

Titre: Une approche géométrique à l'analyse en composantes
Title: indépendantes

Auteur: Wissem Maazoun
Author:

Date: 2008

Type: Mémoire ou thèse / Dissertation or Thesis

Référence: Maazoun, W. (2008). Une approche géométrique à l'analyse en composantes indépendantes [Mémoire de maîtrise, École Polytechnique de Montréal].
Citation: PolyPublie. <https://publications.polymtl.ca/8225/>

Document en libre accès dans PolyPublie

Open Access document in PolyPublie

URL de PolyPublie: <https://publications.polymtl.ca/8225/>
PolyPublie URL:

Directeurs de recherche: Antoine Saucier
Advisors:

Programme: Non spécifié
Program:

UNIVERSITÉ DE MONTRÉAL

UNE APPROCHE GÉOMÉTRIQUE À L'ANALYSE EN COMPOSANTES
INDÉPENDANTES

WISSEM MAAZOUN

DÉPARTEMENT DE MATHÉMATIQUES ET DE GÉNIE INDUSTRIEL
ÉCOLE POLYTECHNIQUE DE MONTRÉAL

MÉMOIRE PRÉSENTÉ EN VUE DE L'OBTENTION
DU DIPLÔME DE MAÎTRISE ÈS SCIENCES APPLIQUÉES
(MATHÉMATIQUES APPLIQUÉES)
DÉCEMBRE 2008



Library and
Archives Canada

Published Heritage
Branch

395 Wellington Street
Ottawa ON K1A 0N4
Canada

Bibliothèque et
Archives Canada

Direction du
Patrimoine de l'édition

395, rue Wellington
Ottawa ON K1A 0N4
Canada

Your file *Votre référence*

ISBN: 978-0-494-48930-7

Our file *Notre référence*

ISBN: 978-0-494-48930-7

NOTICE:

The author has granted a non-exclusive license allowing Library and Archives Canada to reproduce, publish, archive, preserve, conserve, communicate to the public by telecommunication or on the Internet, loan, distribute and sell theses worldwide, for commercial or non-commercial purposes, in microform, paper, electronic and/or any other formats.

The author retains copyright ownership and moral rights in this thesis. Neither the thesis nor substantial extracts from it may be printed or otherwise reproduced without the author's permission.

In compliance with the Canadian Privacy Act some supporting forms may have been removed from this thesis.

While these forms may be included in the document page count, their removal does not represent any loss of content from the thesis.

AVIS:

L'auteur a accordé une licence non exclusive permettant à la Bibliothèque et Archives Canada de reproduire, publier, archiver, sauvegarder, conserver, transmettre au public par télécommunication ou par l'Internet, prêter, distribuer et vendre des thèses partout dans le monde, à des fins commerciales ou autres, sur support microforme, papier, électronique et/ou autres formats.

L'auteur conserve la propriété du droit d'auteur et des droits moraux qui protège cette thèse. Ni la thèse ni des extraits substantiels de celle-ci ne doivent être imprimés ou autrement reproduits sans son autorisation.

Conformément à la loi canadienne sur la protection de la vie privée, quelques formulaires secondaires ont été enlevés de cette thèse.

Bien que ces formulaires aient inclus dans la pagination, il n'y aura aucun contenu manquant.

**
Canada

UNIVERSITÉ DE MONTRÉAL

ÉCOLE POLYTECHNIQUE DE MONTRÉAL

Ce mémoire intitulé :

UNE APPROCHE GÉOMÉTRIQUE À L'ANALYSE EN COMPOSANTES
INDÉPENDANTES

présenté par : MAAZOUN Wissem
en vue de l'obtention du diplôme de : Maîtrise ès sciences appliquées
a été dûment accepté par le jury d'examen constitué de :

M. BOURDEAU Marc, Ph.D., président

M. SAUCIER Antoine, Ph.D., membre et directeur de recherche

Mme LABBÉ Chantal, Ph.D., membre et codirectrice

M. ADJENGUE Luc, Ph.D., membre

À mes parents

À ma famille.

REMERCIEMENTS

J'aimerais remercier M. Antoine Saucier mon directeur de recherche pour le support qu'il m'a apporté et la confiance qu'il m'a témoignée tout au long de ce travail. Sa disponibilité, ses précieux conseils, sa grande rigueur, son aide financière et ses encouragements m'ont permis d'avancer de façon constructive et rapide. Je remercie également Mme Chantal Labbé ma codirectrice pour ses conseils, son aide financière et son aide tout au long de ce mémoire.

Ma reconnaissance va aussi naturellement à mes parents dont l'éducation, la confiance, l'appui moral et financier m'ont permis de m'épanouir, à ma femme qui a fait énormément pour moi, celle qui a supporté mes humeurs changeantes en fonction de l'évolution du projet, à mon frère et ma soeur pour leur soutien permanent et leurs encouragements.

Je remercie également M. Marc Bourdeau de bien vouloir présider mon jury, ainsi que M. Luc Adjengue d'avoir accepté d'être membre de ce jury.

Merci à tous mes collègues de travail, à tous les professeurs et le personnel du département de mathématiques et de génie industriel ainsi que ceux de toute l'école polytechnique de Montréal.

Merci à tous.

RÉSUMÉ

Ce mémoire est consacré à l'analyse en composantes indépendantes. On considère des signaux aléatoires \mathbf{S} de la forme

$$\mathbf{S} = \sum_{i=1}^p a_i \mathbf{V}_i,$$

où les $\mathbf{V}_i \in \mathbb{R}^N$ sont des vecteurs inconnus et linéairement indépendants et les a_i sont des variables aléatoires indépendantes de moyenne nulle, mais dont les lois de probabilité sont inconnues. Le problème que nous étudions est l'estimation des \mathbf{V}_i , appelés composantes indépendantes, à partir de réalisations du signal \mathbf{S} . Ce problème a des applications en analyse des signaux ainsi que pour la synthèse stochastique de signaux aléatoires.

Le signal \mathbf{S} appartient à l'espace $E := \text{span}\{\mathbf{V}_1, \dots, \mathbf{V}_p\}$, qui est de dimension $p \leq N$. Notre approche consiste à chercher des vecteurs $\mathbf{X}_i \in E$, $i = 1, \dots, p$, tels que $\mathbf{X}_i^T \mathbf{V}_j = 0$ si $i \neq j$, $\mathbf{X}_i^T \mathbf{V}_i \neq 0$. Si les \mathbf{X}_i sont connus, alors les $p - 1$ conditions d'orthogonalité $\mathbf{X}_i^T \mathbf{V}_j = 0$, $i = 1, 2, \dots, j - 1, j + 1, \dots, p$ permettent la détermination des $\mathbf{V}_j \in E$ car E est de dimension p . Pour estimer les \mathbf{X}_i , nous exploitons l'indépendance des facteurs $\mathbf{X}_i^T \mathbf{S}$, qui découle de l'identité $\mathbf{X}_i^T \mathbf{S} = a_i \mathbf{X}_i^T \mathbf{V}_i$ et de l'indépendance des a_i . On démontre que l'indépendance des facteurs $\mathbf{X}_i^T \mathbf{S}$ and $\mathbf{X}_j^T \mathbf{S}$ pour $i \neq j$ a deux conséquences. D'une part, les conditions d'orthogonalité

$$\mathbf{X}_i^T \mathbf{Z}_n(\mathbf{X}_j) = 0, \quad i \neq j, \quad \forall n \in \mathbb{N},$$

où $\mathbf{Z}_n(\mathbf{X}) := E((\mathbf{X}^T \mathbf{S})^n \mathbf{S})$, sont satisfaites. D'autre part, ces conditions impliquent la colinéarité des vecteurs de l'ensemble

$$Z(\mathbf{X}_i) := \{\mathbf{Z}_n(\mathbf{X}_i), n \in \mathbb{N}\}.$$

Pour $p = 2$, nous avons développé deux méthodes d'estimation des \mathbf{V}_i basées sur la colinéarités des vecteurs de $Z(\mathbf{X}_i)$. Pour $p = 3$, nous avons développé une méthode basée sur les conditions d'orthogonalité : on minimise une fonction objectif à l'aide d'une procédure de descente cyclique. La convergence de la descente vers une valeur minimale est démontrée.

Pour $p = 2$, nous avons comparé la précision des \mathbf{V}_i estimés obtenus avec nos méthodes et avec la méthode *FastICA*, en utilisant des signaux simulés. Nos méthodes s'avèrent nettement plus précises que la méthode *FastICA*, avec des erreurs d'estimation au moins deux fois plus petites.

ABSTRACT

This master's thesis is devoted to independent component analysis. We consider random signals \mathbf{S} of the form

$$\mathbf{S} = \sum_{i=1}^p a_i \mathbf{V}_i,$$

where the vectors $\mathbf{V}_i \in \mathbb{R}^N$ are unknown and the a_i s are independent zero-mean random variables with unknown probability distributions. The problem that we consider is the estimation of the \mathbf{V}_i s, called independent components, given a set of realizations of the signals \mathbf{S} . This problem has applications in signal analysis and for the synthesis of random signals.

The signal \mathbf{S} belongs to the p -dimensional space, ($p \leq N$), $E := \text{span}\{\mathbf{V}_1, \dots, \mathbf{V}_p\}$. Our approach is to look for vectors $\mathbf{X}_i \in E$, $i = 1, \dots, p$, such that $\mathbf{X}_i^T \mathbf{V}_j = 0$ if $i \neq j$, with $\mathbf{X}_i^T \mathbf{V}_i \neq 0$. If the \mathbf{X}_i s are known, then the $p - 1$ orthogonality conditions $\mathbf{X}_i^T \mathbf{V}_j = 0$, $i = 1, 2, \dots, j-1, j+1, \dots, p$ allow the determination of $\mathbf{V}_j \in E$ because E has dimension p . To estimate the \mathbf{X}_i s, we take advantage of the independence of the factors $\mathbf{X}_i^T \mathbf{S}$, which results from the identity $\mathbf{X}_i^T \mathbf{S} = a_i \mathbf{X}_i^T \mathbf{V}_i$ and the independence of the a_i s. We show that the independence of the factors $\mathbf{X}_i^T \mathbf{S}$ and $\mathbf{X}_j^T \mathbf{S}$ for $i \neq j$ has two consequences. Firstly, the orthogonality conditions

$$\mathbf{X}_i^T \mathbf{Z}_n(\mathbf{X}_j) = 0, \quad i \neq j, \quad \forall n \in \mathbb{N},$$

where $\mathbf{Z}_n(\mathbf{X}) := E((\mathbf{X}^T \mathbf{S})^n \mathbf{S})$, are satisfied. Secondly, these conditions imply the colinearity of the vector collection

$$Z(\mathbf{X}_i) := \{\mathbf{Z}_n(\mathbf{X}_i), n \in \mathbb{N}\}.$$

For $p = 2$, we developed two estimation methods based on the colinearity of the

vectors in $Z(\mathbf{X}_i)$. For $p = 3$, we developed a method based on the orthogonality conditions. We minimise an objective function via a cyclic descent procedure. The convergence of this descent to a minimal value of the objective function is established.

For $p = 2$, we compared the precision of the estimated \mathbf{V}_i s obtained with our methods and with the method *FastICA*, using synthetic signals. Our methods are significantly more accurate than *FastICA*, with estimation errors which are at least twice smaller.

