(t)\mathbf{x}^H(t - \tau)$,

$$\begin{aligned} \mathbf{x}(t)\mathbf{x}^H(t - \tau) &= \frac{1}{4}a(t)a^H(t - \tau)[e^{j(2\pi f_0\tau)} + e^{-j(2\pi f_0\tau)} + e^{j(4\pi f_0t+2\theta)}e^{-j(2\pi f_0\tau)} \\ &\quad + e^{-j(4\pi f_0t+2\theta)}e^{j(2\pi f_0\tau)}]. \end{aligned}$$

La fonction d'auto-corrélation cyclique $C_{\mathbf{x}}^{\alpha}(\tau) = \langle \mathbf{x}(t)\mathbf{x}^H(t - \tau)e^{-j2\pi\alpha t} \rangle$ s'écrit donc comme suit,

$$\begin{aligned} C_{\mathbf{x}}^{\alpha}(\tau) &= \frac{1}{4}e^{j(2\pi f_0\tau)} \langle a(t)a^H(t - \tau)e^{-j2\pi\alpha t} \rangle \\ &+ \frac{1}{4}e^{-j(2\pi f_0\tau)} \langle a(t)a^H(t - \tau)e^{-j2\pi\alpha t} \rangle \\ &+ \frac{1}{4}e^{2\theta}e^{-j(2\pi f_0\tau)} \langle a(t)a^H(t - \tau)e^{-j2\pi(\alpha - 2f_0)t} \rangle \\ &+ \frac{1}{4}e^{-2\theta}e^{j(2\pi f_0\tau)} \langle a(t)a^H(t - \tau)e^{-j2\pi(\alpha + 2f_0)t} \rangle, \end{aligned}$$

ou encore,

$$C_{\mathbf{x}}^{\alpha}(\tau) = \begin{cases} \frac{1}{4}e^{\pm 2\theta}C_a(\tau) & \text{pour } \alpha = \pm 2f_0 \\ \frac{1}{2}C_a(\tau)\cos(2\pi f_0\tau) & \text{pour } \alpha = 0 \\ 0 & \text{sinon} \end{cases}$$

avec, $C_a(\tau) = \langle a(t)a^H(t - \tau) \rangle$. Par conséquent, la fonction d'auto-corrélation cyclique est périodique pour l'ensemble des fréquences cycliques $\{-2f_0, 0, +2f_0\}$ comme nous pouvons également le voir sur les Fig.(1.10) et (1.11), le signal $\mathbf{x}(t)$ est donc cyclo-stationnaire à l'ordre 2.

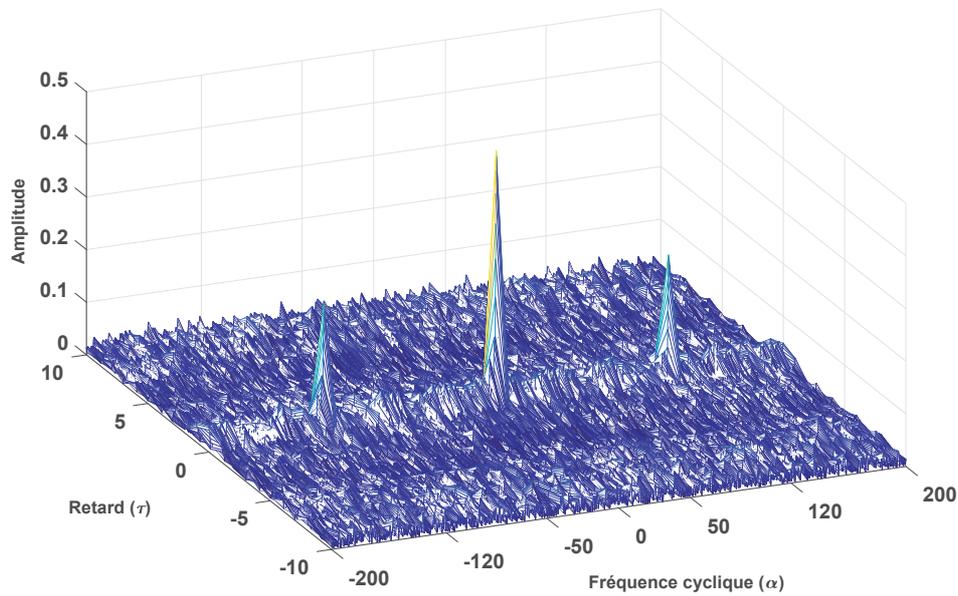


FIGURE 1.10 – Amplitude de la fonction d'auto-corrélation cyclique du signal $\mathbf{x}(t)$ comme une fonction de τ et α .

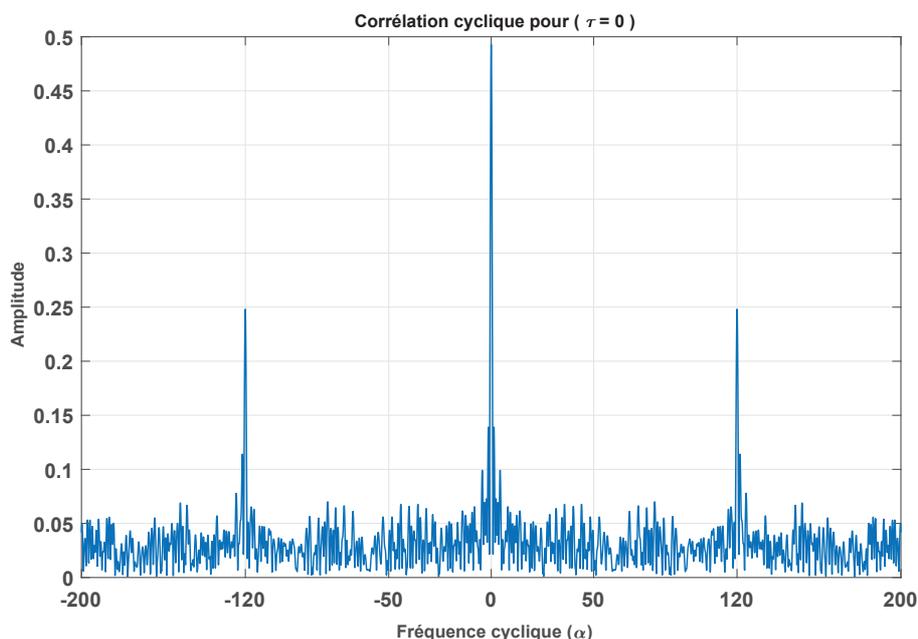


FIGURE 1.11 – Fonction d'auto-corrélation cyclique du signal $\mathbf{x}(t)$ pour un retard nul.

Exemple 2 : modulation numérique linéaire

Le modèle de base utilisé dans les modulations numériques linéaires est celui d'une modulation d'amplitude d'un train d'impulsions (PAM) défini comme suit :

$$\mathbf{x}(t) = \sum_{k \in \mathbb{Z}} a(k)w(t - kT_s),$$

où $a(k)$ est un train binaire aléatoire indépendant et identiquement distribué (i.i.d.) qui représente les symboles transmis appartenant à un alphabet fini, T_s est la période symbole et $w(n)$ est un filtre de mise en forme. Ce modèle présente une périodicité que l'on trouve dans ses propriétés statistiques. En effet, la fonction d'auto-corrélation de ce modèle s'écrit par :

$$\mathbf{C}_{\mathbf{x}}(t + T_s, \tau) = \sum_{k, \ell \in \mathbb{Z}} \mathbf{E}\{a(k)a^*(\ell)\}w(t + T_s - kT_s)w^*(t + T_s - \tau - \ell T_s) \quad (1.40)$$

Compte tenu du fait que les symboles sont i.i.d. alors : $\mathbf{E}\{a(k)a^*(\ell)\} = \sigma_a^2 \delta(k - \ell)$. Donc, l'équation (1.40) devient :

$$\begin{aligned} \mathbf{C}_{\mathbf{x}}(t + T_s, \tau) &= \sum_{k \in \mathbb{Z}} \sigma_a^2 w(t + T_s - kT_s)w^*(t + T_s - \tau - kT_s), \\ &= \sum_{k \in \mathbb{Z}} \sigma_a^2 w(t - kT_s)w^*(t - \tau - kT_s), \quad (\text{changement de variable}) \\ &= \mathbf{C}_{\mathbf{x}}(t, \tau). \end{aligned}$$

Par conséquent, $C_x(t, \tau)$ est périodique de période T_s . Ce qui implique que le signal $x(t)$ est cyclo-stationnaire à l'ordre 2 de fréquences cycliques $\Psi = \{\frac{k}{T_s}, k \in \mathbb{Z}\}$ comme l'indique la Fig.(1.12).

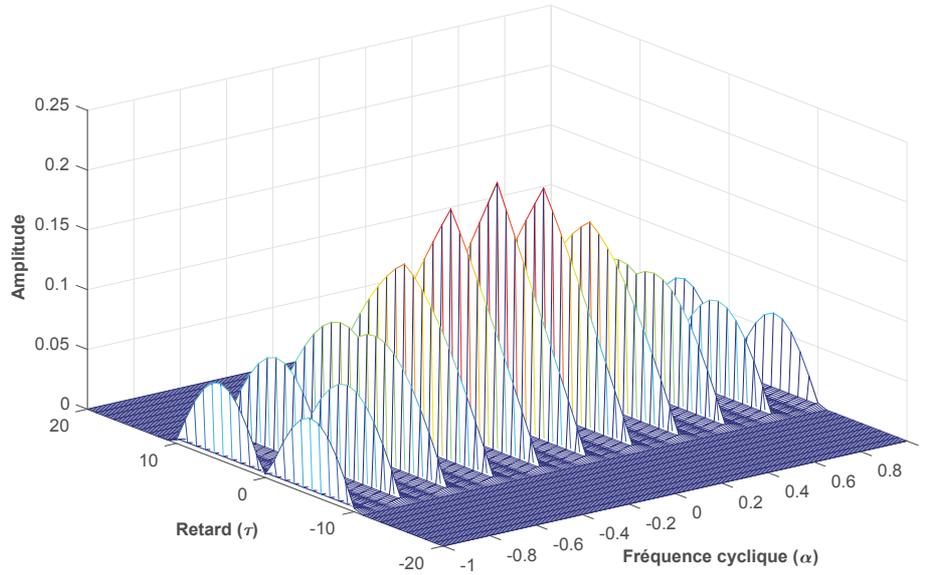


FIGURE 1.12 – Fonction d'auto-corrélation cyclique du signal $\sum_{k \in \mathbb{Z}} a(k)w(t - kT_s)$.

Conclusion

Comme nous l'avons vu dans ce chapitre, la SAS est une thématique commune rencontrée dans de nombreuses applications réelles. Nous avons détaillé quelques généralités liées à ce problème et nous avons évoqué l'usage de la cyclo-stationnarité dans la SAS. Dans le chapitre suivant, nous ferons un tour d'horizon des méthodes de séparation de sources.

Tour d'horizon des techniques de séparation de sources

2

Sommaire

2.1	SAS par analyse en composantes indépendantes	28
2.1.1	Contraste et mesure de l'indépendance	28
2.1.2	Lien avec la bloc diagonalisation conjointe matricielle	31
2.1.3	Présentation de quelques méthodes de SAS	35
2.2	SAS par analyse en composantes parcimonieuses	44
2.2.1	Représentations parcimonieuses	45
2.2.2	Principe des méthodes parcimonieuses	46
2.3	SAS par factorisation en matrices non-négatives	47
2.4	Discussions et conclusion	49

LA classification des techniques de SAS est une tâche très difficile en raison de la variété des hypothèses faites sur les éléments modélisant le problème de SAS et les outils mathématiques employés. Communément, les méthodes de SAS sont découpées en trois grandes familles :

- l'analyse en composantes indépendantes (ICA) et décompositions conjointes de matrices/tenseurs (DCM/T),
- l'analyse en composantes parcimonieuses (SCA),
- la factorisation en matrices non-négatives (NMF).

Après avoir introduit le problème de la SAS dans le premier chapitre, nous dressons, dans ce chapitre, un panorama général des méthodes de SAS. Nous éluciderons dans un premier temps le concept de contraste, son lien avec la diagonalisation conjointe matricielle, les critères de mesure de l'indépendance statistique et quelques méthodes basées sur l'ICA. Dans un second

temps, nous détaillerons l'hypothèse de parcimonie des sources et le principe des méthodes de séparation parcimonieuses. Finalement, nous terminons par examiner brièvement le fondement des méthodes basées sur la NMF.

2.1 SAS par analyse en composantes indépendantes

L'Analyse en Composantes Indépendantes (ICA) a été introduite par C. Jutten dans sa thèse [122] puis reprise par P. Comon dans [74]. Le principe général de la SAS par ICA consiste à appliquer des transformations aux observations pour obtenir des signaux statistiquement indépendants qui seront proportionnels aux signaux sources comme le montre la Fig.(2.1). Nous présentons dans cette section quelques mesures, directes ou indirectes, de l'indépendance statistique ainsi que quelques approches issues de l'emploi de ces mesures.

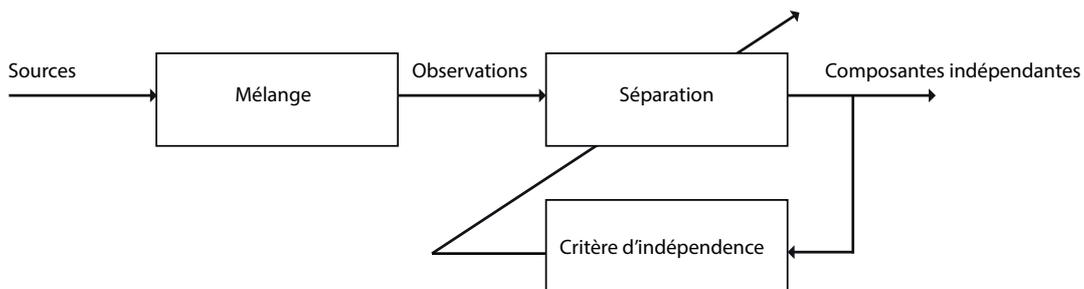


FIGURE 2.1 – Principe de séparation par analyse en composantes indépendantes.

2.1.1 Contraste et mesure de l'indépendance

D'un point de vue pratique, toutes les approches présentées dans le cadre de la SAS par ICA ne se distinguent que par la façon dont elles exploitent l'indépendance statistiques des sources. L'extraction des sources des observations est obtenue par l'optimisation des fonctions dites *contrastes* utilisant les critères de mesure d'indépendance statistique. P. Comon [76] a été le premier à introduire le concept de *fonction de contraste* comme étant une fonction multi-variable $\mathcal{T}(\cdot)$ de \mathcal{Y} (\mathcal{Y} un ensemble de vecteurs aléatoires) dans \mathbb{R} qui ne dépend que de la loi de probabilité d'un vecteur aléatoire y de \mathcal{Y} et qui vérifie les conditions suivantes :

- C_1 . $\forall y \in \mathcal{Y}, \forall \mathbf{D}$ matrice diagonale inversible, $\mathcal{T}(\mathbf{D}y) = \mathcal{T}(y)$.
- C_2 . $\forall y \in \mathcal{Y}, \forall \mathbf{P}$ matrice de permutation inversible, $\mathcal{T}(\mathbf{P}y) = \mathcal{T}(y)$.
- C_3 . $\forall s \in \mathcal{Y}$ de composantes indépendantes et $\forall \mathbf{M}$, tel que

$$\mathbf{M}s \in \mathcal{Y}, \mathcal{T}(\mathbf{M}s) \leq \mathcal{T}(s).$$

C_4 . $\mathcal{T}(Ms) = \mathcal{T}(s) \Leftrightarrow M = PD$.

Les conditions (C_1) et (C_2) signifient que les contrastes sont des fonctions invariantes par changement d'échelle et par permutation. La condition (C_3) indique qu'il s'agit de maximiser la fonction de contraste afin de résoudre le problème de la séparation (C_4) . L'hypothèse d'indépendance statistique est mesurée de différentes façons, nous présenterons dans ce qui suit ces mesures.

Information mutuelle

Soit (X, Y) un couple de variables aléatoires de densité de probabilité jointe données par $P(x, y)$ (pour représenter la probabilité de l'événement $X = x$, nous notons $P(x)$). Notons les distributions marginales $P(x)$ et $P(y)$. Alors l'information mutuelle est donnée par :

$$I(X, Y) = \int_{\mathbb{R}} \int_{\mathbb{R}} p(x, y) \log \frac{p(x, y)}{p(x)p(y)} dx dy. \quad (2.1)$$

où $p(x, y)$, $p(x)$ et $p(y)$ sont respectivement les densités des lois de (X, Y) , X et Y . Les variables aléatoires X et Y sont dites indépendantes si et seulement si $I(X, Y) = 0$.

Entropie relative

L'entropie relative ou la divergence de Kullback-Leibler est une mesure de dissimilarité entre deux distributions de probabilités P et Q . Elle est définie comme suit :

$$D_{\text{KL}}(P||Q) = \int_{-\infty}^{\infty} p(x) \log \frac{p(x)}{q(x)} dx \quad (2.2)$$

où p et q sont les densités respectives de P et Q . En utilisant les propriétés de la fonction \log , nous démontrons que $D_{\text{KL}}(P||Q) \geq 0$ et s'annule si et seulement si $p(x) = q(x)$.

Moments et cumulants

Commençons tout d'abord par rappeler la première et la seconde fonction caractéristique. Soit x une variable aléatoire réelle admettant la densité de probabilité p_x . On appelle la première fonction caractéristique de x la fonction Φ_x définie par,

$$\Phi_x(v) = \mathbf{E}[e^{ivx}] = \int_{\mathbb{R}} e^{ivu} p_x(u) du. \quad (2.3)$$

La deuxième fonction caractéristique Ψ_x définie comme le logarithme népérien de la première caractéristique $\Phi_x = \ln(\Phi_x)$. Les fonctions caractéristiques d'une variable aléatoire, au même titre que la fonction de répartition, décrivent complètement la variable. Néanmoins, en pratique, la loi de probabilité de la variable considérée est inconnue et les fonctions caractéristiques sont difficilement exploitables. On se tourne alors plus aisément vers les moments et cumulants.

Soit x une variable aléatoire réelle admettant la première fonction caractéristique Φ_x . On appelle moment d'ordre q ($q \geq 0$) de x la quantité suivante,

$$\mathbf{M}_q(x) = \mathbf{E}[x^q] = (-i)^q \frac{d^q \Phi_x(v)}{dv^q} \Big|_{v=0}. \quad (2.4)$$

On appelle cumulant d'ordre q ($q \geq 0$) de x la quantité suivante,

$$\kappa_q(x) = \text{Cum} \underbrace{\{x, \dots, x\}}_{q \text{ termes}} = (-i)^q \frac{d^q \Psi_x(v)}{dv^q} \Big|_{v=0}. \quad (2.5)$$

Si x_1 et x_2 sont deux variables aléatoires indépendantes, à valeurs réelles ou complexes, alors les cumulants croisés d'ordre positif quelconque de x_1 et x_2 sont nuls. L'usage de statistiques de tous ordre étant difficile, nous cherchons à cet égard l'ordre suffisant des statistiques, pouvant résoudre le problème de SAS.

Statistiques de second ordre La première solution à laquelle nous pouvons penser consiste à rendre les observations indépendantes au second ordre en tirant profit de la matrice d'auto-corrélation de ces dernières, mais nous nous rendons compte rapidement que celle-ci est insuffisante pour identifier la matrice de mélange. En effet, la matrice d'auto-corrélation des observations est donnée par,

$$\mathbf{C}_X(\tau) = \mathbf{E}\left\{\mathbf{X}\left(t - \frac{\tau}{2}\right)\mathbf{X}^H\left(t + \frac{\tau}{2}\right)\right\} = \mathbf{G}\mathbf{C}_S(\tau)\mathbf{G}^H + \mathbf{C}_R(\tau), \quad (2.6)$$

avec

$$\mathbf{C}_R(\tau) = \sigma_r^2 \delta_0(\tau) \mathbf{I}_P, \quad (2.7)$$

où

$$\delta_0(\tau) = \begin{cases} 1 & \text{si } \tau = 0, \\ 0 & \text{sinon} \end{cases}$$

σ_r^2 est la variance du bruit, τ est le retard, \mathbf{C}_S est la matrice d'auto-corrélation des sources et \mathbf{C}_R est la matrice d'auto-corrélation du bruit. Nous pouvons

convenir que $\mathbf{C}_S(0) = \mathbf{I}$, ceci vient du fait que les sources sont souvent supposées de puissance unité. Ainsi, nous obtenons,

$$\mathbf{C}_X(0) = \mathbf{G}\mathbf{G}^H + \sigma_r^2\mathbf{I}_M, \quad (2.8)$$

$$\mathbf{C}_X(\tau) = \mathbf{G}\mathbf{C}_S(\tau)\mathbf{G}^H, \tau \geq 1. \quad (2.9)$$

En absence de bruit, la matrice d'auto-corrélation de $\mathbf{X}(t)$ se réduit à,

$$\mathbf{C}_x(0) = \mathbf{G}\mathbf{G}^H. \quad (2.10)$$

Il est alors clair qu'à partir de $\mathbf{C}_X(0)$, la matrice \mathbf{G} ne peut être identifiée qu'à une matrice unitaire près. Donc l'utilisation de la seule matrice d'auto-corrélation n'est pas suffisante pour la séparation. Certaines techniques de SAS commencent d'abord par utiliser la matrice d'auto-corrélation des observations pour déterminer \mathbf{G} à une matrice unitaire près à partir de $\mathbf{C}_X(0)$. Ensuite, chercher les rotations manquantes en utilisant d'autres informations.

Statistiques du troisième ordre Notons que les statistiques du troisième ordre sont nulles dans le cas des signaux à densité symétrique¹. Cette restriction stricte exclut généralement leur utilisation pour accomplir la SAS.

Statistiques du quatrième ordre Contrairement aux statistiques du troisième ordre qui sont le plus souvent inutilisable pour réaliser la séparation, les cumulants du quatrième ordre permettent de retrouver les sources recherchées dans les mélanges. Notons que les cumulants d'ordre supérieur à deux d'un signal gaussien sont tous nuls. Ceci constitue un frein pour la SAS gaussiennes, car sans hypothèse supplémentaire, cette dernière est impossible.

2.1.2 Lien avec la bloc diagonalisation conjointe matricielle

L'optimisation d'une fonction de contraste fournit un critère permettant de résoudre le problème de SAS. Pour ce faire, une approche classique repose sur la technique de bloc-diagonalisation conjointe (BDC) d'un ensemble de matrices ou plus généralement de tenseurs [59]. Cette approche est motivée par une préoccupation de la stabilité et de la convergence des algorithmes associés.

1. Les signaux à densité symétrique ont un skewness nul, plus la densité de probabilité est asymétrique, plus le skewness a une grande valeur. Le skewness est défini comme suit : étant donnée une variable aléatoire réelle x de moyenne μ et d'écart type σ , on définit son coefficient d'asymétrie comme le moment d'ordre trois de la variable centrée réduite :

$$\gamma = \mathbb{E} \left[\left(\frac{x-\mu}{\sigma} \right)^3 \right]$$

Formulation du problème de la BDC

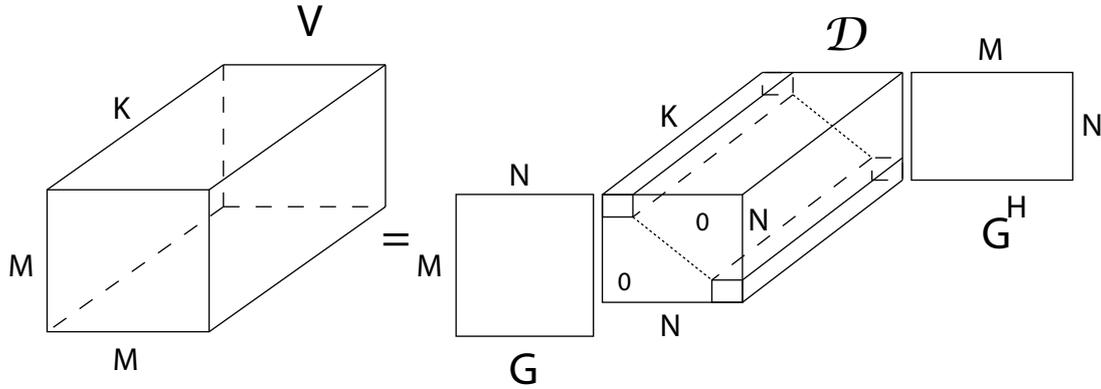


FIGURE 2.2 – Le problème de bloc diagonalisation conjointe en format tensoriel.

Le problème de la BDC d'un ensemble de matrices s'énonce de la manière suivante : soit \mathcal{M} un ensemble de K ($K \in \mathbb{N}^*$) matrices carrées $\mathbf{V}_k \in \mathbb{C}^{M \times M}$, $k \in \{1, \dots, K\}$ admettant toutes la factorisation suivante (voir Fig.(2.2)),

$$\mathbf{V}_k = \mathbf{G} \mathbf{D}_k \mathbf{G}^H, \quad \forall k \in \{1, \dots, K\} \quad (2.11)$$

avec,

$$\mathbf{D}_k = \begin{pmatrix} \mathbf{D}_{k,11} & \mathbf{0}_{12} & \dots & \mathbf{0}_{1r} \\ \mathbf{0}_{21} & \mathbf{D}_{k,22} & \ddots & \vdots \\ \vdots & \ddots & \ddots & \mathbf{0}_{r-1r} \\ \mathbf{0}_{r1} & \dots & \mathbf{0}_{rr-1} & \mathbf{D}_{k,rr} \end{pmatrix}, \quad \forall k \in \{1, \dots, K\} \quad (2.12)$$

où r est le nombre de blocs considérés ($r \in \mathbb{N}^*$), $\mathbf{D}_{k,jj}$, $k \in \{1, \dots, K\}$, $j \in \{1, \dots, r\}$ sont des matrices carrées de dimensions $(n_j \times n_j)$ avec $n_1 + \dots + n_r = N^2$ et $\mathbf{0}_{kj}$ désigne la matrice nulle de dimension $(n_k \times n_j)$. La matrice \mathbf{G} est une matrice de rang plein de dimension $M \times N$ ($M \geq N$).

L'ensemble des K matrices $\mathbf{D}_k \in \mathbb{C}^{N \times N}$ est noté \mathcal{D} . Le problème de la BDC consiste à estimer la matrice \mathbf{G} et l'ensemble \mathcal{D} des matrices bloc-diagonales et cela uniquement à partir des matrices de l'ensemble \mathcal{M} .

La résolution du problème de la BDC passe par la minimisation de l'une des fonctions de coût quadratiques équivalentes suivantes :

$$\mathcal{C}_{BDC}(\{\mathbf{D}_k\}, \mathbf{G}) = \sum_{k=1}^K \|\mathbf{V}_k - \mathbf{G} \mathbf{D}_k \mathbf{G}^H\|_F^2, \quad (2.13)$$

$$\mathcal{C}'_{BDC}(\{\mathbf{D}_k\}, \mathbf{B}) = \sum_{k=1}^K \|\mathbf{B} \mathbf{V}_k \mathbf{B}^H - \mathbf{D}_k\|_F^2. \quad (2.14)$$

2. Dans le cas où nous traitons de matrices $\mathbf{D}_{k,jj}$ ayant les mêmes dimensions alors $N = r \times n_j$, $j \in \{1, \dots, r\}$

Cette équivalence vient du fait que nous avons multiplié à gauche (resp. à droite) la matrice \mathbf{V}_k décrite au niveau de l'équation (2.11) par la pseudo-inverse $\mathbf{G}^\#$ de \mathbf{G} (resp. par $(\mathbf{G}^\#)^H$). Ce qui conduit alors à,

$$\mathbf{G}^\# \mathbf{V}_k (\mathbf{G}^\#)^H = \mathbf{D}_k, \quad \forall k \in \{1, \dots, K\}. \quad (2.15)$$

où encore,

$$\mathbf{B} \mathbf{V}_k \mathbf{B}^H = \mathbf{D}_k, \quad \forall k \in \{1, \dots, K\}. \quad (2.16)$$

où $\mathbf{B} = \mathbf{G}^\#$ (\mathbf{B} n'est alors rien d'autre que la matrice de séparation dans le problème de la SAS). Une autre fonction de coût assez proche de celles présentées précédemment, impliquant le calcul d'une seule matrice \mathbf{B} a été introduite par Ghennioui dans [102] définie comme suit :

$$\mathcal{C}'_{BDC}(\{\widehat{\mathbf{D}}_k\}, \mathbf{B}) = \sum_{k=1}^K \|\text{OffBdiag}_{(n)}\{\mathbf{B} \mathbf{V}_k \mathbf{B}^H\}\|_F^2 = \mathcal{C}'_{BDC}(\mathbf{B}). \quad (2.17)$$

où

$$\widehat{\mathbf{D}}_k = \arg \min_{\mathbf{D}_k} \mathcal{C}'_{BDC}(\{\mathbf{D}_k\}, \mathbf{B}). \quad (2.18)$$

Nous rappelons que les deux opérateurs matriciels $\text{Bdiag}_{(n)}\{\cdot\}$ et $\text{OffBdiag}_{(n)}\{\cdot\}$ sont respectivement définis par :

$$\text{Bdiag}_{(n)}\{\mathbf{V}\} = \begin{pmatrix} \mathbf{V}_{11} & \mathbf{0}_{12} & \dots & \mathbf{0}_{1r} \\ \mathbf{0}_{21} & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & \vdots \\ \mathbf{0}_{r1} & \mathbf{0}_{r2} & \dots & \mathbf{V}_{rr} \end{pmatrix}, \quad (2.19)$$

et

$$\text{OffBdiag}_{(n)}\{\mathbf{V}\} = \mathbf{V} - \text{Bdiag}_{(n)}\{\mathbf{V}\} = \begin{pmatrix} \mathbf{0}_{11} & \mathbf{V}_{12} & \dots & \mathbf{V}_{1r} \\ \mathbf{V}_{21} & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & \vdots \\ \mathbf{V}_{r1} & \mathbf{V}_{r2} & \dots & \mathbf{0}_{rr} \end{pmatrix}, \quad (2.20)$$

où $\mathbf{n} = (n_1, n_2, \dots, n_r)$ et \mathbf{V} est une matrice bloc-diagonale de dimensions $(N \times N)$ ayant la structure suivante :

$$\mathbf{V} = \begin{pmatrix} \mathbf{V}_{11} & \mathbf{V}_{12} & \dots & \mathbf{V}_{1r} \\ \mathbf{V}_{21} & \dots & \dots & \vdots \\ \vdots & \dots & \dots & \vdots \\ \mathbf{V}_{r1} & \mathbf{V}_{r2} & \dots & \mathbf{V}_{rr} \end{pmatrix}, \quad (2.21)$$

avec les \mathbf{V}_{ij} sont des matrices de dimensions $(n_i \times n_j)$ et $n_1 + \dots + n_r = N$. Il est encore possible d'utiliser une autre fonction de coût, telle que celle définie dans [85].

$$C'_{BDC}(\{\mathbf{D}_k\}, \mathbf{B}) = \sum_{k=1}^K \log \frac{\det(\text{Diag}\{\mathbf{B}\mathbf{D}_k\mathbf{B}^H\})}{\det(\mathbf{B}\mathbf{D}_k\mathbf{B}^H)}, \quad (2.22)$$

Cas particulier de la BDC

La diagonalisation conjointe (DC) n'est qu'un cas particulier de la BDC où les matrices $\mathbf{D}_{i,jj}$, $i \in \{1, \dots, K\}$, $j \in \{1, \dots, r\}$ sont réduites à des scalaires ($n_j = 1 \forall j \in \{1, \dots, r\}$ et par conséquent $r = N$). L'opérateur matriciel $\text{OffBdiag}_{(n)}\{\cdot\}$ et $\text{OffBdiag}_{(n)}\{\cdot\}$ devient donc ainsi :

$$\text{OffBdiag}_{(1)}\{\mathbf{M}\} = ((1 - \delta_{ij})M_{ij})\mathbf{1}_N \quad (2.23)$$

$$= \text{OffDiag}\{\mathbf{M}\}, \quad (2.24)$$

avec $\mathbf{M} = (M_{ij})$, $\mathbf{1}_N$ est une matrice de dimensions $(N \times N)$ dont tous les termes valent 1 et $\delta_{ij} = 1$ si $i = j$ et 0 sinon. Tandis que l'opérateur matriciel $\text{OffBdiag}_{(n)}\{\cdot\}$ se simplifie en le notant $\text{Diag}\{\cdot\}$:

$$\text{Bdiag}_{(1)}\{\mathbf{M}\} = ((\delta_{ij})M_{ij})\mathbf{1}_N \quad (2.25)$$

$$= \text{Diag}\{\mathbf{M}\} \quad (2.26)$$

$$= \mathbf{M} - \text{OffDiag}\{\mathbf{M}\}. \quad (2.27)$$

Bref état de l'art des algorithmes de DCM/T

Plusieurs solutions ont été présentées pour résoudre le problème des DCM/T. Les premiers travaux ont porté sur la DC des matrices sous contrainte d'unitarité en passant par un stade de pré-blanchiment, tel est le cas pour les célèbres algorithmes JADE (Diagonalisation Approximative Conjointe des Eigenmatrices) [51] et SOBI (Second Order Blind Identification)[22]. Les travaux suivants ont abordé le problème dans sa globalité en considérant la DC des tenseurs [74], [137], d'autres ont supprimé la contrainte d'unitarité car elle constitue une limitation des performances dans le contexte de la SAS [58], [136], [172], [188]. Pour un panorama assez exhaustif des algorithmes de diagonalisation conjointe matricielle, nous renvoyons le lecteur à la référence [60]. Ces travaux trouvent leur utilité dans le cadre de la localisation des sources, l'estimation de la direction d'arrivée et la séparation aveugle des mélanges instantanés de sources. Récemment, un problème plus général a été traité portant sur la décomposition conjointe d'une combinaison d'ensembles de matrices complexes qui

peuvent suivre des décompositions potentiellement différentes (par exemple lorsque nous considérons des signaux complexes non circulaires ou quand nous voulons exploiter des informations statistiques supplémentaires en utilisant à la fois des matrices symétriques hermitiennes et complexes) [138], [180], [199]. De tels problèmes peuvent survenir dans divers problèmes de traitement du signal, parmi lesquels l'identification aveugle, la séparation ou la déconvolution multidimensionnelle.

Quant à la BDC des matrices, plusieurs travaux ont été proposés [144], [178], [187], [64], [124]. Certains ont considéré des ensembles de matrices définies positives à symétrie hermitienne et des bloc-diagonaliseurs conjoints unitaires tel est le cas des algorithmes de type Jacobi proposés dans [24], [81]. Ensuite, différentes algorithmes alternatifs ont été proposés atteignant le même objectif, mais pour un bloc-diagonaliseur conjoint non-unitaires. Ghennioui et al. ont considéré dans [104], [105], la fonction de coût décrite au niveau de (2.17). Cette fonction a été réécrite afin d'effectuer une optimisation nécessitant une SVD. Notons que les matrices exploitées sont définies positives mais non forcément à symétrie hermitienne. Une approche de type Gradient approché a été proposée par les auteurs de [107]. Par la suite, le calcul exact de ce Gradient a été effectué dans [105] donnant naissance à deux algorithmes dans [102]. Ces algorithmes se caractérisent par le fait qu'ils sont plus généraux étant donné qu'ils peuvent être appliqués à un ensemble quelconque de matrices (celles-ci, réelles ou complexes, peuvent n'être ni définies positives ni à symétrie hermitienne). Cherrak et al. ont présenté dans [67], [65], quatre algorithmes itératifs et passent tous par la minimisation de la fonction de coût décrite au niveau de (2.17). Le premier algorithme, s'appuie sur la minimisation d'un schéma de type gradient conjugué. Le second est de type Levenberg-Marquardt reposant sur un calcul exact des matrices Hessiennes [66].

2.1.3 Présentation de quelques méthodes de SAS

L'année 1985 témoigne du premier algorithme de SAS développé par C. Jutten et J. Héroult [114] pour résoudre le problème de décodage du mouvement en neurobiologie. Depuis, un flot d'algorithmes a été proposé. Nous faisons ici un tour d'horizon des techniques de SAS proposées dans la littérature.

Méthodes utilisant les fonctions de contraste

La première catégorie des méthodes utilise des outils d'ordre supérieur pour construire des fonctions de contraste, dont l'optimisation permet de résoudre le problème de SAS. P. Comon a proposé la méthode COM2 (CONtrast Maximization 2) [74] qui maximise un contraste défini comme la somme des modules au carré des kurtosis³ des sources estimées $\mathbf{y}(t)$:

$$\mathcal{Y}_q^2(\mathbf{y}) = \sum_{i=1}^n |\kappa_q\{\mathbf{y}_i\}|^2 \quad (2.28)$$

où n est le nombre de sources et κ_q représente le cumulante d'ordre q d'un signal stationnaire. J.-F. Cardoso et A. Souloumiac ont présenté dans [51] un contraste baptisé JADE basé sur les cumulants d'ordre quatre :

$$\mathcal{J}(\mathbf{y}) = \sum_{i_1, i_2, i_3, i_4=1}^n |\text{Cum}\{\mathbf{y}_{i_1}, \mathbf{y}_{i_2}, \mathbf{y}_{i_3}, \mathbf{y}_{i_4}\}|^2. \quad (2.29)$$

Maximiser cette fonction revient à diagonaliser conjointement un ensemble de n^2 matrices. J.-F. Cardoso et A. Souloumiac ont proposé également une version non itérative de leur méthode basée sur la DC [54] des matrices propres de la quadri-covariance des observations blanchies. L. De Lathauwer a proposé dans [125] un algorithme similaire baptisé STOTD (Simultaneous Third Order Tensor Diagonalization) qui consiste à diagonaliser conjointement un ensemble de tenseurs d'ordre trois au carré et non plus un ensemble de matrices. Il repose sur la maximisation de la fonction contraste suivante :

$$\mathcal{J}'(\mathbf{y}) = \sum_{i_1, i_2=1}^n |\text{Cum}\{\mathbf{y}_{i_1}, \mathbf{y}_{i_2}, \mathbf{y}_{i_3}, \mathbf{y}_{i_4}\}|^2. \quad (2.30)$$

Une généralisation des trois fonctions de contrastes précédentes a été introduite par E. Moreau dans [137]. Quelques années plus tard, il a été montré dans [141] que la propriété de symétrie des fonctions contrastes pouvait être omise car elle est restrictive dans certains cas, d'où une généralisation vers les fonctions contrastes non-symétriques. Nous citons comme exemple de ces nouveaux contrastes la fonction suivante [141] :

$$\mathcal{Y}_q^f(\mathbf{y}) = \sum_{i=1}^n \gamma_i f(|\kappa_q\{\mathbf{y}_i\}|), \quad q \geq 3 \quad (2.31)$$

3. En théorie des probabilités et statistiques, le kurtosis mesure l'aplatissement de la distribution d'une variable aléatoire réelle. Il est défini comme le rapport entre le moment centré d'ordre 4 et le carré du moment centré d'ordre 2.

où $f(\cdot)$ est une fonction convexe, strictement croissante et $\gamma_i, i \in \{1, \dots, n\}$ vérifiant $\gamma_1 \geq \dots \geq \gamma_n$.

P. Comon a proposé un nouveau contraste nommée COM1 (CONtrast Maxi-mization 1) défini, au signe près, comme somme des kurtosis des sources estimées auquel il a trouvé une solution analytique. Ce contraste a été initialement présenté par E. Moreau et al. dans [77] [139] [175].

Toutes les fonctions contrastes citées précédemment sont étudiées dans le contexte des mélanges instantanés. Ces fonctions ont fait l'objet de plusieurs extensions dans le cas des mélanges convolutifs. Tel est le cas par exemple des travaux suivants : P. Comon dans [75], E. Moreau et J.-C. Pesquet dans [140], ainsi que N. Thirion-Moreau et E. Moreau dans [176]. Plus récemment, P. Comon et L. Rota ont introduit une méthode baptisée PAJOD (PARtial JOint Diagonalization) qui consiste à optimiser une fonction contraste basée sur les cumulants d'ordre quatre en utilisant la DC d'un ensemble de tranches du tenseur cumulant après un blanchiment des observations. Cette méthode permet d'obtenir un banc de filtres égaliseur semi-unitaire et non para-unitaire [162, 163].