TABLE DES MATIÈRES

DÉDICACE	iv
REMERCIEMENTS	v
RÉSUMÉ	vi
ABSTRACT	viii
TABLE DES MATIÈRES	x
LISTE DES FIGURES	xii
INTRODUCTION	1
CHAPITRE 1 DÉFINITION DU PROBLÈME	7
1.1 Représentation du signal dans une base appropriée	7
1.2 Une approche basée sur les facteurs indépendants	11
1.3 Détermination des facteurs indépendants	13
CHAPITRE 2 CAS DE DEUX DIMENSIONS ($p = 2$) : DÉTERMINATION DES COMPOSANTES INDÉPENDANTES EN UTILISANT LA COLINÉARITÉ DES ÉLÉMENTS DE $Z(\mathbf{X})$	17
2.1 La forme particulière de $Z_n(\mathbf{X})$	17
2.2 Analyse des conditions d'orthogonalité	18
2.2.1 Résumé	25
2.2.2 Algorithme de calcul des composantes indépendantes	26
2.3 Expériences numériques	28

CHAPITRE 3 LE CAS 2D : DÉTERMINATION DES COMPOSANTES INDÉPENDANTES PAR LA MINIMISATION D'UNE FONCTION OBJECTIVE	41
3.1 Définition de la méthode II	42
3.2 Analyse de la méthode II	44
3.3 Expériences numériques	48
3.4 Comparaison des deux méthodes	53
3.5 Comparaison de la méthode I avec la méthode <i>FastICA</i>	59
CHAPITRE 4 CAS DE TROIS COMPOSANTES ($p = 3$)	71
4.1 Description de la méthode	72
4.2 Expériences numériques	78
CONCLUSION	85
RÉFÉRENCES	87

LISTE DES FIGURES

Figure 2.1	Graphe de $\hat{f}(\theta) = \ \hat{\mathbf{z}}_3(\mathbf{X}(\theta)) - \hat{\mathbf{z}}_1(\mathbf{X}(\theta))\ $	29
Figure 2.2	Les composantes estimées par notre méthode et les composantes indépendantes.	31
Figure 2.3	La courbe de $\hat{f}(\theta)$ avec $\epsilon = 10$ et $n = 1$	33
Figure 2.4	La courbe de $\hat{f}(\theta)$ avec $\epsilon = 10$ et $n = 4$	33
Figure 2.5	La courbe de $\hat{f}(\theta)$ avec $\epsilon = 10$ et $n = 16$	33
Figure 2.6	La courbe de $\hat{f}(\theta)$ avec $\epsilon = 10$ et $n = 1$	34
Figure 2.7	La courbe de $\hat{f}(\theta)$ avec $\epsilon = 5$ et $n = 1$	34
Figure 2.8	La courbe de $\hat{f}(\theta)$ avec $\epsilon = 1$ et $n = 1$	34
Figure 2.9	Agrandissement de la figure 2.8.	35
Figure 2.10	Agrandissement de la figure 2.5.	35
Figure 2.11	La courbe de SNR en fonction de ϵ , $n = 1$	36
Figure 2.12	La courbe de SNR en fonction de n , $\epsilon = 10$	36
Figure 2.13	Les composantes obtenus avec $\epsilon = 10$ et $n = 1$, SNR=83.	38
Figure 2.14	La courbe de $\hat{f}(\theta)$ avec $\epsilon = 10$ et $n = 32$	38
Figure 2.15	La courbe de $\hat{f}(\theta)$ avec $k_1 k_2 < 0$	40
Figure 2.16	La courbe de $\hat{\mathbf{V}}_1, \hat{\mathbf{V}}_2, \tilde{\mathbf{V}}_1, \tilde{\mathbf{V}}_2$ avec $k_1 k_2 < 0$	40

Figure 3.1	Graphe de $\hat{f}(\theta) = \hat{g}(\hat{g}_1(\mathbf{X}(\theta)), \mathbf{X}(\theta))$ dans un cas où $k_1 k_2 > 0$ pour l'échantillon 1	50
Figure 3.2	Graphe de $\hat{f}(\theta) = \hat{g}(\hat{g}_1(\mathbf{X}(\theta)), \mathbf{X}(\theta))$ dans un cas où $k_1 k_2 > 0$ pour l'échantillon 2	50
Figure 3.3	Les composantes estimées par la deuxième méthode (o) et les composantes indépendantes (*), dans un cas où $k_1 k_2 > 0$ pour l'échantillon 1	51
Figure 3.4	Graphe de $\hat{f}(\theta) = \hat{g}(\hat{g}_1(\mathbf{X}(\theta)), \mathbf{X}(\theta))$ pour un cas où $k_1 k_2 < 0$	51
Figure 3.5	Agrandissement de la figure 3.4	52
Figure 3.6	Les composantes estimées par la deuxième méthode (o) et les composantes indépendantes (*), $k_1 k_2 < 0$	52
Figure 3.7	Graphe de $\hat{f}(\theta) = \hat{g}(\hat{g}_1(\mathbf{X}(\theta)), \mathbf{X}(\theta))$ pour un cas où $\epsilon = 1$	54
Figure 3.8	SNR_{\min} méthode I (o) VS. méthode II (*) en fonction de ϵ , $n = 1, M = 100$	56
Figure 3.9	SNR_{moy} méthode I (o) VS. méthode II (*) en fonction de ϵ , $n = 1, M = 100$	56
Figure 3.10	SNR_{\min} méthode I (o) VS. méthode II (*) en fonction de n , $\epsilon = 10, M = 100$	57
Figure 3.11	SNR_{moy} méthode I (o) VS. méthode II (*) en fonction de n , $\epsilon = 10, M = 100$	57
Figure 3.12	SNR_{\min} méthode I (o) VS. méthode II (*) en fonction de M , $\epsilon = 10, n = 1$	58

Figure 3.13	SNR _{moy} méthode I (o) VS. méthode II (★) en fonction de M , $\epsilon = 10, n = 1$	58
Figure 3.14	Graphe de $\frac{\mathbf{a}_1}{\ \mathbf{a}_1\ }$ (★), $\frac{\hat{\mathbf{a}}_1}{\ \hat{\mathbf{a}}_1\ }$ (o) et $\frac{\hat{\mathbf{a}}_{1FI}}{\ \hat{\mathbf{a}}_{1FI}\ }$ (□), dans le cas où $a_i \sim U[-1, 1], i = 1, 2$	63
Figure 3.15	Graphe de $\frac{\mathbf{a}_2}{\ \mathbf{a}_2\ }$ (★), $\frac{\hat{\mathbf{a}}_2}{\ \hat{\mathbf{a}}_2\ }$ (o) et $\frac{\hat{\mathbf{a}}_{2FI}}{\ \hat{\mathbf{a}}_{2FI}\ }$ (□), dans le cas où $a_i \sim U[-1, 1], i = 1, 2$	64
Figure 3.16	Graphe des résidus $(\frac{\hat{\mathbf{a}}_1}{\ \hat{\mathbf{a}}_1\ } - \frac{\mathbf{a}_1}{\ \mathbf{a}_1\ })$ obtenus avec notre méthode (★) et la méthode <i>FastICA</i> (o), dans le cas où $a_i \sim U[-1, 1], i = 1, 2$	64
Figure 3.17	Graphe des résidus $(\frac{\hat{\mathbf{a}}_2}{\ \hat{\mathbf{a}}_2\ } - \frac{\mathbf{a}_2}{\ \mathbf{a}_2\ })$ obtenus avec notre méthode (★) et la méthode <i>FastICA</i> (o), dans le cas où $a_i \sim U[-1, 1], i = 1, 2$	65
Figure 3.18	Graphe de $\frac{\mathbf{a}_1}{\ \mathbf{a}_1\ }$ (★), $\frac{\hat{\mathbf{a}}_1}{\ \hat{\mathbf{a}}_1\ }$ (o) et $\frac{\hat{\mathbf{a}}_{1FI}}{\ \hat{\mathbf{a}}_{1FI}\ }$ (□), dans le cas où $U_1(k) = k$ et $U_2(k) = k + 1$	66
Figure 3.19	Graphe de $\frac{\mathbf{a}_2}{\ \mathbf{a}_2\ }$ (★), $\frac{\hat{\mathbf{a}}_2}{\ \hat{\mathbf{a}}_2\ }$ (o) et $\frac{\hat{\mathbf{a}}_{2FI}}{\ \hat{\mathbf{a}}_{2FI}\ }$ (□), dans le cas où $U_1(k) = k$ et $U_2(k) = k + 1$	67
Figure 3.20	Graphe des résidus $(\frac{\hat{\mathbf{a}}_1}{\ \hat{\mathbf{a}}_1\ } - \frac{\mathbf{a}_1}{\ \mathbf{a}_1\ })$ obtenus avec notre méthode (★) et la méthode <i>FastICA</i> (o), dans le cas où $U_1(k) = k$ et $U_2(k) = k + 1$	67
Figure 3.21	Graphe des résidus $(\frac{\hat{\mathbf{a}}_2}{\ \hat{\mathbf{a}}_2\ } - \frac{\mathbf{a}_2}{\ \mathbf{a}_2\ })$ obtenus avec notre méthode (★) et la méthode <i>FastICA</i> (o), dans le cas où $U_1(k) = k$ et $U_2(k) = k + 1$	68
Figure 3.22	Graphe de $\frac{\mathbf{a}_1}{\ \mathbf{a}_1\ }$ (★), $\frac{\hat{\mathbf{a}}_1}{\ \hat{\mathbf{a}}_1\ }$ (o) et $\frac{\hat{\mathbf{a}}_{1FI}}{\ \hat{\mathbf{a}}_{1FI}\ }$ (□), dans le cas où $U_1(k) = k$ et $U_2(k) = k^2$	69

Figure 3.23	Graphe de $\frac{\mathbf{a}_2}{\ \mathbf{a}_2\ }$ (\star), $\frac{\hat{\mathbf{a}}_2}{\ \hat{\mathbf{a}}_2\ }$ (\circ) et $\frac{\hat{\mathbf{a}}_{2FI}}{\ \hat{\mathbf{a}}_{2FI}\ }$ (\square), dans le cas où $U_1(k) = k$ et $U_2(k) = k^2$	69
Figure 3.24	Graphe des résidus $(\frac{\hat{\mathbf{a}}_1}{\ \hat{\mathbf{a}}_1\ } - \frac{\mathbf{a}_1}{\ \mathbf{a}_1\ })$ obtenus avec notre méthode (\star) et la méthode <i>FastICA</i> (\circ), dans le cas où $U_1(k) = k$ et $U_2(k) = k^2$	70
Figure 3.25	Graphe des résidus $(\frac{\hat{\mathbf{a}}_2}{\ \hat{\mathbf{a}}_2\ } - \frac{\mathbf{a}_2}{\ \mathbf{a}_2\ })$ obtenus avec notre méthode (\star) et la méthode <i>FastICA</i> (\circ), dans le cas où $U_1(k) = k$ et $U_2(k) = k^2$	70
Figure 4.1	Graphe de $\tilde{f}_3(\theta)$, $\theta \in [0, \pi]$, dans $\text{span}(\tilde{\Phi}_2, \tilde{\Phi}_3)$	79
Figure 4.2	Graphe de $\tilde{f}_3(\theta)$, $\theta \in [0, \pi]$, dans $\text{span}([\hat{\mathbf{C}}\hat{\mathbf{X}}_1^{(1)}]^\perp)$	80
Figure 4.3	Graphe des composantes estimées avec notre méthode (\star) et les composantes indépendantes (\circ), où $W_1(j) = j$, $W_2(j) = j^2$, $W_3(j) = j^3$, $j = 0, \dots, N-1$ et $M = 500$	81
Figure 4.4	Graphe des composantes estimées avec notre méthode (\star) et les composantes indépendantes (\circ), où $W_1(j) = j$, $W_2(j) = j^2$, $W_3(j) = \log(j+1)$, $j = 0, \dots, N-1$ et $M = 5000$	82
Figure 4.5	Graphe des composantes estimées avec notre méthode (\star) et les composantes indépendantes (\circ), où $W_1(j) = j$, $W_2(j) = j^2$, $W_3(j) = j^3$, $j = 0, \dots, N-1$ et $M = 100$	83
Figure 4.6	Spécifications de l'ordinateur utilisé pour l'exécution de nos algorithmes.	84

INTRODUCTION

Analyse en composantes indépendantes

Ce mémoire est consacré au problème de l'analyse en composantes indépendantes, que nous définissons comme suit. Nous considérons des signaux de la forme

$$\mathbf{S} = \sum_{i=1}^p a_i \mathbf{V}_i, \quad (1)$$

où les a_i sont des variables aléatoires mutuellement indépendantes de moyenne nulle, et les $\mathbf{V}_i \in \mathbb{R}^N$ sont des vecteurs unitaires linéairement indépendants les uns des autres. On suppose que les a_i et les \mathbf{V}_i sont inconnus, mais que des réalisations $\{\mathbf{S}^{(1)}, \dots, \mathbf{S}^{(M)}\}$ du signal \mathbf{S} sont données. Le problème est alors d'estimer les vecteurs \mathbf{V}_i à partir des $\{\mathbf{S}^{(1)}, \dots, \mathbf{S}^{(M)}\}$.

Ce problème peut être considéré comme une variante de l'analyse en composantes principales, qui consiste à décomposer un signal aléatoire \mathbf{S} sous la forme

$$\mathbf{S} = \sum_{i=1}^p \alpha_i \Phi_i, \quad (2)$$

où les Φ_i sont des vecteurs orthonormaux et les α_i sont des variables aléatoires *décorrélées* qui ne sont pas nécessairement indépendantes. Dans ce contexte, les vecteurs Φ_i sont appelés les *composantes principales* du signal \mathbf{S} . Les *composantes principales* sont égales aux vecteurs propres de la matrice de corrélation $C := E(\mathbf{S}\mathbf{S}^T)$ du signal \mathbf{S} .

Soulignons tout d'abord les différences principales entre les décompositions (1) et (2). D'une part, les vecteurs \mathbf{V}_i de l'analyse en composantes indépendantes (équation (1)) ne sont pas nécessairement orthogonaux, alors qu'ils le sont toujours pour l'analyse

en composante principales (équation (2)). D'autre part, les variables aléatoires a_i sont mutuellement *indépendantes*, alors que les α_i sont mutuellement *décorrélées*. Notons que deux variables aléatoires peuvent être décorrélées sans être indépendantes.

Une des applications possible de l'analyse en composantes indépendantes définie par (1) est la *synthèse de signaux stochastiques*. En effet, si les composantes indépendantes \mathbf{V}_i sont connues (ou estimées), alors on peut utiliser les signaux $\{\mathbf{S}^{(1)}, \dots, \mathbf{S}^{(M)}\}$ et les \mathbf{V}_i pour calculer les coefficients a_i et estimer leur loi de probabilité. On peut ensuite utiliser (1) directement comme méthode de simulation.

Le problème de l'analyse en composante indépendante a été popularisé par plusieurs auteurs, comme par exemple Hyvarinen et Oja (2000), Comon (1994), Jutten et Héault (1988). Ces auteurs ont défini le problème de l'analyse en composantes indépendantes d'une façon un peu différente de celle que nous utilisons ici. Nous allons d'abord exposer la définition utilisée par ces auteurs, pour pouvoir ensuite la comparer avec la définition (1) utilisée dans ce mémoire.

Dans l'analyse en composantes indépendantes de Comon (1994), on suppose que plusieurs sources indépendantes génèrent des signaux indépendants \mathbf{s}_i , $i = 1, \dots, N_s$. Les signaux sont reçus par plusieurs microphones qui enregistrent des signaux dénotés par \mathbf{x}_i , $i = 1, \dots, N_x$. On construit une matrice \mathbf{s} dont les lignes sont les signaux source \mathbf{s}_i , i.e. $s_{i,j} \equiv s_i(j)$. De même, on construit une matrice \mathbf{x} dont les lignes sont les signaux reçus \mathbf{x}_i , i.e. $x_{i,j} \equiv x_i(j)$. On suppose alors que les signaux reçus et les signaux sources sont reliés par une relation linéaire de la forme

$$\mathbf{x} = A \mathbf{s}, \quad (3)$$

où A est une matrice (appelée *mixing matrix*) aux dimensions appropriées. Dans ce contexte, le problème de l'analyse en composantes indépendantes est de déterminer la matrice A et les signaux \mathbf{s} à partir des signaux reçus \mathbf{x} , en supposant que les signaux

sources s_i sont mutuellement indépendants.

Le modèle (1) que nous étudions dans ce mémoire peut aussi s'écrire sous la forme matricielle (3) en procédant comme suit. On construit une matrice \mathbf{S} dont chaque ligne est une réalisation $\mathbf{S}^{(i)}$ du signal aléatoire \mathbf{S} , i.e. $S_{i,j} \equiv S^{(i)}(j)$. De même, on construit une matrice \mathbf{V} dont les lignes sont les vecteurs \mathbf{V}_i , i.e. $V_{i,j} \equiv V_i(j)$. Finalement, on construit une matrice \mathbf{a} dont la i -ième ligne contient les coefficients $a_1^{(i)} \ a_2^{(i)} \ \dots \ a_p^{(i)}$ de la i -ième réalisation du signal, i.e. $a_{i,j} \equiv a^{(i)}(j)$. On constate alors que la relation (1) peut s'écrire sous la forme équivalente

$$\mathbf{S} = \mathbf{a} \ \mathbf{V}. \quad (4)$$

En transposant (4), on obtient

$$\mathbf{S}^T = \mathbf{V}^T \ \mathbf{a}^T. \quad (5)$$

En comparant (3) et (5), on constate que \mathbf{x} correspond à \mathbf{S}^T , \mathbf{A} correspond à \mathbf{V}^T et que \mathbf{s} correspond à \mathbf{a}^T .

Une approche géométrique à l'estimation des composantes indépendantes

Dans ce mémoire, nous abordons le problème de l'estimation de composantes indépendantes avec une approche *géométrique* très différente de l'approche de Hyvärinen et Oja (2000), qui est fondée sur la minimisation de fonctionnelles mesurant le degré de gaussianité de certaines lois de probabilité.

Pour comprendre le point de départ de notre approche, considérons le cas le plus simple où $p = 2$, pour lequel \mathbf{S} prend la forme

$$\mathbf{S} = a_1 \mathbf{V}_1 + a_2 \mathbf{V}_2. \quad (6)$$

Le signal \mathbf{S} appartient donc à un espace E de dimension deux, i.e. un plan. Considérons deux vecteurs \mathbf{X}_1 et \mathbf{X}_2 dans ce plan et tels que

$$\mathbf{X}_1^T \mathbf{V}_2 = 0 \text{ et } \mathbf{X}_2^T \mathbf{V}_1 = 0. \quad (7)$$

On constate avec (6) et (7) que

$$\mathbf{X}_1^T \mathbf{S} = a_1 \mathbf{X}_1^T \mathbf{V}_1 \text{ et } \mathbf{X}_2^T \mathbf{S} = a_2 \mathbf{X}_2^T \mathbf{V}_2, \quad (8)$$

et donc les facteurs $\mathbf{X}_1^T \mathbf{S}$ et $\mathbf{X}_2^T \mathbf{S}$ sont indépendants puisque a_1 et a_2 sont indépendantes.

Si on pouvait trouver deux vecteurs \mathbf{X}_1 et \mathbf{X}_2 dans le plan E tels que les *facteurs* $\mathbf{X}_1^T \mathbf{S}$ et $\mathbf{X}_2^T \mathbf{S}$ soient indépendants, et si ces vecteurs satisfont effectivement les conditions d'orthogonalité (7), alors \mathbf{X}_1 et \mathbf{X}_2 permettraient de déterminer \mathbf{V}_1 et \mathbf{V}_2 directement via (7).

Si on suppose que les facteurs $\mathbf{X}_1^T \mathbf{S}$ et $\mathbf{X}_2^T \mathbf{S}$ sont indépendants, alors il s'ensuit que

$$E((\mathbf{X}_1^T \mathbf{S})(\mathbf{X}_2^T \mathbf{S})^n) = E((\mathbf{X}_1^T \mathbf{S}))E(\mathbf{X}_2^T \mathbf{S})^n = 0 \quad \forall n \in \mathbb{Z} \quad (9)$$

car $E(\mathbf{X}_1^T \mathbf{S}) = \mathbf{X}_1^T E(\mathbf{S}) = 0$ puisque $E(\mathbf{S}) = 0$ (on suppose $E(a_i) = 0 \forall i$). Par ailleurs, comme $E((\mathbf{X}_1^T \mathbf{S})(\mathbf{X}_2^T \mathbf{S})^n) = \mathbf{X}_1^T E((\mathbf{X}_2^T \mathbf{S})^n \mathbf{S})$, alors on peut toujours écrire (9) sous la forme équivalente

$$\mathbf{X}_1^T \mathbf{Z}_n(\mathbf{X}_2) = 0, \quad \forall n \in \mathbb{Z}, \quad (10)$$

où on définit

$$\mathbf{Z}_n(\mathbf{X}) := E((\mathbf{X}^T \mathbf{S})^n \mathbf{S}), \quad (11)$$

qui est un vecteur qui peut être estimé à partir des réalisations du signal \mathbf{S} . Par

symétrie, on obtient aussi

$$\mathbf{X}_2^T \mathbf{Z}_n(\mathbf{X}_1) = 0, \quad \forall n \in \mathbb{Z}. \quad (12)$$

Les conditions d'orthogonalité (10) et (12) indiquent que \mathbf{X}_1 et \mathbf{X}_2 sont liés par des contraintes géométriques fortes. En particulier, si \mathbf{X}_1 est connu, alors \mathbf{X}_2 est connu grâce à la condition d'orthogonalité (12) appliquée dans un plan. Réciproquement, si \mathbf{X}_2 est connu, alors \mathbf{X}_1 est connu grâce à la condition d'orthogonalité (10).

En résumé, nous avons vu que l'indépendance des facteurs $\mathbf{X}_1^T \mathbf{S}$ et $\mathbf{X}_2^T \mathbf{S}$ implique que les coefficients de corrélations de $\mathbf{X}_1^T \mathbf{S}$ et $(\mathbf{X}_2^T \mathbf{S})^n$ sont nuls, ce qui se traduit par des *contraintes géométriques* liant \mathbf{X}_1 et \mathbf{X}_2 , i.e. les conditions d'orthogonalité (10) et (12). L'objectif principal de ce mémoire est d'examiner ces contraintes géométriques et de les exploiter pour en arriver à développer une méthode d'estimation des composantes indépendantes. Nous comparerons aussi notre méthode avec les méthodes existantes.