Méthodes algébriques

Méthodes basées sur les statistiques de second ordre De nombreuses méthodes de SAS exploitent les statistiques de second ordre dans le cas stationnaire. Ces méthodes se basent généralement sur deux procédés : le blanchiment des observations pour les décorréler spatialement et réduire la dimension du mélange dans le cas sur-déterminé au cas déterminé puis la recherche d'une matrice de rotation qui permet d'estimer les sources. Parmi ces méthodes, nous trouvons la méthode AMUSE (Algorithm for Multiple Unknown Signals Extraction) introduite par Tong et al. dans [179] qui nécessite que les sources soient colorées, autrement dit temporellement corrélées et réalise la DC de deux matrices de corrélation des observations, l'une associée à un retard nul M_0 (définie positive) et la seconde à un retard non nul M_1 . La DC est réalisée de façon exacte pour estimer la matrice unitaire U choisie de sorte à vérifier,

$$UM_0U^H = I, \quad UM_1U^H = D_1, \quad (2.32)$$

où D_1 est une matrice diagonale. La méthode AMUSE opère sur des observations blanchies $\mathbf{x}_w(t) = \mathbf{W}\mathbf{x}(t)$, où la matrice \mathbf{W} est la matrice de blanchiment qui peut être identifiée par la décompositions en valeurs propres (DVP) de la matrice M_0 . La matrice U est ensuite obtenue sous la forme $U = \mathbf{V}\mathbf{W}$ à partir de la décomposition spectrale $\mathbf{W}M_1\mathbf{W}^T = \mathbf{V}^T\mathbf{D}_1\mathbf{V}$, avec \mathbf{V} une

matrice estimée par la DVP de la matrice \mathbf{M}_1 . La matrice de mélange et les sources sont estimées par

$$\hat{\mathbf{A}} = \mathbf{W}^\# \mathbf{V}, \quad (2.33)$$

$$\hat{\mathbf{s}}(t) = \mathbf{V}^H \mathbf{x}_w(t). \quad (2.34)$$

Da ce travail, l'algorithme SOBI est né [22], dans lequel A. Belouchrani et al. exploitent plusieurs matrices de corrélation des observations associées à des retards non nuls. Une DC au moyen d'un algorithme de type Jacobi de ces matrices permet d'estimer le mélange à une matrice triviale près. SOBI opère lui aussi sur des données blanchies. Pour ce faire, toute racine carrée de $\mathbf{C}_X(0)$ dans un contexte non-bruité donne lieu à une matrice de blanchiment, $\mathbf{W} = \mathbf{C}_X^{-\frac{1}{2}}$, et nous obtenons alors $\mathbf{A} = \mathbf{W}^{-1} \mathbf{U}^H$ où \mathbf{U} est une matrice unitaire vérifiant :

$$\mathbf{C}_{\mathbf{x}_w}(\tau) = \mathbf{W} \mathbf{C}_X \mathbf{W}^H = \mathbf{U}^H \mathbf{C}_S(\tau) \mathbf{U}, \quad \tau \geq 1. \quad (2.35)$$

L'étape de blanchiment permet donc de remplacer la DC des matrices $\mathbf{C}_X(\tau)$, au moyen d'une matrice \mathbf{A} quelconque, par celle des matrices $\mathbf{C}_{\mathbf{x}_w}(\tau)$, au moyen d'une matrice unitaire \mathbf{U} . Soit $\mathcal{M} = \{\mathbf{C}_{\mathbf{x}_w}(\tau_k) = \mathbf{U}^H \mathbf{C}_S(\tau_k) \mathbf{U}, k = 1, \dots, K\}$ un ensemble de matrices de corrélation blanchies à plusieurs retards. L'estimation de la matrice de mélange passe par la minimisation du contraste suivant :

$$\mathcal{J}(\mathcal{M}, \mathbf{U}) = \sum_{k=1}^K \text{off}(\mathbf{U} \mathbf{C}_{\mathbf{x}_w}(\tau_k) \mathbf{U}^H), \quad (2.36)$$

où si \mathbf{M} est une matrice diagonale de taille $m \times m$ définie par

$$\mathbf{M} = \begin{pmatrix} m_{11} & m_{12} & \cdots & m_{1m} \\ m_{21} & \cdots & \cdots & \vdots \\ \vdots & \cdots & \cdots & \vdots \\ m_{m1} & m_{m2} & \cdots & m_{mm} \end{pmatrix}.$$

alors $\text{off}(\mathbf{M}) = \sum_{k \neq l} m^2(l, k)$ est la somme des carrés des éléments non diagonaux de \mathbf{M} . Cette approche nécessite néanmoins que les sources aient des spectres distincts. Ziehe et Müller ont proposé durant cette période dans [201] un algorithme très proche de SOBI nommé TDSEP (Temporal Decorrelation Source SEParation). Ultérieurement, A. Belouchrani et A. Cichocki ont modifié l'étape de blanchiment de l'algorithme SOBI en créant un algorithme non perturbé par un bruit temporellement décorrélé et de cohérence spatiale

quelconque appelé RSOBI (Robust SOBI)[25].

Une multitude de travaux a été présentée dans le contexte non-stationnaire en utilisant les matrices de corrélation ou les représentations temps-fréquence. Les matrices de corrélation sont calculées en morceaux où les sources sont quasi-stationnaires *c.à.d* que leur puissance varie très peu sur ce morceau alors qu'elle varie fortement en passant d'un morceau au suivant comme le cas des signaux de parole. Cette variation crée une diversité d'information permettant de résoudre le problème de SAS à l'aide d'une DC des matrices de corrélation estimées à partir de chaque morceau de données. De ce principe naissent les algorithmes BGML (Block-Gaussian Maximum Likelihood) [153] et SEONS (Second-Order Non-stationary source Separation) [71]. Les représentations temps-fréquence-spatiales (RTFS) appartenant à la classe de Cohen constituent un outil robuste pour la caractérisation des signaux non-stationnaires car elles exploitent la diversité temps-fréquence-capteurs au lieu des diversités employées dans les méthodes classiques à savoir la diversité temps-capteurs ou fréquence-capteurs. Ce concept a été introduit en 1996 par A. Belouchrani et M. Amin dans [23] et son application se résume généralement en trois étapes :

1. Composition d'un ensemble de matrices RTFS des observations blanches ou non,
2. Sélection automatique dans le plan temps-fréquence des "bonnes" matrices de RTFS qui présentent une structure algébrique particulière (diagonale ou zéro-diagonale⁴ dans le cas des mélanges instantanés, bloc-diagonale⁵ ou zéro bloc-diagonale dans le cas des mélanges convolutifs). Cette étape est la plus compliquée car elle conditionne l'aboutissement de la séparation.
3. Application des techniques de décompositions tensorielles (DC ou BDC, CPD/PARAFAC (Canonical Polyadic Decomposition ou Candecomp/PARALLEL FACTORIZATION)) sur les matrices sélectionnées afin de répondre au problème de séparation.

Plusieurs travaux ont vu le jour après la première introduction des RTFS et tentent le plus souvent de répondre à la question de choix des matrices RTFS [116] [93] [95]. Quelques travaux se sont affranchis de l'étape de blanchi-

4. Une matrice zéro-diagonale est une matrice carrée dont tous les éléments diagonaux sont nuls.

5. Une matrice bloc-diagonale est une matrice carrée qui possède des blocs matrices carrées sur la diagonale principale, tels que les blocs non diagonaux soient des matrices nulles.

ment préalable des observations [90] [109], tandis que d'autres ont considéré le cas sous-déterminé des mélanges [108] [27].

Méthodes basées sur les statistiques d'ordre supérieur Lorsque les sources ne possèdent pas de cohérence temporelle, il n'est plus possible de les séparer en utilisant uniquement les statistiques de second ordre. Nous employons dans ce cas, les statistiques d'ordre supérieur comme J.-F. Cardoso l'a fait dans son algorithme FOBI (Fourth Order Blind Identification) [52] où il a employé le blanchiment des observations et la diagonalisation d'une seule matrice cumulante afin d'obtenir la matrice de rotation permettant d'assurer l'indépendance des sources estimées. Puis, il a proposé la méthode FOABI (Fourth Order Only Blind Identification) [53] qui opère sans blanchiment préalable des observations. Cette méthode a été améliorée par L. De Lathauwer et al. dans [80] en utilisant la DC. Ensuite, J.-F. Cardoso [51] a proposé la méthode JADE qui tire partie des propriétés algébriques des cumulants d'ordre quatre pour estimer la matrice de mélange en utilisant le blanchiment basé sur les statistiques d'ordre en premier lieu puis les techniques de diagonalisation conjointe matricielle en second lieu. L. Albera et al. ont présenté les méthodes ICAR (Independent Component Analysis using Redundancies) [6, 9, 8] et BIRTH (Blind Identification of mixtures of sources using Redundancies in the daTa Hexacovariance matrix) [10, 7], exploitant la redondance des matrices respectivement à l'ordre quatre et à l'ordre six. Ces deux méthodes font partie de la même famille d'algorithmes baptisée BIOME (Blind Identification of Overcomplete Mixtures of sources) [7]. A. Ferréol et al. ont étendu dans [89] la méthode SOBI à l'ordre quatre sous le nom de FOBIUM (Fourth Order Blind Identification of Underdetermined Mixtures of sources). Ces trois dernières méthodes s'affranchissent de l'étape préalable de blanchiment. En plus les méthodes FOBIUM et BIRTH permettent de considérer le cas de mélanges sous-déterminés.

Méthodes par maximum de vraisemblance

L'approche du maximum de vraisemblance considère le problème d'estimation de la matrice de mélange, en vue de la reconstruction des sources. L'estimation de la matrice de mélange au sens du maximum de vraisemblance est solution de

$$\hat{\mathbf{A}}_{MV} = \arg \max_{\mathbf{A}} p(\mathbf{X}|\mathbf{A}). \quad (2.37)$$

L'approche du maximum de vraisemblance a été employée initialement par M. Gaeta et J. L. Lacoume [97] pour résoudre le problème de la SAS non

gaussiennes. Deux ans plus tard, D. T. Pham, P. Garat dans [154] ont dénoué le problème de séparation par une méthode fondée sur le principe du maximum de vraisemblance en considérant des observations dans un contexte non bruité. F. Harroy et J.-L. Lacoume ont proposé [112] d'approcher la fonction log-vraisemblance des sources en utilisant un développement en série de Gram-Charlier. D.-T. Pham et P. Garrat ont suggéré d'optimiser le critère du maximum de vraisemblance à l'aide de l'algorithme de Newton Raphson [154]. A.-J Bell et T.-J. Sejnowski ont développé la méthode Infomax (maximization of the information flow) [19] [20], basée sur la méthode du maximum de vraisemblance. S. I. Amari ont proposé dans [11] d'améliorer la méthode Infomax en utilisant l'algorithme du gradient naturel afin d'accroître la vitesse de convergence de la méthode.

Méthodes par déflation

Les méthodes de SAS par déflation consistent à trouver un filtre permettant d'extraire une première source puis soustraire sa contribution des observations. Cela forme un nouveau mélange de dimension inférieure. L'étape initiale est répétée jusqu'à séparer toutes les sources. Ces approches ont été initialement introduites par N. Delfosse et Ph. Loubaton dans [82]. Les auteurs ont suggéré une procédure adaptative de séparation nommée DEFA (DEflation Approach for blind source separation) qui consiste à appliquer par déflation, après blanchiment des observations, un filtre ligne pour extraire chacune des sources. Plus tard, A. Hyvärinen et E. Oja ont présenté l'algorithme FastICA qui dans sa version initiale [26] traite de sources à valeurs réelles et tire partie du concept de déflation et la mesure de non-gaussianité approximée par la valeur absolue du kurtosis de la source estimée. Une extension au cas des signaux à valeurs complexes a été introduite dans [26]. A. Hyvärinen a proposé dans [117] une amélioration de son algorithme FastICA en utilisant les moments de fonctions non-quadratiques pour mesurer la non-gaussianité. Dans le cas des mélanges convolutifs, la séparation par déflation revient à un problème de déconvolution SIMO (Single-Input Multiple-Output) comme cela a été montré par J. K. Tugnait dans [181] et C. Simon et al. dans [170]. Par ailleurs, il existe des méthodes dites hiérarchiques [55] [56] [57] qui traitent des signaux non i.i.d. et consistent à maximiser des critères d'une façon itérative. Dans chaque étape, on introduit une contrainte de décorrélation spatiale des sources. Il faut noter que les approches par déflation atteignent rarement les performances des méthodes symétriques⁶ car une légère erreur sur l'estimation d'une source se répercute automatique-

6. Ce sont des méthodes qui estiment les sources simultanément.

ment sur l'estimation des autres. Cependant, elles sont moins coûteuses en termes de complexité [197].

Méthodes cyclo-stationnaires

Une multitude de travaux a été conduite pour le cas des signaux stationnaires et non-stationnaires. Quant à la SAS cyclo-stationnaires, nous trouvons moins de traces dans la littérature. Toutefois, elle est présente dans beaucoup d'applications réelles comme le cas des signaux issus de machines tournantes, dont le fonctionnement est régi par un cycle mécanique au bout duquel le mécanisme aura retrouvé exactement sa configuration initiale. C'est aussi le cas des signaux de télécommunications subissant des modulations directement responsables de leur caractère cyclo-stationnaire. Les premiers travaux proposés dans le cadre de la SAS cyclo-stationnaires, ont été présentés par Liang et al. dans [133, 190]. Ils utilisent une approche basée sur les statistiques cycliques du second ordre. Celle-ci ne fixe aucune restriction sur le choix des fréquences cycliques des sources. Dans l'esprit de l'algorithme JADE, A. Ferréol et P. Chevalier ont proposé dans [92] une version de JADE, nommée JADE cyclique, exploitant les propriétés cyclo-stationnaires des signaux sources. Dapena et al. ont suggéré dans [79] un nouvel algorithme adaptatif qui utilise simultanément les statistiques d'ordre supérieur et des moments cycliques pour des sources de télécommunication. L'approche permet l'exploitation des périodicités intégrées dans les sources telles que la fréquence porteuse ou le débit de symboles. Ferréol et Chevalier ont analysé dans [91] le comportement des estimateurs empiriques des statistiques d'ordre deux et quatre, en présence de sources cyclo-stationnaires à l'ordre deux. Ils ont montré que ces derniers génèrent des estimations biaisées et en conséquence, proposent un estimateur alternatif des statistiques d'ordre quatre pour surmonter ce problème. Abed-Meraim et al. [2] ont proposé deux critères de séparabilité et d'identifiabilité de sources. Ces critères sont basés sur les statistiques cycliques de second ordre et définissent un algorithme itératif permettant de séparer les sources avec ou sans fréquences cycliques distinctes. Ypma et al. [193] ont appliqué la SAS afin d'extraire la signature d'une machine tournante à partir de mesures vibratoires. Ils ont combiné l'algorithme SOBI (Second-Order Blind Identification) [22] avec l'algorithme SCORE (Spectral Self-Coherence Restoral) [3] afin de créer l'algorithme CycloSOBI qui exploite à la fois la cohérence temporelle et la cyclo-stationnarité des sources. Antoni et al. ont formulé une méthode non-paramétrique dans le domaine spectral utilisant les matrices de corrélation spectrale cyclique des observations et les tech-

niques de DC pour donner lieu à des filtres de séparation unique. Bradric et al. ont considéré dans [39][40] le problème d'identification d'un système à entrées multiples/sorties multiples (MIMO) réponse impulsionnelle finie (FIR), excité par des entrées colorées cyclo-stationnaires. Ils ont proposé une méthode employant les spectres de puissance et spectres croisés de la sortie du système. Xiang et al. [191][186] ont présenté un nouvel algorithme pour la SAS, en exploitant la redondance de phase et de fréquence des signaux cyclo-stationnaires de manière complémentaire. Jafari et al. dans [120, 119] ont proposé un algorithme de type gradient naturel pour la SAS cyclo-stationnaires. Ils ont montré que la propriété de la cyclo-stationnarité améliore les performances de séparation en terme de vitesse de convergence. Jallon et al. [121] ont introduit une fonction de contraste basée sur les statistiques d'ordre quatre dont l'estimation ne nécessite pas la connaissance des fréquences cycliques. La maximisation de cette fonction de contraste permet de séparer les sources. Boustany et Antoni [37][38] ont proposé deux algorithmes : CYCLOBLISS (CYCLOstationary BLInd Source Separation) et CYMOD (CYclic Multiple Output Deconvolution) pour la séparation des mélanges convolutifs de signaux cyclo-stationnaires en utilisant les statistiques cycliques d'ordre deux. Bouguerriou et al. [34] ont suggéré une méthode permettant d'extraire un signal d'intérêt cyclo-stationnaire de fréquence cyclique connue en exploitant l'auto-corrélation cyclique dans le cadre d'un mélange instantané. Rhioui et al. ont proposé dans [160], une méthode pour l'identification de la matrice de mélange quelque soit la détermination du système de mélange pour des sources cyclo-stationnaires avec des fréquences cycliques distinctes. Pham a introduit dans [85] une méthode basée sur la DC par bloc des densités interspectrales cycliques. Capdessus et al. ont extrait dans [50] une source cyclo-stationnaire, dont la fréquence cyclique est connue a priori, d'un ensemble d'observations. Ils proposent un nouveau critère basé sur les statistiques de second ordre des observations qui est facile à estimer et conduit à une extraction avec une très bonne précision. Ghennioui et al. ont résolu dans [106] le problème de séparation de mélanges convolutifs de sources cyclo-stationnaires ayant des fréquences cycliques différentes et inconnues, en combinant une procédure de sélection automatique de matrices de corrélation cycliques à la BDC non unitaire basée la méthode de gradient. Cheviet et al. ont proposé dans [69, 68] d'exploiter les caractéristiques des signaux cyclo-stationnaires dans le cadre des modèles de probabilité temporelle non stochastiques afin de séparer simultanément toutes les sources, sans restreindre la distribution ou le nombre de fréquences de cycle de chaque source. Sabri et al. ont introduit dans [165] une approche dans le domaine fréquentiel pour l'identification du système

MIMO de signaux cyclo-stationnaires. Dans ce cas la DC a été appliqué à un ensemble de matrices de densité spectrale cyclique des observations pour identifier le système de mélange. Il a également été proposé un algorithme pour surmonter les ambiguïtés de fréquence et de phase. Ghaderi et al. ont présenté dans [101] une méthode pour l'extraction aveugle de sources cyclo-stationnaires dont les fréquences cycliques sont connues et certaines d'entre elles communes. L'extraction emploie la méthode Jacobi pour la diagonalisation de matrices de corrélation cyclique des observations. Zakaria et al. ont proposé dans [195] d'utiliser l'analyse en composantes morphologiques afin de séparer les contributions cyclo-stationnaires à l'ordre un et deux d'un signal de course à pied. Cette analyse emploie deux dictionnaires, à savoir la transformée en cosinus discrète et l'analyse d'enveloppe. Zhang et al. ont présenté dans [200] une méthode basée sur la cyclo-stationnarité d'électrocardiogramme fœtal (FECG), en particulier la cohérence cyclique intégrée pour estimer un sous-espace FECG à partir d'un ensemble d'enregistrements ECG mixtes provenant des électrodes thoraciques et abdominales placées une femme enceinte. Elia et al. ont combiné dans [88] les techniques de SAS et l'analyse cyclo-stationnaire pour la surveillance d'une *usure distribuée* dans les roulements de boîte de vitesses. Roussel et al. ont proposé dans [164] de décomposer les signaux électromyographiques en utilisant la SAS cyclo-stationnaires basée sur la DC des matrices de corrélation cycliques.

2.2 SAS par analyse en composantes parcimonieuses

Dans de nombreuses applications pratiques, le nombre de sources dépasse le nombre de capteurs. Par exemple en industrie, des contraintes techniques et financières limitent le nombre de capteurs lors de l'acquisition des signaux sur des systèmes mécaniques. Nous finissons donc par des mélanges sous-déterminés où la SAS devient compliquée car le système de mélange n'admet pas d'inverse gauche. Pour estimer les sources, il est nécessaire de compter sur un *préalable* sur les sources comme la parcimonie ou plus rigoureusement l'existence d'un domaine où les différents signaux sont parcimonieux (temporel, spatial, fréquentiel, temps-fréquence, temps-échelle). La SAS par analyse en composantes parcimonieuses (SCA) permet également de contourner l'hypothèse d'indépendance statistique des sources, employée dans les méthodes de l'ICA. Dans cette section nous commençons par définir les critères de mesure de la parcimonie puis nous détaillons le principe des méthodes de séparation par analyse en composantes parcimonieuses.

2.2.1 Représentations parcimonieuses

Un vecteur est parcimonieux au sens stricte lorsque la majorité de ses éléments sont nuls, et seuls quelques uns ont des valeurs non-nulles, tandis qu'au sens large, la parcimonie traduit le fait que la plupart des éléments ont une valeur faible, au moment où le reste prend des valeurs assez grandes. La Fig.(2.3) montre un exemple de signal sinusoïdal bruité, parcimonieux après application d'une transformée en cosinus discrète. Apparue depuis peu, la

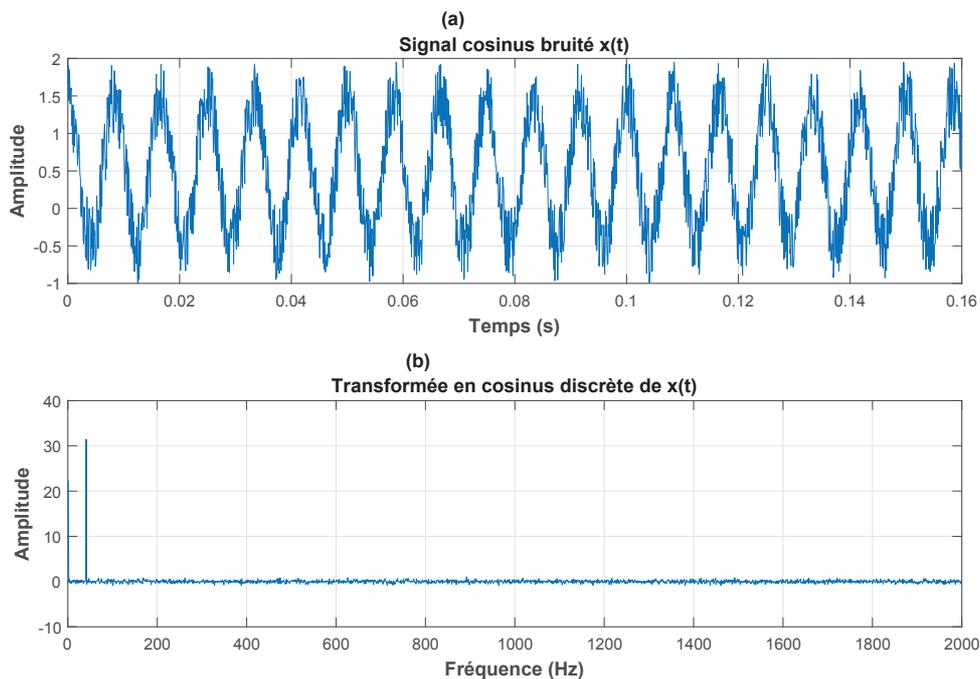


FIGURE 2.3 – (a) Signal cosinus bruité dans le domaine temporel. (b) Une transformée en cosinus discrète révèle le comportement parcimonieux du signal.

SCA ne cesse de susciter l'intérêt de la communauté de traitement de signal au regard de ses avantages. Les représentations parcimonieuses permettent de représenter un signal par une combinaison de signaux élémentaires définis à partir d'un *dictionnaire redondant*. L'objectif est d'estimer un signal parcimonieux \mathbf{x} tel que $\mathbf{y} = \mathbf{D}\mathbf{x}$, où \mathbf{y} représente le signal observé et \mathbf{D} est le dictionnaire. Généralement, la représentation parcimonieuse présente un compromis entre la qualité de l'approximation et la parcimonie du signal. Nous cherchons à minimiser l'erreur entre le signal \mathbf{y} et la représentation $\mathbf{D}\mathbf{x}$ en fixant $\epsilon > 0$ suffisamment petit, selon une norme ℓ^p :

$$\|\mathbf{y} - \mathbf{D}\mathbf{x}\|_p < \epsilon, \quad (2.38)$$

avec

$$\|\mathbf{x}\|_p = \left(\sum_k |\mathbf{x}_k|^p \right)^{\frac{1}{p}}. \quad (2.39)$$

En partant de ces définitions, il est clair que la norme ℓ^0 est la mesure intuitive pour quantifier cette parcimonie. La norme ℓ^0 consiste à compter le nombre éléments non nuls du vecteur \mathbf{x} . Toutefois, la minimisation exacte selon ℓ^0 fait partie de la classe des problèmes NP-complets (NP-hard). La majorité des travaux qui reposent sur la SCA, se contentent d'opter pour la norme ℓ^1 qui offre une solution moins optimale. En contrepartie, la minimisation est calculable par des techniques de programmation linéaire.

Plusieurs algorithmes parcimonieux ont été développés ces vingt dernières années à savoir les algorithmes gloutons (Matching Pursuit [135], Orthogonal Matching Pursuit [151], Orthogonal Least Square [62], ...) qui améliorent la qualité d'approximation itérativement par identification de signaux élémentaires supplémentaires et les algorithmes de relaxation (Basis Pursuit [63], Least Absolute Shrinkage and Selection Operator [177], FOCal Under-determined System Solver [157], ...).

Le choix du dictionnaire \mathbf{D} a un très grand impact sur la qualité de la décomposition \mathbf{x} du signal \mathbf{y} . En effet, si le dictionnaire comporte peu d'atomes adaptés aux structures présentes dans le signal \mathbf{y} , la décomposition \mathbf{x} sera mauvaise. Le choix du dictionnaire peut être orienté selon la connaissance a priori du signal, la coût de calcul, la variation des composantes fréquentielles, etc. . . Un choix typique pour des mélanges audio sera un dictionnaire temps-fréquence de type Gabor, ou bien temps-échelle de type ondelettes, également pour des signaux biologiques tels que les signaux cardiaques ou d'électro-encéphalogrammes. Pour des images naturelles, nous faisons appel à des dictionnaires d'ondelettes, de curvelets ou bien d'atomes de Gabor anisotropes.

2.2.2 Principe des méthodes parcimonieuses

Les méthodes de SAS par SCA suivent généralement les étapes suivantes (voir la Fig.(2.4)) :

1. La recherche d'un domaine de représentation des mélanges de telle manière à ce que les coefficients des sources dans la nouvelle représentation aient des supports aussi disjoints que possible. Nous employons pour cela des techniques de représentation parcimonieuse conjointe à savoir l'application d'une transformée linéaire (transformée de Fourier, transformée de Fourier à court terme, transformée en ondelettes ...)

[100] [113] [189] [5] [132] ou la minimisation d'un critère de mesure de parcimonie comme les normes ℓ^p , pour $0 \leq p \leq \infty$.

2. L'estimation de la matrice de mélange se fait selon la détermination du mélange. Dans le cas déterminé et sur-déterminé, en faisant appel aux méthodes d'ICA, et dans le cas sous-déterminé, le diagramme de dispersion des nouveaux mélanges parcimonieux montre typiquement une densité élevée dans les directions des vecteurs de mélange. Ces densités peuvent être trouvées en utilisant les algorithmes de catégorisation (K-means, C-means, ...) [118] [174] ou encore en étudiant les pics de l'histogramme des directions spatiales [28].
3. La séparation dans le domaine transformé en utilisant le masquage binaire⁷, en supposant que certaines sources sont actives au même moment que d'autres sont passives [159], ou en employant la minimisation d'un critère de parcimonie ℓ^p .
4. La reconstruction des sources.

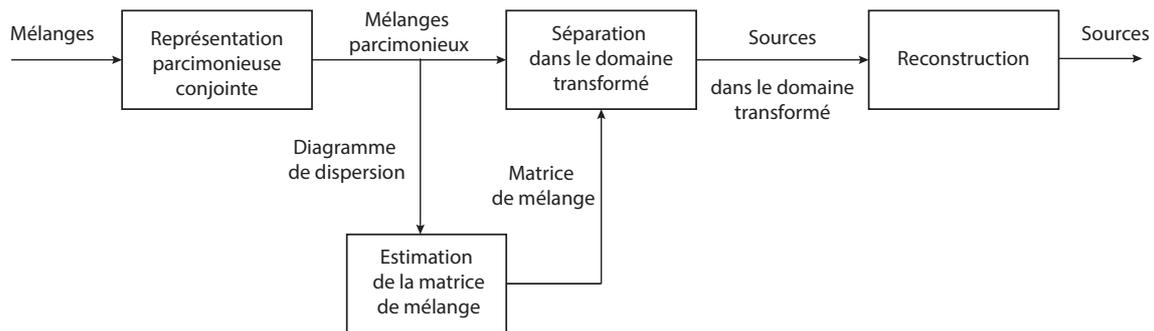


FIGURE 2.4 – Principe de séparation par analyse en composantes parcimonieuses.

2.3 SAS par factorisation en matrices non-négatives

Dans certaines applications réelles comme la transcription automatique des morceaux de musique, la reconnaissance d'instruments, ou l'estimation de la mélodie ..., nous traitons des données ayant uniquement un support positif ou nul. Or, les méthodes précédemment étudiées basées sur l'ICA ou SCA ne peuvent pas garantir une bonne interprétabilité physique des résultats. Ce qui a conduit nombreux travaux à développer d'autres méthodes prenant en considération le caractère non-négatif des données utilisées, comme

7. Seule la source la plus active se voit affecter un coefficient non nul, tandis que les coefficients des autres sources sont mis à zéro, ou "masqués".

celles fondées sur la technique de factorisation en matrices non négatives (NMF). La NMF permet la réduction dimensionnelle d'un ensemble de données afin d'en extraire les éléments représentatifs. Initialement, la NMF a été proposée par Lee et Seung [126] pour la reconnaissance de visages. Ces travaux ont été inspirés par les travaux antérieurs de Paateero [149] sur la factorisation en matrices positives. Mathématiquement, le problème de la NMF est formulé comme suit. Soit \mathbf{X} une matrice de données de taille $p \times q$ dont les éléments sont tous non-négatifs ($x_{ij} \geq 0$), la NMF consiste à retrouver deux matrices non négatives $\mathbf{A} \in \mathbb{R}_+^{p \times r}$ et $\mathbf{S} \in \mathbb{R}_+^{r \times q}$ dont le produit est une approximation de \mathbf{X} tel que,

$$\mathbf{X} \approx \mathbf{A}\mathbf{S}, \quad (2.40)$$

où l'entier naturel r est généralement choisi de telle sorte que $pr + rq \ll pq$, ce qui permet de réduire la dimension des données. Dans le cadre de la SAS, la matrice \mathbf{X} est la matrice composée des vecteurs lignes représentant les observations. De même, \mathbf{S} est la matrice dont les lignes représentent les sources tandis que la matrice \mathbf{A} est la matrice de mélange. La factorisation est résolue par la minimisation de la divergence entre \mathbf{X} et le produit $\mathbf{A}\mathbf{S}$:

$$\min_{\mathbf{A}, \mathbf{S} \geq 0} \mathbf{D}(\mathbf{X} \parallel \mathbf{A}\mathbf{S}). \quad (2.41)$$

Différentes mesures $\mathbf{D}(\mathbf{X} \parallel \mathbf{A}\mathbf{S})$ ont été étudiées dans la littérature. D. D. Lee et H. S. Seung ont proposé deux méthodes basées sur la minimisation de deux fonctions objectives différentes, avec des règles de mise à jour multiplicatives. Elles emploient respectivement la divergence euclidienne [126] et la divergence de Kullback Leibler [127] entre \mathbf{X} et $\mathbf{A}\mathbf{S}$. Les règles de mise à jour multiplicatives peuvent s'écrire de la manière suivante :

— Avec la divergence euclidienne :

$$\mathbf{A} \leftarrow \mathbf{A} \odot \frac{\mathbf{X}\mathbf{S}^T}{\mathbf{A}\mathbf{S}\mathbf{S}^T}, \quad \mathbf{S} \leftarrow \mathbf{S} \odot \frac{\mathbf{A}^T\mathbf{X}}{\mathbf{A}^T\mathbf{A}\mathbf{S}}, \quad (2.42)$$

— Avec la divergence de Kullback Leibler :

$$\mathbf{A} \leftarrow \mathbf{A} \odot \frac{\mathbf{X}}{\mathbf{A}\mathbf{S}} \mathbf{S}^T, \quad \mathbf{S} \leftarrow \mathbf{S} \odot \frac{\mathbf{A}^T \mathbf{X}}{\mathbf{A}^T \mathbf{E}}, \quad (2.43)$$

où l'opérateur \odot représente le produit de Hadamard (multiplication point par point) et $\mathbf{E} \in \mathbb{R}_+^{p \times q}$ est une matrice unitaire, et la division se fait point par point. Contrairement à leurs algorithmes initiaux utilisant des règles de mise à jour multiplicatives, Lee et Seung ont proposé une autre règle de mise à

jour basée sur gradient descendant et minimisant la fonction objective (2.41) au sens des moindres carrés non-négatifs. Depuis, plusieurs algorithmes basés sur la NMF ont vu le jour dans différents domaines d'applications [72] [84] [86] [192]. Elle a suscité particulièrement un grand intérêt dans le traitement des signaux audio, notamment en identification sonore [48] [171], en transcription musicale [182] ou encore en séparation de sources [184] [72] [148] [183]. Pour plus de détails sur algorithmes utilisant la NMF, nous renvoyons le lecteur à l'ouvrage suivant [73].

2.4 Discussions et conclusion

Les méthodes de SAS présentent des limitations qui détériorent leurs performances, bien que la communauté dispose actuellement d'une panoplie d'outils théoriques adaptée. Dans le contexte de mélanges instantanés, l'étape préalable de blanchiment des observations, présente dans certains algorithmes de séparation, peut influencer négativement les performances et la qualité de séparation. En effet, ces algorithmes sont sensibles à tout bruit coloré de cohérence spatiale inconnue. Quant aux algorithmes exploitant la DC des matrices de corrélation retardées, elles ne sont efficaces que dans le cas de sources à spectres distincts. Au sujet des algorithmes par déflation, ces derniers souffrent de l'effet domino, où la moindre erreur d'estimation d'une seule source se répercute automatiquement sur l'estimation des autres, fournissant ainsi des résultats complètement erronés. A l'égard des méthodes utilisant l'estimateur du maximum de vraisemblance, certaines d'entre elles estiment grossièrement les densités de probabilité des sources par une sous ou sur-gaussienne. Cette estimation s'avère parfois incorrecte et détériore considérablement la qualité de séparation.

Dans le contexte de mélanges convolutifs, les méthodes visant à ramener ce mélange à un mélange instantané dans le domaine temporel, nécessite à la fois de connaître la longueur des filtres caractérisant le système et des échantillons considérables. Ces méthodes peuvent s'avérer coûteuses en termes de complexité lorsqu'elles exploitent les statistiques d'ordre supérieur. Tandis que celles qui opèrent dans le domaine fréquentiel en représentant le mélange convolutif par une suite de mélanges instantanés à chaque canal de fréquence, souffrent de deux indéterminations majeures à savoir l'ambiguïté de phase de la réponse en fréquence du système et l'ambiguïté de permutation liée à l'identification de chaque mélange instantané fréquentiel.

Dans le cas de mélanges sous-déterminés, ou si l'hypothèse d'indépendance statistique n'est pas vérifiée, la SAS par SCA améliore généralement la so-

lution a ces problèmes. Cependant, l'hypothèse de parcimonie ne semble pas naturellement adaptée à plusieurs applications réelles. De plus, le choix d'un dictionnaire ajusté lorsque nous cherchons une représentation parcimonieuse conjointe des observations n'est pas trivial. En effet, si le dictionnaire comporte peu d'atomes adaptés aux structures présentes dans le signal, la représentation sera erronée.

Quant aux méthodes reposant sur la NMF, elles souffrent d'une convergence vers un optimum global non garantie. De plus, ces méthodes semblent très coûteuses en temps lors de la mise à jour des éléments du produit matriciel.

Séparation de sources cyclo-stationnaires : application aux signaux de télécommunications

Sommaire

3.1 Séparation de mélanges convolutifs de sources cyclo-stationnaires basée sur la bloc-diagonalisation conjointe matricielle	52
3.1.1 Formulation du problème et hypothèses	52
3.1.2 Préliminaires théoriques	53
3.1.3 Procédure de sélection automatique des points retards-fréquences cycliques	56
3.1.4 Bloc-diagonalisation conjointe des matrices basée sur l'algorithme BFGS	59
3.1.5 Simulations et résultats numériques	63
3.2 Identification de la matrice de mélange dans le cas instantané	67
3.2.1 Modèle des sources et propriété importante	67
3.2.2 Approche proposée combinant la DVP et une méthode de regroupement hiérarchique	71
3.2.3 Simulations et résultats numériques	73
3.3 Conclusion	76

UN tour d'horizon des techniques de SAS, ayant été présenté dans les deux premiers chapitres, nous présenterons dans ce chapitre les contribu-

tions de ce travail de recherche appliqué à la séparation de signaux cyclo-stationnaires pour des systèmes présentant des mélanges sur-déterminés. Nous proposerons à cet égard, des approches qui tirent profit des structures algébriques particulières que la matrice de corrélation cyclique des sources puisse avoir selon la corrélation ou non des sources et leur partage des fréquences cycliques ou non. Un nouvel algorithme de bloc-diagonalisation conjointe combiné avec une procédure de sélection automatique des couples de points (retards-fréquences cycliques) sera introduit. Nous proposerons également une approche qui associe la décomposition en valeurs propres des matrices de corrélation cycliques des observations à une méthode de regroupement hiérarchique afin d'identifier la matrice de mélange puis la restitution des sources. Enfin, nous présenterons des simulations informatiques montrant l'efficacité de ces approches dans le domaine des télécommunications.

3.1 Séparation de mélanges convolutifs de sources cyclo-stationnaires basée sur la bloc-diagonalisation conjointe matricielle

3.1.1 Formulation du problème et hypothèses

La SAS vise à récupérer n sources $\mathbf{s}(t) = [s_1(t), \dots, s_n(t)]^T$ à partir de m observations $\mathbf{x}(t) = [x_1(t), \dots, x_m(t)]^T$ sans aucune pré-connaissance sur le système de mélange. Dans le contexte des systèmes MIMO, nous considérons le modèle donné précédemment par l'équation (1.2) :

$$\forall i = 1, \dots, m, \quad x_i(t) = \sum_{j=1}^n \sum_{k=0}^{\rho} a_{ij}(k) s_j(t-k) + r_i(t), \quad (3.1)$$

où $\mathbf{A}(t) = \sum_{i=1}^m \sum_{j=1}^n a_{ij}(t)$ représente la matrice des réponses impulsionnelles finies à temps discret des filtres d'ordre ρ , associés à chaque relation entre une source et une observation. Nous rappelons que ce système de mélanges peut être réécrit de la façon suivante :

$$\mathbf{X}(t) = \mathbf{G}\mathbf{S}(t) + \mathbf{R}(t), \quad (3.2)$$

où :

- la matrice de mélange \mathbf{G} est une matrice par blocs de taille $(M \times N)$ donnée par $\{\mathbf{G} = (\mathbf{G}_{ij}) / \forall i = 1, \dots, m, \forall j = 1, \dots, n\}$ où les blocs

\mathbf{G}_{ij} sont des matrices de Toeplitz ayant pour dimensions $(\rho' \times \nu)$:

$$\mathbf{G}_{ij} = \begin{bmatrix} a_{ij}(0) & \dots & a_{ij}(\rho) & 0 & \dots & 0 \\ 0 & \ddots & \ddots & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & \ddots & \ddots & 0 \\ 0 & \dots & 0 & a_{ij}(0) & \dots & a_{ij}(\rho) \end{bmatrix},$$

- le vecteur source $\mathbf{S}(t) = [\mathbf{s}_1^T(t), \mathbf{s}_2^T(t), \dots, \mathbf{s}_n^T(t)]^T$ a pour taille $(N \times 1)$ où les vecteurs $\mathbf{s}_i(t)$ ($\forall i = 1, \dots, n$) ont pour taille $(\nu \times 1)$ et sont donnés par $\mathbf{s}_i(t) = [\mathbf{s}_i(t), \mathbf{s}_i(t-1), \dots, \mathbf{s}_i(t-\nu+1)]^T$,
- les vecteurs $\mathbf{X}(t) = [\mathbf{x}_1^T(t), \mathbf{x}_2^T(t), \dots, \mathbf{x}_m^T(t)]^T$ et $\mathbf{R}(t) = [\mathbf{r}_1^T(t), \mathbf{r}_2^T(t), \dots, \mathbf{r}_m^T(t)]^T$ ont pour taille $(M \times 1)$ et contiennent respectivement les observations et les bruits où les vecteurs $\mathbf{x}_i(t)$ et $\mathbf{r}_i(t)$ ($\forall i = 1, \dots, m$) ont pour taille $(\rho' \times 1)$ et sont donnés respectivement par $\mathbf{x}_i(t) = [\mathbf{x}_i(t), \mathbf{x}_i(t-1), \dots, \mathbf{x}_i(t-\rho'+1)]^T$ et $\mathbf{r}_i(t) = [\mathbf{r}_i(t), \mathbf{r}_i(t-1), \dots, \mathbf{r}_i(t-\rho'+1)]^T$,
- $M = m \times \rho'$, $N = n \times (\rho + \rho') = n \times \nu$ avec ρ' est un entier représentant le nombre de retards choisi de telle façon à vérifier l'inégalité $N \leq M \Leftrightarrow n(\rho + \rho') \leq m\rho'$ afin de se placer dans le cas d'un mélange sur-déterminé.