Revue de littérature

L'analyse en composantes indépendantes est une nouvelle méthode statistique d'analyse de données utilisée pour la séparation des sources, plus précisément elle est utilisée pour la séparation aveugle de source (de l'anglais "Blind Source Separation" (BSS)). Ce problème a été formulé initialement par Hérault *et al.* (1985), qui ont traité le problème de l'analyse des messages composites issus d'un ensemble de capteurs sensibles à des mélanges de sources multiples. L'analyse en composantes indépendantes est une extension de l'analyse en composantes principales où on cherche à exprimer les données observées dans un sous-espace sur lequel les projections linéaires conservent le maximum de la variance. Dans le cas de l'analyse en composantes indépendantes, l'hypothèse sous-jacente permettant la meilleure

représentation des données est que les sources sont statistiquement indépendantes les unes des autres. Jutten et Hérault (1988) ont alors adopté l'expression "analyse en composantes indépendantes". Ces deux chercheurs ont effectué d'autres travaux pour améliorer l'algorithme qu'ils ont obtenu, on peut citer Jutten et Hérault (1991) et Comon *et al.* (1991). Ce dernier article a été fait en collaboration avec Pierre Comon qui a redéfini l'algorithme plus précisément Comon (1990b). Pierre Comon a effectué plusieurs travaux dans ce domaine. Parmi ses travaux on peut citer Comon (1990a), Comon (1992), Comon (1994), Cardoso et Comon (1996), Grellier et Comon (1998), De Lathauwer *et al.* (1999). On peut citer aussi les travaux de Hyvärinen, A. et Oja, E. qui ont développé plusieurs algorithmes pour la détermination des composantes indépendantes. On peut citer parmi leurs travaux dans ce domaine, Hyvärinen (1997), Hyvärinen et Oja (1996), Hyvärinen et Oja (1997), Hyvärinen et Oja (1998), Hyvärinen et Oja (2000), Bingham et Hyvärinen (2000), Hyvärinen (2001), Inki et Hyvärinen (2002). Ces travaux ont mené à l'algorithme "FastICA" qui est utilisé par la plupart des applications de l'analyse en composantes indépendantes. Notons que jusqu'à présent Hyvärinen et Oja (1997) a été cité plus de 900 fois dans des articles de journaux et de conférences et Hyvärinen et Oja (2000) a été cité plus de 800 fois. On peut citer aussi plusieurs travaux des applications de l'analyse en composantes indépendantes dans le domaine du traitement des signaux et d'images ou en génie biomédical. Comme applications de l'analyse en composantes indépendantes on peut citer par exemple Kohno *et al.* (2007), Song-yun *et al.* (2007), Yang et Yi (2007), Yang et Yi (2007), Herrmann et Theis (2007), Zewail *et al.* (2007), Chaabouni-Chouayakh et Datcu (2007), Li et Wang (2007), Naik *et al.* (2007), Lai *et al.* (2007), Ou *et al.* (2007), An *et al.* (2007), Zhang *et al.* (2007).

CHAPITRE 1

DÉFINITION DU PROBLÈME

On considère des signaux $\tilde{\mathbf{S}}$ dans \mathbb{R}^N . Nous supposons que ces signaux ont la forme

$$\tilde{\mathbf{S}} = \sum_{j=1}^p a_j \tilde{\mathbf{V}}_j, \quad (1.1)$$

où $p \leq N$. Les a_j sont des variables aléatoires non constantes, mutuellement indépendantes et de moyennes nulles. Il s'ensuit que $E\{\tilde{\mathbf{S}}\} = 0$. De plus on supposera que $E[a_j^2] < \infty, \forall j$. Les $\tilde{\mathbf{V}}_j$ sont des vecteurs linéairement indépendants les uns des autres et de normes unité (ils ne sont pas nécessairement orthogonaux). Tant la distribution des $a_j, j = 1, 2, \dots, p$, que les vecteurs $\tilde{\mathbf{V}}_1, \dots, \tilde{\mathbf{V}}_p$ nous sont inconnus. Pour que les vecteurs $\tilde{\mathbf{V}}_j$ soient définis d'une façon unique, nous supposerons qu'ils sont normés et qu'ils satisfont

$$|\max_{1 \leq k \leq N} \tilde{\mathbf{V}}_j(k)| \geq |\min_{1 \leq k \leq N} \tilde{\mathbf{V}}_j(k)| \text{ et } \max_{1 \leq k \leq N} \tilde{\mathbf{V}}_j(k) > 0, \text{ sinon } \tilde{\mathbf{V}}_j := -\tilde{\mathbf{V}}_j. \quad (1.2)$$

Étant donné un ensemble de M réalisations du signal $\tilde{\mathbf{S}}$, le problème consiste à estimer les $\tilde{\mathbf{V}}_j$.

1.1 Représentation du signal dans une base appropriée

Les signaux $\tilde{\mathbf{S}}$ définis par (1.1) appartiennent au sous-espace de dimension p engendré par les vecteurs $\tilde{\mathbf{V}}_1, \tilde{\mathbf{V}}_2, \dots, \tilde{\mathbf{V}}_p$, soit

$$E := \text{span}\{\tilde{\mathbf{V}}_1, \tilde{\mathbf{V}}_2, \dots, \tilde{\mathbf{V}}_p\}.$$

On définit la matrice d'autocorrélation $\tilde{\mathbf{C}}$ de $\tilde{\mathbf{S}}$ par

$$\begin{aligned}\tilde{\mathbf{C}} &= E\{\tilde{\mathbf{S}}\tilde{\mathbf{S}}^T\} \\ &= \sum_{i=1}^p E\{a_i^2\}\tilde{\mathbf{V}}_i\tilde{\mathbf{V}}_i^T.\end{aligned}$$

La matrice $\tilde{\mathbf{C}}$ est symétrique, donc orthogonalement diagonalisable. Soit $\tilde{\Phi}_1, \tilde{\Phi}_2, \dots, \tilde{\Phi}_N$ les vecteurs propres de la matrice $\tilde{\mathbf{C}}$, triés par ordre décroissant des valeurs propres correspondantes. En d'autres mots, si $\lambda_i, i = 1, \dots, N$ est la valeur propre associée à $\tilde{\Phi}_i$, alors $\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_N \geq 0$.

Lemme 1 La matrice $\tilde{\mathbf{C}}$ possède exactement p valeurs propres non nulles, $\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_p > 0$, pour lesquelles les vecteurs propres associés engendrent E :

$$E = \text{span}\{\tilde{\Phi}_1, \tilde{\Phi}_2, \dots, \tilde{\Phi}_p\}.$$

Preuve Si $\mathbf{U} \in E^\perp$, alors $\tilde{\mathbf{C}}\mathbf{U} = \sum_{i=1}^p E\{a_i^2\}\tilde{\mathbf{V}}_i\tilde{\mathbf{V}}_i^T\mathbf{U} = \mathbf{0}$. Ainsi 0 est une valeur propre de $\tilde{\mathbf{C}}$ (car $\tilde{\mathbf{C}}\mathbf{U} = \mathbf{0}\mathbf{U}$), et $E^\perp \subset \mathcal{V}_0$, où \mathcal{V}_0 dénote le sous-espace des vecteurs propres de 0. Inversement, si $\mathbf{U} \in \mathcal{V}_0$, alors $\tilde{\mathbf{C}}\mathbf{U} = \mathbf{0}$, c'est-à-dire que

$$\sum_{i=1}^p E\{a_i^2\}(\tilde{\mathbf{V}}_i^T\mathbf{U})\tilde{\mathbf{V}}_i = 0.$$

Puisque $\tilde{\mathbf{V}}_1, \tilde{\mathbf{V}}_2, \dots, \tilde{\mathbf{V}}_p$ sont linéairement indépendants, alors

$$\begin{aligned}E\{a_1^2\}(\tilde{\mathbf{V}}_1^T\mathbf{U}) &= \dots = E\{a_p^2\}(\tilde{\mathbf{V}}_p^T\mathbf{U}) = 0 \\ \implies \tilde{\mathbf{V}}_1^T\mathbf{U} &= \dots = \tilde{\mathbf{V}}_p^T\mathbf{U} = 0,\end{aligned}$$

et donc $\mathbf{U} \in E^\perp$. Il s'ensuit que $\mathcal{V}_0 \subset E^\perp$, d'où $\mathcal{V}_0 = E^\perp$ (nous avons auparavant démontré que $E^\perp \subset \mathcal{V}_0$). On a donc

$$\dim(\mathcal{V}_0) = \dim(E^\perp) = N - p,$$

et donc il y a exactement p valeurs propres non nulles. Cela implique $\mathcal{V}_0 = \text{span}\{\tilde{\Phi}_{p+1}, \dots, \tilde{\Phi}_N\}$, d'où

$$E = \mathcal{V}_0^\perp = \text{span}\{\tilde{\Phi}_1, \dots, \tilde{\Phi}_p\}. \quad \blacksquare$$

Comme tous les signaux appartiennent au sous-espace E , alors on peut représenter tous les vecteurs dans une base de E . On choisit la base

$$\mathcal{B} := \{\tilde{\Phi}_1, \tilde{\Phi}_2, \dots, \tilde{\Phi}_p\}.$$

Les coordonnées de $\tilde{\mathbf{S}}$ dans la base \mathcal{B} seront dénotées par $S(i), i = 1, \dots, p$. Les vecteurs $\mathbf{S} = (S(1), S(2), \dots, S(p))^T$ et $\tilde{\mathbf{S}}$ sont reliés par

$$\tilde{\mathbf{S}} = \mathbf{F}^T \mathbf{S},$$

où \mathbf{F} est la matrice de dimension $p \times N$ dont les lignes sont les $\tilde{\Phi}_i^T, i = 1, \dots, p$.

Comme les $\tilde{\Phi}_i$ forment une base orthonormale de E , alors

$$\mathbf{F}\mathbf{F}^T = \mathbf{I}_p,$$

où \mathbf{I}_p est la matrice identité de dimension $p \times p$. Il s'ensuit que

$$\mathbf{S} = \mathbf{F}\tilde{\mathbf{S}}.$$

Dans la base \mathcal{B} , les vecteurs $\tilde{\mathbf{V}}_i$ prennent aussi la forme

$$\mathbf{V}_i = \mathbf{F}\tilde{\mathbf{V}}_i, i = 1, \dots, p,$$

Dans la base \mathcal{B} , $\tilde{\mathbf{C}}$ devient

$$\mathbf{C} = \mathbf{F}\tilde{\mathbf{C}}\mathbf{F}^T,$$

qui est la matrice diagonale donnée par

$$\mathbf{C} = \begin{pmatrix} \lambda_1 & 0 & 0 & \dots & 0 & 0 \\ 0 & \lambda_2 & 0 & \dots & 0 & 0 \\ 0 & 0 & \lambda_3 & \dots & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \dots & \lambda_{p-1} & 0 \\ 0 & 0 & 0 & \dots & 0 & \lambda_p \end{pmatrix}$$

Dans la suite, tous les vecteurs seront exprimés dans la base \mathcal{B} . Dans cette base, le modèle (1.1) prend la forme

$$\mathbf{S} = \sum_{j=1}^p a_j \mathbf{V}_j \quad (1.3)$$

Lemme 2 $\{\mathbf{V}_1, \dots, \mathbf{V}_p\}$ forment une base de \mathbb{R}^p .