En outre, nous supposons que les sources sont de moyenne nulle, décorréelées entre elles, cyclo-stationnaires à l'ordre 2 et les bruits sont blancs, stationnaires, gaussiens et décorréelées des sources.

3.1.2 Préliminaires théoriques

Matrice de corrélation cyclique des sources et son estimateur empirique

Le vecteur des sources comporte N signaux cyclo-stationnaires où la fonction de corrélation $\mathbf{C}_{\mathbf{S}_i \mathbf{S}_i}(t, \tau)$ de chaque source \mathbf{S}_i est par définition périodique de fréquence T_i . Par conséquent $\mathbf{C}_{\mathbf{S}_i \mathbf{S}_i}(t, \tau)$ peut être décomposée en séries de Fourier :

$$\mathbf{C}_{\mathbf{S}_i \mathbf{S}_i}(t, \tau) = \mathbf{E}\left\{\mathbf{S}_i\left(t - \frac{\tau}{2}\right)\mathbf{S}_i^H\left(t + \frac{\tau}{2}\right)\right\} \quad (3.3)$$

$$= \sum_{\xi_i \in \Psi_i} \mathbf{C}_{\mathbf{S}_i \mathbf{S}_i}^{\xi_i}(\tau) e^{j2\pi\xi_i t}, \quad (3.4)$$

où $\mathbf{C}_{\mathbf{S}_i \mathbf{S}_i}^{\xi_i}(\tau)$ est la fonction de corrélation cyclique définie par

$$\mathbf{C}_{\mathbf{S}_i \mathbf{S}_i}^{\xi_i}(\tau) = \lim_{T_i \rightarrow \infty} \frac{1}{T_i} \int_{-\frac{T_i}{2}}^{\frac{T_i}{2}} \mathbf{C}_{\mathbf{S}_i \mathbf{S}_i}(t, \tau) e^{-j2\pi\xi_i t} dt. \quad (3.5)$$

Si le vecteur des sources comporte N_s échantillons alors la fonction de corrélation cyclique d'une source \mathbf{S}_i est estimée en substituant l'intégrale dans l'équation (3.5) par une somme de la façon suivante :

$$\hat{\mathbf{C}}_{\mathbf{S}_i \mathbf{S}_i}^{\xi_i}(\tau) = \frac{1}{N_s} \sum_{n=1}^{N_s} \mathbf{S}_i(n) \mathbf{S}_i(n - \tau) e^{-j2\pi\xi_i n}. \quad (3.6)$$

Nous définissons l'ensemble Ψ_i des fréquences cycliques relatif à la source \mathbf{S}_i comme suit :

$$\Psi_i = \{\xi_i = \frac{k}{T_i}, k \in \mathbb{Z}\}. \quad (3.7)$$

À la lumière de l'hypothèse de la décorrélation des sources, la matrice de corrélation cyclique des sources est donc une matrice diagonale de taille $(N \times N)$ et s'écrit de la façon suivante :

$$\mathbf{C}_{\mathbf{S}}^{\xi}(\tau) = \begin{bmatrix} \mathbf{C}_{\mathbf{S}_1 \mathbf{S}_1}^{\xi}(\tau) & 0 & \dots & 0 \\ 0 & \mathbf{C}_{\mathbf{S}_2 \mathbf{S}_2}^{\xi}(\tau) & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \dots & 0 & \mathbf{C}_{\mathbf{S}_N \mathbf{S}_N}^{\xi}(\tau) \end{bmatrix},$$

où l'une des structures de la matrice $\mathbf{C}_{\mathbf{S}}^{\xi}(\tau)$ s'écrit :

- Si les sources ont des fréquences cycliques distinctes alors les matrices $\mathbf{C}_{\mathbf{S}}^{\xi}(\tau)$ sont bloc-diagonales avec un seul bloc non nul de taille $(\nu \times \nu)$ sur leur diagonale situé à la position (i, i) correspondant à la i -ème source de fréquence ξ_i . En effet, $\forall \xi_j \in \Psi_j, \forall \tau \in \mathcal{Z}, \forall i \in \{1, \dots, N\}$, nous avons :

$$\begin{aligned} \mathbf{C}_{\mathbf{S}_i \mathbf{S}_i}^{\xi_j}(\tau) &= \lim_{T \rightarrow \infty} \frac{1}{T} \int_{-\frac{T}{2}}^{\frac{T}{2}} \mathbf{C}_{\mathbf{S}_i \mathbf{S}_i}(t, \tau) e^{-j2\pi\xi_j t} dt, \\ &= \lim_{T \rightarrow \infty} \frac{1}{T} \int_{-\frac{T}{2}}^{\frac{T}{2}} \sum_{\xi_i \in \Psi_i} \mathbf{C}_{\mathbf{S}_i \mathbf{S}_i}^{\xi_i}(\tau) e^{j2\pi\xi_i t} e^{-j2\pi\xi_j t}, \\ &= \sum_{\xi_i \in \Psi_i} \mathbf{C}_{\mathbf{S}_i \mathbf{S}_i}^{\xi_i}(\tau) \lim_{T \rightarrow \infty} \frac{1}{T} \int_{-\frac{T}{2}}^{\frac{T}{2}} e^{j2\pi(\xi_i - \xi_j)t}, \\ &= \sum_{\xi_j} \mathbf{C}_{\mathbf{S}_i \mathbf{S}_i}^{\xi_i}(\tau) \delta(\xi_i - \xi_j), \end{aligned}$$

où

$$\delta(\xi) = \lim_{T \rightarrow \infty} \frac{1}{T} \int_{-\frac{T}{2}}^{\frac{T}{2}} e^{-j2\pi\xi t} dt, \quad (3.8)$$

entraînant :

$$\delta(\xi) = \begin{cases} 1 & \text{si } \xi = 0 \\ 0 & \text{sinon} \end{cases},$$

dans la mesure où les sources ont des fréquences cycliques distinctes ($\xi_j \neq \xi_i$) alors $\mathbf{C}_{\mathbf{S}_i \mathbf{S}_j}^{\xi_j}(\tau)$ est nulle.

- b. Si les sources partagent une ou plusieurs fréquences cycliques alors les matrices $\mathbf{C}_{\mathbf{S}}^{\xi}(\tau)$ sont bloc-diagonales avec n blocs de taille $(\nu \times \nu)$ à ces fréquences communes.

Matrice de corrélation cyclique des observations et bruits

La matrice de corrélation des observations $\mathbf{C}_{\mathbf{X}}(t, \tau)$ est définie par :

$$\mathbf{C}_{\mathbf{X}}(t, \tau) = \mathbf{E}\{\mathbf{X}(t - \frac{\tau}{2})\mathbf{X}^H(t + \frac{\tau}{2})\}. \quad (3.9)$$

A partir de l'équation (1.5), il est facile de démontrer que la matrice $\mathbf{C}_{\mathbf{X}}(t, \tau)$ peut s'écrire de la manière suivante :

$$\mathbf{C}_{\mathbf{X}}(t, \tau) = \mathbf{G}\mathbf{C}_{\mathbf{S}}(t, \tau)\mathbf{G}^H + \mathbf{C}_{\mathbf{R}}(\tau). \quad (3.10)$$

où $\mathbf{C}_{\mathbf{S}}(\tau)$ (resp. $\mathbf{C}_{\mathbf{R}}(\tau)$) est la matrice de corrélation des sources (resp. bruits). Si nous considérons la décomposition en série de Fourier de la matrice $\mathbf{C}_{\mathbf{X}}(t, \tau)$, nous avons donc :

$$\mathbf{C}_{\mathbf{X}}^{\xi}(\tau) = \lim_{T \rightarrow \infty} \frac{1}{T} \int_{-\frac{T}{2}}^{\frac{T}{2}} \mathbf{C}_{\mathbf{X}}(t, \tau) e^{-j2\pi\xi t} dt. \quad (3.11)$$

Compte tenu de la linéarité du coefficient de Fourier, nous pouvons écrire :

$$\mathbf{C}_{\mathbf{X}}^{\xi}(\tau) = \mathbf{G}\mathbf{C}_{\mathbf{S}}^{\xi}(\tau)\mathbf{G}^H + \mathbf{C}_{\mathbf{R}}^{\xi}(\tau), \quad (3.12)$$

où $\mathbf{C}_{\mathbf{R}}^{\xi}(\tau)$ est la matrice de corrélation cyclique bruits. L'expression de $\mathbf{C}_{\mathbf{R}}^{\xi}(\tau)$ peut être plus développée. Sachant que les bruits sont stationnaires, alors :

$$\mathbf{C}_{\mathbf{R}}^{\xi}(\tau) = \mathbf{C}_{\mathbf{R}}(\tau)\delta(\xi), \quad (3.13)$$

si $\xi \neq 0$, nous obtenons ainsi :

$$\mathbf{C}_{\mathbf{X}}^{\xi}(\tau) = \mathbf{G}\mathbf{C}_{\mathbf{S}}^{\xi}(\tau)\mathbf{G}^H. \quad (3.14)$$

3.1.3 Procédure de sélection automatique des points retards- fréquences cycliques

Un des paramètres clé dans un signal cyclo-stationnaire est la fréquence cyclique. Sa connaissance *a priori* est souvent nécessaire pour résoudre plusieurs problèmes réels comme ceux rencontrés dans le domaine des télécommunications tels que l'égalisation aveugle des canaux, la détection des symboles et la synchronisation. Cependant, en pratique, la fréquence cyclique n'est pas toujours connue et nécessite d'être estimée à l'aide d'un détecteur automatique des matrices de corrélation cyclique des sources, possédant une des structures citées dans (§3.1.2, a et b).

Fréquences cycliques communes aux sources

Comme nous l'avons précisé précédemment, lorsque les sources ont des fréquences cycliques communes alors leurs matrices de corrélation $\mathbf{C}_S^\xi(\tau)$ sont bloc-diagonales avec n blocs de taille $(\nu \times \nu)$ à ces fréquences communes, autrement dit $\mathbf{C}_S^\xi(\tau)$ comportent N termes non nuls. Cette structure sera une cible à traquer par une procédure automatique de détection de points (retards-fréquences cycliques) relatifs aux matrices $\mathbf{C}_X^\xi(\tau)$. La procédure consiste en 2 étapes :

1. Calculer $\mathbf{C}_X^\xi(\tau)$ pour un nombre suffisant de fréquences afin d'assurer le balayage de presque toutes les fréquences cycliques des sources.
2. Détecter les matrices $\mathbf{C}_{X_w}^\xi(\tau) = \mathbf{W}\mathbf{C}_X^\xi(\tau)\mathbf{W}^H$ ¹ ayant le même nombre N de valeurs propres que $\mathbf{C}_S^\xi(\tau)$ puisque toutes les sources sont présentes à une fréquence cyclique commune donnée. Afin d'accéder aveuglement aux termes diagonaux de $\mathbf{C}_S^\xi(\tau)$, nous calculons les valeurs propres de $\mathbf{C}_{X_w}^\xi(\tau)$ comme suit :

$$\text{eig}(\mathbf{C}_{X_w}^\xi(\tau)) = \text{eig}(\mathbf{W}\mathbf{G}\mathbf{C}_S^\xi(\tau)\mathbf{G}^H\mathbf{W}^H) \quad (3.15)$$

$$= \text{eig}(\mathbf{C}_S^\xi(\tau)) = [\theta_1, \dots, \theta_N]^T, \quad (3.16)$$

où le vecteur $\text{eig}(\mathbf{M})$ contient les valeurs propres de la matrice carrée \mathbf{M} , $\mathbf{W}\mathbf{G}$ est une matrice unitaire et θ_i pour $i = 1, \dots, N$ sont les valeurs propres de $\mathbf{C}_S^\xi(\tau)$. Par conséquent, nous proposons le premier critère défini par :

$$\mathbf{O}_1 = \frac{\sum_{i=1}^N (\theta_i)^2}{\|\mathbf{C}_{X_w}^\xi(\tau)\|_F^2} \geq 1 - \epsilon, \quad (3.17)$$

1. \mathbf{W} est une matrice de blanchiment de taille $(N \times M)$ telle que $\mathbf{W}\mathbf{G}\mathbf{G}^H\mathbf{W}^H = \mathbf{I}_N$.

où $\|\cdot\|_F^2$ désigne la norme de Frobenius et ϵ est une constante positive de faible valeur. Nous avons également l'égalité suivante $\|\mathbf{C}_{\mathbf{x}_w}^\xi(\tau)\|_F^2 = \|\mathbf{C}_S^\xi(\tau)\|_F^2$ car la norme de Frobenius est invariante sous une transformation unitaire. Idéalement, \mathbf{O}_1 est égal à 1. Cependant, en pratique, les matrices $\mathbf{C}_S^\xi(\tau)$ ne sont pas exactement bloc-diagonales à cause des erreurs d'estimation de ces matrices et du nombre fini des échantillons des signaux.

Ainsi, les matrices $\mathbf{C}_{\mathbf{x}_w}^\xi(\tau)$ doivent être sélectionnées de telle sorte à vérifier $\mathbf{O}_1 \geq 1 - \epsilon$ avec ϵ proche de zéro. En outre, afin d'éviter de sélectionner des matrices $\mathbf{C}_{\mathbf{x}_b}^\xi(\tau)$ pour de faibles valeurs, celles-ci $\mathbf{C}_{\mathbf{x}_w}^\xi(\tau)$ doivent vérifier le deuxième critère défini par :

$$\mathbf{O}_2 = \det(\mathbf{C}_{\mathbf{x}_w}^\xi(\tau)) = \det(\mathbf{C}_S^\xi(\tau)) \geq \eta, \quad (3.18)$$

où $\det(\cdot)$ désigne le déterminant matriciel, et η une constante positive assez grande.

Enfin, si une matrice $\mathbf{C}_{\mathbf{x}_w}^\xi(\tau)$ satisfait les critères (3.17) et (3.18) alors elle est retenue dans l'ensemble des matrices à bloc-diagonaliser conjointement. Le choix des valeurs de ϵ et η dépend de la nature des signaux traités et de leurs énergies. Généralement, si nous souhaitons réduire le nombre de matrices sélectionnées, ϵ doit être quasiment nulle et η doit être de l'ordre de la valeur maximale des tous les déterminants matriciels de toutes les matrices $\mathbf{C}_{\mathbf{x}_w}^\xi(\tau)$ balayées.

Fréquences cycliques distinctes des sources

Dans le cas où les sources ont des fréquences cycliques distinctes, Giulieri et al. [109] ont proposé une procédure automatique de sélection des matrices $\mathbf{C}_S^\xi(\tau)$ de rang 1 pour les mélanges instantanés, puis Ghennioui et al. ont généralisé [103], [102] cette procédure dans le cadre des mélanges convolutifs pour sélectionner les matrices $\mathbf{C}_S^\xi(\tau)$ de rang ν , car les matrices $\mathbf{C}_S^\xi(\tau)$ le sont également. Ces matrices ont une structure bloc-diagonale avec un seul bloc non nul de taille $(\nu \times \nu)$ sur la diagonale situé à la position (i, i) correspondant à la i -ème source. En effet,

$$\begin{aligned} \text{rank}(\mathbf{C}_S^\xi(\tau)) &= \text{rank}(\mathbf{G}\mathbf{C}_S^\xi(\tau)\mathbf{G}^H) \\ &\leq \min(\text{rank}(\mathbf{G}), \text{rank}(\mathbf{C}_S^\xi(\tau)), \text{rank}(\mathbf{G}^H)), \end{aligned}$$

car

$$\text{rank}(\mathbf{G}) = \text{rank}(\mathbf{G}^H),$$

il s'ensuit que

$$\text{rank}(\mathbf{C}_{\mathbf{X}}^{\xi}(\tau)) \leq \min(\text{rank}(\mathbf{G}), \text{rank}(\mathbf{C}_{\mathbf{S}}^{\xi}(\tau))),$$

et par conséquent

$$\text{rank}(\mathbf{C}_{\mathbf{S}}^{\xi}(\tau)) = \nu \Rightarrow \text{rank}(\mathbf{C}_{\mathbf{X}}^{\xi}(\tau)) = \nu,$$

en sachant que \mathbf{G} est une matrice de rang plein, alors :

$$\text{rank}(\mathbf{C}_{\mathbf{X}}^{\xi}(\tau)) = \text{rank}(\mathbf{C}_{\mathbf{S}}^{\xi}(\tau)).$$

Cette procédure consiste en deux étapes :

1. Calculer $\mathbf{C}_{\mathbf{X}}^{\xi}(\tau)$ pour un nombre suffisant de bancs de fréquences afin d'assurer le balayage de presque toutes les fréquences cycliques des sources.
2. Décomposer en éléments singuliers les matrices $\mathbf{C}_{\mathbf{X}}^{\xi}(\tau)$ telles que :

$$\mathbf{C}_{\mathbf{X}}^{\xi}(\tau) = \mathbf{U}\mathbf{\Delta}\mathbf{V}^H, \quad (3.19)$$

où \mathbf{V} et \mathbf{U} sont des matrices unitaires de taille $(M \times M)$ et $\mathbf{\Delta} = \text{diag}\{\theta\} = [\theta_1, \dots, \theta_M]^T$ une matrice diagonale dont les éléments diagonaux sont positifs et triés selon un ordre descendant. Pour vérifier si une matrice $\mathbf{C}_{\mathbf{X}}^{\xi}(\tau)$ est de rang ν alors elle doit vérifier le troisième critère défini par :

$$\mathbf{O}_3 = \frac{\sum_{i=1}^{\nu} \theta_i^2}{\|\mathbf{C}_{\mathbf{X}}^{\xi}(\tau)\|_F^2} \geq 1 - \epsilon_1. \quad (3.20)$$

avec ϵ_1 est un scalaire positif très proche de zéro. En conclusion, lorsqu'une matrice $\mathbf{C}_{\mathbf{X}}^{\xi}(\tau)$ vérifie (3.20), elle est retenue dans l'ensemble des matrices à bloc-diagonaliser conjointement.

Les critères de sélection présentés dans cette section permettent en dehors de la composition de l'ensemble des matrices à (bloc) diagonaliser conjointement, l'estimation des fréquences cycliques des sources communes ou distinctes et leur nombre lorsque celles-ci ne partagent pas de fréquences cycliques. Cela peut être fait en comptant le nombre de fréquences cycliques non-multiples différentes.

3.1.4 Bloc-diagonalisation conjointe des matrices basée sur l'algorithme BFGS

La matrice de corrélation cyclique des observations admet comme nous l'avons montré dans l'équation (3.14) pour $\xi \neq 0$ la décomposition suivante :

$$\mathbf{C}_X^\xi(\tau) = \mathbf{G} \mathbf{C}_S^\xi(\tau) \mathbf{G}^H, \quad (3.21)$$

où les matrices $\mathbf{C}_S^\xi(\tau)$ sont bloc-diagonales ayant une des structures évoquées en (3.1.2). Nous proposons alors de construire un ensemble de matrices $\mathbf{C}_S^\xi(\tau)$ pour différentes valeurs de ξ et τ , à l'aide d'un des détecteurs automatiques présenté en (§3.1.3) puis de le bloc-diagonaliser afin d'estimer la matrice de mélange et finalement les sources.

Comme nous l'avons précisé dans (2.17), l'estimation de la matrice \mathbf{B} nécessite la minimisation sans contraintes de la fonction de coût suivante :

$$\mathcal{F}_{\text{BDC}}(\mathbf{B}) = \sum_{i=1}^K \|\text{OffBdiag}_{(n)}\{\mathbf{B} \mathbf{C}_X^{\xi_i}(\tau_i) \mathbf{B}^H\}\|_F^2 \quad (3.22)$$

Pour ce faire, nous proposons d'utiliser un schéma de minimisation basé sur l'algorithme quasi-newtonien de Broyden-Fletcher-Goldfarb-Shanno (BFGS)[96] [130] faisant partie de la classe des algorithmes à directions de descente² avec un calcul exact du pas d'adaptation. L'algorithme nécessite le calcul du gradient matriciel et l'approximation de la matrice Hessienne.

Principe de l'algorithme

A partir d'une matrice de séparation initiale $\mathbf{B}_{(0)}$, d'une matrice approchée $\mathbf{H}_{(0)}$ et d'un nombre donné d'itérations n_{iter} , les étapes suivantes sont répétées jusqu'à ce que $\mathbf{B}_{(k)}$ converge vers une solution.

S1. Trouver la direction de quasi-Newton $\mathbf{d}_{(k)}$ en résolvant l'équation suivante :

$$\mathbf{H}_{(k)} \mathbf{d}_{(k)} = -\nabla_a \mathcal{F}_{\text{BDC}}(\mathbf{B}_{(k)});$$

où $\nabla_a \mathcal{F}_{\text{BDC}}(\mathbf{B})$ représente le gradient matriciel complexe de la fonction de coût \mathcal{F}_{BDC} . Il est défini par (voir [115]) :

$$\nabla_a \mathcal{F}_{\text{BDC}}(\mathbf{B}) = 2 \frac{\partial \mathcal{F}_{\text{BDC}}(\mathbf{B})}{\partial \mathbf{B}^*}, \quad (3.23)$$

2. Les algorithmes à directions de descente sont des algorithmes itératifs qui commencent par une solution arbitraire, puis tentent de trouver une meilleure solution en changeant progressivement un seul élément de la solution. Si la modification produit une meilleure solution, un changement supplémentaire est apporté à la nouvelle solution, en répétant jusqu'à ce que d'autres améliorations puissent être trouvées.

avec \mathbf{B}^* représentant la conjuguée de la matrice \mathbf{B} et $\frac{\partial}{\partial \mathbf{B}^*}$ l'opérateur dérivée partielle. Le calcul de $\nabla_a \mathcal{F}_{\text{BDC}}(\mathbf{B})$ a été déjà accompli dans [105] et devient :

$$\begin{aligned} \nabla_a \mathcal{F}_{\text{BDC}}(\mathbf{B}) = & 2 \sum_{i=1}^K \left[\text{OffBdiag}_{(n)} \{ \mathbf{B} \mathbf{C}_X^{\xi_i}(\tau_i) \mathbf{B}^H \} \mathbf{B} \mathbf{C}_X^{\xi_i}(\tau_i)^H \right] \\ & + 2 \sum_{i=1}^K \left[\left(\text{OffBdiag}_{(n)} \{ \mathbf{B} \mathbf{C}_X^{\xi_i}(\tau_i) \mathbf{B}^H \} \right)^H \mathbf{B} \mathbf{C}_X^{\xi_i}(\tau_i) \right]. \end{aligned} \quad (3.24)$$

S2. Effectuer une recherche linéaire pour trouver le pas d'adaptation optimal $\alpha_{(k)}$ à valeur réelle positive dans la direction trouvée dans **S1** garantissant une courbure positive à chaque itération ;

S3. Mettre à jour $\mathbf{B}_{(k+1)}$:

$$\mathbf{B}_{(k+1)} = \mathbf{B}_{(k)} + \alpha_{(k)} \mathbf{d}_{(k)};$$

S4. Affecter les expressions ci-dessous à $\mathbf{p}_{(k)}$ et $\mathbf{y}_{(k)}$

$$\begin{aligned} \mathbf{p}_{(k)} &= \mathbf{B}_{(k+1)} - \mathbf{B}_{(k)}; \\ \mathbf{y}_{(k)} &= \nabla_a \mathcal{F}_{\text{BDC}}(\mathbf{B}_{(k+1)}) - \nabla_a \mathcal{F}_{\text{BDC}}(\mathbf{B}_{(k)}); \end{aligned}$$

S5. Au lieu de calculer d'une façon directe la matrice Hessienne, l'algorithme BFGS approxime cette dernière par une matrice $\mathbf{H}_{(k)}$ en utilisant les mises à jour du premier rang spécifiées par les évaluations du gradient :

$$\mathbf{H}_{(k+1)} = \mathbf{H}_{(k)} + \frac{\mathbf{y}_{(k)} \mathbf{y}_{(k)}^T}{\mathbf{y}_{(k)}^T \mathbf{p}_{(k)}} - \frac{\mathbf{H}_{(k)} \mathbf{p}_{(k)} \mathbf{p}_{(k)}^T \mathbf{H}_{(k)}}{\mathbf{p}_{(k)}^T \mathbf{H}_{(k)} \mathbf{p}_{(k)}}, \quad k = 1, \dots, n_{iter} \quad (3.25)$$

Certaines contraintes s'appliquent sur le choix de la matrice $\mathbf{B}_{(0)}$. En effet, cette dernière doit être de rang plein et différente de la matrice nulle car elle représente une solution triviale de (2.17), tandis que $\mathbf{H}_{(0)}$ doit être une matrice définie positive et peut être initialisée avec la matrice identité de telle façon qu'elle soit équivalente à celle de l'algorithme du gradient. La première étape de l'algorithme s'effectue à l'aide de l'inverse de la matrice $\mathbf{H}_{(k)}^{-1}$, obtenue en appliquant la formule Sherman-Morrison [16] à la matrice donnée par l'étape **S5**, c.à.d. :

$$\mathbf{H}_{(k+1)}^{-1} = \left(I - \frac{\mathbf{p}_{(k)} \mathbf{y}_{(k)}^T}{\mathbf{y}_{(k)}^T \mathbf{p}_{(k)}} \right) \mathbf{H}_{(k)}^{-1} \left(I - \frac{\mathbf{y}_{(k)} \mathbf{p}_{(k)}^T}{\mathbf{y}_{(k)}^T \mathbf{p}_{(k)}} \right) + \frac{\mathbf{p}_{(k)} \mathbf{p}_{(k)}^T}{\mathbf{y}_{(k)}^T \mathbf{p}_{(k)}}. \quad (3.26)$$

La formule de mise à jour (3.25) préserve la propriété définie positive pour la matrice $\mathbf{H}_{(k+1)}$ si la matrice $\mathbf{H}_{(k)}$ l'est aussi et $\mathbf{y}_{(k)}^T \mathbf{p}_{(k)} > 0$. Plusieurs travaux ont montré la convergence de l'algorithme BFGS pour des problèmes

de minimisation convexe (voir, par exemple, [131],[49],[158]). Pourtant, on ne sait pas si cet algorithme converge globalement lorsqu'il est appliqué à des problèmes de minimisation non convexe, même si la recherche linéaire exacte est utilisée.

Recherche linéaire du pas d'adaptation

Il est clair qu'un choix convenable du pas d'adaptation $\alpha_{(k)}$ pour chaque itération k dans la direction de quasi-Newton permet de diminuer le nombre d'itérations globales pour atteindre la convergence. Une recherche linéaire de ce pas consiste à minimiser par rapport à α_k le critère suivant :

$$\min_{\alpha_{(k)} \in \mathbb{R}} \mathcal{F}_{\text{JBD}}(\Lambda_{(k)}), \quad (3.27)$$

où $\Lambda_{(k)} = \mathbf{B}_{(k)} - \alpha_{(k)} \mathbf{H}_{(k)}^{-1} \nabla_{\alpha} \mathcal{F}_{\text{JBD}}(\mathbf{B}_{(k)})$. Nous montrons que la fonction de coût $\mathcal{F}_{\text{JBD}}(\Lambda_{(k)})$ est un polynôme d'ordre 4 de variable $\alpha_{(k)}$ ayant pour expression (voir Annexe (A.3)) :

$$\mathcal{F}_{\text{JBD}}(\Lambda_{(k)}) = a\alpha_{(k)}^4 + b\alpha_{(k)}^3 + c\alpha_{(k)}^2 + d\alpha_{(k)} + e. \quad (3.28)$$

dont les coefficients sont donnés par :

$$\begin{aligned} a &= \sum_{i=1}^K \text{tr} [\mathbf{V}_3 \text{OffBdiag}\{\mathbf{V}_3^H\}], \\ b &= - \sum_{i=1}^K \text{tr} [\mathbf{V}_0 \text{OffBdiag}\{\mathbf{V}_3^H\}] - \text{tr} [\mathbf{V}_3 \mathbf{V}_2], \\ c &= \sum_{i=1}^K \text{tr} [\mathbf{V}_1^H \text{OffBdiag}\{\mathbf{V}_3^H\} + \mathbf{V}_3 \text{OffBdiag}\{\mathbf{V}_1\}] \\ &\quad + \text{tr} [\mathbf{V}_0 \mathbf{V}_2], \\ d &= - \sum_{i=1}^K \text{tr} [\mathbf{V}_1^H \mathbf{V}_2] - \text{tr} [\mathbf{V}_0 \text{OffBdiag}\{\mathbf{V}_1\}], \\ e &= \sum_{i=1}^K \text{tr} [\mathbf{V}_1^H \text{OffBdiag}\{\mathbf{V}_1\}], \end{aligned} \quad (3.29)$$

où,

$$\begin{aligned} \mathbf{V}_0 &= \mathbf{C}_{\mathbf{x}}^{\xi_i}(\tau_i)^H \mathbf{\Gamma}_{(\mathbf{k})}^H \mathbf{B} + \mathbf{C}_{\mathbf{x}}^{\xi_i}(\tau_i)^H \mathbf{B}^H \mathbf{\Gamma}_{(\mathbf{k})}, \\ \mathbf{V}_1 &= \mathbf{B} \mathbf{C}_{\mathbf{x}}^{\xi_i}(\tau_i) \mathbf{B}^H, \\ \mathbf{V}_2 &= \text{OffBdiag}\{\mathbf{\Gamma}_{(\mathbf{k})} \mathbf{C}_{\mathbf{x}}^{\xi_i}(\tau_i) \mathbf{B}^H\} + \text{OffBdiag}\{\mathbf{B} \mathbf{C}_{\mathbf{x}}^{\xi_i}(\tau_i) \mathbf{\Gamma}_{(\mathbf{k})}^H\}, \\ \mathbf{V}_3 &= \mathbf{\Gamma}_{(\mathbf{k})} \mathbf{C}_{\mathbf{x}}^{\xi_i}(\tau_i)^H \mathbf{\Gamma}_{(\mathbf{k})}^H, \end{aligned}$$

avec $\text{tr}[\cdot]$ représente l'opérateur trace matricielle et $\Gamma_{(\mathbf{k})} = \mathbf{H}_{(k)}^{-1} \nabla_a \mathcal{F}_{\text{JBD}}(\mathbf{B}_{(\mathbf{k})})$. Le pas d'adaptation optimal est déterminé par la recherche des racines d'un polynôme d'ordre trois obtenu comme la dérivée du polynôme d'ordre quatre précédent,

$$\frac{\partial \mathcal{F}_{\text{JBD}}(\Lambda_{(k)})}{\partial \alpha_{(k)}} = 0 \Leftrightarrow 4a\alpha_{(k)}^3 + 3b\alpha_{(k)}^2 + 2c\alpha_{(k)} + d = 0.$$

pour lequel il existe trois racines. Le minimum peut être calculé en remplaçant chaque racine par le polynôme donné dans l'équation (3.28) et choisir la racine qui donne la plus petite valeur.

Complexité de l'algorithme

Pour la matrice de gradient dont l'expression est donnée par (3.24), le coût de calcul s'élève approximativement à :

Cas	Coût par itération
$M \neq N$	$4KNM(M + N) + 2KN^2$
$M \gg N$	$4KNM^2$
$M = N$	$8KN^3 + 2KN^2$

Le coût de calcul du pas d'adaptation est donné par le calcul des cinq coefficients du polynôme d'ordre. Il correspond approximativement à :

Cas	Coût par itération
$M \neq N$	$24KMN(M + N) + 9KN^2(1 + N)$
$M \gg N$	$24KM^2N$
$M = N$	$57KN^3 + 9KN^2$

Par conséquent, le coût de calcul de l'algorithme BFGS est :

Cas	Coût par itération
$M \neq N$	$52KNM(M + N) + 4N^3M^3 + 20N^2K + 18N^3K + 10N^2M^2 + 4NM + 2NM(M + N)$
$M \gg N$	$52KNM^2 + 4N^3M^3 + 20N^2K + 18N^3K + 10N^2M^2 + 4NM + 2NM^2 + N$
$M = N$	$122KN^3 + 4N^6 + 20N^2K + 10N^4 + 4N^2 + 4N^3$

Enfin, il faut noter que la complexité globale de l'algorithme BFGS doit être multipliée par le nombre d'itérations globales n_{iter} requis pour atteindre la convergence. En plus, il faut tenir compte également du temps de calcul nécessaire pour composer l'ensemble des K matrices à bloc-diagonaliser.

Résumé de la méthode proposée

Le principe de la méthode proposée est résumé dans l'algorithme (1).

Algorithme 1 Méthode proposée

Composition des matrices :

-Construire un ensemble de matrices \mathcal{M}_{JBD} composé des matrices $\mathbf{C}_{\mathbf{x}}^{\xi_i}(\tau_i)$ aux fréquences cycliques des sources et à plusieurs retards à l'aide d'un détecteur automatique de points retards-fréquences cycliques ;

Initialisation :

- Initialiser $\mathbf{B}_{(0)}$ et $\mathbf{H}_{(0)}$;
- Donner un nombre d'itérations maximal n_{iter} ;
- Fixer un seuil positif assez petit ϵ ;

pour $k \leftarrow 0, n_{iter} - 1$ **faire**

Calculer $\nabla_{\mathbf{a}} \mathcal{F}_{\text{JBD}}(\mathbf{B}_{(k)})$ dont l'expression est donnée par (3.24) ;

Rechercher linéairement le pas d'adaptation optimal $\alpha_{(k)}$;

$\mathbf{B}_{(k+1)} \leftarrow \mathbf{B}_{(k)} - \alpha_{(k)} \mathbf{H}_{(k)}^{-1} \nabla_{\mathbf{a}} \mathcal{F}_{\text{JBD}}(\mathbf{B}_{(k)})$;

Calculer $\mathbf{H}_{(k)}$ dont l'expression est donnée par (3.25) ;

si $\|\mathbf{B}_{(k+1)} - \mathbf{B}_{(k)}\| \leq \epsilon$ **alors**

break ;

fin si

fin pour

3.1.5 Simulations et résultats numériques

Afin de montrer la qualité d'estimation de la matrice de mélange, nous utilisons pour les deux simulations l'indice de performance $I_{\text{perf}}(\cdot)$ (1.25) en décibel et estimé en moyennant 100 tirages de Monte-Carlo. Le rapport signal sur bruit est donné par $\text{RSB} = -10 \log_{10}(\sigma_b^2)$.

Étude de simulation 1

Pour illustrer l'efficacité de la méthode proposée dans le cas des mélanges instantanés de sources cyclo-stationnaires de fréquences cycliques communes et inconnues, nous proposons de considérer un mélange de deux sources modulées en amplitude ayant chacune 2000 échantillons définies par :

$$\mathbf{s}_i(t) = \sum_{k \in \mathbb{Z}} a_i(k) w(t - kT_i) \cos(2\pi f_i t + \phi_i),$$

où $a_{i=1,2}(k)$ sont deux trains binaires i.i.d et de moyenne nulle, $T_{i=1,2}$ représentent les périodes symboles et ont respectivement pour valeurs $T_1 = 10$, $T_2 = 8$, $f_{i=1,2} = 0.2$ sont les fréquences porteuses normalisées, $\phi_{i=1,2}$ sont les phases porteuses et ont respectivement pour valeurs $\phi_1 = \frac{\pi}{6}$, $\phi_2 = \frac{\pi}{8}$ et $w(n)$ est un filtre de mise en forme s'écrivant comme suit :

$$w(n) = \begin{cases} \frac{2}{T}n & \text{si } 0 \leq n \leq \frac{T}{2} \\ -\frac{2}{T}n + 2 & \text{si } \frac{T}{2} + 1 \leq n \leq T - 1 \\ 0 & \text{sinon.} \end{cases}$$

Les deux sources partageant les fréquences cycliques suivantes $0, \pm 2f_1 = \pm 0.4$ et $\pm 5f_1 = \frac{10}{T_1} = \pm \frac{8}{T_2} = \pm 1$ et leurs harmoniques. La Fig.(3.1) montre le contour de la matrice de corrélation cyclique des sources $C_s^\xi(\tau)$. Elle révèle peu de corrélation entre les termes hors diagonales de la matrices $C_s^\xi(\tau)$ étant donné que leurs amplitudes sont faibles. Cette corrélation peut s'expliquer par le fait que nous traitons des signaux de taille finie. Nous étudions deux

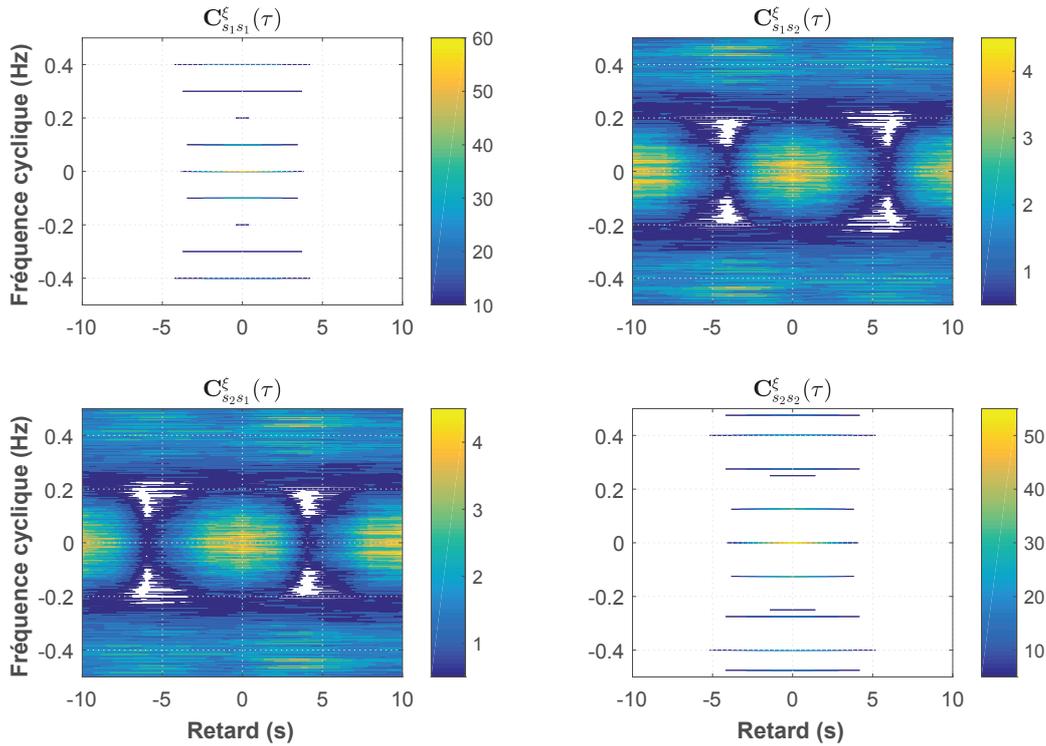


FIGURE 3.1 – Contour de la matrice de corrélation cyclique des sources $C_s^\xi(\tau)$.

situations : le cas déterminé et le cas sur-déterminé. Pour cela, deux matrices de mélanges A_1 et A_2 sont générées aléatoirement suivant la loi uniforme,

respectivement pour les cas déterminé et sur-déterminé de la façon suivante :

$$\mathbf{A}_1 = \begin{pmatrix} 0.5599 & 0.4070 \\ 0.2899 & 0.8431 \\ 0.9415 & 0.3397 \end{pmatrix}, \quad \mathbf{A}_2 = \begin{pmatrix} 0.8400 & 0.9971 \\ 0.7035 & 0.1965 \end{pmatrix}. \quad (3.30)$$

Nous supposons que les fréquences cycliques sont inconnues et nous utilisons la procédure de détection assurée par les critères (3.17) et (3.18) avec $\epsilon = 10^{-3}$ et $\eta = 10^{-1}$ pour toutes les simulations afin de composer l'ensemble des matrices de corrélation cyclique des observations à diagonaliser conjointement par l'algorithme BFGS. Nous comparons notre méthode proposée BFGS à deux méthodes classiques à savoir la méthode JADE et la méthode GRAD basée sur une diagonalisation non-unitaire utilisant le gradient descendant et la recherche linéaire du pas d'adaptation introduite dans [106].

Étude de simulation 2

Dans le cas des mélanges convolutifs de sources cyclo-stationnaires de fréquences cycliques distinctes et inconnues, nous proposons d'étudier le système suivant : 4 mélanges de 2 sources de télécommunications comportant 15200 échantillons définies par :

$$s_i(t) = \text{Re} \left[\sum_{k \in \mathbb{Z}} a_i(k) w(t - kT_i) e^{j2\pi f_i t} \right]$$

où $a_{i=1,2}(k)$ sont deux trains binaires i.i.d et de moyenne nulle, $T_{i=1,2}$ représentent les périodes symboles et ont respectivement pour valeurs $T_1 = 10$, $T_2 = 8$, $f_{i=1,2} = (0.2, 0.25)$ sont les fréquences porteuses normalisées et $w(n)$ est un filtre de mise en forme triangulaire. La fonction de transfert du canal avec $\rho = \rho' = 3$, $\mathbf{H}[z] = (\mathbf{H}_1[z] \mathbf{H}_2[z])$ est donnée par,

$$\mathbf{H}_1[z] = \begin{bmatrix} -0.0856 - 0.9858z^{-1} + 0.1446z^{-2} \\ 0.5082 - 0.1630z^{-1} - 0.8457z^{-2} \\ 0.8261 - 0.2216z^{-1} + 0.5182z^{-2} \\ -0.5399 - 0.5938z^{-1} + 0.5966z^{-2} \end{bmatrix},$$

$$\mathbf{H}_2[z] = \begin{bmatrix} -0.6960 - 0.6032z^{-1} - 0.3896z^{-2} \\ -0.4964 + 0.3306z^{-1} - 0.8027z^{-2} \\ -0.1404 - 0.9725z^{-1} - 0.1858z^{-2} \\ 0.2198 - 0.5750z^{-1} + 0.7881z^{-2} \end{bmatrix}.$$

Les deux sources ont des fréquences cycliques distinctes comme le montre la Fig.(3.2). La méthode proposée est comparée à la méthode GRAD. Nous

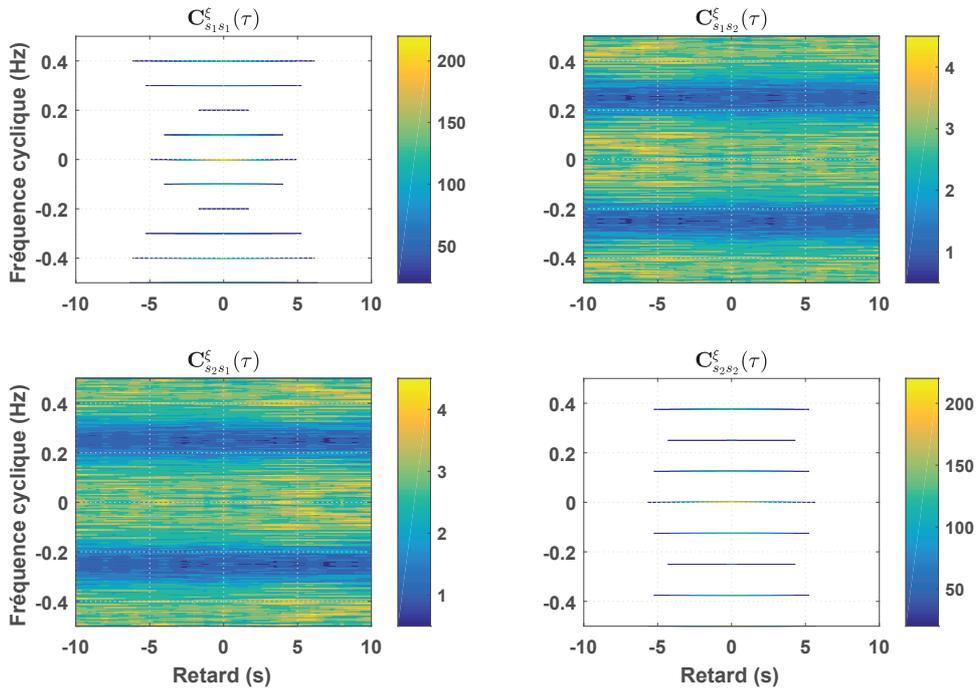


FIGURE 3.2 – Contour de la matrice de corrélation cyclique des sources $C_s^\xi(\tau)$.

supposons que les fréquences cycliques des sources sont inconnues et nous employons la procédure de détection automatique décrite par (3.20) afin de construire l'ensemble de matrices $C_x^\xi(\tau)$ à bloc-diagonaliser conjointement.

Résultats des simulations

Dans le cas d'un mélange instantané, les Fig.(3.4) et (3.3) montrent que le nouvel algorithme BFGS dépasse en termes de performances les méthodes JADE et GRAD dans les cas des systèmes de mélange instantané déterminé et sur-déterminé ($m > n$ or $m = n$). Les performances s'améliorent naturellement lorsque :

- le rapport signal sur bruit augmente,
- le nombre de matrice à bloc diagonaliser conjointement augmente,
- la taille des signaux augmente (3.6).

Dans les deux derniers cas, la complexité de calcul devient plus coûteuse.

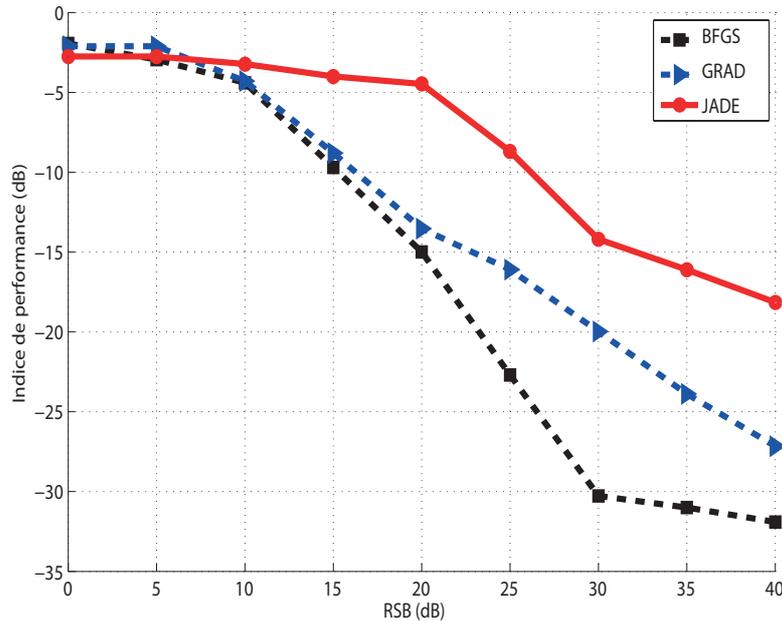


FIGURE 3.3 – Évolution de l'indice de performances en fonction du RSB avec la matrice A_2 .

3.2 Identification de la matrice de mélange dans le cas instantané

3.2.1 Modèle des sources et propriété importante

Nous considérons dans cette partie le modèle de mélange instantané sur-déterminé précédemment donné par l'équation (1.4) où les sources sont de moyenne nulle, cyclo-stationnaires à l'ordre 2 de fréquences cycliques distinctes et inconnues et les bruits sont blancs, gaussiens et décorrélés des sources. Nous nous intéressons particulièrement aux signaux de télécommunication issus des modulations analogiques qui se modélisent par le produit d'un signal harmonique à support borné et d'une porteuse sinusoïdale comme suit :

$$s_i(t) = e_i(t) \cos(2\pi f_i t), \quad \forall i \in \{1, \dots, n\} \quad (3.31)$$

où f_i est la fréquence porteuse, $e_i(t)$ sont des signaux stationnaires de moyenne nulle et mutuellement corrélés. Les signaux $s_i(t)$ ont un caractère cyclo-stationnaire à l'ordre 2 provenant des opérations de codage et modulation où la fréquence cyclique est liée à la fréquence porteuse. Nous supposons que $e_i(t)$ n'ont pas de composante fréquentielle au-delà de la fréquence de la porteuse, f_i , dans ce cas là, la transformée de Hilbert (voir Annexe

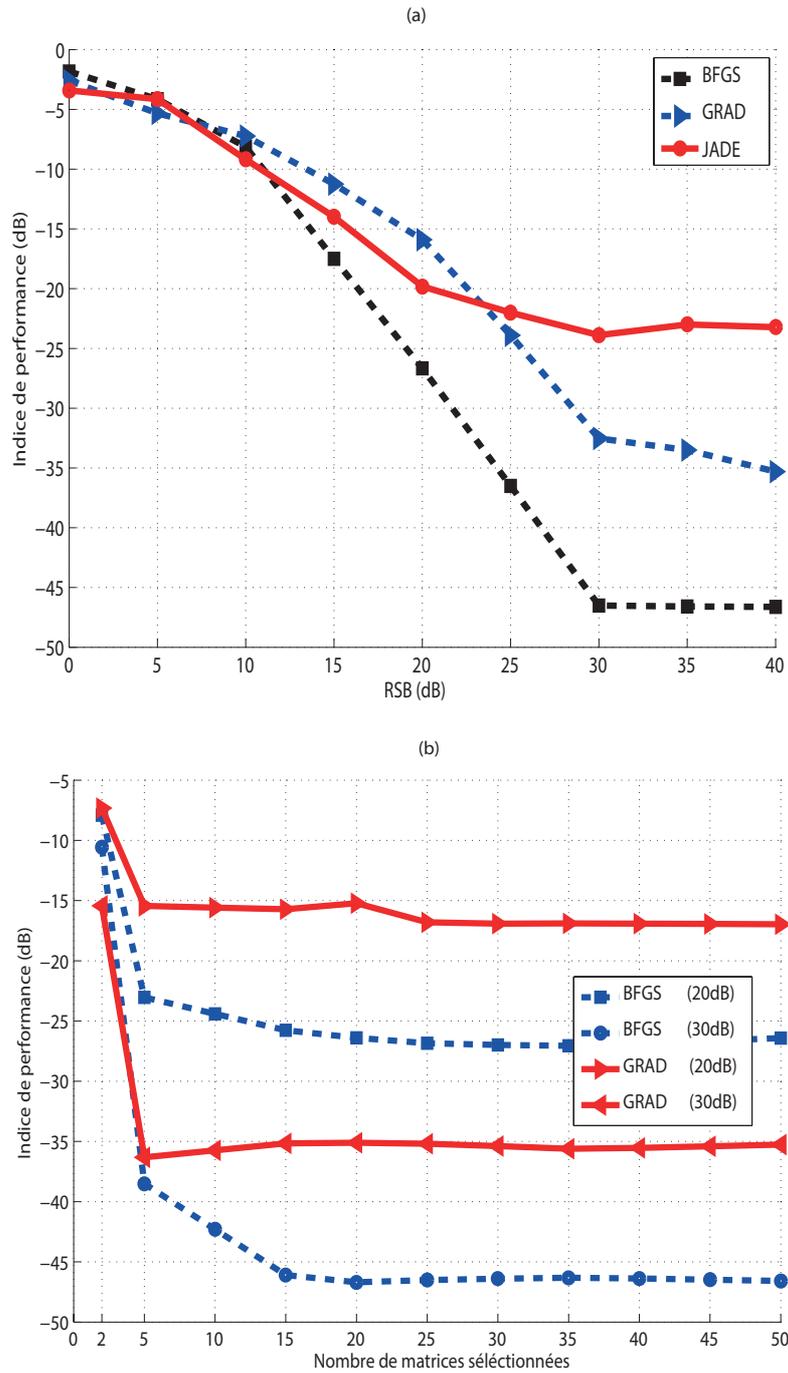


FIGURE 3.4 – Évolution de l'indice de performances : (a) en fonction du RSB, (b) en fonction du nombre des matrices sélectionnées avec la matrice \mathbf{A}_1 .

(A.4)) de $s_i(t)$ notée $s_i^{(h)}(t)$ est donnée par :

$$s_i^{(h)}(t) = (h * s)(t) \quad (3.32)$$

$$= e_i(t) \sin(2\pi f_i t) \quad (3.33)$$

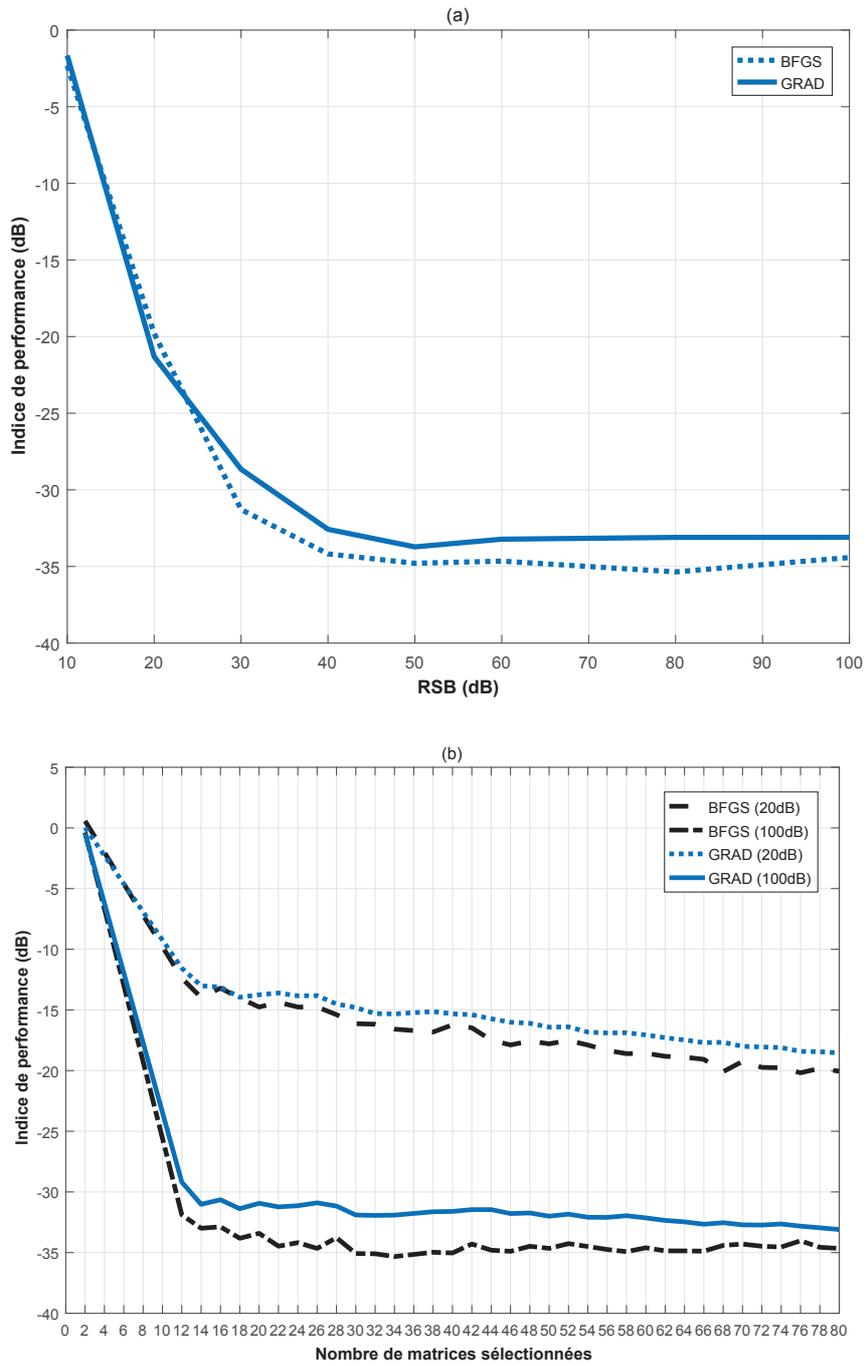


FIGURE 3.5 – Évolution de l'indice de performances : (a) en fonction du RSB, (b) fonction du nombre des matrices sélectionnées.

où

$$h(t) = \frac{1}{\pi t} \tag{3.34}$$

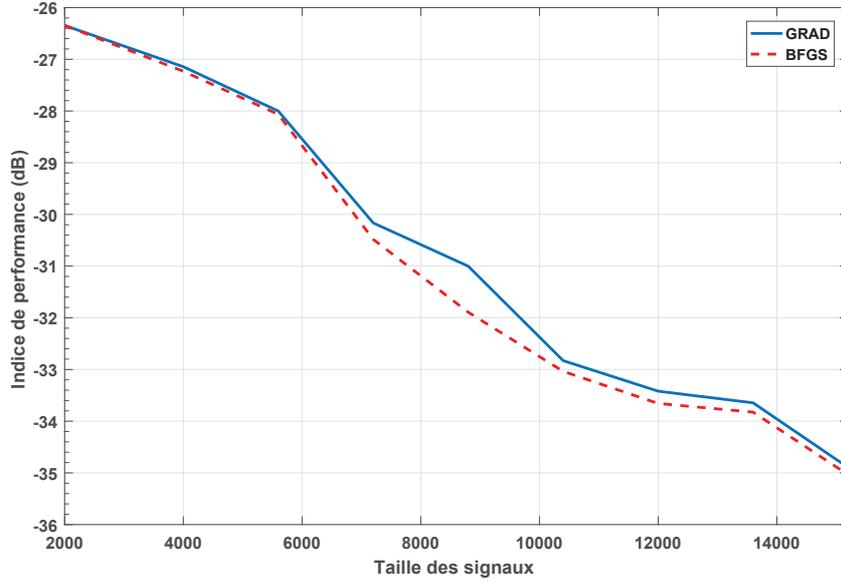


FIGURE 3.6 – Évolution de l'indice de performances en fonction de la taille des signaux.

Une représentation analytique de $s_i(t)$ notée $s_i^{(a)}(t)$ est définie ainsi à l'aide de la transformée de Hilbert :

$$s_i^{(a)}(t) = s(t) + js_i^{(h)}(t) \quad (3.35)$$

$$= e_i(t) \cos(2\pi f_i t) + je_i(t) \sin(2\pi f_i t) \quad (3.36)$$

$$= e_i(t) e^{j2\pi f_i t} \quad (3.37)$$

A partir de (1.4) et des hypothèses \mathcal{H}_1 et \mathcal{H}_9 , nous démontrons que la matrice de corrélation cyclique des signaux analytiques des observations $\mathbf{x}^{(a)}(t) = \mathbf{x}(t) + j\mathbf{x}^{(h)}(t)$ ($\mathbf{x}^{(h)}(t)$ est la transformée de Hilbert de $\mathbf{x}(t)$) pour $\xi \neq 0$ s'écrit :

$$\mathbf{C}_{\mathbf{x}^{(a)}}^\xi(\tau) = \mathbf{A} \mathbf{C}_{\mathbf{s}^{(a)}}^\xi(\tau) \mathbf{A}^H, \quad (3.38)$$

avec

$$\mathbf{C}_{\mathbf{s}^{(a)}}^\xi(\tau) = \begin{bmatrix} \mathbf{C}_{s_1^{(a)} s_1^{(a)}}^\xi(\tau) & \mathbf{C}_{s_1^{(a)} s_2^{(a)}}^\xi(\tau) & \cdots & \mathbf{C}_{s_1^{(a)} s_n^{(a)}}^\xi(\tau) \\ \mathbf{C}_{s_2^{(a)} s_1^{(a)}}^\xi(\tau) & \mathbf{C}_{s_2^{(a)} s_2^{(a)}}^\xi(\tau) & \cdots & \mathbf{C}_{s_2^{(a)} s_n^{(a)}}^\xi(\tau) \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{C}_{s_n^{(a)} s_1^{(a)}}^\xi(\tau) & \mathbf{C}_{s_n^{(a)} s_2^{(a)}}^\xi(\tau) & \cdots & \mathbf{C}_{s_n^{(a)} s_n^{(a)}}^\xi(\tau) \end{bmatrix}, \quad (3.39)$$

où $s_\ell^{(a)}(t)$ est le signal analytique associé à la ℓ -ième source $s_\ell(t)$ donnée par (3.31).

Proposition : La matrice $\mathbf{C}_{\mathbf{s}^{(a)}}^\xi(\tau)$ est zéro diagonale avec un seul terme non nul hors diagonale et ceci pour $\xi \neq 0$ et $\xi = \xi_\ell - \xi_k, \forall \ell, k = 1, 2, \dots, n$ de telle sorte que

$$\mathbf{C}_{s_\ell^{(a)} s_k^{(a)}}^\xi(\tau) = \begin{cases} 0 & \text{si } \ell = k \\ \mathbf{C}_{e_\ell e_k}(\tau) e^{j\pi(f_\ell + f_k)\tau} & \text{sinon} \end{cases} \quad (3.40)$$

Preuve : pour $\xi \neq 0$ et $\xi = \xi_\ell - \xi_k, \forall \tau \in \mathcal{Z}, \forall \ell, k = 1, 2, \dots, n$, nous avons :

$$\begin{aligned} \mathbf{C}_{s_\ell^{(a)} s_k^{(a)}}^\xi(\tau) &= \lim_{T \rightarrow \infty} \frac{1}{T} \int_{-\frac{T}{2}}^{\frac{T}{2}} \mathbf{C}_{s_\ell^{(a)} s_k^{(a)}}(t, \tau) e^{-j2\pi\xi t} dt, \\ &= \lim_{T \rightarrow \infty} \frac{1}{T} \int_{-\frac{T}{2}}^{\frac{T}{2}} \mathbf{E} \left[s_\ell^{(a)}\left(t + \frac{\tau}{2}\right) s_k^{(a)H}\left(t - \frac{\tau}{2}\right) \right] e^{-j2\pi\xi t} dt, \\ &= \lim_{T \rightarrow \infty} \frac{1}{T} \int_{-\frac{T}{2}}^{\frac{T}{2}} \mathbf{E} \left[e_\ell\left(t + \frac{\tau}{2}\right) e^{j2\pi\xi_\ell\left(t + \frac{\tau}{2}\right)} e_k^H\left(t - \frac{\tau}{2}\right) e^{-j2\pi\xi_k\left(t - \frac{\tau}{2}\right)} \right] e^{-j2\pi\xi t} dt, \\ &= \mathbf{C}_{e_\ell e_k}(t, \tau) e^{j\pi(\xi_\ell + \xi_k)\tau} \lim_{T \rightarrow \infty} \frac{1}{T} \int_{-\frac{T}{2}}^{\frac{T}{2}} e^{j2\pi(\xi_\ell - \xi_k)t} e^{-j2\pi\xi t} dt, \\ &= \mathbf{C}_{e_\ell e_k}(\tau) e^{j\pi(\xi_\ell + \xi_k)\tau} \delta(\xi - (\xi_\ell - \xi_k)), \quad (\text{car } e(t) \text{ est stationnaire}) \end{aligned}$$

en remplaçant ξ par $\xi_\ell - \xi_k$, nous obtenons :

$$\mathbf{C}_{s_\ell^{(a)} s_k^{(a)}}^\xi(\tau) = \mathbf{C}_{e_\ell e_k}(\tau) e^{j\pi(\xi_\ell + \xi_k)\tau} \delta(0) \neq 0,$$

de la même façon, nous montrons que :

$$\mathbf{C}_{s_\ell^{(a)} s_\ell^{(a)}}^\xi(\tau) = \mathbf{C}_{e_\ell e_\ell}(\tau) e^{2j\pi\xi_\ell\tau} \delta(\xi) = 0,$$

car $\xi \neq 0$, où $\mathbf{C}_{e_\ell e_\ell}(\tau) \neq 0$ est l'inter-corrélation des signaux corrélés $e_\ell(t)$ et $e_k(t)$.

$$\mathbf{C}_{\mathbf{x}^{(a)}}^{(\xi_\ell - \xi_k)}(\tau) = \mathbf{C}_{s_\ell^{(a)} s_k^{(a)}}^{(\xi_\ell - \xi_k)}(\tau) \mathbf{A}_\ell \mathbf{A}_k^H, \quad (3.41)$$

où $\mathbf{A}_\ell, \mathbf{A}_k$ sont la ℓ -ième and k -ième colonnes de la matrice de mélange \mathbf{A} . Ce qui implique que la matrice $\mathbf{C}_{\mathbf{x}^{(a)}}^{(\xi_\ell - \xi_k)}(\tau)$ est de rang un.

3.2.2 Approche proposée combinant la DVP et une méthode de regroupement hiérarchique

En notant que $\mathbf{P}_k = \mathbf{C}_{\mathbf{x}^{(a)}}^{(\xi_\ell - \xi_k)}(\tau)^H \mathbf{C}_{\mathbf{x}^{(a)}}^{(\xi_\ell - \xi_k)}(\tau)$ est une matrice de rang un car $\mathbf{C}_{s_\ell^{(a)} s_k^{(a)}}^{(\xi_\ell - \xi_k)}(\tau)$ l'est aussi, nous établissons notre méthode. La matrice \mathbf{P}_k est

exprimée par :

$$\begin{aligned}
\mathbf{P}_k &= \mathbf{C}_{\mathbf{x}^{(a)}}^{(\xi_\ell - \xi_k)}(\tau)^H \mathbf{C}_{\mathbf{x}^{(a)}}^{(\xi_\ell - \xi_k)}(\tau), \\
&= \mathbf{A}_k \mathbf{A}_\ell^H \mathbf{C}_{s_\ell^{(a)} s_k^{(a)}}^{(\xi_\ell - \xi_k)}(\tau)^H \mathbf{C}_{s_\ell^{(a)} s_k^{(a)}}^{(\xi_\ell - \xi_k)}(\tau) \mathbf{A}_\ell \mathbf{A}_k^H, \\
&= \beta_{ij}^2 \gamma \mathbf{A}_k \mathbf{A}_k^H,
\end{aligned} \tag{3.42}$$

où $\beta_{ij}^2 = \mathbf{C}_{s_\ell^{(a)} s_k^{(a)}}^{(\xi_\ell - \xi_k)}(\tau)^2$ et $\gamma = \mathbf{A}_\ell^H \mathbf{A}_\ell$ sont deux scalaires. Il s'ensuit que \mathbf{P}_k a une seule valeur propre non nulle. Soit λ_1 la valeur propre de \mathbf{P}_k et \mathbf{v}_1 son vecteur propre correspondant. La trace de \mathbf{P}_k est donnée par :

$$\text{tr}\{\mathbf{P}_k\} = \text{tr}\{\beta_{ij}^2 \gamma \mathbf{A}_k \mathbf{A}_k^H\} = \beta_{ij}^2 \gamma \|\mathbf{A}_k\|_F^2 = \lambda_1.$$

De même \mathbf{P}_k est aussi donnée également par $\lambda_1 \mathbf{v}_1 \mathbf{v}_1^H$. Elle représente sa décomposition en éléments propres, par conséquent :

$$\lambda_1 \mathbf{v}_1 \mathbf{v}_1^H = \beta_{ij}^2 \gamma \mathbf{A}_k \mathbf{A}_k^H \Rightarrow \beta_{ij}^2 \gamma \|\mathbf{A}_k\|_F^2 \mathbf{v}_1 \mathbf{v}_1^H = \beta_{ij}^2 \gamma \mathbf{A}_k \mathbf{A}_k^H.$$

Avec la contrainte $\mathbf{v}_1 \mathbf{v}_1^H = 1$, nous avons finalement :

$$\mathbf{v}_1 \mathbf{v}_1^H = \frac{\mathbf{A}_k \mathbf{A}_k^H}{\|\mathbf{A}_k\|_F^2} \Rightarrow \mathbf{v}_1 = \pm \frac{\mathbf{A}_k}{\|\mathbf{A}_k\|_F}.$$

Le vecteur propre \mathbf{v}_1 lié à la plus grande valeur propre λ_1 de \mathbf{P}_k est obtenu à un signe \pm et une multiplication par $e^{j\phi_k}$ près où la phase ϕ est inconnue. De plus, le vecteur propre \mathbf{v}_1 est colinéaire à la k -ième colonne de la matrice \mathbf{A} . Nous obtenons alors la colonne \mathbf{A}_k , mais ceci reste insuffisant pour séparer les mélanges. La séparation requiert la connaissance de toutes les colonnes de \mathbf{A} .

A cet effet, l'identification de la matrice de mélange est réalisée selon les quatre étapes suivantes :

- Etape 1. Calcul des matrices $\mathbf{C}_{\mathbf{x}^{(a)}}^\xi(\tau)$ pour un nombre suffisamment grand de fréquences cycliques et de retards.
- Etape 2. Sélection des matrices $\mathbf{C}_{\mathbf{x}^{(a)}}^{(\xi_\ell - \xi_k)}(\tau)$ qui ont une structure zéro diagonale avec un seul terme non nul hors diagonale. Nous utilisons alors deux critères de détection de matrices de rang un, inspirés du travail proposé dans [109], à savoir :

$$\begin{cases} \mathbf{C}_1 = \frac{\theta_1^2}{\|\mathbf{C}_{\mathbf{x}^{(a)}}^\xi(\tau)\|_F^2} \geq 1 - \epsilon_1, \\ \mathbf{C}_2 = \theta_1 \geq \epsilon_2 \end{cases}, \tag{3.43}$$

où $\|\cdot\|_F$ est la norme de Frobenius, θ_1 est la plus grande valeur singulière de $\mathbf{C}_{\mathbf{x}^{(a)}}^\xi(\tau)$, ϵ_1 est un scalaire positif très proche de zéro et ϵ_2 est un

scalaire suffisamment grand pour éviter la sélection des matrices de faible valeur. Ainsi, une des matrices calculées dans l'étape précédente est zéro diagonale avec une seul terme non nul hors diagonale, si et seulement si elle vérifie les deux critères donnés par l'équation (3.43).

Etape 3. Formation des matrices \mathbf{P}_k à partir des matrices sélectionnées et calcul de leurs éléments propres pour identifier les colonnes de la matrice de mélange.

Etape 4. Après la procédure de détection, nous nous retrouvons avec un ensemble de matrices sélectionnées dont la taille dépend des valeurs ϵ_1 et ϵ_2 . Ensuite, nous calculons leurs éléments propres, ce qui nous donne plusieurs vecteurs. Selon un aspect théorique, les différents vecteurs qui identifient la k -ième colonne de \mathbf{A} devraient être égaux. Néanmoins, en pratique, et en raison d'un certain nombre d'erreurs numériques dans le calcul de $\mathbf{C}_{\mathbf{x}^{(a)}}^{\xi}(\tau)$ et l'incertitude associée aux valeurs ϵ_1 et ϵ_2 , certaines matrices n'auraient pas dû être sélectionnées. Ces vecteurs sont donc pas tous égaux et ils forment un nuage de points autour de la valeur théorique. Afin de palier à ces contraintes, nous utilisons l'algorithme de regroupement hiérarchique agglomératif *Neighbor Joining* introduit dans [168] pour estimer le centre de masse de chaque nuage de points comme suit : dans chaque étape, deux points sont fusionnés lorsque la distance euclidienne entre eux a la plus petite valeur par rapport à un seuil fixé à 0,1 parmi tous les autres points. Ces deux points donnent naissance à un nouveau point qui est la moyenne des deux points considérés. Les deux points sont alors retirés et remplacés par le nouveau point calculé. Quand un point n'a plus de voisin, il est retenu comme un centre de masse. L'opération est itérée tant que l'ensemble de points contient plus d'un élément. Les n points ayant le poids le plus significatif sont alors sélectionnés, leurs coordonnées représentent l'estimation de chaque colonne de la matrice de mélange.

Le principe de l'approche proposée est résumé dans l'algorithme (2).

3.2.3 Simulations et résultats numériques

Nous considérons trois sources issues des modulations analogiques d'amplitude d'une taille de $N = 2000$ échantillons chacune, dont les fréquences porteuses sont respectivement $f_1 = 1$ kHz, $f_2 = 1.8$ kHz et $f_3 = 0.6$ kHz échantillonnées à $f_s = 20$ kHz. Ces sources sont mélangées par la matrice

Algorithme 2 Approche proposée**Composition des matrices :**

- Calculer $\mathbf{C}_{\mathbf{x}^{(a)}}^{\xi_i}$ pour un nombre suffisamment grand de fréquences cycliques et retards et détecter les matrices $\mathbf{C}_{\mathbf{x}^{(a)}}^{\xi_i}$ de rang 1 ;
- Composer un ensemble de k matrices ;

commencer :

- Composer un ensemble de k matrices ;

pour $\tau = 0, i \leftarrow 1 : k$ **faire**

- Calculer $\mathbf{P} = \mathbf{C}_{\mathbf{x}^{(a)}}^{\xi_i}(\tau)^H \mathbf{C}_{\mathbf{x}^{(a)}}^{\xi_i}(\tau)$;
- Décomposer \mathbf{P} en valeurs propres ;
- Associer le vecteur propre lié à la plus grande valeur propre de \mathbf{P} à i -ème colonne de $\hat{\mathbf{A}}$;

fin pour

- Utiliser l'algorithme de regroupement hiérarchique agglomératif *Neighbor Joining* pour réduire la dimension de $\hat{\mathbf{A}}$ à $(m \times n)$.

fin

suivante générée d'une façon aléatoire par une loi uniforme :

$$\mathbf{A} = \begin{pmatrix} 0.8367 & 0.9452 & 0.2992 \\ 0.4488 & 0.7331 & 0.9508 \\ 0.2664 & 0.4761 & 0.1583 \end{pmatrix}. \quad (3.44)$$

La méthode proposée est comparée aux méthodes SOBI et JADE. Tout d'abord, nous vérifions la structure de la matrice de corrélation cyclique des signaux analytiques des sources. Pour cela, nous traçons sur la Fig.(3.7) le contour de la matrice $\mathbf{C}_{\mathbf{s}^{(a)}}^{\xi}(\tau)$ pour différentes fréquences cycliques et retards. Nous remarquons que la structure zéro-diagonale avec un seul terme hors la diagonale principale est bien vérifiée pour $-1200, -800, -400, 400, 800, 1200$ Hz. Ces fréquences représentent les différences des fréquences cycliques des sources. Comme nous l'avons évoqué précédemment, les fréquences cycliques sources sont inconnues aux niveau des antennes de réception. Par conséquent, nous utilisons la procédure de détection avec les seuils $\epsilon_1 = 10^{-2}$, $\epsilon_2 = 0.1$.

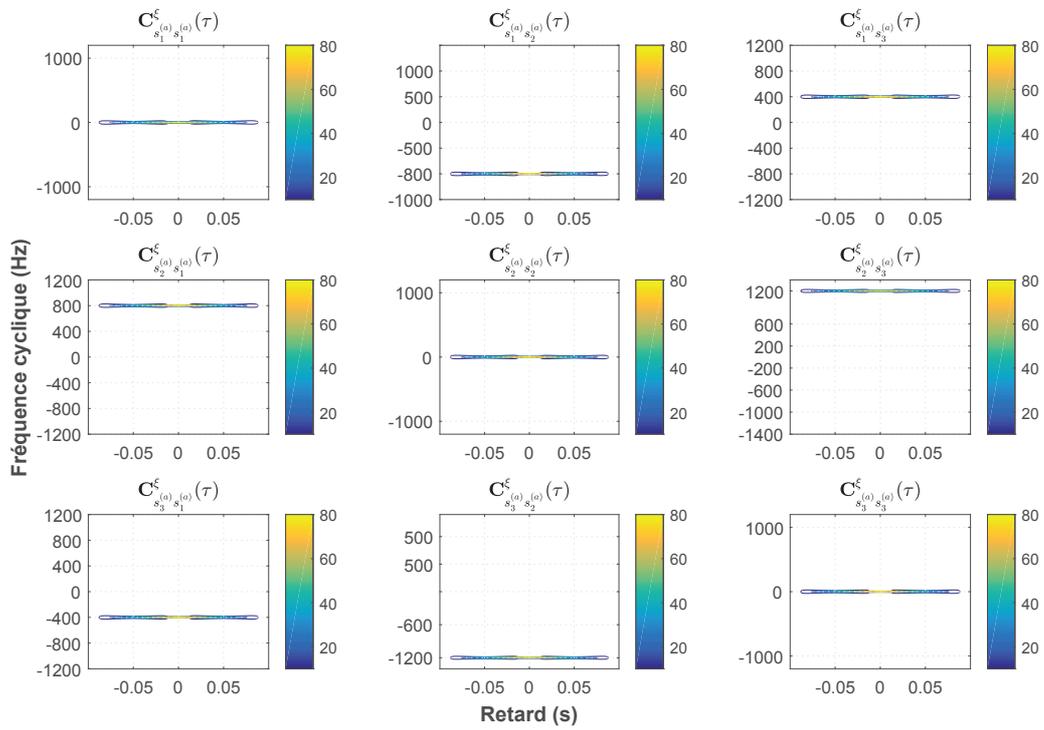


FIGURE 3.7 – Contour de la matrice de corrélation cyclique des sources analytiques $C_{s^{(a)}}^\xi(\tau)$.

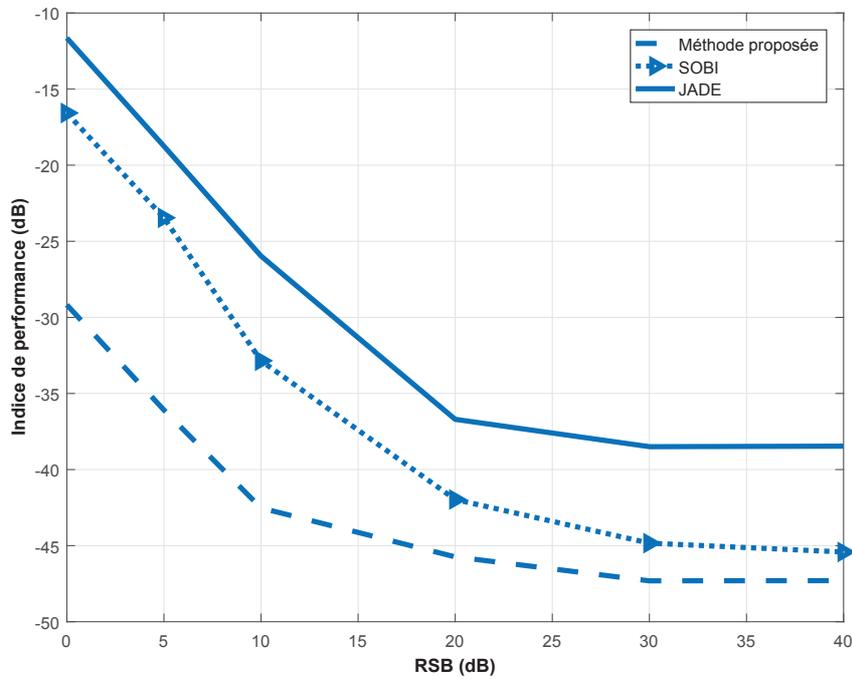


FIGURE 3.8 – Evolution de l'indice de performances en fonction du RSB.

La Fig.(3.8) montre que la méthode proposée dépasse les méthodes JADE et SOBI en terme d'indice de performance. Notons encore une fois que les performance s'améliorent considérablement lorsque le RSB augmente.

3.3 Conclusion

Dans ce chapitre, deux nouvelles solutions ont été proposées pour le problème de séparation de mélanges linéaires sur-déterminés de sources de nature cyclo-stationnaire dans le cas où les sources ont des fréquences distinctes et communes. La première approche est basée sur un nouvel algorithme de bloc-diagonalisation conjointe non unitaire d'un ensemble de matrices construit à partir des statistiques cycliques d'ordre deux par détection automatique. Cette approche opère sur des données non blanchies et emploie un schéma d'optimisation qui se base sur la méthode de BFGS avec un pas d'adaptation optimal obtenu en résolvant une équation de troisième ordre. La deuxième approche cherche à identifier la matrice de mélange en combinant la décomposition en éléments propres des matrices de corrélation cyclique des observations et une méthode regroupement hiérarchique. L'approche proposée est appliquée à des signaux de communication analogique. Il faut rappeler que les deux approches souffrent des ambiguïtés classiques inhérentes au problème de la SAS.

Séparation de sources cyclo-stationnaires : application aux signaux mécaniques et biomécaniques

Sommaire

4.1	Séparation de signaux vibratoires mécaniques, application au cas de roulements	78
4.1.1	Généralités sur les roulements	78
4.1.2	Etude de la cyclo-stationnarité d'un roulement endommagé	81
4.1.3	Séparation de roulements endommagés réels	84
4.2	Séparation des contributions cyclo-stationnaires aux ordres 1 et 2, application aux signaux biomécaniques	92
4.2.1	Introduction	92
4.2.2	Formulation du problème	95
4.2.3	Préliminaires théoriques	95
4.2.4	Méthode proposée	97
4.2.5	Application sur des signaux FRS synthétiques	99
4.2.6	Application sur des signaux FRS réels	99
4.3	Conclusion	102

DANS le domaine des télécommunications, les signaux étudiés sont supposés centrés avec une composante cyclo-stationnaire à l'ordre 1 nulle. Ceci n'est pas le cas, pour les signaux mécaniques ou biomécaniques car cette composante est généralement présente et souvent loin d'être négligeable.

En effet, elle peut représenter une bonne partie de l'énergie du signal. Il est donc indispensable dans ce cas de prendre en compte cette hypothèse.

Ce chapitre sera donc consacré à la SAS appliquée à ce type de signaux provenant d'expérimentation en mécanique et biomécaniques. Les signaux mécaniques et biomécaniques présente en effet un caractère cyclostationnaire résultant d'un couplage entre les phénomènes cycliques étudiées et certains aléas pouvant avoir plusieurs origines. Dans le cas de machines tournantes, le cycle est inhérents à la machines, les aléas peuvent provenir de fluctuation d'amplitude de phase ou de fréquence résultant de variation de vitesse, de charges, de chocs mécaniques répétitifs. . . Dans le cas de signaux biomécaniques tel que l'analyse de la course, le cycle est lié à la foulée, l'aléa résulte de la variation de rythme, ou surtout de variabilité dans la maîtrise de l'attaque au sol. Les méthodes proposées en mécanique et biomécanique sont un peu différentes. Nous présenterons de manière séparé chaque problématique tout en rappelant les spécificités, et les objectifs attendus.