Preuve Puisque $\mathbf{V}_1, \mathbf{V}_2, \dots, \mathbf{V}_p$ sont des vecteurs de \mathbb{R}^p , il suffit de montrer qu'ils sont linéairement indépendants. Soient $c_1, \dots, c_p \in \mathbb{R}$, tels que $c_1\mathbf{V}_1 + \dots + c_p\mathbf{V}_p = 0$. Alors $\mathbf{F}(c_1\tilde{\mathbf{V}}_1 + \dots + c_p\tilde{\mathbf{V}}_p) = 0$. Donc $c_1\tilde{\mathbf{V}}_1 + \dots + c_p\tilde{\mathbf{V}}_p$ est orthogonal à $\tilde{\Phi}_i, i = 1, \dots, p$. Puisque $E = \text{span}\{\Phi_1, \dots, \Phi_p\}$, alors $c_1\tilde{\mathbf{V}}_1 + \dots + c_p\tilde{\mathbf{V}}_p \in E^\perp$. De plus, puisque $E = \text{span}\{\tilde{\mathbf{V}}_1, \dots, \tilde{\mathbf{V}}_p\}$, alors $c_1\tilde{\mathbf{V}}_1 + \dots + c_p\tilde{\mathbf{V}}_p \in E$. Donc $c_1\tilde{\mathbf{V}}_1 + \dots + c_p\tilde{\mathbf{V}}_p = 0$, or $\tilde{\mathbf{V}}_1, \dots, \tilde{\mathbf{V}}_p$ sont linéairement indépendants, alors $c_1 = \dots = c_p = 0$.

L'indépendance linéaire des $\mathbf{V}_i, i = 1, \dots, p$, suit. ■

1.2 Une approche basée sur les facteurs indépendants

Considérons des combinaisons linéaires (aussi appelées facteurs dans la suite) de la forme $\mathbf{X}_i^T \mathbf{S}$, $i = 1, 2, \dots, p$, où les $\mathbf{X}_i \in \mathbb{R}^p$ sont des vecteurs déterministes de norme unité qui satisfont la condition (1.2). En utilisant (1.3), ces combinaisons linéaires prennent la forme

$$\mathbf{X}_i^T \mathbf{S} = \sum_{j=1}^p a_j \mathbf{X}_i^T \mathbf{V}_j, \quad i = 1, 2, \dots, p. \quad (1.4)$$

Cherchons des vecteurs \mathbf{X}_i qui sont tels que les facteurs $\mathbf{X}_i^T \mathbf{S}$ soient statistiquement indépendants les uns des autres. Considérons les vecteurs non nuls $\bar{\mathbf{X}}_1, \bar{\mathbf{X}}_2, \dots, \bar{\mathbf{X}}_p \in \mathbb{R}^p$ qui satisfont

$$\bar{\mathbf{X}}_i^T \mathbf{V}_j = 0, \quad \forall j \neq i. \quad (1.5)$$

On remarque que si (1.5) est satisfait, alors

$$\bar{\mathbf{X}}_i^T \mathbf{S} = a_i \bar{\mathbf{X}}_i^T \mathbf{V}_i$$

pour chaque i et donc chacun des $\bar{\mathbf{X}}_i^T \mathbf{S}$ est exprimé en terme d'une seule variable aléatoire a_i , et par conséquent les $\bar{\mathbf{X}}_i^T \mathbf{S}$ sont mutuellement indépendants. Notons qu'il existe peut-être des vecteurs $\mathbf{X}_1, \dots, \mathbf{X}_p$, pour lesquels les facteurs $\mathbf{X}_i^T \mathbf{S}$ sont indépendants sans que la condition (1.5) soit satisfaite.

Lemme 3 Les vecteurs $\bar{\mathbf{X}}_1, \bar{\mathbf{X}}_2, \dots, \bar{\mathbf{X}}_p$ définis par les conditions d'orthogonalités (1.5) sont linéairement indépendants et $\bar{\mathbf{X}}_i^T \mathbf{V}_i \neq 0$ pour $i = 1, 2, \dots, p$.

Preuve Rappelons que $\{\mathbf{V}_1, \dots, \mathbf{V}_p\}$ forment une base de \mathbb{R}^p par le lemme 2. Si $\bar{\mathbf{X}}_i \neq 0$ et que (1.5) tient, alors on doit nécessairement avoir $\bar{\mathbf{X}}_i^T \mathbf{V}_i \neq 0, \forall i = 1, \dots, p$. Soient $c_1, c_2, \dots, c_p \in \mathbb{R}$. Supposons que

$$c_1 \bar{\mathbf{X}}_1 + c_2 \bar{\mathbf{X}}_2 + \dots + c_p \bar{\mathbf{X}}_p = 0.$$

Donc pour $j = 1, 2, \dots, p$

$$0 = (c_1 \bar{\mathbf{X}}_1 + c_2 \bar{\mathbf{X}}_2 + \dots + c_p \bar{\mathbf{X}}_p)^T \mathbf{V}_j = c_j \bar{\mathbf{X}}_j^T \mathbf{V}_j,$$

où la seconde égalité découle de (1.5). Puisque $\bar{\mathbf{X}}_j^T \mathbf{V}_j \neq 0$, alors $c_j = 0$.

L'indépendance linéaire des $\bar{\mathbf{X}}_j, j = 1, \dots, p$, suit. ■

Si les $\bar{\mathbf{X}}_i$ satisfont les conditions d'orthogonalité (1.5) alors on peut déterminer les \mathbf{V}_j complètement. Les \mathbf{V}_j sont obtenus avec

$$\mathbf{V}_j \in E_j := \ker \begin{pmatrix} \bar{\mathbf{X}}_1^T \\ \bar{\mathbf{X}}_2^T \\ \vdots \\ \bar{\mathbf{X}}_{j-1}^T \\ \bar{\mathbf{X}}_{j+1}^T \\ \vdots \\ \bar{\mathbf{X}}_p^T \end{pmatrix}, \quad (1.6)$$

où E_j est un sous-espace de dimension un par le lemme 3 et le fait que les $\bar{\mathbf{X}}_i \in \mathbb{R}^p, i = 1, 2, \dots, p$.

La question qui subsiste est alors la suivante : pouvons-nous trouver les p vecteurs $\bar{\mathbf{X}}_i$ qui font en sorte que les conditions d'orthogonalité (1.5) sont satisfaites, et ce malgré le fait que nous ne connaissons pas $\mathbf{V}_1, \mathbf{V}_2, \dots, \mathbf{V}_p$? En d'autres mots, pouvons-nous déterminer $\bar{\mathbf{X}}_1, \bar{\mathbf{X}}_2, \dots, \bar{\mathbf{X}}_p$ en utilisant seulement le signal observé \mathbf{S} ?

1.3 Détermination des facteurs indépendants

Soit m le plus grand entier tel que $E\{|a_i|^{m+1}\} < \infty$ pour $i = 1, 2, \dots, p$. Si tous les moments de tous les a_i existent, alors on pose $m = \infty$. Si $\mathbf{X}_i^T \mathbf{S}$ est indépendant de tous les $\mathbf{X}_j^T \mathbf{S}$ avec $j \neq i$, alors il s'ensuit que

$$E\{(\mathbf{X}_i^T \mathbf{S})(\mathbf{X}_j^T \mathbf{S})^n\} = 0 \quad (1.7)$$

pour tout $j \neq i$ et pour $n = 0, 1, 2, \dots, m$. Le résultat (1.7) résulte de l'indépendance des deux facteurs et du fait que $E\{\mathbf{X}_i^T \mathbf{S}\} = \mathbf{X}_i^T E\{\mathbf{S}\} = 0$, car $E\{\mathbf{S}\} = \mathbf{0}$.

Remarquons que la covariance (1.7) peut aussi s'écrire sous la forme équivalente

$$E\{(\mathbf{X}_i^T \mathbf{S})(\mathbf{X}_j^T \mathbf{S})^n\} = \mathbf{X}_i^T (E\{(\mathbf{X}_j^T \mathbf{S})^n \mathbf{S}\}), \quad (1.8)$$

ce qui nous amène à définir pour tout $\mathbf{X} \in \mathbb{R}^p$ le vecteur

$$\mathbf{Z}_n(\mathbf{X}) = E\{(\mathbf{X}^T \mathbf{S})^n \mathbf{S}\}. \quad (1.9)$$

Les conditions de corrélation nulle (1.7) se traduisent donc par les conditions d'orthogonalité suivantes :

$$\mathbf{X}_i^T \mathbf{Z}_n(\mathbf{X}_j) = 0, \forall j \neq i, n = 0, 1, 2, \dots, m, \quad i, j = 1, 2, \dots, p. \quad (1.10)$$

Si on fixe j dans la condition d'orthogonalité (1.10), alors on obtient

$$\mathbf{X}_i^T \mathbf{Z}_n(\mathbf{X}_j) = 0, i = 1, 2, \dots, p, i \neq j$$

$$\implies \mathbf{Z}_n(\mathbf{X}_j) \in \Gamma_j := \ker \begin{pmatrix} \bar{\mathbf{X}}_1^T \\ \bar{\mathbf{X}}_2^T \\ \vdots \\ \bar{\mathbf{X}}_{j-1}^T \\ \bar{\mathbf{X}}_{j+1}^T \\ \vdots \\ \bar{\mathbf{X}}_p^T \end{pmatrix}, \quad (1.11)$$

pour $n = 0, 1, 2, \dots, m$. Si les $p - 1$ vecteurs $\mathbf{X}_i \in \mathbb{R}^p, i \neq j$ sont linéairement indépendants, alors $\mathbf{Z}_n(\mathbf{X}_j), n = 0, 1, \dots, m$, appartient à un espace de dimension un dénoté par Γ_j . Les vecteurs $\mathbf{Z}_n(\mathbf{X}_j), n = 0, 1, 2, \dots, m$ sont donc colinéaires pour chaque j .

Si on substitue (1.3) dans (1.9), alors on obtient

$$\begin{aligned} \mathbf{Z}_n(\mathbf{X}) &= E\{(\mathbf{X}^T \mathbf{S})^n \sum_{j=1}^p a_j \mathbf{V}_j\} \\ &= \sum_{j=1}^p E\{a_j (\mathbf{X}^T \mathbf{S})^n\} \mathbf{V}_j. \end{aligned} \quad (1.12)$$

Définissons l'ensemble de vecteurs de \mathbb{R}^p

$$Z(\mathbf{X}) := \{\mathbf{Z}_n(\mathbf{X}), n = 1, 2, \dots, m\}.$$

L'ensemble $Z(\mathbf{X})$ ne peut donc contenir qu'un maximum de p vecteurs linéairement indépendants, et ce pour tout $\mathbf{X} \in \mathbb{R}^p$. De plus, si $\mathbf{X}_1^T \mathbf{S}, \dots, \mathbf{X}_p^T \mathbf{S}$ sont indépendants, alors $Z(\mathbf{X}_i), i = 1, 2, \dots, p$, sont de dimension au plus 1.

Si on compare (1.6) et (1.11), alors on voit que $E_j = \Gamma_j$ si les conditions

d'orthogonalité (1.5) sont satisfaites, auquel cas on peut affirmer que pour tout j

$$\mathbf{V}_j = \varepsilon_n \mathbf{Z}_n(\bar{\mathbf{X}}_j), n = 0, 1, 2, \dots, m,$$

où ε_n est une constante réelle.