4.1 Séparation de signaux vibratoires mécaniques, application au cas de roulements

4.1.1 Généralités sur les roulements

Le roulement est un dispositif mécanique omniprésent dans l'industrie permettant la rotation entre deux pièces, tout en optimisant le frottement et la précision de la liaison cinématique. Le roulement est généralement composé d'une bague intérieure et d'une bague extérieure, d'éléments roulants (billes/rouleaux) et d'une cage. Les éléments roulants interposés entre les deux bagues du roulement sont maintenus à égale distance les uns des autres par la cage qui les guide et facilite leur rotation (voir la Fig.(4.1) ¹). En raison de sa structure compliquée, de ses conditions de travail hostiles, le roulement est sans doute l'organe le plus sensible dans un parc de machines. Par conséquent, il est très important de développer une bonne surveillance de son état et des méthodes efficaces de diagnostic en vue d'une productivité et d'un gain économique élevés.

1. Figure sous licence ouverte CC BY-SA 3.0 accessible depuis <https://fr.wikipedia.org/wiki/Utilisateur:YannickPatois>

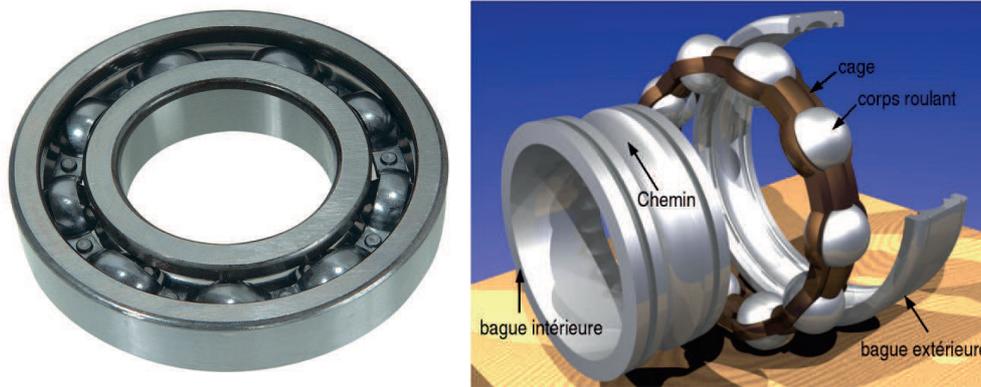


FIGURE 4.1 – Éclaté d'un roulement à billes.

Origines des défauts des roulements

La rupture par fatigue constitue la cause courante de dégradation des roulements dans les conditions normales de fonctionnement d'un roulement correctement monté. Cette rupture débute par des petites fissures qui se propagent au fil du temps et donnent lieu à un écaillage qui mène à la défaillance. D'où la nécessité dans certains cas de remplacer les roulements.

Il existe d'autres circonstances de défaillance des roulements en dehors de la rupture par fatigue. Nous recensons principalement quatre origines affectant fréquemment un roulement [161], [110], [146] :

1. La corrosion altère le roulement par réaction chimique avec un oxydant comme un lubrifiant contaminé par des matières étrangères qui sont souvent présentes dans les milieux industriels (sable, poussière). La nature abrasive de ces minuscules particules provoque des actions de piqûres et de ponçage qui cèdent à l'usure mesurable des billes et des bagues.
2. La lubrification excessive ou insuffisante donne lieu à une mauvaise condition de rotation de billes causant une augmentation de température. La chaleur excessive réduit la capacité de lubrification, et donc augmente le contact avec le métal, d'où un enlèvement de matière à la surface de la bague.
3. Les défauts de montage sont souvent causés par un forçement du roulement dans le pallier ou sur l'arbre. Ce forçement crée des dégradations physiques et des dés-alignements conduisant à une défaillance prématurée.
4. Les roulements sont souvent sujets de contraintes mécaniques radiales et axiales. Les contraintes axiales sont liées à la présence d'une charge appliquée à l'axe de l'arbre de rotation. Tandis que les contraintes ra-

diales s'exercent sur la bague interne du roulement qui les transmet à la bague externe par l'intermédiaire des billes.

Types des défauts des roulements

Les défauts de roulements se divisent en deux catégories :

1. Les défauts localisés sont des défauts qui apparaissent sur un seul élément du roulement. La zone défaillante provoque un choc mécanique à chaque fois qu'il y a un contact avec cette zone et les billes, ce qui se traduit par des vibrations à une période fixe et donc provoque la présence de signaux à des fréquences particulières appelées fréquences caractéristiques de bille ou de bague [147],[173] :

$$f_{\text{int}} = \frac{f_r N_b}{2} \left(1 + \frac{D_b \cos(\theta)}{D_c} \right) \quad (4.1)$$

$$f_{\text{ext}} = \frac{f_r N_b}{2} \left(1 - \frac{D_b \cos(\theta)}{D_c} \right) \quad (4.2)$$

$$f_{\text{bille}} = \frac{f_r D_c}{2 D_b} \left(1 - \left[\frac{D_b \cos(\theta)}{D_c} \right]^2 \right) \quad (4.3)$$

$$f_{\text{cage}} = \frac{f_r}{2} \left(1 - \frac{D_b \cos(\theta)}{D_c} \right) \quad (4.4)$$

$$f_r = \frac{rpm}{60} \quad (4.5)$$

où :

- rpm : nombre de tours par minute,
- N_b : nombre de billes,
- D_b : diamètre d'une bille,
- D_c : diamètre de la cage,
- θ : angle de contact indiquant la direction de la force que les billes exercent sur la bague extérieure,
- f_{int} : fréquence caractéristique de la bague interne,
- f_{ext} : fréquence caractéristique de la bague externe,
- f_{bille} : fréquence caractéristique des billes,
- f_r : fréquence de rotation,
- f_{cage} : fréquence caractéristique de la cage.

Il faut noter que la plupart des fabricants de roulements fournissent des tables de référence qui permettent de s'affranchir de ces calculs. Généralement ces fréquences de passage sont des fréquences théoriques sensiblement respectées tant qu'il y a un roulement parfait des billes sans glissement. En pratique, le passage d'un élément roulant sur le défaut, a tendance à générer des fluctuations locales de

vitesse avec des caractéristiques assez aléatoire. Ce phénomène se caractérise par la présence d'un bruit cyclo-stationnaire (souvent à l'ordre 2) à la fréquence du défaut.

2. Les défauts non localisés sont caractérisés par une dégradation globale qui mène à une large bande de changements dans le spectre de vibrations. Ce type de défaut affecte toute la surface du roulement et le palier devient rugueux, irrégulier, déformé et la surface du roulement entier est détériorée. Les défauts non localisés ne produisent pas de fréquences détectables dans les signaux de vibrations [146], mais produisent des changements imprévisibles dans les vibrations des machines et dans les courants statoriques s'il s'agit de roulements montés sur des moteurs.

4.1.2 Etude de la cyclo-stationnarité d'un roulement endommagé

Pour surveiller un roulement, nous pouvons faire appel soit à une analyse de la température et de la composition de l'huile, soit à une analyse de vitesse soit une analyse vibratoire. Dans ce dernier type d'analyse, plusieurs travaux ont montré la nature cyclo-stationnaire des vibrations issues de capteurs accélérométriques pour un roulement contenant un défaut localisé [33], [12], [30]. Afin d'illustrer cette cyclo-stationnarité, nous proposons d'étudier un signal² d'un roulement comportant un défaut de bague interne (Fig.4.2). Les caractéristiques du roulement sont données par le tableau (4.1) :

N_b	D_b	D_c	θ
8	0.235	1.245	0°

TABLE 4.1 – Caractéristiques du roulement.

La vitesse du moteur d'entraînement est de 1500trs/mn soit $f_r = 25Hz$ sans charge initiale. La fréquence de bague interne théorique f_{int} est donc égale $118.8755Hz$. Le signal est échantillonné à $48828Hz$. La Fig.(4.3) montre le signal accélérométrique de roulement dans le domaine temporel.

2. Ce signal est extrait d'un jeu de données de défauts fourni par MFPT Society (Machinery Failure Prevention Technology). Les signaux sont accessibles depuis <https://www.mfpt.org/FaultData/FaultData.htm>



FIGURE 4.2 – Défaut de bague interne.

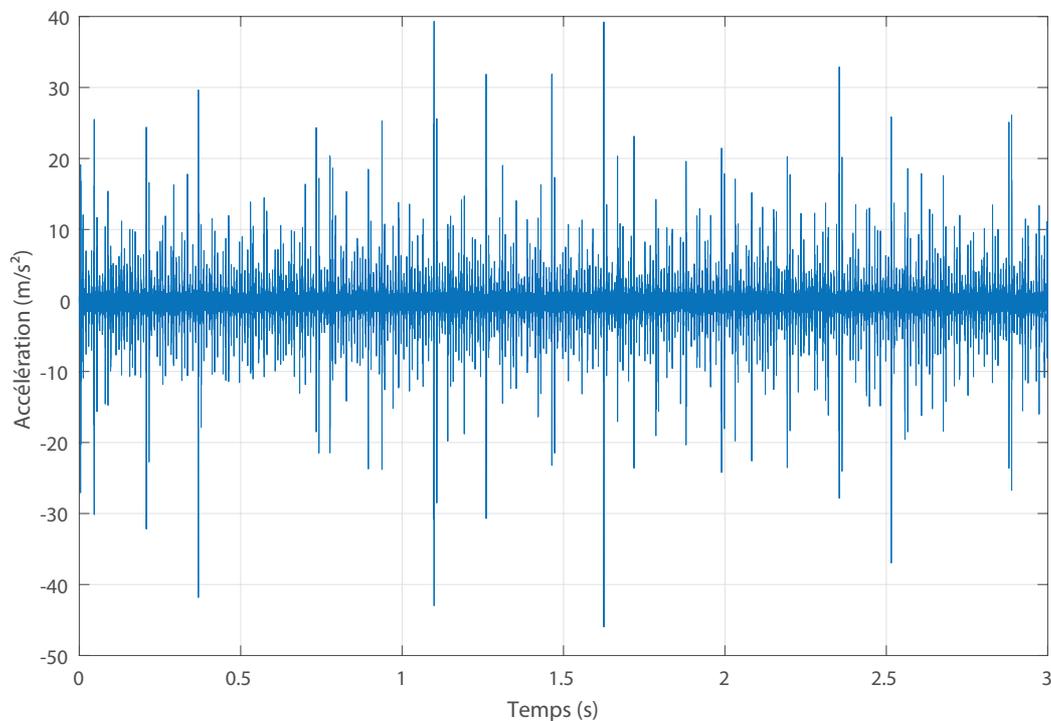


FIGURE 4.3 – Signal vibratoire acquis d'un accéléromètre placé sur le palier comportant le roulement défaillant.

Nous nous intéressons dans un premier temps à examiner la transformée de Fourier du signal vibratoire. La Fig.(4.4) montre la présence d'un spectre de raies autour de $25Hz$ correspondant à la fréquence de rotation du moteur auquel se superposent des fréquences multiples du réseau électrique. La périodicité du phénomène de rotation du moteur indique la présence d'une composante cyclo-stationnaire à l'ordre 1.

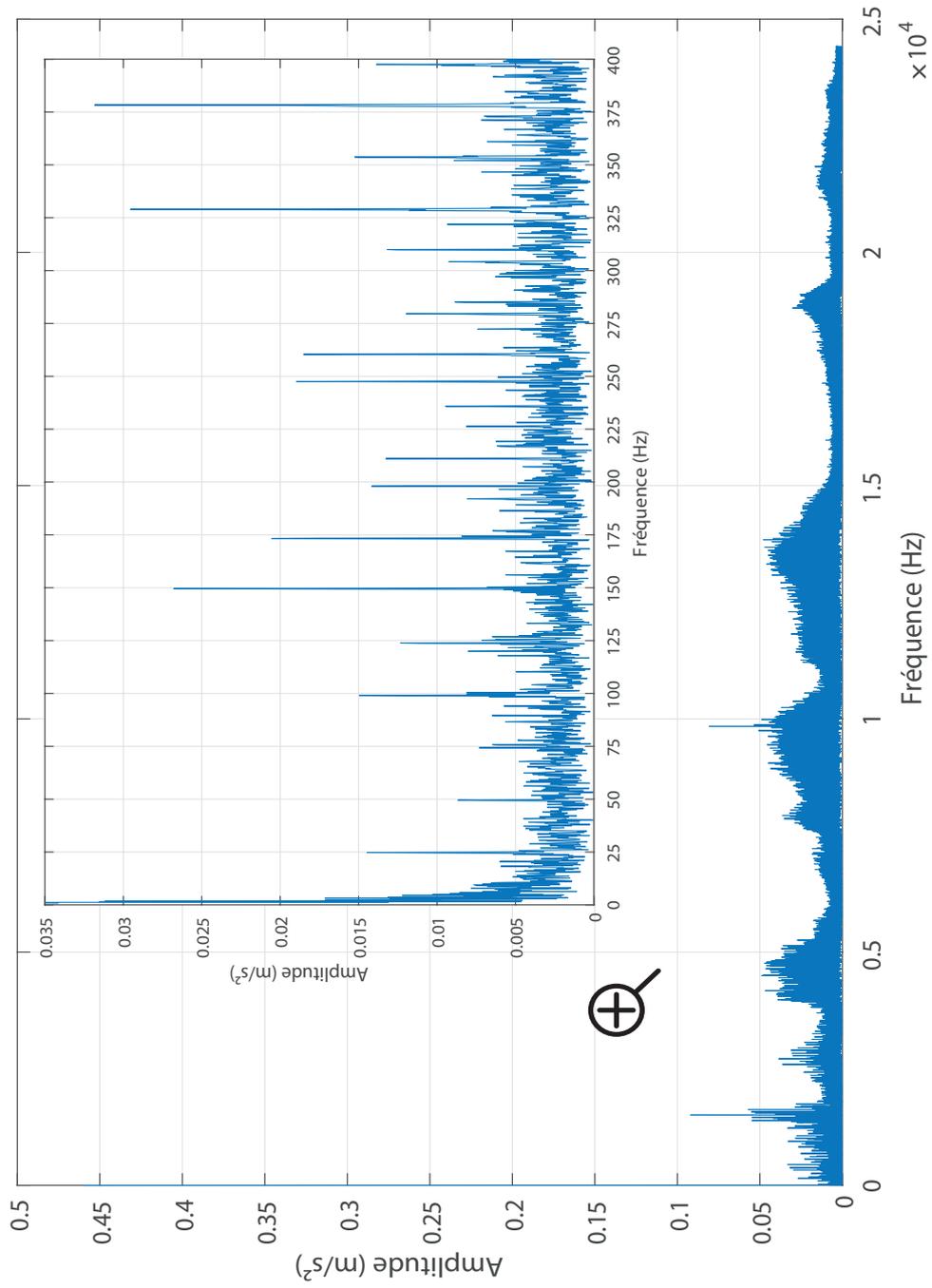


FIGURE 4.4 – Spectre d'amplitude du signal de roulement.

Afin d'étudier la composante cyclo-stationnaire à l'ordre 2, nous examinons la DSC donnée par l'équation (1.36) et la DSCi donnée par l'équation (1.37) évaluées sur le plan bi-fréquentiel (f, α) . Cette analyse permet de détecter et localiser le défaut de bague interne³. Les résultats donnés par les Fig.(4.6) et (4.5) font apparaître des pics visibles aux multiples entiers de la fréquence de la bague interne soit environ $f_{\text{int}} = 117.82\text{Hz}$ avec une fréquence de rotation égale à 24.77Hz . Les pics qui se trouvent de part et d'autre de la fréquence f_{int} et ses harmoniques sont espacés de la fréquence de rotation.

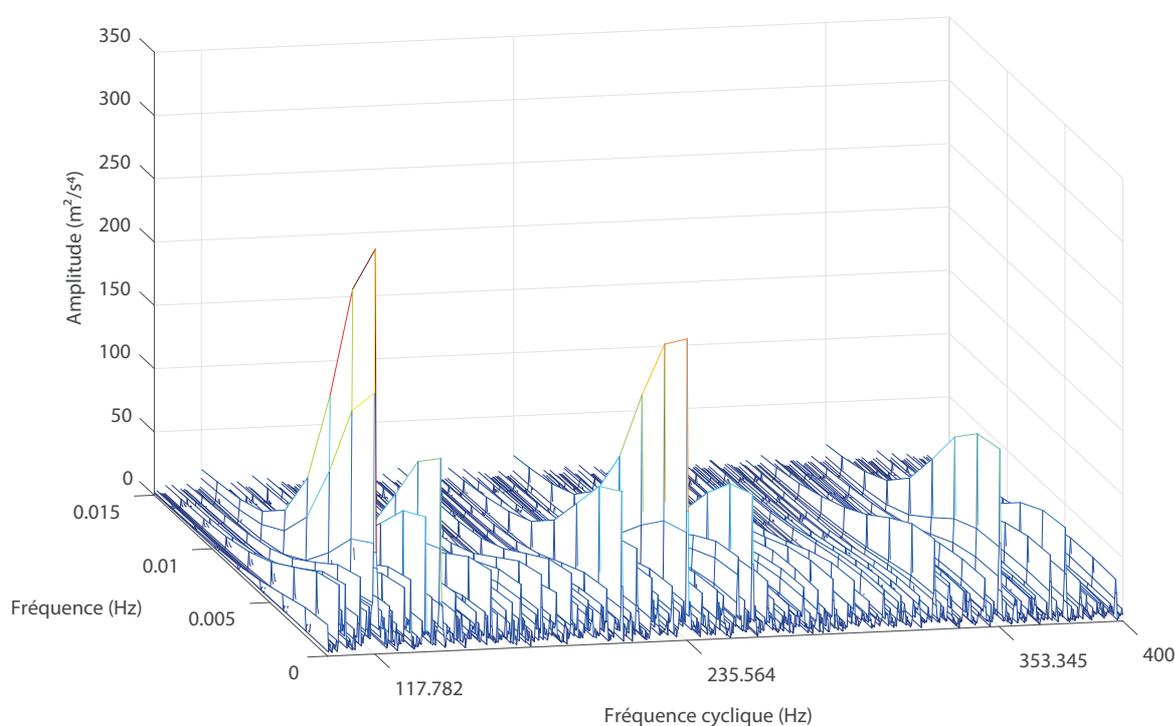


FIGURE 4.5 – Densité spectrale cyclique du roulement.

4.1.3 Séparation de roulements endommagés réels

Introduction

Nous proposons d'utiliser la méthode présentée dans (§3.1.4) afin de séparer différentes sources cyclo-stationnaires issues de roulement, nous allons présenter à travers cette partie les résultats de séparation obtenus. Nous donnerons un descriptif du banc d'essai ainsi que les caractéristiques du système

3. Pour plus de détails sur l'aptitude des représentations DSC et DSCi à détecter les défauts localisés dans un roulement, nous renvoyons le lecteur à lire la référence [14].

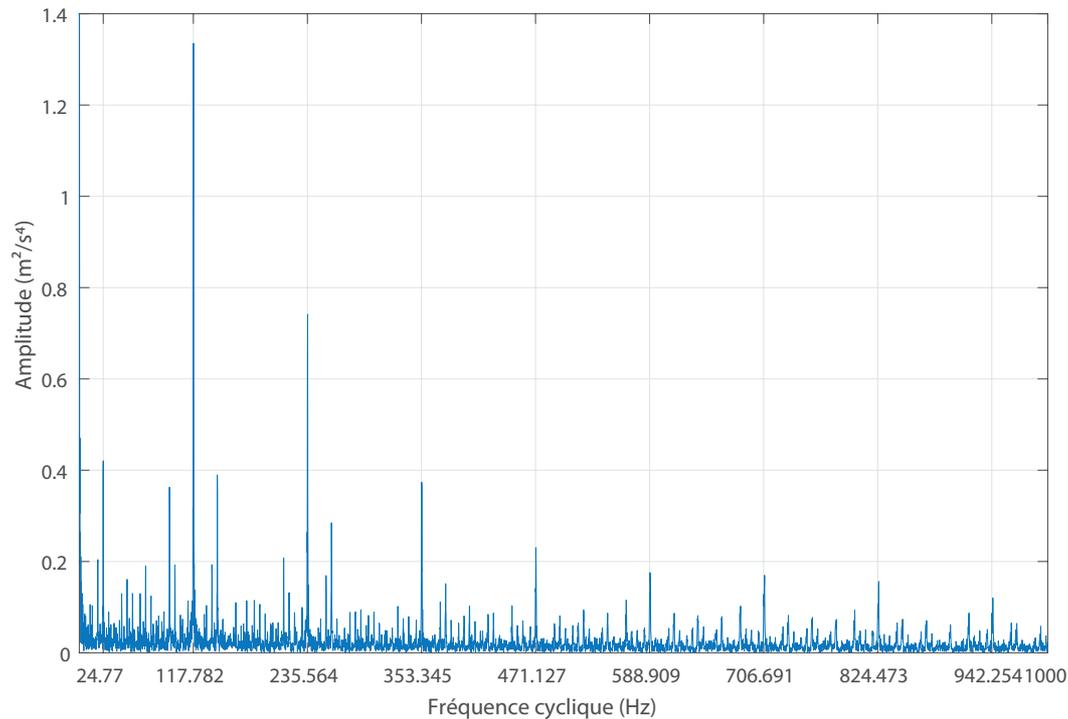


FIGURE 4.6 – Densité spectrale cyclique intégrée du roulement.

de mesure. Nous analyserons les mesures enregistrées à partir des capteurs accélérométriques. Enfin, les résultats obtenus seront analysés et discutés.

Présentation du banc d'essai et des conditions d'acquisition

Les signaux proviennent d'un banc d'essai développé par le département *Ingénierie Mécanique* de l'Université KU Leuven en Belgique. Ce banc est constitué de deux roulements à rotules sur rouleaux de type *22208-EK-C3-SKF* (voir Fig.(4.7)) et d'un moteur asynchrone lié à un réducteur à charge variable. Chaque roulement présente un défaut localisé : un sur la bague interne et l'autre sur la bague externe. Deux signaux vibratoires ont été acquis à partir de deux capteurs accélérométriques triaxiaux *PCB* du modèle *356A16* placés à distances des paliers monoblocs qui servent à supporter et à guider l'arbre du moteur tournant. L'acquisition des signaux a été réalisée en utilisant une carte de *National Instruments* et le logiciel d'instrumentation *LabVIEW* avec une fréquence d'échantillonnage de 10kHz avec une durée d'acquisition de 1 seconde. Ces signaux ont été ensuite traités par le logiciel *MATLAB*. La vitesse du moteur d'entraînement est de 960trs/mn soit $f_r = 16\text{Hz}$. Une photo du banc d'essai est donnée par la Fig.(4.8).



FIGURE 4.7 – Roulement à rotules sur rouleaux de type 22208-EK-C3-SKF.

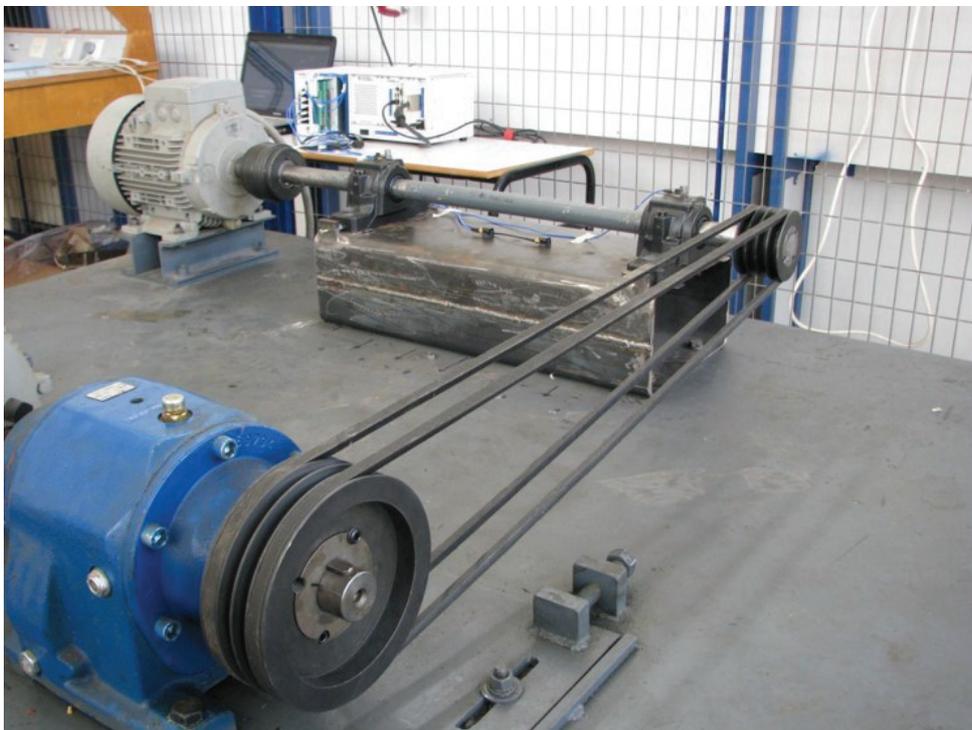


FIGURE 4.8 – Banc expérimental de roulement.

Le tableau (4.2) donne les fréquences caractéristiques fournies par le constructeur SKF :

f_{int}	f_{ext}	f_{rouleaux}	f_{cage}
$149.7Hz$	$106.3Hz$	$45Hz$	$6.64Hz$

TABLE 4.2 – Fréquences caractéristiques du roulement 22208-EK-C3-SKF.

Analyse des mesures acquises et résultats de séparation

L'objectif de cette étude expérimentale est de séparer un mélange de signatures de roulements comportant deux défauts localisés différents provenant de deux capteurs accélérométriques (voir Fig.(4.9)). Il s'agit donc d'un

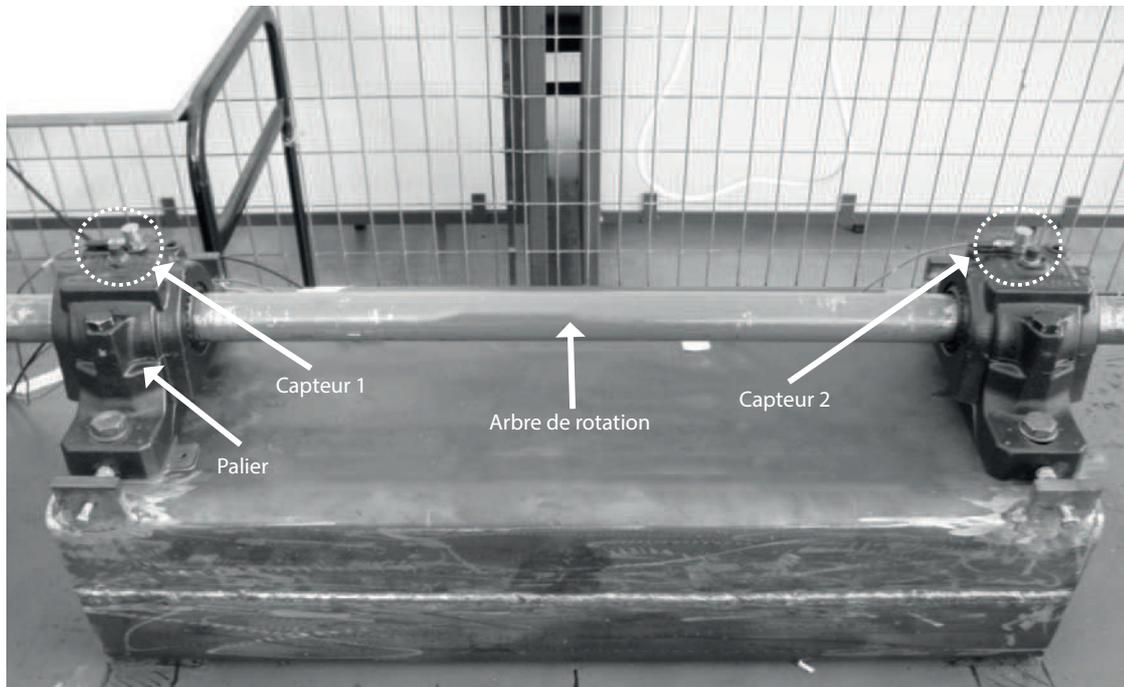


FIGURE 4.9 – Position des capteurs accélérométriques.

système de mélange déterminé de sources cyclo-stationnaires de fréquences cycliques différentes. Comme les capteurs sont placés à proximité des roulements, nous supposons que le mélange est linéaire instantané. La Fig.(4.10) présente les mesures acquises dans le domaine temporel. Nous pouvons constater des signaux fortement bruités. Aucune information ne peut être tirée sans examen supplémentaire. Nous proposons d'examiner leurs DSCI afin de diagnostiquer l'état des roulements. La Fig.(4.12) révèle la présence des pics à la fréquence de rotation de $f_r = 16.48Hz$, aux fréquences caractéristiques et à ses harmoniques du défaut de bague interne à $f_{int} = 148.9Hz$ et de bague externe à $f_{ext} = 108Hz$. Avec des pics de faibles amplitudes, nous concluons que les défauts ne sont pas assez prononcés et qu'il s'agit bien d'un mélange de deux sources.

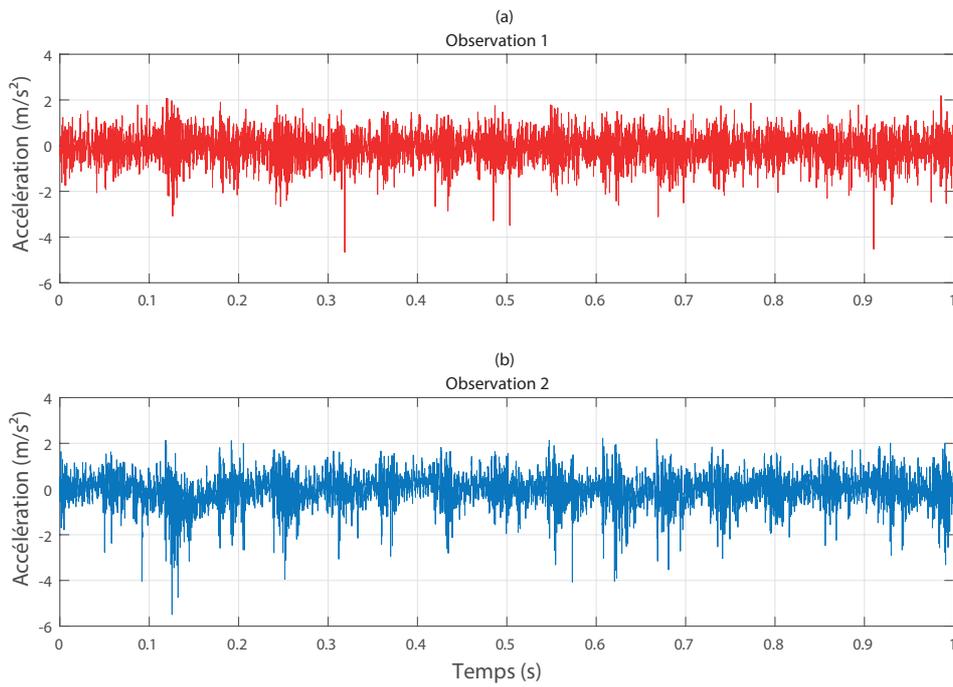


FIGURE 4.10 – Signaux vibratoires des observations : (a) observation 1, (b) observation 2.

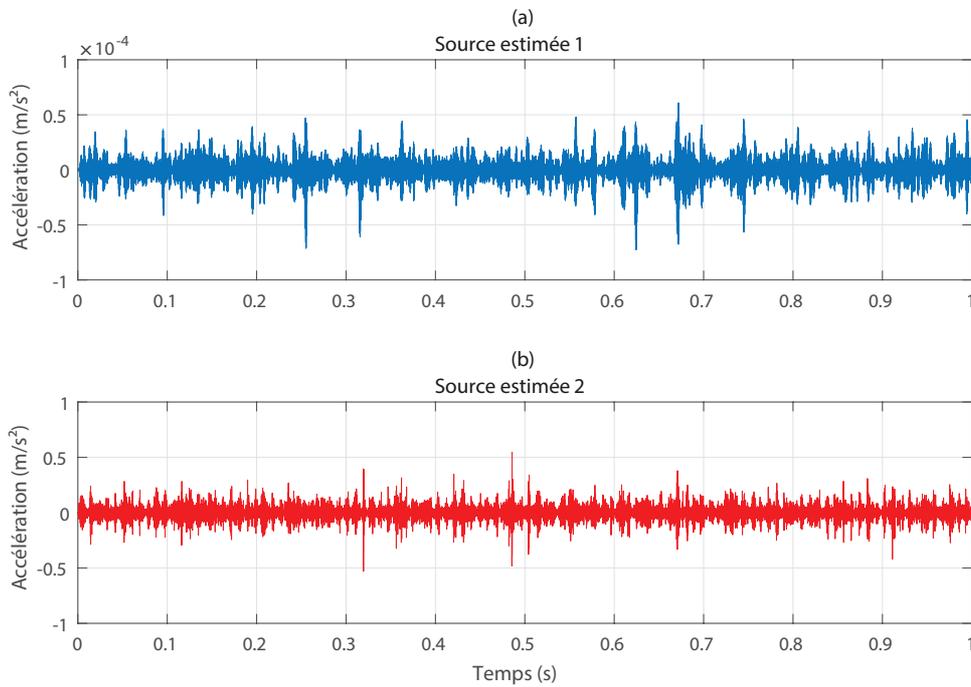


FIGURE 4.11 – Estimation des signaux sources : (a) source 1, (b) source 2.

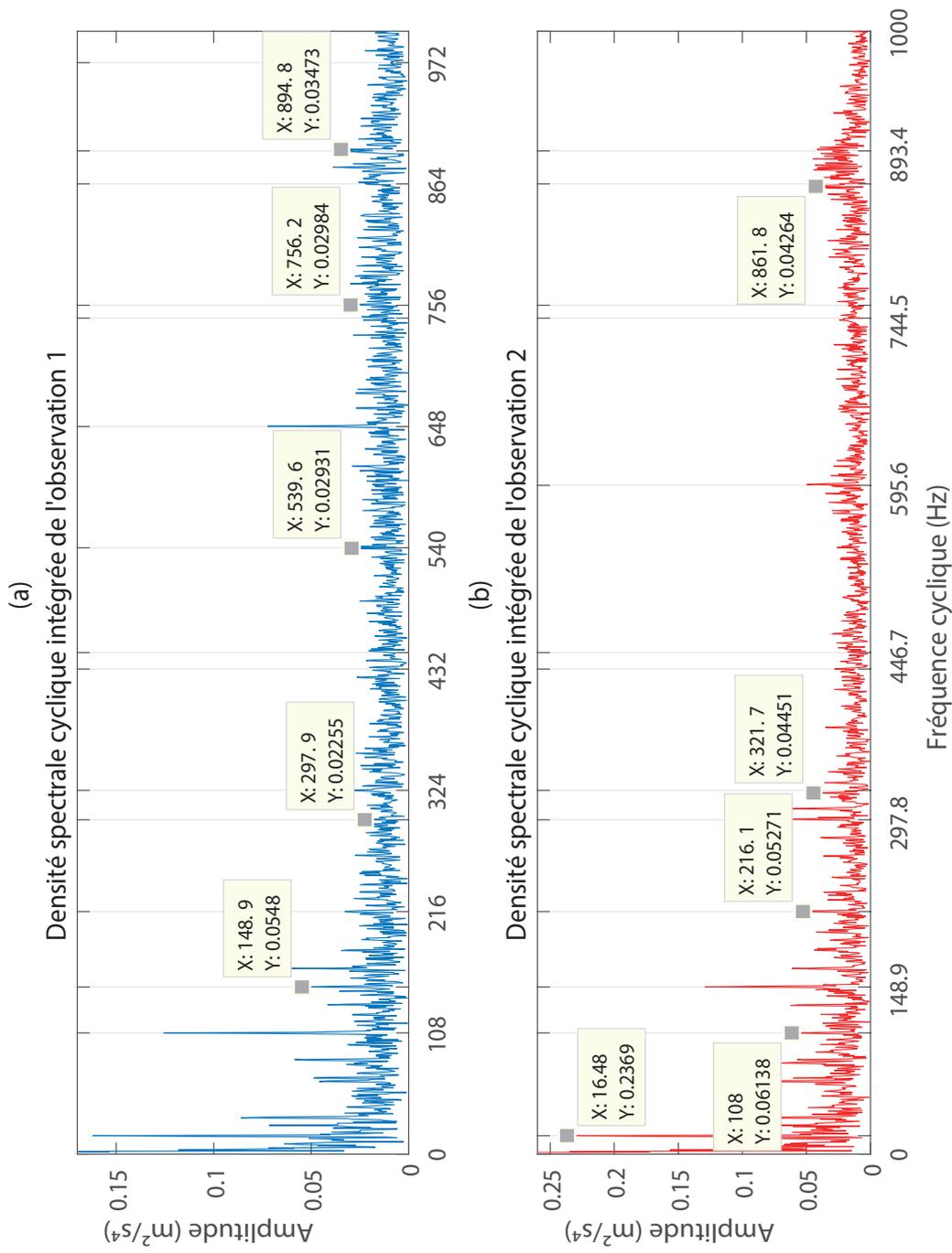


FIGURE 4.12 – Densité spectrale cyclique intégrée : (a) observation 1, (b) observation 2.

Afin de séparer ces mélanges, nous appliquons une diagonalisation conjointe d'un ensemble de matrices de corrélation cyclique indexées par différents retards et par les fréquences cycliques des défauts et leurs harmoniques. La diagonalisation conjointe est assurée par la méthode BFGS. Avant de traiter les observations, nous appliquons un filtre passe bande en haut fréquence afin d'éliminer la contribution cyclo-stationnaire à l'ordre 1 (basses fréquences). Il faut rappeler que la méthode souffre de deux indéterminations à savoir l'ambiguïté de permutation et l'ambiguïté d'amplitude. Nous récapitulons les étapes de la Fig.(4.13).

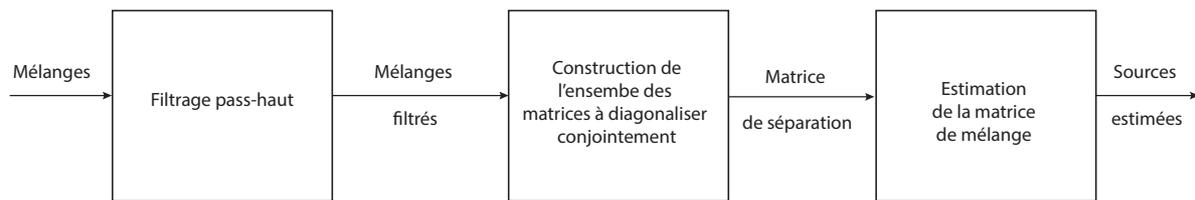


FIGURE 4.13 – Principe de la méthode proposée.

La résultats de la séparation sont illustrés par la Fig.(4.11). Nous constatons la possibilité de retrouver la période d'apparition des défauts dans certains intervalles temporels. Voyons à présent si ces résultats ont pu restituer les deux sources en analysant leurs DSCi. Nous constatons effectivement dans chaque DSCi la présence des pics à la fréquence de rotation, à la fréquence caractéristique et quelques harmoniques du défaut de bague externe et l'atténuation ou l'absence de la fréquence caractéristique du défaut de bague interne dans la Fig.(4.14 (b)) et inversement dans la Fig.(4.14 (a)). Ce qui montre que la méthode a pu estimer convenablement les signatures de chaque défaut.