Les résultats de cette section peuvent se résumer comme suit :

$$\begin{aligned} \bar{\mathbf{X}}_i^T \mathbf{V}_j &= 0 \quad \forall i \neq j, \\ \implies \text{les } \bar{\mathbf{X}}_i^T \mathbf{S} &\text{ sont mutuellement indépendants.} \\ \implies \bar{\mathbf{X}}_i^T \mathbf{Z}_n(\bar{\mathbf{X}}_j) &= 0, \forall j \neq i, n = 0, 1, \dots, m. \\ \implies \text{Pour chaque } i, \text{ les vecteurs } \mathbf{Z}_n(\bar{\mathbf{X}}_i), n &= 0, 1, \dots, m, \text{ sont colinéaires.} \end{aligned} \quad (1.13)$$

Conclusion

Les $\mathbf{V}_j, j = 1, 2, \dots, p$, sont entièrement déterminés si on trouve des vecteurs non nuls $\bar{\mathbf{X}}_1, \bar{\mathbf{X}}_2, \dots, \bar{\mathbf{X}}_p \in \mathbb{R}^p$ satisfaisant (1.5). Ces vecteurs recherchés se trouvent parmi l'ensemble de tous les vecteurs $\mathbf{X} \in \mathbb{R}^p$ tels que $Z(\mathbf{X})$ est contenu dans un sous-espace de dimension un. Notons qu'il est possible que ce critère fournit plus de p candidats.

L'avantage de rechercher des candidats $\mathbf{X} \in \mathbb{R}^p$ parmi les vecteurs pour lesquels tous les vecteurs de $Z(\mathbf{X})$ sont colinéaires est que l'ensemble $Z(\mathbf{X})$ est entièrement défini en termes de \mathbf{X} et du signal observable \mathbf{S} . Ainsi, tous les vecteurs dans $Z(\mathbf{X})$ peuvent être estimés dès que l'on observe M réalisations du signal. De même, il faut réaliser qu'une fois les vecteurs $\mathbf{V}_1, \dots, \mathbf{V}_p \in \mathbb{R}^p$ identifiés, ou du moins estimés, le travail n'est pas terminé. Il faut en effet utiliser la relation $\tilde{\mathbf{V}}_j = \mathbf{F}^T \mathbf{V}_j, j = 1, \dots, p$, pour recouvrir les vecteurs $\tilde{\mathbf{V}}_1, \dots, \tilde{\mathbf{V}}_p \in \mathbb{R}^N$, ce qui constitue notre objectif fondamental. Pour ce faire, la matrice \mathbf{F} est requise. Or, cette matrice contient les vecteurs propres de la matrice $\tilde{\mathbf{C}}$ de corrélation du signal observable $\tilde{\mathbf{S}}$, et peuvent donc être estimés

facilement.

Comme notre but est de déterminer $\mathbf{V}_1, \mathbf{V}_2, \dots, \mathbf{V}_p$, il est raisonnable de chercher $\bar{\mathbf{X}}_1, \dots, \bar{\mathbf{X}}_p \neq 0$ tels que (1.5) est satisfait.

Pour les vecteurs $\bar{\mathbf{X}}_1, \dots, \bar{\mathbf{X}}_p \neq 0$ recherchés, (ceux qui satisfont (1.5)), on a que $Z(\bar{\mathbf{X}}_j)$, est contenu dans un espace de dimension 1 pour tout $j = 1, \dots, p$. Nous allons donc chercher $\bar{\mathbf{X}}_1, \dots, \bar{\mathbf{X}}_p$ parmi l'ensemble des vecteurs $\mathbf{X} \in \mathbb{R}^p$ tels que $Z(\mathbf{X})$ est contenu dans un espace de dimension 1.

Dans la prochaine section, nous verrons comment dans le cas $p = 2$ les candidats superflus peuvent être éliminés pour ne conserver que les vecteurs $\bar{\mathbf{X}}_1$ et $\bar{\mathbf{X}}_2 \in \mathbb{R}^2$ satisfaisant les conditions d'orthogonalité (1.5).

CHAPITRE 2

CAS DE DEUX DIMENSIONS ($p = 2$) : DÉTERMINATION DES COMPOSANTES INDÉPENDANTES EN UTILISANT LA COLINÉARITÉ DES ÉLÉMENTS DE $Z(\mathbf{X})$

Ici $E = \text{span}\{\tilde{\mathbf{V}}_1, \tilde{\mathbf{V}}_2\} = \text{span}\{\tilde{\Phi}_1, \tilde{\Phi}_2\}$. Puisque la dimension de l'espace E est deux, l'ensemble $Z(\mathbf{X})$ contient un maximum de deux éléments linéairement indépendants.

2.1 La forme particulière de $Z_n(\mathbf{X})$

Si on pose $p = 2$ et qu'on substitue (1.3) dans (1.12), alors on obtient par indépendance des a_i et a_j pour $n = 0, 1, 2, \dots, m$:

$$\begin{aligned}
 \mathbf{Z}_n(\mathbf{X}) &= \sum_{j=1}^2 E\{a_j(\mathbf{X}^T(a_1\mathbf{V}_1 + a_2\mathbf{V}_2))^n\}\mathbf{V}_j \\
 &= \sum_{j=1}^2 E\{a_j(a_1\mathbf{X}^T\mathbf{V}_1 + a_2\mathbf{X}^T\mathbf{V}_2)^n\}\mathbf{V}_j \\
 &= \sum_{j=1}^2 E\{a_j \sum_{k=0}^n \binom{n}{k} (a_1^k (\mathbf{X}^T\mathbf{V}_1)^k a_2^{n-k} (\mathbf{X}^T\mathbf{V}_2)^{n-k})\}\mathbf{V}_j \\
 &= E\{a_1 \sum_{k=0}^n \binom{n}{k} (a_1^k (\mathbf{X}^T\mathbf{V}_1)^k a_2^{n-k} (\mathbf{X}^T\mathbf{V}_2)^{n-k})\}\mathbf{V}_1 \\
 &\quad + E\{a_2 \sum_{k=0}^n \binom{n}{k} (a_1^k (\mathbf{X}^T\mathbf{V}_1)^k a_2^{n-k} (\mathbf{X}^T\mathbf{V}_2)^{n-k})\}\mathbf{V}_2 \quad (2.1) \\
 &= E\{\sum_{k=0}^n \binom{n}{k} (a_1^{k+1} (\mathbf{X}^T\mathbf{V}_1)^k a_2^{n-k} (\mathbf{X}^T\mathbf{V}_2)^{n-k})\}\mathbf{V}_1 \\
 &\quad + E\{\sum_{k=0}^n \binom{n}{k} (a_1^k (\mathbf{X}^T\mathbf{V}_1)^k a_2^{n-k+1} (\mathbf{X}^T\mathbf{V}_2)^{n-k})\}\mathbf{V}_2 \\
 &= \sum_{k=0}^n \binom{n}{k} (\mathbf{X}^T\mathbf{V}_1)^k (\mathbf{X}^T\mathbf{V}_2)^{n-k} E\{a_1^{k+1}\} E\{a_2^{n-k}\}\mathbf{V}_1 \\
 &\quad + \sum_{k=0}^n \binom{n}{k} (\mathbf{X}^T\mathbf{V}_1)^k (\mathbf{X}^T\mathbf{V}_2)^{n-k} E\{a_1^k\} E\{a_2^{n-k+1}\}\mathbf{V}_2.
 \end{aligned}$$

Si les a_i , en plus d'être des variables aléatoires mutuellement indépendantes et

de moyennes nulles, sont de distributions symétriques, alors pour tout k impair, $E\{a_i^k\} = 0$. Par conséquent, on peut déduire de (2.1) que $\mathbf{Z}_n(\mathbf{X}) = 0$, pour tout n pair.

En général les a_i ne sont pas forcément symétriquement distribuées, mais elles sont de moyennes nulles, donc les moments impairs de a_i ont tendance à être plus petits que les moments pairs. Comme les $\mathbf{Z}_n(\mathbf{X})$ seront plus tard estimés, l'erreur relative sur $\mathbf{Z}_n(\mathbf{X})$ sera plus grande pour n pair que pour n impair. Remarquons que pour n pair, les termes $E\{a_1^{k+1}\}E\{a_2^{n-k}\}$ et $E\{a_1^k\}E\{a_2^{n-k+1}\}$ sont plutôt petits pour tout $k \leq n$, puisque les termes $k+1$ et $n-k$ (respectivement k et $n-k+1$) doivent être l'un pair et l'autre impair lorsque n est pair. Tandis que si n est impair alors les termes $k+1$ et $n-k$ (respectivement k et $n-k+1$) sont tous les deux pairs ou bien les deux impairs. Dû à ces considérations, il est préférable de porter notre attention sur les $\mathbf{Z}_n(\mathbf{X})$ avec n impair dans la suite.

2.2 Analyse des conditions d'orthogonalité

Supposons que $m \geq 4$, c'est-à-dire que $E\{a_1^4\} < \infty$ et $E\{a_2^4\} < \infty$. Puisqu'on considère $\mathbf{Z}_n(\mathbf{X})$ pour n impair et qu'au maximum deux vecteurs linéairement indépendants se trouvent dans $Z(\mathbf{X})$, seulement $\mathbf{Z}_1(\mathbf{X})$ et $\mathbf{Z}_3(\mathbf{X})$ seront considérés au sein de $Z(\mathbf{X})$. D'après les conditions d'orthogonalité (1.10), les vecteurs \mathbf{X}_1 et \mathbf{X}_2 satisfont :

$$\begin{cases} \mathbf{X}_1^T \mathbf{Z}_1(\mathbf{X}_2) = 0 \\ \mathbf{X}_1^T \mathbf{Z}_3(\mathbf{X}_2) = 0 \\ \mathbf{X}_2^T \mathbf{Z}_1(\mathbf{X}_1) = 0 \\ \mathbf{X}_2^T \mathbf{Z}_3(\mathbf{X}_1) = 0 \end{cases} \quad (2.2)$$

D'après (2.2) le vecteur \mathbf{X}_1 , qui appartient à l'espace E de dimension deux, doit être orthogonal à deux vecteurs ($\mathbf{Z}_1(\mathbf{X}_2)$ et $\mathbf{Z}_3(\mathbf{X}_2)$) appartenant au même espace. Si $\mathbf{Z}_1(\mathbf{X}_2)$ et $\mathbf{Z}_3(\mathbf{X}_2)$ ne sont pas colinéaires, alors le seul vecteur \mathbf{X}_1 qui est orthogonal

à ces deux vecteurs simultanément est le vecteur nul. Pour obtenir des solutions non-nulles au problème (2.2), nous chercherons donc les vecteurs \mathbf{X} qui sont tels que $\mathbf{Z}_3(\mathbf{X})$ et $\mathbf{Z}_1(\mathbf{X})$ sont colinéaires, i.e. qui satisfont la condition

$$\mathbf{Z}_3(\mathbf{X}) = \lambda \mathbf{Z}_1(\mathbf{X}), \quad (2.3)$$

pour un certain $\lambda \in \mathbb{R}$, ce qui est un cas particulier du résultat (1.13). On sait qu'au moins deux solutions $\tilde{\mathbf{X}}_1$ et $\tilde{\mathbf{X}}_2$ de (2.3) existent.