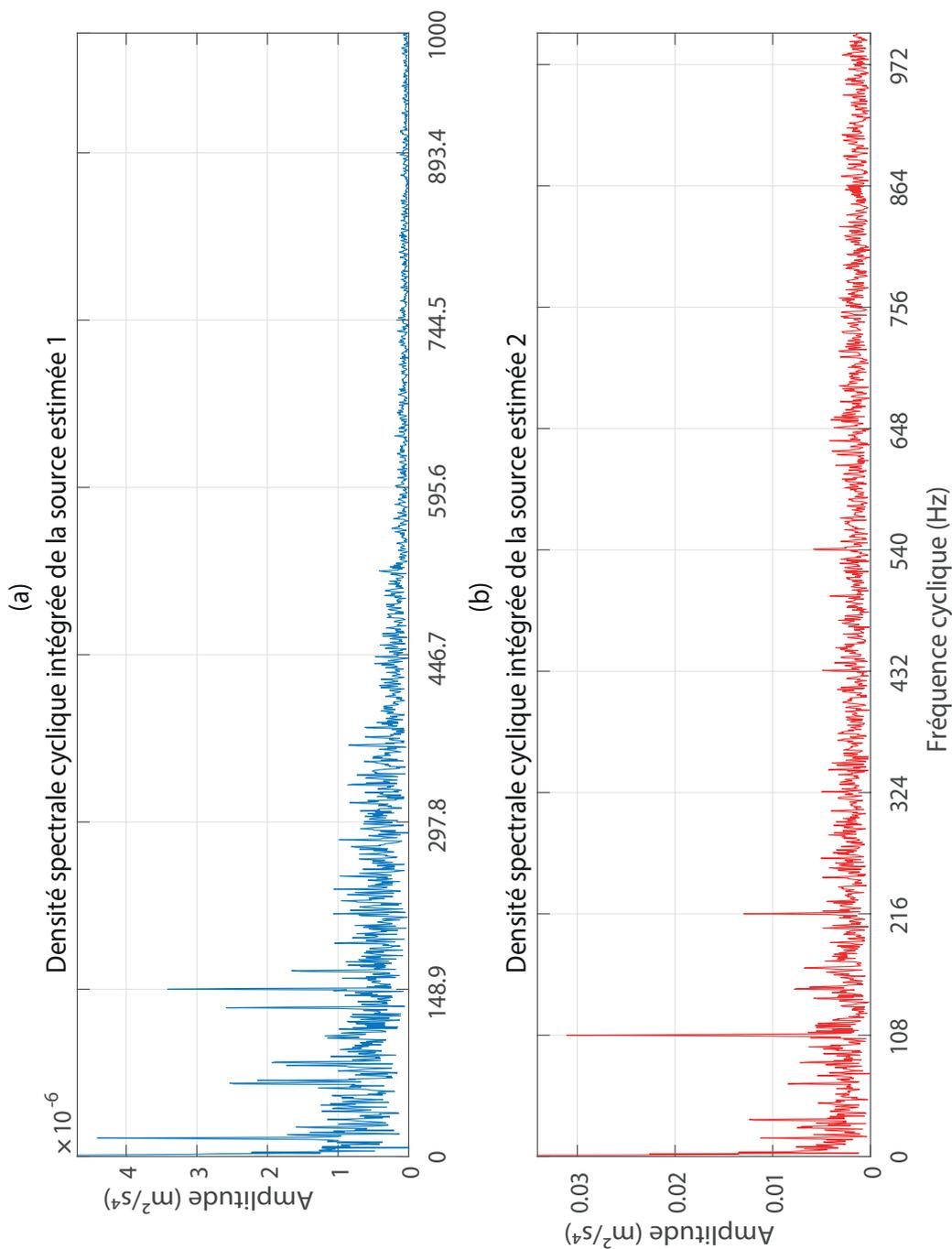


FIGURE 4.14 – Densité spectrale cyclique intégrée : (a) source estimée 1, (b) source estimée 2.

4.2 Séparation des contributions cyclo-stationnaires aux ordres 1 et 2, application aux signaux biomécaniques

4.2.1 Introduction

En biomécanique, la FRS est la force exercée par le sol en contact avec un corps. Elle s'oppose à la force exercée par le corps sur le sol qui n'est rien d'autre que son poids. L'emploi du mot réaction dérive de la troisième loi de Newton, qui stipule que si une force, appelée *action*, agit sur un corps, alors une force égale et opposée, appelée *réaction*, doit agir sur ce corps. La FRS est souvent observée pour évaluer la production de force dans divers groupes d'athlètes au sein de la communauté sportive. l'objectif est d'aider ces athlètes à estimer leurs capacités à exercer une performance sportive. Cela peut conduire à proposer des exercices de musculation adaptés, évaluer les changements dans la fatigue du muscle pendant la locomotion (marche ou course). L'analyse de la FRS peut également apporter des informations utiles sur le diagnostic clinique, l'évaluation des gestes sportifs, la réadaptation et d'extraire les caractéristiques mécaniques et cinématiques des membres inférieurs et d'analyser les impacts au sol [29, 185].

Les signaux FRS sont généralement composés de deux parties distinctes : une force active représentant la force de propulsion en plus d'une force passive représentant la force d'impact. La force d'impact, principalement générée passivement, est due à la décélération de la masse corporelle à l'instant du toucher. Cela implique qu'une partie de l'énergie a été transférée passivement dans le système os-ligament. Ainsi, la force d'impact est un facteur majeur indiquant la réaction du muscle qui peut refléter l'état de fatigue et la performance du muscle. La Fig.(4.15) montre un coureur sur un tapis dynamométrique et un exemple de signal mesuré pendant deux pas consécutifs. La majorité des personnes courent avec un appui talon. C'est une attaque du sol avec le talon du pied, suivie d'un déroulement progressif de la plante qui va se finir par une impulsion pour effectuer une nouvelle foulée et ainsi de suite. Le reste des personnes court avec un appui en plante de pied qui consiste à amortir en premier lieu avec l'avant du pied le contact avec le sol puis le talon vient toucher très légèrement le sol dans un second temps. Ce type d'appui se caractérise par une force d'impact amortie et masquée dans la poussée musculaire comme le montre la Fig.(4.16)⁴. Pour étudier les variations des

4. La figure reprend les résultats d'une expérience réalisée dans Harvard University Skeletal Biology and Biomechanics Lab par Daniel E. Lieberman et al. ([1])

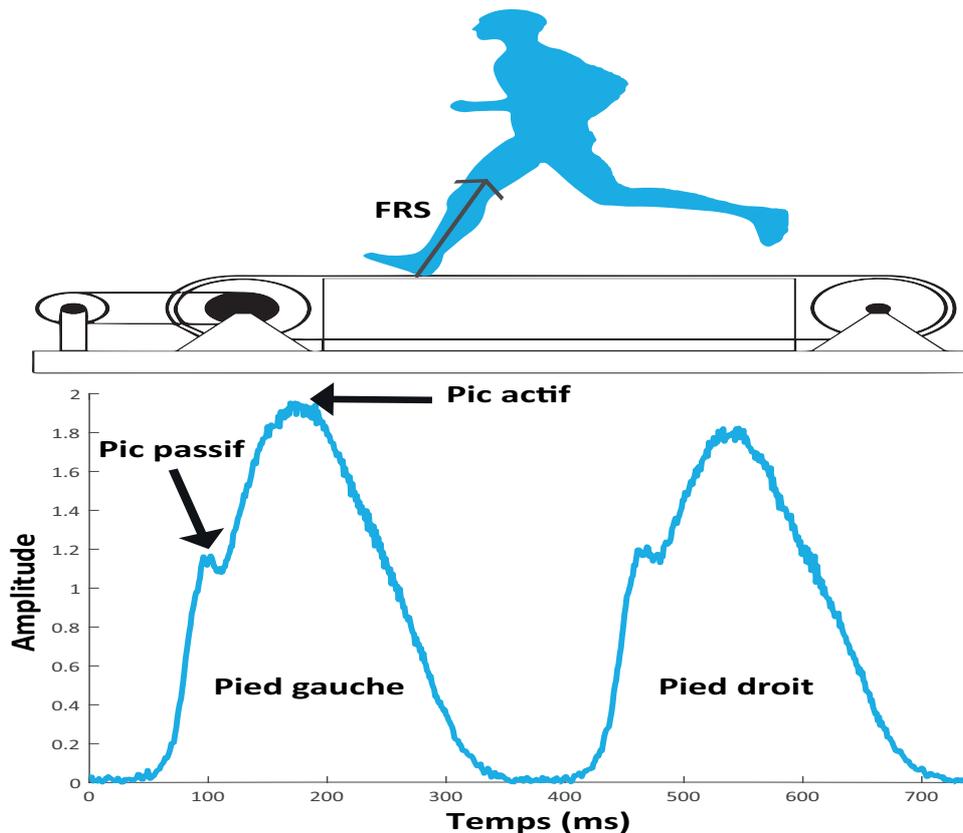


FIGURE 4.15 – Coureur sur le tapis roulant et exemple de signal mesuré

signaux de FRS dans le temps, il est nécessaire de séparer les contributions active et passive. Sabri et al. ont montré dans [167] que la cyclo-stationnarité définit correctement la nature des signaux FRS. Cette propriété se caractérise principalement par le fait que l'on a un couplage d'un phénomène périodique (cycle de la foulée) avec un phénomène aléatoire stationnaire (amplitude aléatoire de la force d'une foulée à l'autre). Sabri et al. ont utilisé dans [167] plusieurs techniques de séparation de sources basées sur l'indépendance statistique entre les parties passives et actives pour chaque jambe à savoir les algorithmes SOBI, Cyclo-SOBI et JADE. Ils ont également employé la décomposition en valeurs singulières et la diagonalisation conjointe des matrices spectrales cycliques des observations dans [166]. Parmi toutes ces techniques, seul JADE a pu donner des résultats convenables car les sources estimées ont un Kurtosis supérieur à 0. Autrement dit, les sources ne sont pas de distributions gaussiennes. F.A. Zakaria et al. considèrent dans [195] que la composante active du signal FRS suit une distribution gaussienne et par suite une détection des vrais extrémums permet, par le biais de la méthode d'estimation des moindres carrés non linéaires, de séparer la partie active (cyclo-stationnaire à l'ordre 1 basse fréquence) de celle passive (cyclo-stationnaire

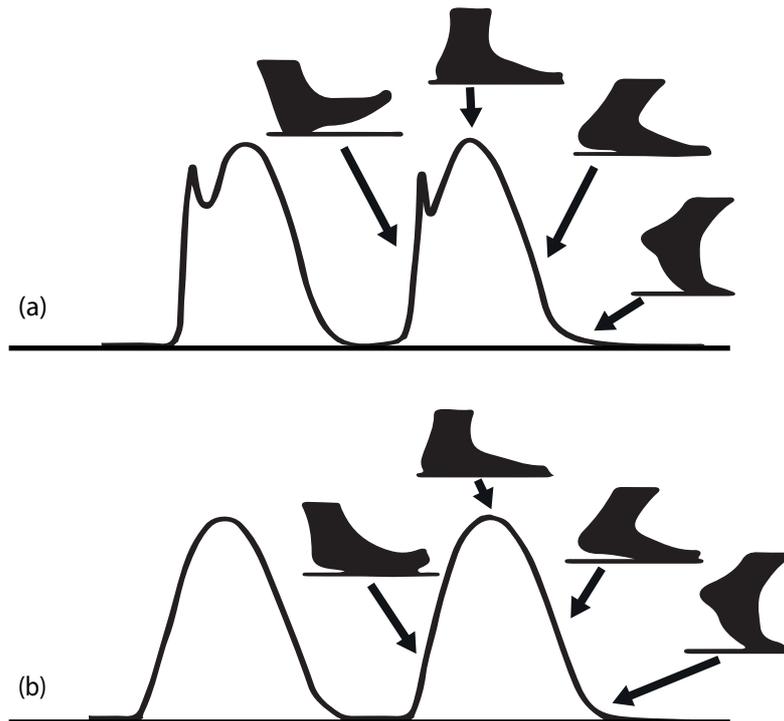


FIGURE 4.16 – Évolution de la FRS et cinématique des pieds pendant la course (a) avec l'appui talon (b) et l'appui en plante de pied.

à l'ordre 2 large bande). F.A. Zakaria et al. ont proposé dans [196] d'utiliser l'analyse en composantes morphologiques afin de décomposer un signal cyclo-stationnaire en deux composantes périodiques et aléatoires. S. Maiz a identifié dans sa thèse [134] les caractéristiques de la variabilité liée à la fatigue en étudiant la FRS lors d'une course de 24 heures. La fatigue du coureur peut être associée à différentes signatures fréquentielles des statistiques du second ordre de la composante passive des signaux FRS. F. Bonnardot et al. ont réalisé un système d'acquisition portable par des coureurs occasionnels pour mesurer les vibrations produites lors d'un chemin varié (goudron, terre, ...). L'analyse cyclo-stationnaire de la composante passive de ces signaux a permis de détecter la fatigue chez ces coureurs [31].

Considérons maintenant, le problème de la séparation des contributions cyclo-stationnaires aux ordres 1 et 2 appliqué au signaux FRS de course à pied, à savoir, la force d'impact et la force de propulsion, en exploitant l'analyse en composantes parcimonieuses utilisant l'un des algorithmes suivants OMP et LASSO. Nous traitons en premier lieu des signaux synthétiques modélisant les signaux FRS, puis des signaux biomécaniques expérimentés par des coureurs sur une ultra-longue distance.

4.2.2 Formulation du problème

Le signal FRS peut être modélisé mathématiquement par deux composantes distinctes comme suit :

$$\mathbf{y}(t) = \mathbf{p}(t) + \mathbf{a}(t), \quad (4.6)$$

où

- $\mathbf{p}(t)$ est la composante passive, cyclo-stationnaire à l'ordre 2⁵ traduisant l'aléa et le manque de maîtrise qui caractérise la force d'appui d'une jambe d'une foulée à l'autre. Il apparaît donc une fluctuation aléatoire cyclique au moment du contact entre le talon et le sol. Cette composante est déterminée par la masse effective du corps, la vitesse et la rigidité de la jambe,
- $\mathbf{a}(t)$ est la composante active, cyclo-stationnaire à l'ordre 1, plus large que la composante passive représentant la force propulsive appliquée par le système musculosquelettique. Contrairement à la composante passive, la composante active est caractérisée par la détermination et la maîtrise du pied par rapport au sol.

Les deux composantes se produisent de manière indépendante sans qu'il y ait de corrélation entre ces actions et ce pour chaque jambe. L'objectif principal étant d'extraire la composante active de celle passive à partir d'une observation mono-canal d'un signal FRS.

4.2.3 Préliminaires théoriques

Pour résoudre notre problème d'intérêt, nous choisissons d'utiliser les deux algorithmes suivants : Orthogonal Matching Pursuit (cet algorithme ne fait appel à aucune mesure de parcimonie) et Least Absolute Shrinkage and Selection Operator (cet algorithme utilise une norme mixte $\ell_1 - \ell_2$). Dans les deux sous-sections suivantes, nous détaillerons le fonctionnement de ces derniers.

Algorithme OMP

L'OMP introduit dans [152] est un algorithme glouton itératif qui choisit à chaque étape l'atome qui maximise le produit scalaire avec le résidu. Cet atome est alors ajouté dans l'ensemble d'atomes choisis. La mise à jour correspond à une projection orthogonale des données sur tous les atomes

5. Pour une analyse détaillée de la cyclo-stationnarité de ces signaux, nous renvoyons le lecteur aux travaux de Sabri et al. [167], S. Maiz [134] et F.A. Zakaria et al. [195].

choisis, ce qui évite l'ajout des atomes déjà sélectionnés. Après, l'algorithme s'exécute en boucle jusqu'à ce que le critère d'arrêt soit atteint. Le principe de l'algorithme OMP est résumé dans l'algorithme (3).

Algorithme 3 OMP

Entrées :

-Un signal \mathbf{y} , un dictionnaire \mathbf{A} et un critère d'arrêt ;

Initiation :

1- $\mathbf{r}^{(0)} = \mathbf{y}$, $\Lambda_0 = \emptyset$, $k = 0$;

2- Trouver un indice $w^{(k)}$ qui résout : $w^{(k)} = \arg \max_w |\mathbf{A}_w^T \mathbf{r}^{(k-1)}|$;

3- Ajouter $w^{(k)}$ aux indices sélectionnés $\Lambda_k = \Lambda_{k-1} \cup \{w^{(k)}\}$;

4- Calculer $\hat{\mathbf{x}}^{(k)} = \mathbf{A}_{\Lambda_k}^\# \mathbf{y}$ et $\mathbf{r}^{(k)} = \mathbf{y} - \mathbf{A}_{\Lambda_k} \hat{\mathbf{x}}^{(k)}$;

si critère d'arrêt est atteint **alors**

-Arrêter l'algorithme ;

sinon

-Incrémenter k et retourner à l'étape 2 ;

fin si

Sorties :

- Vecteur des coefficients parcimonieux $\hat{\mathbf{x}}$.

- Résidu \mathbf{r} .

Algorithme LASSO

Développé par Robert Tibshirani dans [87], le *Least Absolute Shrinkage and Selection Operator* (LASSO) est une approche de contraction des valeurs de la régression. Dans le cas d'un modèle linéaire standard, les coefficients sont calculés par la minimisation de la somme des carrés des résidus. Le LASSO est formulé comme suit : soit un ensemble de mesures composé des vecteurs suivants $\mathbf{x}_1, \dots, \mathbf{x}_p$ et \mathbf{y} une mesure de sortie, le LASSO correspond à un modèle linéaire s'écrivant comme $\hat{\mathbf{y}} = b_0 + b_1 \mathbf{x}_1 + b_2 \mathbf{x}_2 + \dots + b_p \mathbf{x}_p$. Retrouver ce modèle passe par la minimisation de la fonction de cout suivante :

$$\min \sum |\mathbf{y} - \hat{\mathbf{y}}|^2 \text{ sous contrainte } \sum_{j=0}^p |b_j| \leq s \quad (4.7)$$

où s est un paramètre de réglage. Si s est assez grand alors la contrainte n'a aucun effet sur le résultat et la solution n'est que la régression linéaire des moindres carrés multiple de \mathbf{y} sur $\mathbf{x}_1, \dots, \mathbf{x}_p$. Cependant, lorsque s prend des valeurs faibles ($s \geq 0$), alors les solutions sont des versions rétrécies des estimations des moindres carrés. Le choix de s conditionne le nombre de

prédicteurs à utiliser dans le modèle de régression. Le calcul des solutions du LASSO est un problème de programmation quadratique et peut être abordé par des algorithmes d'analyse numérique standard. Cependant, l'algorithme le plus utilisé est le Least Angle Regression (LARS) vu son efficacité par rapport aux autres méthodes. Le LARS suit les étapes suivantes :

- Initialiser tous les coefficients à zéro et trouver le prédicteur \mathbf{x}_j le plus corrélé avec \mathbf{y} ,
- Déplacer le coefficient \mathbf{b}_j dans la direction du signe de sa corrélation avec \mathbf{y} en mettant à jour le résiduel $r = \mathbf{y} - \hat{\mathbf{y}}$. Arrêter le déplacement quand nous trouvons un autre prédicteur \mathbf{x}_k plus corrélé que \mathbf{x}_j avec \mathbf{r} .
- Déplacer le couple $(\mathbf{b}_j, \mathbf{b}_k)$ dans leur direction conjointe des moindres carrés, jusqu'à ce qu'un autre prédicteur \mathbf{x}_m a plus de corrélation avec le résidu \mathbf{r} que \mathbf{x}_j et \mathbf{x}_k .
- Continuer jusqu'à retrouver tous les prédicteurs composant le modèle linéaire.

4.2.4 Méthode proposée

L'objectif étant de séparer les deux composantes active et passive du signal FRS. Nous proposons de chercher une représentation parcimonieuse de l'une des deux composantes du signal puis de la retrancher au signal initial afin d'obtenir l'autre composante. Cela demande de construire un dictionnaire adéquat avec l'une des composantes.

En partant des travaux de Sabri et al. dans [167], où il a été montré que la composante active a un caractère cyclo-stationnaire à l'ordre 1, bien que la composante passive soit cyclo-stationnaire d'ordre 2 et à la même fréquence. Nous suggérons d'utiliser un dictionnaire qui reflète la nature cyclo-stationnaire à l'ordre 1 de la composante active du signal FRS, à savoir la transformée en sinus discrète (DST). Ce choix se justifie par le fait que la DST représente d'une manière adaptée les parties lisses et périodiques d'un signal, ainsi les informations utiles seront contenues dans un nombre limité de coefficients parcimonieux. La multiplication de ses coefficients par la DST permet de reconstruire la composante périodique dans le domaine temporel. La DST se caractérise par la création de coefficients réels, car son noyau de projection est un sinus et elle est calculée de la façon suivante :

$$\text{DST}(n) = \sqrt{\frac{2}{N}} \sum_{k=1}^N \sin \left[\frac{\pi}{N} \left(k + \frac{1}{2} \right) n \right], \quad k = 0, \dots, N \quad (4.8)$$

où n est l'indice ligne de la DST et N sa taille. La partie passive est obtenue en retranchant la composante active du signal initial. L'organigramme de la méthode proposée est donné par la Fig.(4.17). Il est à noter que le calcul

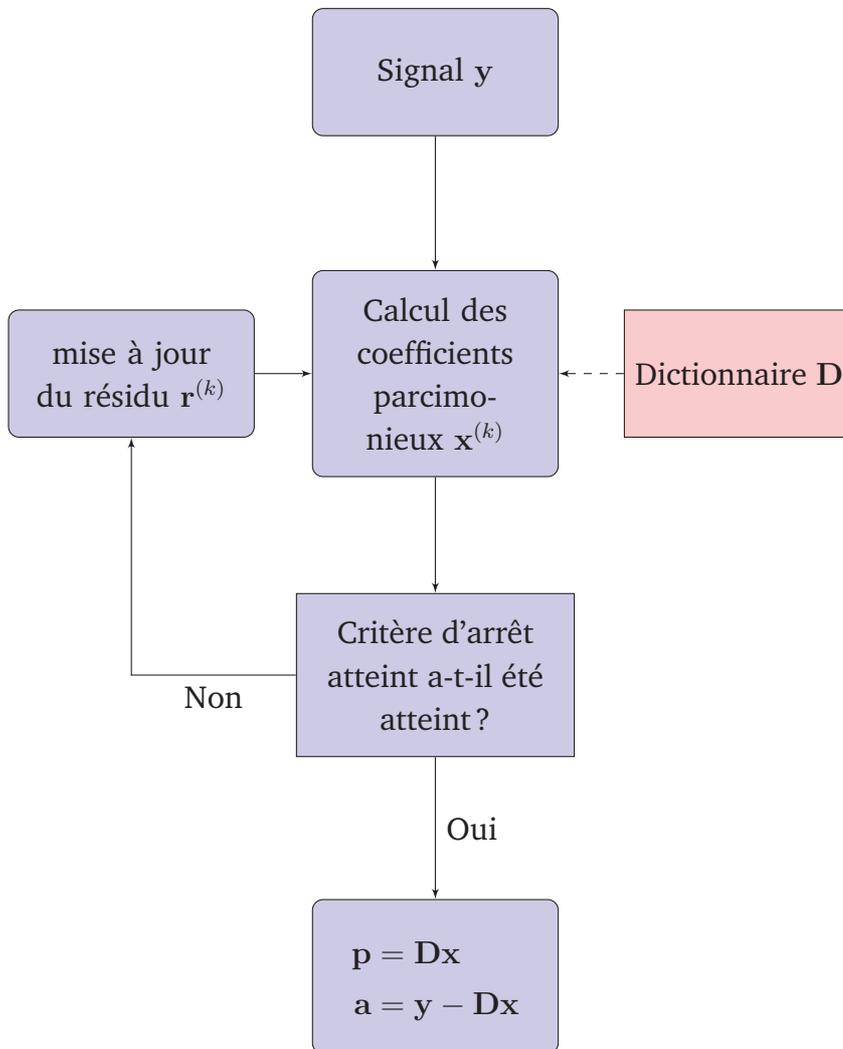


FIGURE 4.17 – Organigramme de la méthode proposée pour la séparation des composantes de la FRS.

des coefficients parcimonieux se fait par l'un des deux algorithmes présentés dans la section précédente. Notre méthode dépend principalement du nombre d'itérations. En effet, durant les premières itérations, nous cherchons à extraire les coefficients parcimonieux les plus corrélés avec le dictionnaire DST. Au fur et à mesure, cette corrélation diminue suite au prélèvement de la composante active du signal FRS. Par conséquent, le reste est lié particulièrement à la composante passive. Nous proposons alors un critère d'arrêt : si la variation de l'énergie du résidu pour chaque itération n'est plus signifi-

cative, nous arrêtons l'algorithme. Cela se traduit par :

$$v = \{\mathbf{E}\{\mathbf{r}^{(k+1)}(\mathbf{r}^{(k+1)})^H\} - \mathbf{E}\{\mathbf{r}^{(k)}(\mathbf{r}^{(k)})^H\}\} \leq \epsilon \quad (4.9)$$

où ϵ est un scalaire positif suffisamment petit et N est le nombre d'échantillons.

4.2.5 Application sur des signaux FRS synthétiques

Afin de valider la méthode proposée, nous allons tout d'abord traiter un exemple à vocation didactique avec un signal $\mathbf{x}(t)$ très simple modélisant la FRS. Le signal $\mathbf{x}(t)$ est composé d'une sinusoïde bruitée et d'une modulation d'amplitude des impulsions (PAM) échantillonné à 1kHz. La sinusoïde décrit la partie active, tandis que la PAM modélise la partie passive de la FRS. Le signal $\mathbf{x}(t)$ s'écrit donc :

$$\mathbf{y}(t) = \mathbf{p}(t) + \mathbf{a}(t), \quad (4.10)$$

où

$$\mathbf{a}(t) = \sin(2\pi ft) + \mathbf{r}(t) \quad (4.11)$$

$$\mathbf{p}(t) = \sum_i \alpha_i \delta(t - \frac{i}{f}) \quad (4.12)$$

avec $\mathbf{p}(t)$ un vecteur parcimonieux contenant des permutations aléatoires α_n qui suivent une loi normale $\alpha_n \sim \mathcal{N}(0.7, 0.2)$ répétées périodiquement avec une fréquence $f = 20\text{Hz}$ et $\mathbf{r}(t)$ un bruit blanc stationnaire indépendant de $\mathbf{p}(t)$. La Fig.(4.18) montre le signal $\mathbf{x}(t)$ avec ses composantes. Nous constatons d'après les Fig.(4.19) et Fig.(4.20) que la méthode proposée en utilisant l'OMP et le LASSO discrimine correctement les deux composantes du signal FRS avec un léger avantage pour l'OMP compte tenu de ses erreurs quadratiques moyennes inférieures par rapport à celles données par le LASSO.

4.2.6 Application sur des signaux FRS réels

Paradigme expérimental

Le paradigme expérimental a été réalisé au sein du Laboratoire Interuniversitaire de Biologie et de la Motricité (LIBM) de l'Université Jean Monnet Saint-Etienne par A. Belli et al.[21]. Pour plus de détails sur le banc expérimental et ses caractéristiques, ainsi que les informations liées à la physiologie des sujets, nous renvoyons le lecteur à la référence [21]. Les mesures ont été effectuées sur un tapis dynamométrique équipé de 12 accéléromètres répartis sur les 4 coins du tapis (3 accéléromètres sont placés sur chaque coin du

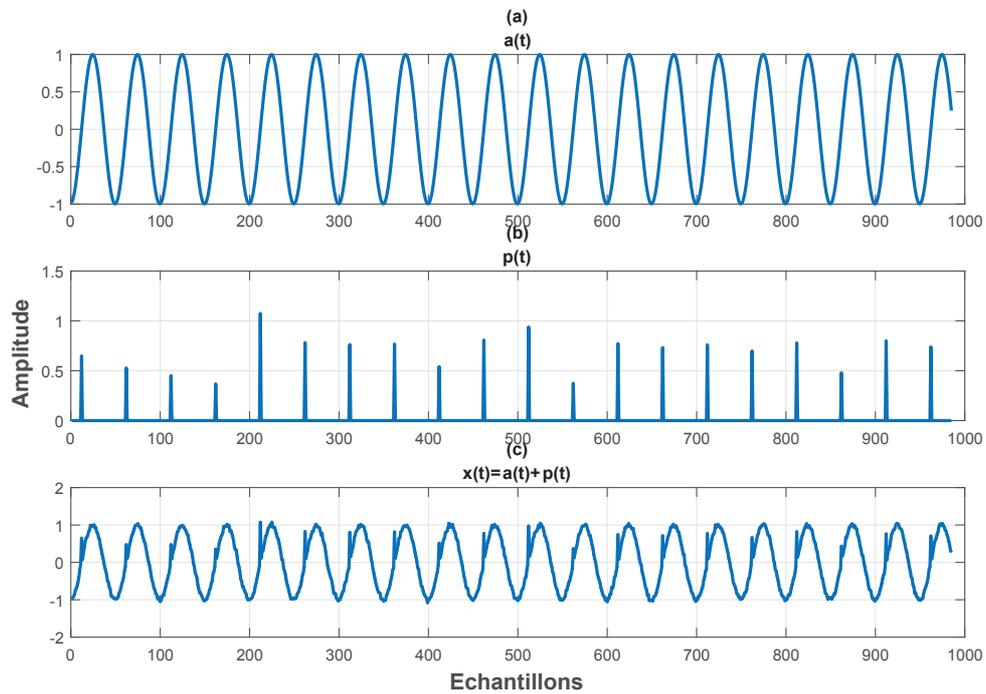


FIGURE 4.18 – (c) Signal $x(t)$ modélisant la FRS composé (a) d'une sinusoïde et (b) d'une modulation PAM.

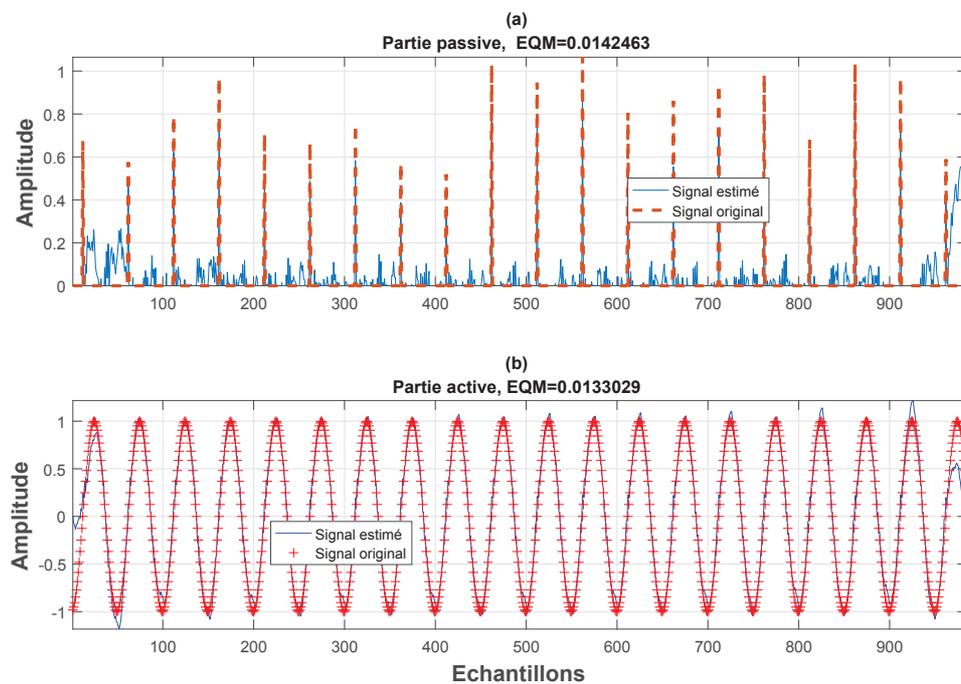


FIGURE 4.19 – Résultats de la séparation par l'OMP avec l'erreur quadratique moyenne : (a) partie active, (b) partie passive.

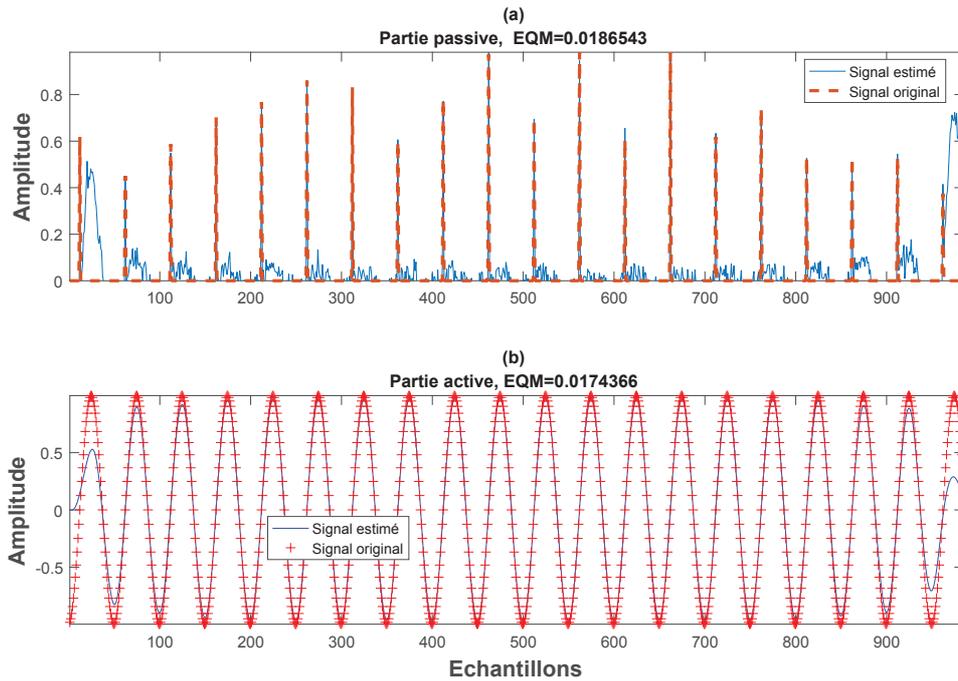
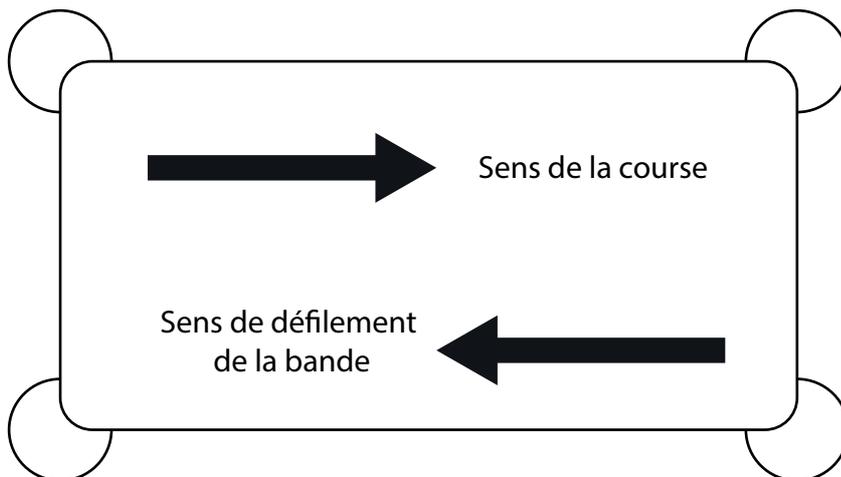


FIGURE 4.20 – Résultats de la séparation par le LASSO avec l’erreur quadratique moyenne : (a) partie active, (b) partie passive.

tapis dynamométrique voir la Fig.(4.21). Ces derniers permettent la mesure et l’enregistrement des signaux de force de réaction au sol issus d’une course d’un sujet sur le tapis. Les caractéristiques de la manipulation sont résumées

Capteurs 1, 2 et 3

Capteurs 7, 8 et 9



Capteurs 4, 5 et 6

Capteurs 10, 11 et 12

FIGURE 4.21 – Localisation des capteurs.

comme suit :

- 10 sportifs masculins de haut niveau ayant une bonne santé se sont portés volontaires à cette étude.
- L'expérience consistait en une course de 24 heures découpée en avec courtes périodes de repos toutes les 2 heures pendant lesquelles le coureur se ravitaillait.
- Aucune blessure n'a été signalée par les coureurs au moment de l'expérience.
- Les signaux issus du tapis dynamométrique étaient échantillonnés à 1kHz.
- La vitesse du tapis roulant a été réglée à 10km/h.
- Le nombre d'échantillons de chaque signal est 20000 et la durée d'acquisition en temps est de 20 secondes chaque 2 heures.

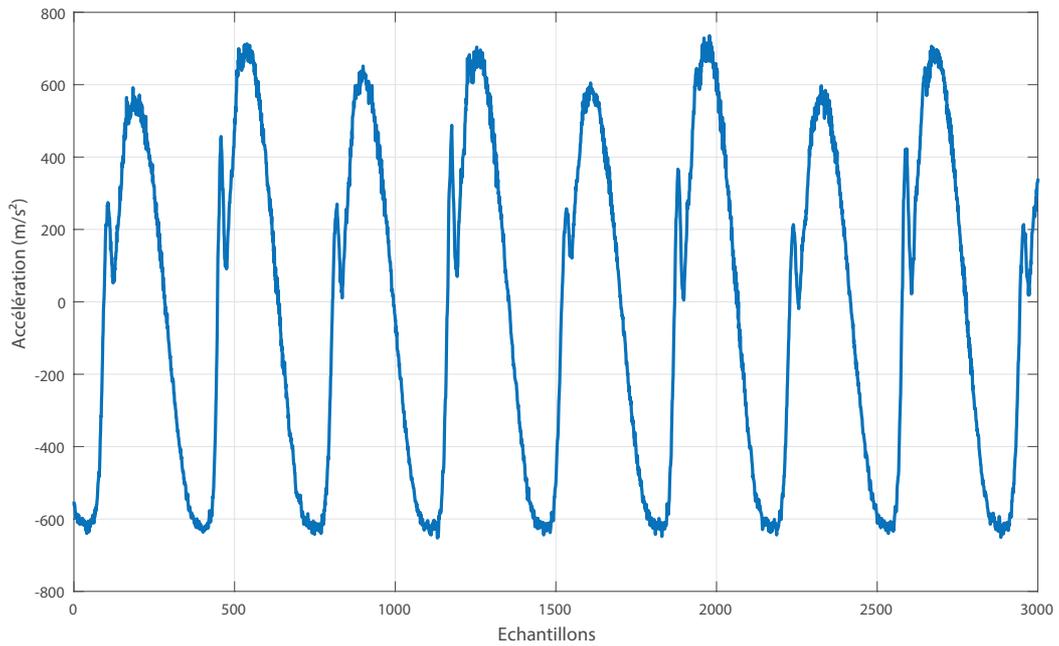
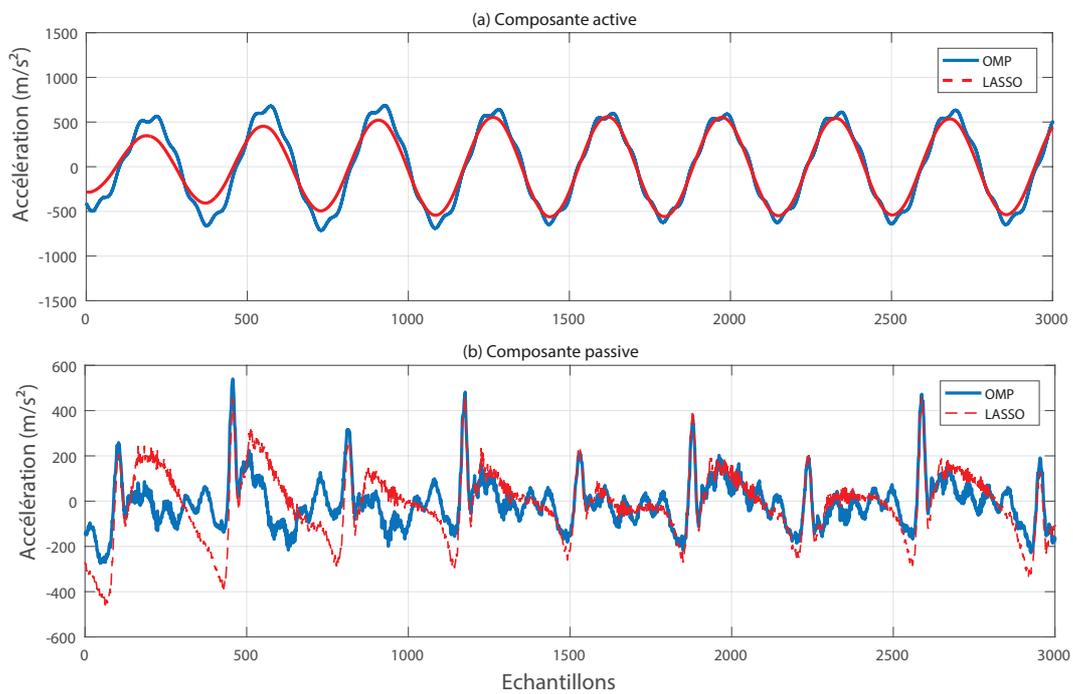
Nous cherchons à extraire la composante active de celle passive à partir du signal FRS illustré par la Fig.(4.22) en utilisant les deux algorithmes OMP et LASSO. Le critère d'arrêt est choisi de tel façon à ce que le seuil ϵ soit égal à 0.5% de l'énergie initiale du signal FRS.

Résultats

Les résultats de la séparation sont donnés par la Fig.(4.23). Nous constatons dans la Fig.(4.24) que les deux algorithmes extraient les coefficients parcimonieux les plus corrélés avec le dictionnaire DST durant les premières itérations. Nous remarquons également que l'algorithme OMP retire d'une itération à l'autre plus d'informations relatives à la composante active comparant au LASSO car l'énergie du résidu diminue dans le cas où nous utilisons l'OMP plus que dans le cas où nous utilisons le LASSO.

4.3 Conclusion

Dans ce chapitre, nous avons présenté dans un premier temps quelques généralités sur les roulements mécaniques et leur nature cyclo-stationnaire puis nous avons testé la méthode proposée (§3.1.4) dans le cas concret de diagnostic de roulements provenant d'un banc d'essai expérimental. Les résultats montrent que la méthode a séparé les signatures de chaque défaut localisé. Dans un second temps, nous avons proposé une solution basée sur l'analyse en composantes parcimonieuses utilisant la DST comme dictionnaire pour la séparation des composantes du signal FRS. La méthode a été

FIGURE 4.22 – *Signal FRS réel.*FIGURE 4.23 – *Résultats de la séparation des composantes de la FRS : (a) composante active, (b) composante passive.*

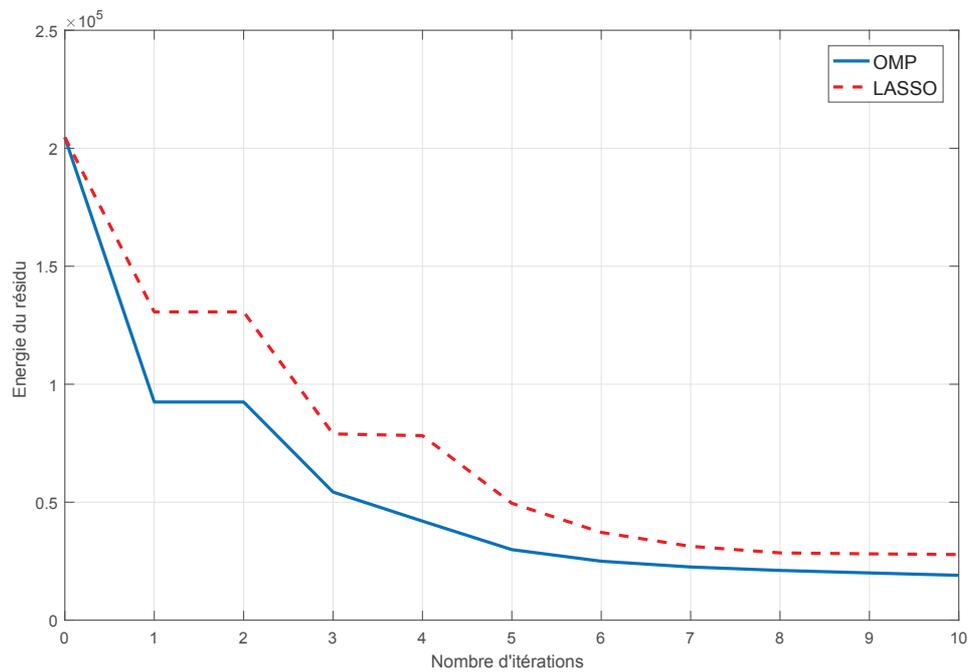


FIGURE 4.24 – Évolution de l'énergie du signal résiduel selon le nombre d'itérations.

concluante sur des signaux synthétiques et réels. La solution proposée pourrait être une piste de caractérisation de la fatigue ou l'amélioration des performances sportives chez un coureur de haut niveau en étudiant par exemple l'évolution en fonction du temps de la DSC de la partie passive lors d'une course à pied [134].

Conclusion générale

Conclusions

La cyclo-stationnarité est une propriété propre aux processus dont les statistiques évoluent cycliquement au cours du temps. Elle offre un cadre particulièrement adapté à l'analyse de nombreuses applications réelles comme les communications radios, les vibrations issus de machines tournantes, les signaux biomécaniques de marche ou course à pieds. Au cours de ce travail de recherche, nous nous sommes focalisés sur le problème de la séparation aveugle dans le cas des systèmes de mélanges linéaires de sources cyclo-stationnaires. Nous avons cherché à tirer parti de la cyclo-stationnarité afin d'extraire l'information utile présente dans les matrices de corrélation cyclique des observations. Cela nous a amené à développer différentes méthodes d'identification aveugle de mélange selon certains aprioris et hypothèses sur les fréquences cycliques des sources, la nature du mélange et la corrélation ou la décorrélation des sources.

Nous avons présenté en premier lieu le problème de la SAS dans sa globalité ainsi que la base théorique de la cyclo-stationnarité illustrée par des exemples. Nous avons ensuite passé en revue la littérature qui donne un compte rendu exhaustif des approches de SAS. Cette étude bibliographique a été accompagnée d'une discussion sur les avantages et inconvénients de chaque méthode.

En second lieu, nous avons proposé deux nouvelles méthodes de SAS. La première traite du cas où les sources partagent une ou plusieurs fréquences cycliques sans qu'elles soient a priori connues. La méthode s'appuie sur un nouvel algorithme de bloc-diagonalisation conjointe non unitaire d'un ensemble de matrices construit à partir des statistiques cycliques d'ordre deux. L'ensemble étant construit par le biais d'une nouvelle procédure de sélection automatique. L'avantage principal de cette approche est le fait qu'elle est directe dans le sens qu'elle ne requiert pas de blanchiment préalable des observations. Cette approche utilise un schéma d'optimisation qui se base

sur la méthode de BFGS avec une recherche linéaire du pas d'adaptation. L'idée principale de cette méthode est d'éviter de construire explicitement la Hessienne et de construire à la place une approximation de l'inverse de la dérivée seconde de la fonction à minimiser, en analysant les différents gradients successifs. La deuxième méthode traite le cas d'un mélange instantané de sources cyclo-stationnaires de fréquences cycliques inconnues et différentes. Une première étape permet de détecter les matrices de rang un. Le produit des matrices ainsi sélectionnées est décomposé en éléments propres. Cette décomposition permet de conserver les vecteurs pertinents selon une méthode de regroupement hiérarchique. Ce processus itératif nous conduit à terme à l'obtention de la matrice de mélange. Les deux solutions sont appliquées aux signaux de télécommunications et présentent de bonnes performances dans plusieurs contextes de rapport signal sur bruit.

En troisième lieu, nous avons abordé deux autres applications mécaniques et biomécaniques. Les signaux qui en découlent, ont un caractère cyclo-stationnaire du à la variabilité autour du cycle de fonctionnement de la machine dans le cas des signaux vibratoires et au fait que la force d'appui d'une jambe à l'autre et d'une foulée à l'autre n'est pas identique dans le cas des signaux biomécaniques. En guise de validation de la première méthode, nous l'avons appliquée à des mesures vibratoires acquises à partir d'un banc expérimental de roulements possédant deux défauts localisés différents. La méthode a montré son aptitude à extraire la signature de chaque défaut. Nous nous sommes intéressés également au traitement des signaux biomécaniques dans le but de séparer leurs composantes active (cyclo-stationnaire à l'ordre 1) et passive (cyclo-stationnaire à l'ordre 2). Pour ce faire, nous avons proposé une approche basée sur la parcimonie utilisant deux familles d'algorithmes, un de type *glouton* et l'autre reposant sur la relaxation à savoir l'OMP et Le LASSO. Nous employons la transformée en sinus discrète comme dictionnaire afin d'extraire la composante active du signal de réaction au sol et nous avons appliqué l'approche sur des signaux biomécaniques synthétiques et réels issus du CHU de Saint-Etienne. Cette étude peut aider à caractériser les états de fatigue musculaire afin d'améliorer les performances physiques chez le sportif de haut niveau.

Perspectives

Cette thèse ouvre de nombreuses perspectives sur de futures travaux que nous allons énumérer selon une liste non exhaustive.

1. Il nous semblerait intéressant d'étendre les résultats obtenus dans le cadre de la solution semi-analytique pour l'identification de la matrice de mélange de sources cyclo-stationnaires aux mélanges convolutifs.
2. Les algorithmes proposés dans cette thèse opèrent dans le domaine retard - fréquence cyclique. Nous pourrions nous intéresser à étudier les performances de ces algorithmes dans le domaine fréquence-fréquence cyclique.
3. Les résultats du troisième chapitre seront focalisés sur les mélanges de signaux synthétiques issus de système de communication. L'application à des signaux de radio fréquence réels ou d'autres signaux physiques pourrait être une perspective de travail afin de valider l'aptitude de ces algorithmes à séparer les sources.
4. Il serait pertinent d'appliquer les méthodes de séparation développées à d'autres configurations de roulements avec des fréquences de défauts communes ou distinctes d'abord sur des signaux synthétiques puis réels. Les résultats de séparation permettraient un meilleur diagnostic mécanique.
5. L'approche proposée pour séparer les composantes de la FRS, l'estimation de la composante cyclo-stationnaire à l'ordre 2 est conditionnée par une estimation robuste de la composante cyclo-stationnaire à l'ordre 1. Il serait pertinent de proposer une approche multi-dictionnaire qui estime les deux composantes simultanément.
6. Dans le domaine des machines tournantes, enfin, il y a toujours un couplage entre les roulements et les engrenages. Ces derniers se caractérisent par une cyclo-stationnarité d'ordre 1 qui masque la contribution des roulements. Il serait tout à fait envisageable d'appliquer l'analyse en composantes parcimonieuses afin d'extraire la contribution d'engrenage dans le cas d'un éventuel défaut de roulement. Dans ce cas, il serait possible de choisir la DST comme dictionnaire pour cette extraction.

Annexes



Sommaire

A.1	Matrice de corrélation des observations	109
A.1.1	Matrice de corrélation de mélanges de sources aléatoires . .	109
A.1.2	Matrice de corrélation de mélanges de sources déterministes	109
A.2	Relation entre la densité spectrale cyclique intégrée et l'analyse d'enveloppe	111
A.3	Expression des coefficients du polynôme d'ordre 4 impliqué dans le calcul du pas d'adaptation	112
A.3.1	Quelques propriétés et définitions importantes	112
A.3.2	Coefficients du polynôme d'ordre 4	113
A.4	Rappel sur la transformée de Hilbert	114
A.4.1	Définition	114
A.4.2	Propriétés	114

A.1 Matrice de corrélation des observations

A.1.1 Matrice de corrélation de mélanges de sources aléatoires

Sources aléatoires stationnaires

Dans le cas de la séparation des sources aléatoires stationnaires centrées, la matrice de corrélation des signaux observés est définie par (version asymptotique) :

$$\begin{aligned} \mathbf{C}_X(t, \tau) &= \mathbf{E}\{\mathbf{X}(t)\mathbf{X}^H(t + \tau)\} \\ &= \mathbf{G}\mathbf{C}_S(\tau)\mathbf{G}^H + \mathbf{C}_R(\tau). \end{aligned} \quad (\text{A.1})$$

Comme les bruits sont supposés stationnaires, gaussiens et centrés de variance σ_R^2 , nous obtenons en $\tau = 0$ la relation suivante :

$$\mathbf{C}_X(0) = \mathbf{G}\mathbf{C}_S(0)\mathbf{G}^H + \sigma_R^2\delta(\tau)\mathbf{I}_M. \quad (\text{A.2})$$

Sources aléatoires non-stationnaires

Lorsque les sources sont aléatoires non-stationnaires (dans ce cas là, la matrice de corrélation dépend à la fois du temps et du retard), nous introduisons alors la matrice de corrélation moyenne $\bar{\mathbf{C}}_S(\tau)$:

$$\bar{\mathbf{C}}_S(\tau) = \langle \mathbf{C}_S(t, \tau) \rangle, \quad (\text{A.3})$$

où $\langle \cdot \rangle$ désigne l'opérateur de moyennage temporel défini pour un vecteur $\mathbf{z}(t)$ comme :

$$\langle \mathbf{z}(t) \rangle = \lim_{T \rightarrow \infty} \frac{1}{T} \int_{-\frac{T}{2}}^{\frac{T}{2}} \mathbf{z}(t) dt, \quad (\text{A.4})$$

ainsi,

$$\bar{\mathbf{C}}_X(0) = \mathbf{G}\bar{\mathbf{C}}_S(0)\mathbf{G}^H + \sigma_R^2\delta(\tau)\mathbf{I}_M. \quad (\text{A.5})$$

A.1.2 Matrice de corrélation de mélanges de sources déterministes

Sources déterministes à puissance moyenne finie

Dans le cas où les sources sont déterministes centrées à puissance moyenne finie, nous faisons appel à l'opérateur de moyennage temporel. Ainsi, la matrice de corrélation du vecteur $\mathbf{S}(t)$ est définie par :

$$\mathbf{C}_S(\tau) = \langle \mathbf{S}(t)\mathbf{S}^H(t + \tau) \rangle, \quad (\text{A.6})$$

et par conséquent :

$$\bar{\mathbf{C}}_{\mathbf{X}}(0) = \mathbf{G}\mathbf{C}_{\mathbf{S}}(0)\mathbf{G}^H + \sigma_{\mathbf{R}}^2\delta(\tau)\mathbf{I}_{\mathbf{M}}. \quad (\text{A.7})$$

Sources déterministes à énergie finie

La définition précédente de la matrice de corrélation reste valable pour les sources déterministes à énergie finie, en changeant uniquement la définition de l'opérateur de moyennage par :

$$\langle \mathbf{z}(t) \rangle = \frac{1}{T} \int_{-\frac{T}{2}}^{\frac{T}{2}} \mathbf{z}(t) dt. \quad (\text{A.8})$$

A.2 Relation entre la densité spectrale cyclique intégrée et l'analyse d'enveloppe

Nous montrons dans cette partie le lien entre la DSCi et l'analyse spectrale de l'enveloppe du signal étudié $\mathbf{X}(t)$. Comme nous l'avons défini dans l'Eq. (1.37), la DSCi est donnée par :

$$\text{DSCi}_{\mathbf{X}}(\alpha) = \int_{\mathbb{R}} \text{DSC}_{\mathbf{X}}^{\alpha}(f) df. \quad (\text{A.9})$$

où $\text{DSC}_{\mathbf{X}}^{\alpha}(f)$ est la densité spectrale cyclique définie par :

$$\text{DSC}_{\mathbf{X}}^{\alpha}(f) = \int \int_{\mathbb{R}^2} \mathbf{C}_{\mathbf{X}}(t, \tau) e^{-2j\pi(f\tau + \alpha t)} dt d\tau. \quad (\text{A.10})$$

Par substitution avec la définition de l'Eq. (A.10), nous obtenons :

$$\text{DSCi}_{\mathbf{X}}(\alpha) = \int \int \int_{\mathbb{R}^3} \mathbf{C}_{\mathbf{X}}(t, \tau) e^{-2j\pi(f\tau + \alpha t)} dt d\tau df \quad (\text{A.11})$$

$$= \underbrace{\int_{\mathbb{R}} e^{-2j\pi f\tau} df}_{\text{fonction delta-Dirac centrée en } \tau = 0} \int \int_{\mathbb{R}^2} \mathbf{C}_{\mathbf{X}}(t, \tau) e^{-2j\pi\alpha t} dt d\tau \quad (\text{A.12})$$

$$= \int \int_{\mathbb{R}^2} \mathbf{C}_{\mathbf{X}}(t, \tau) \delta(\tau) d\tau e^{-2j\pi\alpha t} dt. \quad (\text{A.13})$$

$$(\text{A.14})$$

Il en résulte que la DSCi est équivalente à la transformée de Fourier sur le temps t de la fonction d'auto-corrélation à $\tau = 0$:

$$\text{DSCi}_{\mathbf{X}}(\alpha) = \int_{\mathbb{R}} \mathbf{C}_{\mathbf{X}}(t, 0) e^{-2j\pi\alpha t} dt \quad (\text{A.15})$$

$$= \int_{\mathbb{R}} \mathbf{E}[|\mathbf{X}(t)|^2] e^{-j2\pi\alpha t} dt. \quad (\text{A.16})$$

Or pour un retard τ nul, la fonction d'auto-corrélation équivaut au module carré de $\mathbf{X}(t)$. Nous en déduisons que la DSCi est également la transformée de Fourier du module au carré du signal étudié. Nous notons que ce résultat est uniquement valable pour les signaux à moyenne nulle. L'estimation et l'interprétation de la DSCi peuvent être erronées de par la présence de repliement de spectre, même lorsque le théorème de Nyquist-Shannon est respecté. Il convient donc de prendre certaines précautions pour la calculer.

A.3 Expression des coefficients du polynôme d'ordre 4 impliqué dans le calcul du pas d'adaptation

A.3.1 Quelques propriétés et définitions importantes

Considérons trois matrices carrées $\Theta_1, \Theta_2, \Theta_3$ de dimension $(M \times M)$ et deux matrices rectangulaires $\Theta_4 (M \times N)$ et $\Theta_5 (N \times M)$ et une matrice carrée Θ_6 de dimension $N \times N$, les propriétés suivantes introduites dans [115], [47], [78], [129] sont utilisées dans les développements à venir :

- P₁. $\|\Theta_1\|_F^2 = \text{tr} \{ \Theta_1^H \Theta_1 \}$
- P₂. $\|\text{OffBdiag}_{(n)}\{\Theta_1\}\|_F^2 = \text{tr} \{ \Theta_1^H \text{OffBdiag}_{(n)}\{\Theta_1\} \}$.
- P₃. $\text{tr} \{ \Theta_1 \} = \text{tr} \{ \Theta_1^T \}$.
- P₄. $\text{tr} \{ \Theta_1 + \Theta_2 \} = \text{tr} \{ \Theta_1 \} + \text{tr} \{ \Theta_2 \}$.
- P₅. $\text{tr} \{ \Theta_1 \Theta_2 \Theta_3 \} = \text{tr} \{ \Theta_3 \Theta_1 \Theta_2 \} = \text{tr} \{ \Theta_2 \Theta_3 \Theta_1 \}$.
- P₆. $\text{tr} \{ \Theta_1^H \Theta_2 \} = (\text{vec} \{ \Theta_1 \})^H \text{vec} \{ \Theta_2 \}$.
- P₇. $\text{vec} \{ \text{OffBdiag}_{(n)}\{\Theta_6\} \} = \mathbf{T}_{\text{Off}} \text{vec} \{ \Theta_6 \}$.
- P₈. $\text{vec} \{ \Theta_1 \Theta_2 \Theta_3 \} = \Theta_3^T \otimes \Theta_1 \text{vec} \{ \Theta_2 \}$.
- P₉. $(\Theta_1 \otimes \Theta_2)^H = \Theta_1^H \otimes \Theta_2^H$.

Où \otimes désigne le produit de Kronecker, $\text{vec} \{ \cdot \}$ est l'opérateur de vectorisation que s'il est appliqué à une matrice carrée $\mathbf{V} \in \mathbb{C}^{M \times M}$, alors il permet la concaténation de ses colonnes en un vecteur colonne de dimension $M^2 \times 1$ tel que :

$$\mathbf{V} = \begin{pmatrix} V_{11} & V_{12} & \dots & V_{1M} \\ V_{21} & \dots & \dots & \vdots \\ \vdots & \dots & \dots & \vdots \\ V_{M1} & V_{M2} & \dots & V_{MM} \end{pmatrix} \Rightarrow \text{vec}(\mathbf{V}) = \begin{pmatrix} V_{11} \\ \vdots \\ V_{M1} \\ V_{12} \\ \vdots \\ V_{M2} \\ \vdots \\ V_{MM} \end{pmatrix}, \quad (\text{A.17})$$

\mathbf{T}_{Boff} désigne la matrice de transformation de dimension $(N^2 \times N^2)$ définie à partir de la matrice de transformation \mathbf{T}_{Diag} qui a pour dimension $(N^2 \times N^2)$, comme suit :

$$\mathbf{T}_{\text{Diag}} = \text{diag} \{ \text{vec} \{ \text{Bdiag}_{(n)} \{ \mathbf{1}_N \} \} \}, \quad (\text{A.18})$$

$$\mathbf{T}_{\text{Off}} = \mathbf{I}_{N^2} - \mathbf{T}_{\text{Diag}}, \quad (\text{A.19})$$

où $\mathbf{1}_N$ est la matrice carrée de dimension $(N \times N)$ dont tous ses éléments sont égaux à 1, \mathbf{I}_{N^2} est la matrice identité de dimension $(N^2 \times N^2)$, $\text{Diag}_{(n)}\{\cdot\}$ et $\text{OffBdiag}_{(n)}\{\cdot\}$ sont deux opérateurs matriciels respectivement définis au niveau de (2.19) et (2.20).

A.3.2 Coefficients du polynôme d'ordre 4

En utilisant les propriétés \mathbf{P}_2 et \mathbf{P}_6 , la fonction de coût $\mathcal{F}_{\text{JBD}}(\mathbf{B})$ s'écrit de la façon suivante :

$$\mathcal{F}_{\text{JBD}}(\mathbf{B}) = \sum_{i=1}^K (\text{vec} \{ \mathbf{B} \mathbf{C}_{\mathbf{x}}^{\xi_i}(\tau_i) \mathbf{B}^H \})^H \text{vec} \{ \text{OffBdiag}_{(n)} \{ \mathbf{B} \mathbf{C}_{\mathbf{x}}^{\xi_i}(\tau_i) \mathbf{B}^H \} \}.$$

A partir des propriétés \mathbf{P}_7 , \mathbf{P}_8 et \mathbf{P}_9 , nous trouvons :

$$\mathcal{F}_{\text{JBD}}(\mathbf{B}) = \sum_{i=1}^K (\text{vec} \{ \mathbf{C}_{\mathbf{x}}^{\xi_i}(\tau_i) \})^H \mathbf{B}^{\otimes} \mathbf{B}^H \mathbf{T}_{\text{Boff}} \mathbf{B}^* \otimes \text{Bvec} \{ \mathbf{C}_{\mathbf{x}}^{\xi_i}(\tau_i) \}.$$

De la même façon, nous remplaçons \mathbf{B} par $\mathbf{B} - \alpha_{(k)} \mathbf{H} e^{-1} \nabla_a \mathcal{F}_{\text{JBD}}(\mathbf{B})$ et nous obtenons un polynôme d'ordre 4 en $\alpha_{(k)}$:

$$\begin{aligned} & \mathcal{F}_{\text{JBD}}(\mathbf{B} - \alpha_{(k)} \mathbf{H} e^{-1} \nabla_a \mathcal{F}_{\text{JBD}}(\mathbf{B})) && \text{(A.20)} \\ &= \alpha_{(k)}^4 \sum_{i=1}^K \text{tr} [\mathbf{V}_3 \text{OffBdiag}_{(n)} \{ \mathbf{V}_3^H \}] \\ &- \alpha_{(k)}^3 \sum_{i=1}^K \text{tr} [\mathbf{V}_0 \text{OffBdiag}_{(n)} \{ \mathbf{V}_3^H \}] - \text{tr} [\mathbf{V}_3 \mathbf{V}_2] \\ &+ \alpha_{(k)}^2 \sum_{i=1}^K \text{tr} [\mathbf{V}_1^H \text{OffBdiag}_{(n)} \{ \mathbf{V}_3^H \} + \mathbf{V}_3 \text{OffBdiag}_{(n)} \{ \mathbf{V}_1 \}] + \text{tr} [\mathbf{V}_0 \mathbf{V}_2] \\ &- \alpha_{(k)} \sum_{i=1}^K \text{tr} [\mathbf{V}_1^H \mathbf{V}_2] - \text{tr} [\mathbf{V}_0 \text{OffBdiag}_{(n)} \{ \mathbf{V}_1 \}] \\ &+ \sum_{i=1}^K \text{tr} [\mathbf{V}_1^H \text{OffBdiag}_{(n)} \{ \mathbf{V}_1 \}] \\ &= a\alpha_{(k)}^4 + b\alpha_{(k)}^3 + c\alpha_{(k)}^2 + d\alpha_{(k)} + e \end{aligned}$$

où les matrices \mathbf{V}_0 , \mathbf{V}_1 , \mathbf{V}_2 et \mathbf{V}_3 sont définies comme suit :

$$\begin{aligned} \mathbf{V}_0 &= \mathbf{C}_{\mathbf{x}}^{\xi_i}(\tau_i)^H \mathbf{\Gamma}_{(k)}^H \mathbf{B} + \mathbf{C}_{\mathbf{x}}^{\xi_i}(\tau_i)^H \mathbf{B}^H \mathbf{\Gamma}_{(k)}, \\ \mathbf{V}_1 &= \mathbf{B} \mathbf{C}_{\mathbf{x}}^{\xi_i}(\tau_i) \mathbf{B}^H, \\ \mathbf{V}_2 &= \text{OffBdiag}_{(n)} \{ \mathbf{\Gamma}_{(k)} \mathbf{C}_{\mathbf{x}}^{\xi_i}(\tau_i) \mathbf{B}^H \} + \text{OffBdiag}_{(n)} \{ \mathbf{B} \mathbf{C}_{\mathbf{x}}^{\xi_i}(\tau_i) \mathbf{\Gamma}_{(k)}^H \}, \\ \mathbf{V}_3 &= \mathbf{\Gamma}_{(k)} \mathbf{C}_{\mathbf{x}}^{\xi_i}(\tau_i)^H \mathbf{\Gamma}_{(k)}^H, \end{aligned}$$

avec $\mathbf{\Gamma}_{(k)} = \mathbf{H}_{(k)}^{-1} \nabla_a \mathcal{F}_{\text{JBD}}(\mathbf{B}_{(k)})$.

A.4 Rappel sur la transformée de Hilbert

A.4.1 Définition

La transformée de Hilbert d'un signal $x(t)$ s'écrit comme :

$$y(t) = \mathsf{H}\{x(t)\} = \frac{1}{\pi} \text{vp} \left(\int \frac{x(\tau)}{t - \tau} d\tau \right) \quad (\text{A.21})$$

où vp est la valeur principale de Cauchy. Si le signal x n'est pas intégrable sur un intervalle $[a, b]$, mais intégrable sur tout compact qui ne contient pas $c \in [a, b]$, alors nous pouvons écrire :

$$\text{vp} \left(\int_a^b x(t) dt \right) = \lim_{\epsilon \rightarrow 0} \left(\int_a^{c-\epsilon} x(t) dt + \int_{c+\epsilon}^b x(t) dt \right) \quad (\text{A.22})$$

A.4.2 Propriétés

Soit $x^{(a)}$ une fonction appartenant à l'ensemble des fonctions de carré sommable L^2 , $x^{(a)}$ est dit analytique si sa transformée de Fourier $X^{(a)}$ est nulle pour les fréquences négatives.

$$X^{(a)}(\omega) = 0 \quad \text{si } \omega < 0. \quad (\text{A.23})$$

Une fonction analytique (complexe par définition) est entièrement caractérisée par sa partie réelle ; Si on pose $X_r = \Re\{X^{(a)}\}$, alors

$$X_r(\omega) = \frac{1}{2} \left(X^{(a)}(\omega) + \bar{X}^{(a)}(-\omega) \right) \quad (\text{A.24})$$

et donc :

$$X^{(a)}(\omega) = \begin{cases} 2X_r(\omega) & \text{si } \omega \geq 0 \\ 0 & \text{si } \omega < 0 \end{cases}$$

Étant donné un signal $x(t)$ et sa transformée de Hilbert $y(t)$, nous pouvons calculer sa fonction analytique par :

$$x^{(a)}(t) = x(t) + jy(t). \quad (\text{A.25})$$

Nous pouvons écrire $x^{(a)}(t)$ sous forme polaire :

$$x^{(a)}(t) = \|x^{(a)}(t)\| e^{j\theta(t)}, \quad (\text{A.26})$$

avec $\|x^{(a)}(t)\| = \sqrt{x^2(t) + y^2(t)}$ et $\theta(t) = \arctan\left(\frac{y(t)}{x(t)}\right)$.

Liste des publications associée à la thèse

- [1] A. Brahmi, H. Ghennioui, C. Corbier, and F. Guillet. Non-unitary joint block diagonalization of matrices using bfgs method. In *Surveillance 8 : International Conference, Roanne, France, October 20-21, 2015, Proceedings*. (Cité page 3.)
- [2] A. Brahmi, H. Ghennioui, M. Lahbabi, C. Corbier, and F. Guillet. Blind identification of fir mimo systems driven by cyclostationary inputs in wireless communications using joint block diagonalization. In *Wireless Networks and Mobile Communications (WINCOM), Fez, 2016 International Conference on*, pages 17–22. IEEE. doi : 10.1109/WINCOM.2016.7777184 (Cité page 3.)
- [3] A. Brahmi, H. Ghennioui, C. Corbier, M. Lahbabi, and F. Guillet. *Blind Separation of Cyclostationary Sources with Common Cyclic Frequencies*, In : Tichavský P., Babaie-Zadeh M., Michel O., Thirion-Moreau N.. (eds) Latent Variable Analysis and Signal Separation. LVA/ICA 2017. Lecture Notes in Computer Science, vol 10169, pages 449–458. Springer, Cham. doi=10.1007/978-3-319-53547-0_42 (Cité page 3.)
- [4] A. Brahmi, H. Ghennioui, C. Corbier, F. Guillet, et al. Blind separation of cyclostationary sources sharing common cyclic frequencies using joint diagonalization algorithm. *Mathematical Problems in Engineering*, vol. 2017, Article ID 2546838, 9 pages, 2017. doi :10.1155/2017/2546838 (Cité page 3.)
- [5] A. Brahmi, H. Ghennioui, C. Corbier, F. Guillet, and M. Lahbabi. A semi-analytical method for blind separation of cyclostationary sources in digital communications. In *Advanced Technologies for Signal and Image Processing (ATSIP), Fez, 2017 International Conference on*, pages 1–6. IEEE. doi : 10.1109/ATSIP.2017.8075607 (Cité page 3.)
- [6] A. Brahmi, H. Ghennioui, C. Corbier, F. Guillet, and M. Lahbabi. Une nouvelle solution pour l'identification aveugle de mélanges de sources cyclostationnaires appliquée aux signaux de télécommunications. In

26 °Colloque GRETSI, Juan-les-Pins, France, 5-8 septembre 2017. GRETSI, Groupe d'Etudes du Traitement du Signal et des Images, 2017. (Cité page 3.)

- [7] A. Had, A. Brahmi, K. Sabri, H. Ghennioui, and M. El badaoui. Séparation de composantes des signaux biomécaniques par une approche parcimonieuse. In *26 °Colloque GRETSI, Juan-les-Pins, France, 5-8 septembre 2017.* GRETSI, Groupe d'Etudes du Traitement du Signal et des Images, 2017. (Cité page 4.)

Bibliographie

- [1] Foot kinematics for two foot strikes. <http://barefootrunning.fas.harvard.edu/4BiomechanicsofFootStrike.html>. Accessed : 2017-09-05. (Cité page 92.)
- [2] K. Abed-Meraim, J. H. J. Manton, Y. Xiang, J. H. J. Manton, and Y. Hua. Blind source separation using second-order cyclostationary statistics. *IEEE - Transactions on Signal Processing*, 49(4) :694–701, apr 2001. (Cité page 42.)
- [3] B. Agee, S. Schell, W. Gardner, and C. Science. Spectral self-coherence restoral : A new approach to blind adaptive signal extraction using antenna arrays. *Proceedings of the IEEE*, 78(4) :753–767, 1990. (Cité page 42.)
- [4] Y. Ainhoren, S. Engelberg, and S. Friedman. The cocktail party problem. *IEEE Instrumentation and Measurement Magazine*, 11(3) :44–48, 2008. (Cité page 2.)
- [5] A. Aïssa-El-Bey, N. Linh-Trung, K. Abed-Meraim, A. Belouchrani, Y. Grenier, A. Aïssa-El-Bey, N. Linh-Trung, K. Abed-Meraim, A. Belouchrani, and Y. Grenier. Underdetermined blind separation of non-disjoint sources in the time-frequency domain. *IEEE Transactions on Signal Processing*, 55(3) :897–907, 2007. (Cité page 47.)
- [6] L. Albera. *Identification autodidacte de m{é}langes potentiellement sous-d{é}termin{é}s*. PhD thesis, Ph. D. dissertation, University of Nice-Sophia Antipolis, 2003. (Cité page 40.)
- [7] L. Albera, P. Comon, P. Chevalier, and A. Ferréol. Blind identification of underdetermined mixtures based on the hexacovariance. In *Acoustics, Speech, and Signal Processing, 2004. Proceedings.(ICASSP'04). IEEE International Conference on*, volume 2, pages ii—29. IEEE, 2004. (Cité page 40.)
- [8] L. Albera, A. Ferréol, P. Chevalier, and P. Comon. ICAR, un algorithme d'ICA {à} convergence rapide, robuste au bruit. In *19^o Colloque sur le traitement du signal et des images, FRA, 2003*. GRETSI, Groupe d'Etudes du Traitement du Signal et des Images, 2003. (Cité page 40.)

- [9] L. Albera, A. Ferréol, P. Chevalier, and P. Comon. ICAR : Independent component analysis using redundancies. In *Circuits and Systems, 2004. ISCAS'04. Proceedings of the 2004 International Symposium on*, volume 5, pages V—V. IEEE, 2004. (Cité page 40.)
- [10] L. Albera, A. Ferréol, P. Comon, and P. Chevalier. Sixth order blind identification of underdetermined mixtures (BIRTH) of sources. In *ICA*, volume 3, pages 1–4, 2003. (Cité page 40.)
- [11] S.-I. Amari. Natural gradient works efficiently in learning. *Neural computation*, 10(2) :251–276, 1998. (Cité page 41.)
- [12] J. Antoni, F. Bonnardot, A. Raad, and M. El Badaoui. Cyclostationary modelling of rotating machine vibration signals. *Mechanical systems and signal processing*, 18(6) :1285–1314, 2004. (Cité page 81.)
- [13] J. Antoni, F. Guillet, M. El Badaoui, and F. Bonnardot. Blind separation of convolved cyclostationary processes. *Signal Processing*, 85(1) :51–66, jan 2005. (Cité page 4.)
- [14] J. Antoni and R. Randall. On the use of the cyclic power spectrum in rolling element bearings diagnostics. *Journal of Sound and Vibration*, 281(1) :463–468, 2005. (Cité page 84.)
- [15] H. Asari, B. A. Pearlmutter, and A. M. Zador. Sparse Representations for the Cocktail Party Problem. *Journal of Neuroscience*, 26(28) :7477–7490, 2006. (Cité page 2.)
- [16] M. Author, J. Sherman, and W. J. Morrison. Adjustment of an Inverse Matrix Corresponding to a Change in One Element of a Given Matrix. *Source : The Annals of Mathematical Statistics*, 21(1) :124–127, 1950. (Cité page 60.)
- [17] M. S. Back, M. J. Kim, Y. H. You, and H. K. Song. Semi-blind channel estimation and PAR reduction for MIMO-OFDM system with multiple antennas. *IEEE Transactions on Broadcasting*, 50(4) :414–424, 2004. (Cité page 4.)
- [18] G. Bedoya, C. Jutten, S. Bermejo, and J. Cabestany. Improving semiconductor-based chemical sensor arrays using advanced algorithms for blind source separation. In *ISA/IEEE Sensors for Industry Conference, 2004. Proceedings the*, pages 149–154. IEEE. (Cité page 9.)
- [19] A. J. Bell and T. J. Sejnowski. An information-maximization approach to blind separation and blind deconvolution. *Neural computation*, 7(6) :1129–1159, 1995. (Cité page 41.)

- [20] A. J. Bell and T. J. Sejnowski. The "independent components" of natural scenes are edge filters. *Vision research*, 37(23) :3327–3338, 1997. (Cité page 41.)
- [21] A. Belli, P. Bui, A. Berger, A. Geysant, and J.-R. Lacour. A treadmill ergometer for three-dimensional ground reaction forces measurement during walking. *Journal of Biomechanics*, 34(1) :105–112, jan 2001. (Cité page 99.)
- [22] A. Belouchrani, K. Abed-Meraim, J.-F. Cardoso, and E. Moulines. A blind source separation technique using second-order statistics. *IEEE Transactions on Signal Processing*, 45(2) :434–444, 1997. (Cité pages 34, 38 et 42.)
- [23] A. Belouchrani and M. G. Amin. Blind source separation based on time-frequency signal representations. *IEEE Transactions on Signal Processing*, 46(11) :2888–2897, 1998. (Cité page 39.)
- [24] A. Belouchrani, M. G. Amin, and K. Abed-Meraim. Direction finding in correlated noise fields based on joint block-diagonalization of spatio-temporal correlation matrices. *IEEE Signal Processing Letters*, 4(9) :266–268, 1997. (Cité page 35.)
- [25] A. Belouchrani and A. Cichocki. Robust whitening procedure in blind source separation context. *Electronics Letters*, 36(24) :2050–2051(1), 2000. (Cité page 39.)
- [26] E. Bingham and A. Hyvärinen. A fast fixed-point algorithm for independent component analysis of complex valued signals. *International journal of neural systems*, 10(01) :1–8, 2000. (Cité page 41.)
- [27] B. Boashash. *Time-frequency signal analysis and processing : a comprehensive reference*. Academic Press, 2015. (Cité page 40.)
- [28] P. Bofill and M. Zibulevsky. Underdetermined blind source separation using sparse representations. *Signal Processing*, 81(11) :2353–2362, nov 2001. (Cité page 47.)
- [29] M. Boham, M. DeBeliso, C. Harris, R. Pfeiffer, J. McChesney, and J. M. Berning. The effects of functional fatigue on ground reaction forces of a jump, land, and cut task. *International Journal of Science and Engineering Investigations*, 2013. (Cité page 92.)
- [30] F. Bonnardot. *Comparaison entre les analyses angulaire et temporelle des signaux vibratoires de machines tournantes. Etude du concept de cyclostationnarité floue*. PhD thesis, Institut National Polytechnique de Grenoble-INPG, 2004. (Cité page 81.)

- [31] F. Bonnardot and M. El Badaoui. Etude de la fatigue d'un coureur, de l'instrumentation à l'analyse vibratoire. In *10ème Congrès Français d'Acoustique*, 2010. (Cité page 94.)
- [32] F. Bonnardot, R. B. Randall, and F. Guillet. Extraction of second-order cyclostationary sources - Application to vibration analysis. *Mechanical Systems and Signal Processing*, 19(6) :1230–1244, nov 2005. (Cité page 4.)
- [33] N. Bouguerriou. *Caractérisation et séparation de sources de nature cyclostationnaire : application aux roulements*. PhD thesis, Orléans, 2006. (Cité page 81.)
- [34] N. Bouguerriou, M. Haritopoulos, C. Capdessus, and L. Allam. Novel cyclostationarity-based blind source separation algorithm using second order statistical properties : Theory and application to the bearing defect diagnosis. *Mechanical Systems and Signal Processing*, 19(6) :1260–1281, nov 2005. (Cité page 43.)
- [35] H. Bousbia-Salah, A. Belouchrani, and K. Abed-Meraim. Blind separation of convolutive mixtures using joint block diagonalization. In *Signal Processing and its Applications, Sixth International, Symposium on. 2001*, volume 1, pages 13–16. IEEE, 2001. (Cité page 8.)
- [36] H. Bousbia-Salah, A. Belouchrani, and K. Abed-Meraim. Jacobi-like algorithm for blind signal separation of convolutive mixtures. *Electronics Letters*, 37(16) :1049, 2001. (Cité page 8.)
- [37] R. Boustany and J. Antoni. A subspace method for the blind extraction of a cyclostationary source : Application to rolling element bearing diagnostics. *Mechanical Systems and Signal Processing*, 19(6) :1245–1259, nov 2005. (Cité page 43.)
- [38] R. Boustany and J. Antoni. Blind extraction of a cyclostationary signal using reduced-rank cyclic regression-A unifying approach. *Mechanical Systems and Signal Processing*, 22(3) :520–541, apr 2008. (Cité pages 4 et 43.)
- [39] I. Bradaric, A. Petropulu, and K. Diamantaras. Blind MIMO FIR channel identification based on second-order spectra correlations. *IEEE Transactions on Signal Processing*, 51(6) :1668–1674, jun 2003. (Cité page 43.)
- [40] I. Bradaric, A. Petropulu, and K. Diamantaras. On blind identifiability of FIR-MIMO systems with cyclostationary inputs using second order statistics. *IEEE Transactions on Signal Processing*, 51(2) :434–441, feb 2003. (Cité page 43.)

- [41] A. Brahmi, H. Ghennioui, C. Corbier, and F. Guillet. Non-unitary joint block diagonalization of matrices using bfgs method. In *Surveillance 8 : International Conference, Roanne, France, October 20-21, 2015, Proceedings*. (Cité page 3.)
- [42] A. Brahmi, H. Ghennioui, C. Corbier, F. Guillet, et al. Blind separation of cyclostationary sources sharing common cyclic frequencies using joint diagonalization algorithm. *Mathematical Problems in Engineering*, 2017, 2017. (Cité page 3.)
- [43] A. Brahmi, H. Ghennioui, C. Corbier, F. Guillet, and M. Lahbabi. A semi-analytical method for blind separation of cyclostationary sources in digital communications. In *2016 3rd International Conference on Advanced Technologies for Signal and Image Processing (ATSIP)*. IEEE, May 2017. (Cité page 3.)
- [44] A. Brahmi, H. Ghennioui, C. Corbier, F. Guillet, and M. Lahbabi. Une nouvelle solution pour l'identification aveugle de mélanges de sources cyclostationnaires appliquée aux signaux de télécommunications. In *26° Colloque GRETSI, Juan-les-Pins, France, 5-8 septembre 2017*. GRETSI, Groupe d'Etudes du Traitement du Signal et des Images, 2017. (Cité page 3.)
- [45] A. Brahmi, H. Ghennioui, C. Corbier, M. Lahbabi, and F. Guillet. *Blind Separation of Cyclostationary Sources with Common Cyclic Frequencies*, pages 449–458. Springer International Publishing, Cham, 2017. (Cité page 3.)
- [46] A. Brahmi, H. Ghennioui, M. Lahbabi, C. Corbier, and F. Guillet. Blind identification of fir mimo systems driven by cyclostationary inputs in wireless communications using joint block diagonalization. In *Wireless Networks and Mobile Communications (WINCOM), 2016 International Conference on*, pages 17–22. IEEE, 2016. (Cité page 3.)
- [47] J. Brewer. Kronecker products and matrix calculus in system theory. *IEEE Transactions on circuits and systems*, 25(9) :772–781, 1978. (Cité page 112.)
- [48] J. C. Brown. Non-Negative Matrix Factorization for Polyphonic Music Transcription. In *IEEE Workshop on Applications of Signal Processing to Audio and Acoustics*, number 3, pages 177–180, 2003. (Cité page 49.)
- [49] R. H. Byrd and J. Nocedal. A tool for the analysis of quasi-Newton methods with application to unconstrained minimization. *SIAM Journal on Numerical Analysis*, 26(3) :727–739, 1989. (Cité page 61.)