Si on pose $n = 1$ dans (2.1), alors on obtient

$$\mathbf{Z}_1(\mathbf{X}) = E\{a_1^2\}(\mathbf{X}^T \mathbf{V}_1) \mathbf{V}_1 + E\{a_2^2\}(\mathbf{X}^T \mathbf{V}_2) \mathbf{V}_2. \quad (2.4)$$

Si on pose $n = 3$ dans (2.1), alors on obtient

$$\begin{aligned} \mathbf{Z}_3(\mathbf{X}) = & [(\mathbf{X}^T \mathbf{V}_1)^2 E\{a_1^4\} + 3(\mathbf{X}^T \mathbf{V}_2)^2 E\{a_1^2\} E\{a_2^2\}] (\mathbf{X}^T \mathbf{V}_1) \mathbf{V}_1 \\ & + [(\mathbf{X}^T \mathbf{V}_2)^2 E\{a_2^4\} + 3(\mathbf{X}^T \mathbf{V}_1)^2 E\{a_1^2\} E\{a_2^2\}] (\mathbf{X}^T \mathbf{V}_2) \mathbf{V}_2. \end{aligned} \quad (2.5)$$

Comme \mathbf{V}_1 et \mathbf{V}_2 sont linéairement indépendants, la condition (2.3) implique :

$$\left\{ \begin{array}{l} \lambda E\{a_1^2\}(\mathbf{X}^T \mathbf{V}_1) = \\ (\mathbf{X}^T \mathbf{V}_1)[(\mathbf{X}^T \mathbf{V}_1)^2 E\{a_1^4\} + 3(\mathbf{X}^T \mathbf{V}_2)^2 E\{a_1^2\} E\{a_2^2\}] \\ \lambda E\{a_2^2\}(\mathbf{X}^T \mathbf{V}_2) = \\ (\mathbf{X}^T \mathbf{V}_2)[(\mathbf{X}^T \mathbf{V}_2)^2 E\{a_2^4\} + 3(\mathbf{X}^T \mathbf{V}_1)^2 E\{a_1^2\} E\{a_2^2\}]. \end{array} \right. \quad (2.6)$$

On peut voir directement que le système (2.6) est satisfait si

$$\mathbf{X}^T \mathbf{V}_1 = 0 \text{ et } \lambda = \frac{(\mathbf{X}^T \mathbf{V}_2)^2 E\{a_2^4\}}{E\{a_2^2\}} \quad (2.7)$$

ou bien si

$$\mathbf{X}^T \mathbf{V}_2 = 0 \text{ et } \lambda = \frac{(\mathbf{X}^T \mathbf{V}_1)^2 E\{a_1^4\}}{E\{a_1^2\}}. \quad (2.8)$$

Si $(\mathbf{X}^T \mathbf{V}_1) \neq 0$, alors la première ligne de (2.6) nous donne

$$\lambda = \frac{(\mathbf{X}^T \mathbf{V}_1)^2 E\{a_1^4\} + 3(\mathbf{X}^T \mathbf{V}_2)^2 E\{a_1^2\} E\{a_2^2\}}{E\{a_1^2\}}. \quad (2.9)$$

En substituant (2.9) dans la deuxième ligne de (2.6), on obtient

$$[(\mathbf{X}^T \mathbf{V}_1)^2 E\{a_1^4\} + 3(\mathbf{X}^T \mathbf{V}_2)^2 E\{a_1^2\} E\{a_2^2\}] E\{a_2^2\} (\mathbf{X}^T \mathbf{V}_2) = \\ E\{a_1^2\} (\mathbf{X}^T \mathbf{V}_2) [(\mathbf{X}^T \mathbf{V}_2)^2 E\{a_2^4\} + 3(\mathbf{X}^T \mathbf{V}_1)^2 E\{a_2^2\} E\{a_1^2\}]. \quad (2.10)$$

Si $(\mathbf{X}^T \mathbf{V}_2) \neq 0$, alors (2.10) implique :

$$(\mathbf{X}^T \mathbf{V}_1)^2 E\{a_2^2\} [E\{a_1^4\} - 3E\{a_1^2\}^2] = (\mathbf{X}^T \mathbf{V}_2)^2 E\{a_1^2\} [E\{a_2^4\} - 3E\{a_2^2\}^2]. \quad (2.11)$$

On remarque que si les a_i suivent une loi normale centrée, alors $E\{a_i^4\} - 3E\{a_i^2\}^2 = 0$, $i = 1, 2$, et donc l'équation (2.11) est vérifiée pour tout \mathbf{X} . Dorénavant, nous supposerons que les a_i ont un coefficient d'aplatissement différent de 3, c'est-à-dire

$$k_i := \frac{E\{a_i^4\}}{E\{a_i^2\}^2} - 3 \neq 0, i = 1, 2, \quad (2.12)$$

ce qui exclut le cas gaussien.

Si $k_1 k_2 < 0$, c'est à dire que k_1 et k_2 sont de signes contraires, alors (2.11) n'a pas de solutions. Autrement, il y a exactement deux vecteurs linéairement indépendants satisfaisant (2.3) (correspondant à (2.7) et (2.8)) et ces deux vecteurs correspondent donc aux deux vecteurs $\bar{\mathbf{X}}_1$ et $\bar{\mathbf{X}}_2$ recherchés. Dans la suite, nous traiterons le cas $k_1 k_2 > 0$.

Si $k_1 k_2 > 0$, alors (2.11) peut s'écrire sous la forme

$$(\mathbf{X}^T \mathbf{V}_2)^2 - (\alpha \mathbf{X}^T \mathbf{V}_1)^2 = 0, \quad (2.13)$$

où α satisfait

$$\alpha^2 = \frac{\frac{E\{a_1^4\} - 3E\{a_1^2\}^2}{E\{a_1^2\}}}{\frac{E\{a_2^4\} - 3E\{a_2^2\}^2}{E\{a_2^2\}}} = \frac{k_1}{k_2} \frac{E\{a_1^2\}}{E\{a_2^2\}} > 0. \quad (2.14)$$

Remarquons que α dépend des coefficients d'aplatissement k_i des a_i , qui s'annulent si les a_i sont des variables aléatoires gaussiennes. Sous la condition (2.12) on a $\alpha \neq 0$. Si a_1 et a_2 sont identiquement distribuées alors $\alpha = 1$. L'équation (2.13) implique successivement

$$(\mathbf{X}^T \mathbf{V}_2 - \alpha \mathbf{X}^T \mathbf{V}_1)(\mathbf{X}^T \mathbf{V}_2 + \alpha \mathbf{X}^T \mathbf{V}_1) = 0 \quad (2.15)$$

$$\Rightarrow \begin{cases} \mathbf{X}^T \mathbf{V}_2 - \alpha \mathbf{X}^T \mathbf{V}_1 = 0 \\ \text{ou} \\ \mathbf{X}^T \mathbf{V}_2 + \alpha \mathbf{X}^T \mathbf{V}_1 = 0 \end{cases} \quad (2.16)$$

$$\Rightarrow \begin{cases} \mathbf{X}^T (\mathbf{V}_2 - \alpha \mathbf{V}_1) = 0 \\ \text{ou} \\ \mathbf{X}^T (\mathbf{V}_2 + \alpha \mathbf{V}_1) = 0 \end{cases} \quad (2.17)$$

Selon les équations (2.7), (2.8) et (2.17), l'équation (2.3) est satisfaite si

$$\left\{ \begin{array}{l} \mathbf{X}^T \mathbf{V}_1 = 0 \\ \text{ou} \\ \mathbf{X}^T \mathbf{V}_2 = 0 \\ \text{ou} \\ \mathbf{X}^T (\mathbf{V}_2 - \alpha \mathbf{V}_1) = 0 \\ \text{ou} \\ \mathbf{X}^T (\mathbf{V}_2 + \alpha \mathbf{V}_1) = 0. \end{array} \right. \quad (2.18)$$

Puisque $\mathbf{X} \in \mathbb{R}^2$ et que $\mathbf{V}_1, \mathbf{V}_2 \in \mathbb{R}^2$, chacune des quatre équations de (2.18) définit sans ambiguïté un candidat satisfaisant la condition de colinéarité (2.3) et la condition (1.2), disons $\mathbf{X}_1, \mathbf{X}_2, \mathbf{X}_3, \mathbf{X}_4 \in \mathbb{R}^2$, on a donc

$$\left\{ \begin{array}{l} \mathbf{X}_1^T \mathbf{V}_1 = 0 \\ \mathbf{X}_2^T \mathbf{V}_2 = 0 \\ \mathbf{X}_3^T (\mathbf{V}_2 - \alpha \mathbf{V}_1) = 0 \\ \mathbf{X}_4^T (\mathbf{V}_2 + \alpha \mathbf{V}_1) = 0. \end{array} \right. \quad (2.19)$$

Lemme 4 Chaque paire de vecteurs parmi $\mathbf{V}_1, \mathbf{V}_2, \mathbf{V}_3 := \mathbf{V}_2 - \alpha \mathbf{V}_1, \mathbf{V}_4 := \mathbf{V}_2 + \alpha \mathbf{V}_1$ est linéairement indépendante. De plus,

$$\mathbf{X}_i^T \mathbf{V}_j \neq 0, \quad i \neq j, \quad i, j = 1, 2, 3, 4.$$

Preuve Prenons $c_1, c_2 \in \mathbb{R}$ et supposons que

$$c_1 \mathbf{V}_1 + c_2 \mathbf{V}_3 = 0.$$

Alors $c_1\mathbf{V}_1 + c_2(\mathbf{V}_2 - \alpha\mathbf{V}_1) = (c_1 - \alpha c_2)\mathbf{V}_1 + c_2\mathbf{V}_2 = 0$. Comme $\mathbf{V}_1, \mathbf{V}_2$ sont linéairement indépendants par hypothèse, alors

$$c_1 - \alpha c_2 = 0, c_2 = 0 \Rightarrow c_1 = c_2 = 0.$$

Donc \mathbf{V}_1 et \mathbf{V}_3 sont linéairement indépendants. On établit le résultat similairement pour les autres paires. Par exemple,

$$\begin{aligned} c_1\mathbf{V}_3 + c_2\mathbf{V}_4 &= 0 \\ \implies c_1(\mathbf{V}_2 - \alpha\mathbf{V}_1) + c_2(\mathbf{V}_2 + \alpha\mathbf{V}_1) &= 0 \\ \implies \alpha(c_2 - c_1)\mathbf{V}_1 + (c_2 + c_1)\mathbf{V}_2 &= 0 \end{aligned}$$

Puisque $\alpha \neq 0$ et que \mathbf{V}_1 et \mathbf{V}_2 sont linéairement indépendants, alors

$$c_1 - c_2 = c_1 + c_2 = 0 \implies c_1 = c_2 = 0.$$

Le fait que $\mathbf{X}_i^T\mathbf{V}_j \neq 0$ si $i \neq j$ découle de l'indépendance linéaire entre \mathbf{V}_i et \mathbf{V}_j dans \mathbb{R}^2 . ■

Le système (2.19) nous indique qu'on a trouvé quatre solutions, alors qu'on en cherche deux. En fait, on sait que les vecteurs cherchés \mathbf{X}_i et \mathbf{X}_j doivent être tels que $\mathbf{X}_i^T\mathbf{S}$ et $\mathbf{X}_j^T\mathbf{S}$ soient statistiquement indépendants.