- [50] C. Capdessus, A. K. Nandi, and N. Bouguerriou. A New Source Extraction Algorithm for Cyclostationary Sources. In *LVA/ICA - Independent Component Analysis and Blind Signal Separation*, volume 4666, pages 145–151, Berlin, Heidelberg, 2007. Springer Berlin Heidelberg. (Cité page 43.)
- [51] J. Cardoso and A. Souloumiac. Blind beamforming for non-gaussian signals. *IEE Proceedings F Radar and Signal Processing*, 140(6) :362, 1993. (Cité pages 34, 36 et 40.)
- [52] J. F. Cardoso. Blind identification of independent signals. In *Workshop on Higher-Order Spectral Analysis*, pages 569–572, 1989. (Cité page 40.)
- [53] J.-F. Cardoso. Super-symmetric decomposition of the fourth-order cumulant tensor. Blind identification of more sources than sensors. In *Acoustics, Speech, and Signal Processing, 1991. ICASSP-91., 1991 International Conference on*, pages 3109–3112. IEEE, 1991. (Cité page 40.)
- [54] J.-F. Cardoso and A. Souloumiac. Jacobi Angles for Simultaneous Diagonalization. *SIAM Journal on Matrix Analysis and Applications*, 17(1) :161–164, 1996. (Cité page 36.)
- [55] M. Castella. *Séparation de sources non linéaires dans le cas des mélanges convolutifs*. PhD thesis, Université de Marne la Vallée, 2004. (Cité page 41.)
- [56] M. Castella, A. Chevreuil, and J.-C. Pesquet. Séparation aveugle d'un mélange convolutif de sources non linéaires par une approche hiérarchique. In *19th Colloque sur le traitement du signal et des images, FRA, 2003*. GRETSI, Groupe d'Etudes du Traitement du Signal et des Images, 2003. (Cité page 41.)
- [57] M. Castella and J.-C. Pesquet. Méthodes séquentielles de séparation aveugle de sources en présence de bruit. In *20th Colloque sur le traitement du signal et des images, FRA, 2005*. GRETSI, Groupe d'Etudes du Traitement du Signal et des Images, 2005. (Cité page 41.)
- [58] G. Chabriel and J. Barrere. A direct algorithm for nonorthogonal approximate joint diagonalization. *IEEE Transactions on Signal Processing*, 60(1) :39–47, 2012. (Cité page 34.)
- [59] G. Chabriel, M. Kleinsteuber, E. Moreau, H. Shen, P. Tichavsky, and A. Yeredor. Joint Matrices Decompositions and Blind Source Separation : A survey of methods, identification, and applications. *IEEE Signal Processing Magazine*, 31(3) :34–43, may 2014. (Cité page 31.)

- [60] G. Chabriel, M. Kleinsteuber, E. Moreau, H. Shen, P. Tichavsky, and A. Yeredor. Joint matrices decompositions and blind source separation : A survey of methods, identification, and applications. *IEEE Signal Processing Magazine*, 31(3) :34–43, 2014. (Cité page 34.)
- [61] C. Chaouchi. Méthodes de séparation aveugle de sources non linéaires, étude du modèle quadratique 2×2 . jun 2011. (Cité page 9.)
- [62] S. Chen, S. A. Billings, and W. Luo. Orthogonal least squares methods and their application to non-linear system identification. *International Journal of Control*, 50(5) :1873–1896, 1989. (Cité page 46.)
- [63] S. S. Chen, D. L. Donoho, and M. A. Saunders. Atomic Decomposition by Basis Pursuit. *SIAM Journal on Scientific Computing*, 20(1) :33–61, 1998. (Cité page 46.)
- [64] G.-H. Cheng, S.-M. Li, and E. Moreau. New jacobi-like algorithms for non-orthogonal joint diagonalization of hermitian matrices. *Signal Processing*, 128 :440 – 448, 2016. (Cité page 35.)
- [65] O. Cherrak. *Blind sources separation : application to marine surveillance*. Theses, Université de Toulon, Mar. 2016. (Cité page 35.)
- [66] O. Cherrak, H. Ghennioui, E.-H. Abarkan, and N. Thirion-Moreau. Non-unitary joint block diagonalization of matrices using a levenberg-marquardt algorithm. In *Signal Processing Conference (EUSIPCO), 2013 Proceedings of the 21st European*, pages 1–5. IEEE, 2013. (Cité page 35.)
- [67] O. Cherrak, H. Ghennioui, N. Thirion-Moreau, et al. Preconditioned optimization algorithms solving the problem of the non unitary joint block diagonalization : application to blind separation of convolutive mixtures. *Multidimensional Systems and Signal Processing*, pages 1–24, 2017. (Cité page 35.)
- [68] N. A. CheViet, M. El Badaoui, A. Belouchrani, and F. Guillet. Blind separation of cyclostationary sources using non-orthogonal approximate joint diagonalization. In *SAM 2008 - 5th IEEE Sensor Array and Multichannel Signal Processing Workshop*, pages 492–495, 2008. (Cité page 43.)
- [69] N. A. Cheviet, M. El Badaoui, A. Belouchrani, F. Guillet, M. E. Badaoui, A. Belouchrani, and F. Guillet. Blind separation of cyclostationary signals. In *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, volume 5441, pages 25–33. Springer, Berlin, Heidelberg, 2009. (Cité page 43.)

- [70] C. Y. Chi and C. H. Chen. Cumulant-based inverse filter criteria for MIMO blind deconvolution : Properties, algorithms, and application to DS/CDMA systems in multipath. *IEEE Transactions on Signal Processing*, 49(7) :1282–1299, 2001. (Cité page 4.)
- [71] S. Choi. Blind separation of second-order nonstationary and temporally colored sources. ... *Signal Processing, 2001. ...*, pages 6–9, 2001. (Cité page 39.)
- [72] A. Cichocki, R. Zdunek, and S.-I. Amari. New Algorithms for Non-Negative Matrix Factorization in Applications to Blind Source Separation. *2006 IEEE International Conference on Acoustics Speed and Signal Processing*, pages V–621–V–624, 2006. (Cité page 49.)
- [73] A. Cichocki, R. Zdunek, A. H. Phan, and S. I. Amari. *Nonnegative Matrix and Tensor Factorizations : Applications to Exploratory Multi-Way Data Analysis and Blind Source Separation*. 2009. (Cité page 49.)
- [74] P. Comon. Independent component analysis, A new concept? *Signal Processing*, 36(3) :287–314, 1994. (Cité pages 28, 34 et 36.)
- [75] P. Comon. Contrasts for multichannel blind deconvolution. *IEEE Signal Processing Letters*, 3(7) :209–211, 1996. (Cité page 37.)
- [76] P. Comon, C. Jutten, and J. Herault. Blind separation of sources, Part II : Problems statement. *Signal processing*, 24(1) :11–20, 1991. (Cité page 28.)
- [77] P. Comon and E. Moreau. Improved contrast dedicated to blind separation in communications. In *Acoustics, Speech, and Signal Processing, 1997. ICASSP-97., 1997 IEEE International Conference on*, volume 5, pages 3453–3456. IEEE, 1997. (Cité page 37.)
- [78] M. Computation. The johns hopkins university press. *Baltimore and London*, 1989. (Cité page 112.)
- [79] L. Dapena, Adriana and Iglesia, Daniel and Castedo. Blind separation of cyclostationary signals. In *Signal Processing Conference (EUSIPCO 1998), 9th European*, Rhodes, Greece, 1998. IEEE. (Cité page 42.)
- [80] L. De Lathauwer, B. De Moor, J. Vandewalle, and J. F. Cardoso. Independent component analysis of largely underdetermined mixtures. In *Proc. 4th Int. Symp. on Independent Component Analysis and Blind Signal Separation (ICA 2003)*, pages 29–34, 2003. (Cité page 40.)
- [81] L. De Lathauwer, C. Févotte, B. De Moor, and J. Vandewalle. Jacobi algorithm for block diagonalization in blind identification. In *Proc. Symp. Inform. Theory in the Benelux*, pages 155–162. (Cité page 35.)

- [82] N. Delfosse and P. Loubaton. Adaptive blind separation of independent sources : a deflation approach. *Signal processing*, 45(1) :59–83, 1995. (Cité page 41.)
- [83] Y. Deville. Panorama des applications biomédicales des méthodes de séparation aveugle de sources. *GRETSI*, pages 2–5, 2003. (Cité page 5.)
- [84] I. S. Dhillon and S. Sra. Generalized Nonnegative Matrix Approximations with Bregman Divergences. *Advances in neural information processing systems*, 19 :283–290, 2005. (Cité page 49.)
- [85] P. Dinh-Tuan and D. T. Pham. Blind Separation of Cyclostationary Sources Using Joint Block Approximate Diagonalization. *IVA/ICA - Independent Component Analysis and Blind Signal Separation*, 4666 :244–251, 2007. (Cité pages 34 et 43.)
- [86] K. Drakakis, S. Rickard, R. De Fréin, and A. Cichocki. Analysis of Financial Data Using Non-Negative Matrix Factorization. *International Mathematical Forum*, 3(38) :1853–1870, 2008. (Cité page 49.)
- [87] B. Efron, T. Hastie, I. Johnstone, R. Tibshirani, and Others. Least angle regression. *The Annals of Statistics*, 32(2) :407–499, 2004. (Cité page 96.)
- [88] G. Elia, M. Cocconcelli, E. Mucchi, and G. Dalpiaz. Combining blind separation and cyclostationary techniques for monitoring distributed wear in gearbox rolling bearings. *Proceedings of the Institution of Mechanical Engineers, Part C : Journal of Mechanical Engineering Science*, 231(6) :1113–1128, mar 2017. (Cité page 44.)
- [89] A. F., L. Albera, and P. Chevalier. Fourth-order blind identification of underdetermined mixtures of sources (FOBIUM). *IEEE Transactions on Signal Processing*, 53(5) :1640–1653, 2005. (Cité page 40.)
- [90] E. H. Fadaili and Others. *D{é}compositions Matricielles Conjointes et S{é}paration Aveugle de Sources*. PhD thesis, Universit{é} du Sud Toulon-Var. UFR de Sciences et Techniques, 2006. (Cité page 40.)
- [91] A. Ferreol and P. Chevalier. On the behavior of current second and higher order blind source separation methods for cyclostationary sources. *IEEE Transactions on Signal Processing*, 48(6) :1712–1725, jun 2000. (Cité page 42.)
- [92] A. Ferreol, P. Chevalier, A. F{é}rreol, P. Chevalier, and A. Ferreol. Higher order blind source separation using the cyclostationarity property of the signals. In *1997 IEEE International Conference on Acoustics, Speech,*

- and Signal Processing*, volume 5, pages 4061–4064. IEEE Comput. Soc. Press, 1997. (Cité page 42.)
- [93] C. Févotte. Approche temps-fréquence pour la séparation aveugle de sources non-stationnaires. oct 2003. (Cité pages 8 et 39.)
- [94] C. Févotte and C. Doncarli. A unified presentation of blind separation methods for convolutive mixtures using block-diagonalization. *Proc. ICA 2003*, 2003. (Cité page 8.)
- [95] C. Févotte, C. Doncarli, C. Févotte, and C. Doncarli. Two contributions to blind source separation using time-frequency distributions. *IEEE Signal Processing Letters*, 11(3) :386–389, 2004. (Cité page 39.)
- [96] R. Fletcher. *Practical Methods of Optimization*, volume 1 : Unconst. 1987. (Cité page 59.)
- [97] J.-L. Gaeta, M., Lacoume. Sources Separation without a priori Knowledge : the Maximum Likelihood Solution. In *Fifth European Signal Processing Conference*, Barcelona, Spain, 1990. (Cité page 40.)
- [98] W. A. Gardner. Cyclostationarity in communications and signal processing. Technical report, STATISTICAL SIGNAL PROCESSING INC YOUNTVILLE CA, 1994. (Cité page 21.)
- [99] W. A. Gardner, A. Napolitano, and L. Paura. Cyclostationarity : Half a century of research. *Signal Processing*, 86(4) :639–697, 2006. (Cité page 18.)
- [100] P. Georgiev, F. Theis, and A. Cichocki. Sparse component analysis and blind source separation of underdetermined mixtures. *IEEE Transactions on Neural Networks*, 16(4) :992–996, 2005. (Cité page 47.)
- [101] F. Ghaderi, B. Makkiabadi, J. G. McWhirter, and S. Sanei. Blind source extraction of cyclostationary sources with common cyclic frequencies. In *ICASSP, IEEE International Conference on Acoustics, Speech and Signal Processing - Proceedings*, pages 4146–4149. IEEE, 2010. (Cité page 44.)
- [102] H. Ghennioui. *Blind separation of linear convolutive mixtures of correlated sources*. Theses, Université du Sud Toulon Var, July 2008. (Cité pages 17, 33, 35 et 57.)
- [103] H. Ghennioui, E. M. Fadaili, N. Thirion-Moreau, A. Adib, and E. Moreau. Une nouvelle méthode sans blanchiment pour la séparation aveugle de mélanges convolutifs de sources basée sur l'utilisation de distributions temps-fréquence spatiales quadratiques. In *21^e Colloque GRETSI, Troyes, France, 11-14 septembre 2007*. GRETSI, Groupe d'Etudes du Traitement du Signal et des Images, 2007. (Cité page 57.)

- [104] H. Ghennioui, E. M. Fadaili, N. N. Thirion-Moreau, A. Adib, and E. Moreau. A nonunitary joint block diagonalization algorithm for blind separation of convolutive mixtures of sources. *IEEE Signal Processing Letters*, 14(11) :860–863, 2007. (Cité pages 17 et 35.)
- [105] H. Ghennioui, N. Thirion-Moreau, E. Moreau, and D. Aboutajdine. Gradient-based joint block diagonalization algorithms : Application to blind separation of FIR convolutive mixtures. *Signal Processing*, 90(6) :1836–1849, 2010. (Cité pages 35 et 60.)
- [106] H. Ghennioui, N. Thirion-Moreau, E. Moreau, D. Aboutajdine, and A. Adib. A novel approach based on non-unitary joint block-diagonalization for the blind MIMO equalization of cyclo-stationary signals. In *European Signal Processing Conference*, pages 1–5, 2008. (Cité pages 43 et 65.)
- [107] H. Ghennioui, N. Thirion-Moreau, E. Moreau, A. Adib, and D. Aboutajdine. Non unitary joint block diagonalization of complex matrices using a gradient approach. *Independent Component Analysis and Signal Separation*, pages 201–208, 2007. (Cité page 35.)
- [108] L. Giulieri. *S{é}paration aveugle de sources bas{é}e sur l'utilisation des transform{é}es spatiales quadratiques*. PhD thesis, Université du Sud Toulon-Var. UFR de Sciences et Techniques, 2003. (Cité pages 17 et 40.)
- [109] L. Giulieri, H. Ghennioui, N. Thirion-Moreau, and E. Moreau. Nonorthogonal joint diagonalization of spatial quadratic time-frequency matrices for source separation. *IEEE Signal Processing Letters*, 12(5) :415–418, may 2005. (Cité pages 40, 57 et 72.)
- [110] R. A. Guyer Jr. *Rolling bearings handbook and troubleshooting guide*. CRC Press, 1996. (Cité page 79.)
- [111] A. Had, A. Brahmi, K. Sabri, H. Ghennioui, and M. El badaoui. Séparation de composantes des signaux biomécaniques par une approche parcimonieuse. In *26 ° Colloque GRETSI, Juan-les-Pins, France, 5-8 septembre 2017*. GRETSI, Groupe d'Etudes du Traitement du Signal et des Images, 2017. (Cité page 4.)
- [112] F. Harroy and J.-L. Lacoume. Maximum likelihood estimators and Cramer-Rao bounds in source separation. *Signal processing*, 55(2) :167–177, 1996. (Cité page 41.)
- [113] Z. He, S. Xie, S. Ding, and A. Cichocki. Convolutive blind source separation in the frequency domain based on sparse representation. *IEEE*

- Transactions on Audio, Speech and Language Processing*, 15(5) :1551–1563, 2007. (Cité page 47.)
- [114] J. Héroult, C. Jutten, and B. Ans. Détection de grandeurs primitives dans un message composite par une architecture de calcul neuromimétique en apprentissage non supervisé. In *10^o Colloque sur le traitement du signal et des images, FRA, 1985*. GRETSI, Groupe d'Etudes du Traitement du Signal et des Images, 1985. (Cité page 35.)
- [115] A. Hjørungnes. *Complex-valued matrix derivatives : with applications in signal processing and communications*. Cambridge University Press, 2011. (Cité pages 59 et 112.)
- [116] A. Holobar, C. Fevotte, C. Doncarli, and D. Zazula. Single autoterms separation based on bilinear timefrequency representations. *EUSIPCO, Toulouse, France, 2002*. (Cité page 39.)
- [117] A. Hyvärinen, J. Karhunen, and E. Oja. *Independent component analysis*, volume 46. John Wiley & Sons, 2004. (Cité page 41.)
- [118] J. J. Thiagarajan, K. Natesan Ramamurthy, A. Spanias, J. J. Thiagarajan, K. N. Ramamurthy, and A. Spanias. Mixing matrix estimation using discriminative clustering for blind source separation. *Digital Signal Processing : A Review Journal*, 23(1) :9–18, 2013. (Cité page 47.)
- [119] M. Jafari, J. Chambers, and D. Mandic. Natural gradient algorithm for cyclostationary sources. *Electronics Letters*, 38(14) :758–759, 2002. (Cité page 43.)
- [120] M. G. Jafari, S. R. Alty, and J. A. Chambers. New natural gradient algorithm for cyclostationary sources, 2004. (Cité page 43.)
- [121] P. Jallon, A. Chevreuil, P. Loubaton, and P. Chevalier. Separation of Convulsive Mixtures of Cyclostationary Sources : A Contrast Function Based Approach. pages 508–515. Springer, Berlin, Heidelberg, 2004. (Cité page 43.)
- [122] C. Jutten. *Calcul neuromimétique et traitement du signal : analyse en composantes indépendantes*. PhD thesis, jan 1987. (Cité page 28.)
- [123] M. Kanaujia. ECG Signal Decomposition Using PCA and ICA. *National Conference on Recent Advances in Electronics & Computer Engineering*, pages 301–305, 2015. (Cité page 5.)
- [124] D. Lahat, J.-F. Cardoso, H. Messer, et al. Joint block diagonalization algorithms for optimal separation of multidimensional components. In *LVA/ICA*, pages 155–162. Springer, 2012. (Cité page 35.)

- [125] L. D. Lathauwer. Signal processing based on multilinear algebra. 1997. (Cité page 36.)
- [126] D. D. Lee and H. S. Seung. Learning the parts of objects by non-negative matrix factorization. *Nature*, 401(6755) :788–91, 1999. (Cité page 48.)
- [127] D. D. Lee and H. S. Seung. Algorithms for Non-negative Matrix Factorization. *Advances in Neural Information Processing Systems*, (1) :556–562, 2001. (Cité page 48.)
- [128] T. W. Lee, A. Ziehe, R. Orglmeister, and T. Sejnowski. Combining time-delayed decorrelation and ICA : Towards solving the cocktail party problem. In *ICASSP, IEEE International Conference on Acoustics, Speech and Signal Processing - Proceedings*, volume 2, pages 1249–1252, 1998. (Cité page 2.)
- [129] H. Lev-Ari et al. Efficient solution of linear matrix equations with application to multistatic antenna array processing. *Communications in Information & Systems*, 5(1) :123–130, 2005. (Cité page 112.)
- [130] A. Lewis and M. Overton. Nonsmooth optimization via BFGS. *Submitted to SIAM J. Optimiz*, pages 1–35, 2009. (Cité page 59.)
- [131] D. H. Li and M. Fukushima. A modified BFGS method and its global convergence in nonconvex minimization. *Journal of Computational and Applied Mathematics*, 129(1-2) :15–35, 2001. (Cité page 61.)
- [132] Y. Li, A. Cichocki, and S.-i. Amari. Sparse component analysis for blind source separation with less sensors than sources. *Ica2003*, (April) :89–94, 2003. (Cité page 47.)
- [133] Y.-C. Liang, A. Leyman, and B.-H. Soong. Blind source separation using second-order cyclic-statistics. In *First IEEE Signal Processing Workshop on Signal Processing Advances in Wireless Communications*, pages 57–60. IEEE. (Cité page 42.)
- [134] S. Maiz. Estimation et détection de signaux cyclostationnaires par les méthodes de ré-échantillonnage statistique : applications à l’analyse des signaux biomécaniques. 2014. (Cité pages 94, 95 et 104.)
- [135] S. G. Mallat and Z. Zhang. Matching Pursuits With Time-Frequency Dictionaries. *IEEE Transactions on Signal Processing*, 41(12) :3397–3415, 1993. (Cité page 46.)
- [136] V. Maurandi, C. De Luigi, and E. Moreau. Fast jacobi like algorithms for joint diagonalization of complex symmetric matrices. In *Signal Processing Conference (EUSIPCO), 2013 Proceedings of the 21st European*, pages 1–5. IEEE, 2013. (Cité page 34.)

- [137] E. Moreau. A generalization of joint-diagonalization criteria for source separation. *IEEE Transactions on Signal Processing*, 49(3) :530–541, 2001. (Cité pages 17, 34 et 36.)
- [138] E. Moreau and T. Adali. *Blind Identification and Separation of Complex-valued Signals*. Wiley Online Library, 2013. (Cité page 35.)
- [139] E. Moreau and O. Macchi. New self-adaptative algorithms for source separation based on contrast functions. In *[1993 Proceedings] IEEE Signal Processing Workshop on Higher-Order Statistics*, pages 215–219. IEEE. (Cité page 37.)
- [140] E. Moreau and J. C. Pesquet. Generalized contrasts for multichannel blind deconvolution of linear systems. *IEEE Signal Processing Letters*, 4(6) :182–183, 1997. (Cité page 37.)
- [141] E. Moreau and N. Thirion-Moreau. Non symmetrical contrasts for source separation. *IEEE Transactions on Signal Processing*, 47(8) :2241–2253, 1999. (Cité page 36.)
- [142] A. Napolitano. Cyclostationarity : New trends and applications. *Signal Processing*, 120 :385–408, 2016. (Cité page 18.)
- [143] M. Niegowski and M. Zivanovic. ECG-EMG Separation by Using Enhanced Non-Negative Matrix Factorization. In *36th Annual International Conference of the IEEE Engineering in Medicine and Biology Society*, pages 4212–4215, 2014. (Cité page 5.)
- [144] D. Nion. A tensor framework for nonunitary joint block diagonalization. *IEEE Transactions on Signal Processing*, 59(10) :4585–4594, 2011. (Cité page 35.)
- [145] T. Nishikawa, H. Saruwatari, K. Shikano, S. Araki, and S. Makino. Multistage ICA for Blind Source Separation of Real Acoustic Convolutional Mixture. 2003. (Cité page 9.)
- [146] Z. Obeid. *Mise au point d’algorithmes pour la détection de dégradations de roulements d’actionneurs synchrones à aimants permanents. Application dans le domaine aéronautique sur des ventilateurs embarqués*. PhD thesis, École Doctorale Génie Électrique, Électronique et Télécommunications : du système au nanosystème (Toulouse) ; 154235504, 2012. (Cité pages 79 et 81.)
- [147] Z. Obeid, S. Poignant, J. Régner, and P. Maussion. Stator current based indicators for bearing fault detection in synchronous machine by statistical frequency selection. In *IECON 2011-37th Annual Conference on IEEE Industrial Electronics Society*, pages 2036–2041. IEEE, 2011. (Cité page 80.)

- [148] A. Ozerov and C. Fevotte. Multichannel nonnegative matrix factorization in convolutive mixtures for audio source separation. *IEEE Transactions on Audio, Speech and Language Processing*, 18(3) :550–563, 2010. (Cité page 49.)
- [149] P. Paatero. Least squares formulation of robust non-negative factor analysis. In *Chemometrics and Intelligent Laboratory Systems*, volume 37, pages 23–35, 1997. (Cité page 48.)
- [150] L. Parra and C. Spence. Convolutive blind separation of non-stationary sources. *IEEE Transactions on Speech and Audio Processing*, 8(3) :320–327, may 2000. (Cité page 9.)
- [151] Y. Pati, R. Rezaifar, and P. Krishnaprasad. Orthogonal matching pursuit : recursive function approximation with applications to wavelet decomposition. *Proceedings of 27th Asilomar Conference on Signals, Systems and Computers*, pages 40–44, 1993. (Cité page 46.)
- [152] Y. C. Pati, R. Rezaifar, and P. S. Krishnaprasad. Orthogonal matching pursuit : recursive function approximation with applications to wavelet decomposition. In *Proceedings of 27th {Asilomar} {Conference} on {Signals}, {Systems} and {Computers}*, pages 40—44 vol.1, 1993. (Cité page 95.)
- [153] D. T. Pham and J. F. Cardoso. Blind separation of instantaneous mixtures of nonstationary sources. *IEEE Transactions on Signal Processing*, 49(9) :1837–1848, 2001. (Cité page 39.)
- [154] D. T. Pham and P. Garat. Blind separation of mixture of independent sources through a maximum likelihood approach. In *Proc. EUSIPCO*, 45(7) :1712–1725, 1997. (Cité page 41.)
- [155] K. Rahbar and J. P. J. Reilly. A frequency domain method for blind source separation of convolutive audio mixtures. *IEEE Transactions on Speech and Audio Processing*, 13(5) :832–844, sep 2005. (Cité page 9.)
- [156] R. RANDALL, J. ANTONI, and S. CHOBSAARD. THE RELATIONSHIP BETWEEN SPECTRAL CORRELATION AND ENVELOPE ANALYSIS IN THE DIAGNOSTICS OF BEARING FAULTS AND OTHER CYCLOSTATIONARY MACHINE SIGNALS. *Mechanical Systems and Signal Processing*, 15(5) :945–962, sep 2001. (Cité page 20.)
- [157] B. D. Rao and K. Kreutz-Delgado. An affine scaling methodology for best basis selection. *IEEE Transactions on Signal Processing*, 47(1) :187–200, 1999. (Cité page 46.)

- [158] A. I. Rauf, M. Fukushima, and C. P. by Tseng. Globally Convergent BFGS Method for Nonsmooth Convex Optimization. *JOURNAL OF OPTIMIZATION THEORY AND APPLICATIONS*, 104(3) :539–558, 2000. (Cité page 61.)
- [159] V. G. Reju, S. N. Koh, and I. Y. Soon. Underdetermined Convolutional Blind Source Separation via Time-Frequency Masking. *IEEE Transactions on Audio, Speech, and Language Processing*, 18(1) :101–116, jan 2010. (Cité page 47.)
- [160] S. Rhioui, N. Thirion-Moreau, and E. Moreau. Under-Determined Blind Identification of Cyclo-Stationary Signals with Unknown Cyclic Frequencies. *2006 IEEE International Conference on Acoustics Speech and Signal Processing Proceedings*, 4 :553–556, 2006. (Cité page 43.)
- [161] J. Riddle. *Ball bearing maintenance*. University of Oklahoma Press, 1955. (Cité page 79.)
- [162] L. Rota, P. Comon, and S. Icart. Blind equalization of MIMO channels. In *Signal Processing Advances in Wireless Communications, 2003. SPAWC 2003. 4th IEEE Workshop on*, pages 314–318. IEEE, 2003. (Cité page 37.)
- [163] L. Rota, P. Comon, and S. Icart. Égalisation aveugle de m{é}langes convolutifs MIMO. In *19^o Colloque sur le traitement du signal et des images, FRA, 2003*. GRETSI, Groupe d'Etudes du Traitement du Signal et des Images, 2003. (Cité page 37.)
- [164] J. Roussel, P. Ravier, M. Haritopoulos, D. Farina, and O. Buttelli. Decomposition of Multi-Channel Intramuscular EMG Signals by Cyclostationary-Based Blind Source Separation. *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, pages 1–1, 2017. (Cité page 44.)
- [165] K. Sabri, M. El Badaoui, F. Guillet, A. Adib, and D. Aboutajdine. A frequency domain-based approach for blind MIMO system identification using second-order cyclic statistics. *Signal Processing*, 89(1) :77–86, 2009. (Cité pages 9 et 43.)
- [166] K. Sabri, M. El Badaoui, F. Guillet, A. Belli, and G. Millet. Blind source separation based on cyclic spectra : Application to biomechanical signals. In *European Signal Processing Conference, 2008*. (Cité page 93.)
- [167] K. Sabri, M. El Badaoui, F. Guillet, A. Belli, G. Millet, and J. Benoit Morin. Cyclostationary modeling of ground reaction force signals. *Signal Processing*, 90(4) :1146–1152, 2010. (Cité pages 93, 95 et 97.)

- [168] N. M. Saitou N. The Neighbor-joining Method : A New Method for Reconstructing Phylogenetic Trees'. *Molecular Biology and Evolution*, 4(4) :406–425, 1987. (Cité page 73.)
- [169] C. Shin, R. W. Heath, and E. J. Powers. Blind Channel Estimation for MIMO-OFDM Systems. *IEEE Transactions on Vehicular Technology*, 56(2) :670–685, 2007. (Cité page 4.)
- [170] C. Simon, P. Loubaton, and C. Jutten. Separation of a class of convolutive mixtures : a contrast function approach. *Signal Processing*, 81(4) :883–887, 2001. (Cité page 41.)
- [171] P. Smaragdis. Non-negative matrix factor deconvolution ; extraction of multiple sound sources from monophonic inputs. *Independent Component Analysis and Blind Signal Separation*, (2) :494–499, 2004. (Cité page 49.)
- [172] A. Souloumiac. Nonorthogonal joint diagonalization by combining givens and hyperbolic rotations. *IEEE Transactions on Signal Processing*, 57(6) :2222–2231, 2009. (Cité page 34.)
- [173] J. R. Stack, R. G. Harley, and T. G. Habetler. An amplitude modulation detector for fault diagnosis in rolling element bearings. *IEEE Transactions on Industrial Electronics*, 51(5) :1097–1102, 2004. (Cité page 80.)
- [174] J. Sun, Y. Li, J. Wen, and S. Yan. Novel mixing matrix estimation approach in underdetermined blind source separation. *Neurocomputing*, 173, Part :623–632, 2016. (Cité page 47.)
- [175] N. Thirion-Moreau and E. Moreau. New criteria for blind signal separation. In *Statistical Signal and Array Processing, 2000. Proceedings of the Tenth IEEE Workshop on*, pages 344–348. IEEE, 2000. (Cité page 37.)
- [176] N. Thirion-Moreau and E. Moreau. Generalized criteria for blind multivariate signal equalization. *IEEE Signal Processing Letters*, 9(2) :72–74, 2002. (Cité page 37.)
- [177] R. Tibshirani. Regression selection and shrinkage via the lasso. *Journal of the Royal Statistical Society B*, 58(1) :267–288, 1996. (Cité page 46.)
- [178] P. Tichavsky and Z. Koldovsky. Algorithms for nonorthogonal approximate joint block-diagonalization. In *Signal Processing Conference (EU-SIPCO), 2012 Proceedings of the 20th European*, pages 2094–2098. IEEE, 2012. (Cité page 35.)

- [179] L. Tong, V. C. Soon, Y. F. Huang, and R. Liu. AMUSE : a new blind identification algorithm. In *Circuits and Systems, 1990., IEEE International Symposium on*, pages 1784–1787. IEEE, 1990. (Cité page 37.)
- [180] T. Trainini and E. Moreau. A coordinate descent algorithm for complex joint diagonalization under hermitian and transpose congruences. *IEEE Transactions on Signal Processing*, 62(19) :4974–4983, 2014. (Cité page 35.)
- [181] J. K. Tugnait. Identification and deconvolution of multichannel linear non-Gaussian processes using higher order statistics and inverse filter criteria. *IEEE Transactions on Signal Processing*, 45(3) :658–672, 1997. (Cité page 41.)
- [182] E. Vincent, N. Benin, and R. Badeau. Harmonic and inharmonic Non-negative Matrix Factorization for polyphonic pitch transcription. In *ICASSP, IEEE International Conference on Acoustics, Speech and Signal Processing - Proceedings*, pages 109–112, 2008. (Cité page 49.)
- [183] T. Virtanen. Monaural sound source separation by nonnegative matrix factorization with temporal continuity and sparseness criteria. *IEEE Transactions on Audio, Speech and Language Processing*, 15(3) :1066–1074, 2007. (Cité page 49.)
- [184] B. Wang and M. Plumbley. Investigating single-channel audio source separation methods based on non-negative matrix factorization. *Proc. ICA Research Network International ...*, pages 0–3, 2006. (Cité page 49.)
- [185] M. W. Whittle. Generation and attenuation of transient impulsive forces beneath the foot : a review. *Gait & Posture*, 10(3) :264–275, dec 1999. (Cité page 92.)
- [186] Y. Xiang, D. Peng, I. Ubhayaratne, B. Rolfe, and M. Pereira. Second-Order Cyclostationary Statistics-Based Blind Source Extraction From Convolutional Mixtures. *IEEE Access*, 5 :2011–2019, 2017. (Cité page 43.)
- [187] X.-F. Xu, D.-Z. Feng, W. X. Zheng, and H. Zhang. Convolutional blind source separation based on joint block toeplitzation and block-inner diagonalization. *Signal Processing*, 90(1) :119–133, 2010. (Cité page 35.)
- [188] A. Yeredor. Non-orthogonal joint diagonalization in the least-squares sense with application in blind source separation. *IEEE Transactions on signal processing*, 50(7) :1545–1553, 2002. (Cité page 34.)

- [189] Ö. Yilmaz and S. Rickard. Blind separation of speech mixtures via time-frequency masking. *IEEE Transactions on Signal Processing*, 52(7) :1830–1846, 2004. (Cité page 47.)
- [190] Ying-Chang Liang, A. Leyman, and Boon-Hee Song. A network structure approach to blind source separation using second order cyclic statistics. In *Proceedings of 1997 IEEE International Symposium on Circuits and Systems. Circuits and Systems in the Information Age ISCAS '97*, volume 4, pages 2549–2551. IEEE. (Cité page 42.)
- [191] Yong Xiang, Wensheng Yu, and Jingxin Zhang. Blind separation of cyclostationary sources based on phase-frequency redundancy. In *Proceedings 7th International Conference on Signal Processing, 2004. Proceedings. ICSP '04. 2004.*, volume 1, pages 403–406. IEEE. (Cité page 43.)
- [192] S. S. Young, P. Fogel, and D. M. Hawkins. Clustering Scotch Whiskies using Non-Negative Matrix Factorization, 2006. (Cité page 49.)
- [193] A. Ypma, A. Leshem, and R. P. W. Duin. Blind separation of rotating machine sources : Bilinear forms and convolutive mixtures. *Neurocomputing*, 49 :349–368, 2002. (Cité page 42.)
- [194] Y. Yu and A. P. Petropulu. PARAFAC-based blind estimation of possibly underdetermined convolutive MIMO systems. *IEEE Transactions on Signal Processing*, 56(1) :111–124, 2008. (Cité page 4.)
- [195] F. Zakaria, M. El Badaoui, M. Lamraoui, and M. Khalil. Fatigue study of ultra-runners : Presentation of a new approach for the separation of GRF signals components. *Mechanical Systems and Signal Processing*, 75 :648–667, 2016. (Cité pages 44, 93 et 95.)
- [196] F. Zakaria, S. Maiz, M. El Badaoui, and M. Khalil. First- and second-order cyclostationary signal separation using morphological component analysis. 58 :134–144, nov 2016. (Cité page 94.)
- [197] V. Zarzoso, P. Comon, and M. Kallel. How fast is FastICA? In *Signal Processing Conference, 2006 14th European*, pages 1–5. IEEE, 2006. (Cité page 42.)
- [198] V. Zarzoso and A. K. Nandi. Noninvasive fetal electrocardiogram extraction : Blind separation versus adaptive noise cancellation. *IEEE Transactions on Biomedical Engineering*, 48(1) :12–18, 2001. (Cité page 5.)
- [199] W.-J. Zeng, X.-L. Li, X.-D. Zhang, and X. Jiang. An improved signal-selective direction finding algorithm using second-order cyclic statistics. In *Acoustics, Speech and Signal Processing, 2009. ICASSP 2009.*

- IEEE International Conference on*, pages 2141–2144. IEEE, 2009. (Cité page 35.)
- [200] M. Zhang, M. Haritopoulos, and A. K. Nandi. Fetal ECG subspace estimation based on cyclostationarity. In *2016 24th European Signal Processing Conference (EUSIPCO)*, pages 2060–2064. IEEE, aug 2016. (Cité pages 5 et 44.)
- [201] A. Ziehe and K. Müller. TDSEP-an efficient algorithm for blind separation using time structure. *Proceedings of the 8th International Conference on Artificial Neural Networks (ICANN98)*, (September) :675–680, 1998. (Cité page 38.)

Titre Contribution à la séparation de sources cyclo-stationnaires : application aux signaux de télécommunications, mécaniques et biomécaniques

Résumé Dans cette thèse, nous nous sommes attaqués au problème de séparation aveugle de mélanges linéaires de sources ayant des propriétés de cyclo-stationnarité. Trois applications ont été abordées à savoir : télécommunications, vibrations mécaniques et biomécaniques. Dans un premier temps, deux nouvelles méthodes ont été proposées, la première a pour but de séparer aveuglement des sources cyclo-stationnaires partageant une ou plusieurs fréquences cycliques inconnues. Elle combine la diagonalisation conjointe à un nouveau détecteur de points utiles (retard-fréquence cyclique) permettant de composer l'ensemble de matrices de corrélation cyclique devant être diagonalisées conjointement. Quant à la deuxième méthode, elle vise à identifier la matrice de mélange de sources cyclo-stationnaires de fréquences cycliques inconnues et différentes. L'identification commence par une étape de détection des matrices de rang un, puis décompose en éléments propres le produit de matrices sélectionnées, enfin une méthode de regroupement hiérarchique restitue les colonnes de notre matrice recherchée. Les deux solutions ont été appliquées aux signaux de télécommunications. Dans un second temps, nous avons appliqué d'abord la première méthode proposée sur des signaux mécaniques issus d'un banc de roulements défailants afin de tester son aptitude à séparer les sources. Ensuite, nous avons proposé une approche qui s'appuie sur l'analyse en composantes parcimonieuses pour séparer les composantes de la force de réaction au sol ayant des propriétés cyclo-stationnaires à l'ordre 1 et 2.

Mots-clés séparation aveugle de sources, mélanges linéaires, cyclo-stationnarité, diagonalisation conjointe, corrélation cyclique, télécommunications, vibrations mécaniques et biomécaniques.

Title Contribution to the separation of cyclo-stationary sources : application to telecommunications, mechanical and biomechanical signals

Abstract In this thesis, we have tackled the problem of blind separation of linear mixtures of sources with cyclo-stationarity properties. Three applications were studied : telecommunications, mechanical vibrations and biomechanics. First, two new methods have been proposed, the first one aims to blindly separate cyclo-stationary sources sharing one or more unknown cyclic frequencies. It combines the joint diagonalization with a new useful point detector (time lag-cyclic frequency) to compose the set of cyclic correlation matrices to be jointly diagonalized. As for the second method, it aims to identify the mixture matrix of cyclo-stationary sources of unknown and different cyclic frequencies. The identification begins with a step of detecting the matrices of rank one, then the product of selected matrices is decomposed into eigen-elements, and finally a hierarchical regrouping method returns the columns of our sought matrix. Both solutions have been applied to telecommunications signals. In a second step, we first applied the first proposed method on mechanical signals coming from a bank of faulty bearings in order to test its ability to separate the sources. Next, we proposed an approach based on sparse component analysis to separate the components of the ground reaction force with cyclo-stationary properties at order 1 and 2.

Keywords blind separation of sources, linear mixtures, cyclo-stationarity, joint diagonalisation, cyclic correlation, telecommunications, mechanical and biomechanical vibrations.