En utilisant $\mathbf{S} = a_1\mathbf{V}_1 + a_2\mathbf{V}_2$ et le résultat (2.19), les facteurs $\mathbf{X}_i^T\mathbf{S}$, $i = 1, 2, 3, 4$ prennent la forme suivante :

$$\begin{aligned} \mathbf{X}_1^T\mathbf{S} &= a_1\mathbf{X}_1^T\mathbf{V}_1 + a_2\mathbf{X}_1^T\mathbf{V}_2 \\ &= a_2\mathbf{X}_1^T\mathbf{V}_2 \end{aligned} \tag{2.20}$$

$$\begin{aligned}\mathbf{X}_2^T \mathbf{S} &= a_1 \mathbf{X}_2^T \mathbf{V}_1 + a_2 \mathbf{X}_2^T \mathbf{V}_2 \\ &= a_1 \mathbf{X}_2^T \mathbf{V}_1\end{aligned}\tag{2.21}$$

$$\begin{aligned}\mathbf{X}_3^T \mathbf{S} &= a_1 \mathbf{X}_3^T \mathbf{V}_1 + a_2 \mathbf{X}_3^T \mathbf{V}_2 \\ &= (a_1 + \alpha a_2) \mathbf{X}_3^T \mathbf{V}_1\end{aligned}\tag{2.22}$$

$$\begin{aligned}\mathbf{X}_4^T \mathbf{S} &= a_1 \mathbf{X}_4^T \mathbf{V}_1 + a_2 \mathbf{X}_4^T \mathbf{V}_2 \\ &= (a_1 - \alpha a_2) \mathbf{X}_4^T \mathbf{V}_1\end{aligned}\tag{2.23}$$

Par le lemme 4, $\mathbf{X}_1^T \mathbf{V}_2 \neq 0$, $\mathbf{X}_2^T \mathbf{V}_1 \neq 0$, $\mathbf{X}_3^T \mathbf{V}_1 \neq 0$, $\mathbf{X}_4^T \mathbf{V}_1 \neq 0$. L'indépendance statistique entre les $\mathbf{X}_i^T \mathbf{S}$, $i = 1, 2, 3$ et 4 , est donc équivalente à celle des quatre variables aléatoires a_1 , a_2 , $a_3 := (a_1 + \alpha a_2)$ et $a_4 := (a_1 - \alpha a_2)$. Par hypothèse, a_1 et a_2 sont des variables aléatoires indépendantes, ce qui implique l'indépendance de $\mathbf{X}_1^T \mathbf{S}$ et $\mathbf{X}_2^T \mathbf{S}$, alors que les autres paires (a_i, a_j) ne sont pas nécessairement indépendantes.

Approfondissons la structure de dépendance des $\mathbf{X}_i^T \mathbf{S}$. Puisque a_1 et a_2 sont indépendantes, alors :

$$E\{a_1^n a_2^m\} = E\{a_1^k\} E\{a_2^l\}, \forall k, l \in \{1, 2, 3, 4\}.\tag{2.24}$$

On définit $\mathcal{C}_{i,j}$, $i, j = 1, 2, 3, 4$ par

$$\mathcal{C}_{i,j} := E\{(\mathbf{X}_i^T \mathbf{S})^2 (\mathbf{X}_j^T \mathbf{S})^2\} - E\{(\mathbf{X}_i^T \mathbf{S})^2\} E\{(\mathbf{X}_j^T \mathbf{S})^2\}.\tag{2.25}$$

En utilisant les résultats (2.20) à (2.23), le calcul des $\mathcal{C}_{i,j}$, $i, j = 1, 2, 3, 4$, donne les

résultats suivants

$$\left\{ \begin{array}{l} \mathcal{C}_{1,2} = 0 \\ \mathcal{C}_{1,3} = c_{1,3} \text{Var}(a_1^2) \\ \mathcal{C}_{1,4} = c_{1,4} \text{Var}(a_1^2) \\ \mathcal{C}_{2,3} = \alpha^2 c_{2,3} \text{Var}(a_2^2) \\ \mathcal{C}_{2,4} = \alpha^2 c_{2,4} \text{Var}(a_2^2) \\ \mathcal{C}_{3,4} = c_{3,4} [\text{Var}(a_1^2) + \alpha^4 \text{Var}(a_2^2) - 4\alpha^2 \text{Var}(a_1) \text{Var}(a_2)]. \end{array} \right. \quad (2.26)$$

où $c_{1,3} = (\mathbf{X}_1^T \mathbf{V}_2)^2 (\mathbf{X}_3^T \mathbf{V}_1)^2$, $c_{1,4} = (\mathbf{X}_1^T \mathbf{V}_2)^2 (\mathbf{X}_4^T \mathbf{V}_1)^2$, $c_{2,3} = (\mathbf{X}_2^T \mathbf{V}_1)^2 (\mathbf{X}_3^T \mathbf{V}_1)^2$, $c_{2,4} = (\mathbf{X}_2^T \mathbf{V}_1)^2 (\mathbf{X}_4^T \mathbf{V}_1)^2$, $c_{3,4} = (\mathbf{X}_3^T \mathbf{V}_1)^2 (\mathbf{X}_4^T \mathbf{V}_1)^2$. Comme $\alpha \neq 0$, nous remarquons que $\mathcal{C}_{1,3} \neq 0$, $\mathcal{C}_{1,4} \neq 0$, $\mathcal{C}_{2,3} \neq 0$, $\mathcal{C}_{2,4} \neq 0$. Quant à $\mathcal{C}_{3,4}$, on trouve, en substituant (2.14) dans la dernière équation de (2.26), que $\mathcal{C}_{3,4} \neq 0$ si et seulement si

$$k_1^2 E\{a_1^2\}^2 \text{Var}(a_2^2) + k_2^2 E\{a_2^2\}^2 \text{Var}(a_1^2) - 4k_1 k_2 E\{a_1^2\}^2 E\{a_2^2\}^2 \neq 0. \quad (2.27)$$

Nous supposons dorénavant que (2.27) est vérifiée. Par exemple, si a_1 et a_2 sont identiquement distribuées, alors (2.27) se réduit à

$$k_1^3 E\{a_1^2\}^4 \neq 0,$$

ce qui est nécessairement vrai si $k_1 \neq 0$.

2.2.1 Résumé

Les résultats de cette section peuvent se résumer comme suit :

Supposons que $p = 2$ et que $0 < E\{a_i^4\} < \infty$, $i = 1, 2$.

Si $k_1 k_2 = 0$, alors la condition de colinéarité (2.3) sera satisfaite pour tout $\mathbf{X} \in \mathbb{R}^2$.

Si $k_1 k_2 < 0$, alors la condition de colinéarité (2.3) sera satisfaite seulement pour les deux vecteurs cherchés, soient $\bar{\mathbf{X}}_1$ et $\bar{\mathbf{X}}_2$.

Si $k_1 k_2 > 0$, alors la condition de colinéarité (2.3) sera satisfaite pour quatre vecteurs qui sont $\mathbf{X}_1 = \bar{\mathbf{X}}_1, \mathbf{X}_2 = \bar{\mathbf{X}}_2, \mathbf{X}_3, \mathbf{X}_4 \in \mathbb{R}^2$. Si nous supposons que la condition (2.27) est satisfaite, alors on peut distinguer les deux vecteurs recherchés $\bar{\mathbf{X}}_1$ et $\bar{\mathbf{X}}_2$ des quatre vecteurs trouvés $\mathbf{X}_1, \mathbf{X}_2, \mathbf{X}_3, \mathbf{X}_4$ en utilisant les coefficients de corrélation $\mathcal{C}_{i,j}$. Les vecteurs cherchés correspondent à \mathbf{X}_i et $\mathbf{X}_j, i, j = 1, 2, 3, 4$, tels que $\mathcal{C}_{i,j} = 0$.

2.2.2 Algorithme de calcul des composantes indépendantes

Si nous supposons que M observations du signal $\tilde{\mathbf{S}}$ sont disponibles, alors notre algorithme d'estimation des $\bar{\mathbf{X}}_i$ se résume comme suit :

(1) Estimer la matrice de covariance $\tilde{\mathbf{C}}$ du signal avec

$$\hat{\mathbf{C}}_{i,j} = \frac{1}{M} \sum_{k=1}^M \tilde{S}^{(k)}(i) \tilde{S}^{(k)}(j), \forall i, j \in \{1, 2, \dots, N\}$$

où $\tilde{S}^{(k)}(i)$ désigne la i -ème composante de la k -ième réalisation de $\tilde{\mathbf{S}}$. Calculer les valeurs propres et les vecteurs propres de $\hat{\mathbf{C}}$. Par le lemme 1, exactement deux valeurs propres devraient être significativement différentes de zéro. Posons $\hat{E} = \text{span}\{\hat{\Phi}_1, \hat{\Phi}_2\}$ l'espace propre associé.

(2) Remarquons que tout vecteur $\mathbf{X} \in \hat{E}$ s'écrit sous la forme $\mathbf{X} = x_1 \hat{\Phi}_1 + x_2 \hat{\Phi}_2$. Comme $\|\mathbf{X}\| = 1$ et $\hat{\Phi}_1^T \hat{\Phi}_2 = 0$, car $\tilde{\mathbf{C}}^T = \tilde{\mathbf{C}}$, alors $x_1^2 + x_2^2 = 1$ et donc on peut

toujours écrire x_1 et x_2 sous la forme :

$$\begin{cases} x_1 = \cos(\theta) \\ x_2 = \sin(\theta), \end{cases}$$

où $\theta \in [0, 2\pi]$. On cherche donc \mathbf{X} sous la forme

$$\mathbf{X}(\theta) = \cos(\theta) \hat{\Phi}_1 + \sin(\theta) \hat{\Phi}_2, \quad (2.28)$$

pour $\theta \in [0, \pi]$ pour éliminer l'ambiguité du signe.

Posons $f(\theta) := \|\mathbf{z}_3(\mathbf{X}(\theta)) - \mathbf{z}_1(\mathbf{X}(\theta))\|$, où

$$\mathbf{z}_n(\mathbf{X}) := \epsilon \frac{\mathbf{Z}_n(\mathbf{X})}{\|\mathbf{Z}_n(\mathbf{X})\|},$$

pour $n = 1, 3$, où $\epsilon = \pm 1$ est ajusté pour que $\mathbf{z}_n(\mathbf{X})$ satisfasse la condition (1.2). Les vecteurs \mathbf{X} qui vérifient la condition (2.3) satisfont $\mathbf{z}_3(\mathbf{X}) - \mathbf{z}_1(\mathbf{X}) = 0$, d'où $\|\mathbf{z}_3(\mathbf{X}) - \mathbf{z}_1(\mathbf{X})\| = 0$.

On définit

$$\hat{f}(\theta) = \|\hat{\mathbf{z}}_3(\mathbf{X}(\theta)) - \hat{\mathbf{z}}_1(\mathbf{X}(\theta))\|,$$

où $\hat{\mathbf{z}}_n(\mathbf{X}) := \epsilon \frac{\hat{\mathbf{Z}}_n(\mathbf{X})}{\|\hat{\mathbf{Z}}_n(\mathbf{X})\|}$, pour $n = 1, 3$, $\epsilon = \pm 1$ est ajusté pour que $\hat{\mathbf{z}}_n(\mathbf{X})$ satisfasse la condition (1.2) et $\hat{\mathbf{Z}}_n(\mathbf{X}) = \frac{1}{M} \sum_{k=1}^M (\mathbf{X}^T \mathbf{S}^{(k)})^n \mathbf{S}^{(k)}$.

On recherche les zéros de la fonction $\hat{f}(\theta)$ sur $[0, \pi]$. On peut trouver jusqu'à quatre zéros, disons $\theta_1, \theta_2, \theta_3, \theta_4$. Posons $\hat{\mathbf{X}}(\theta_i) = \cos(\theta_i) \hat{\Phi}_1 + \sin(\theta_i) \hat{\Phi}_2$.

(3) Si plus de deux vecteurs $\hat{\mathbf{X}}(\theta_i)$ sont déterminés, estimer $\mathcal{C}_{i,j}$ au moyen de

$$\hat{C}_{i,j} = \frac{1}{M} \sum_{k=1}^M \left(\hat{\mathbf{X}}(\theta_i)^T \mathbf{S}^{(k)} \right)^2 \left(\hat{\mathbf{X}}(\theta_j)^T \mathbf{S}^{(k)} \right)^2 - \left(\frac{1}{M} \sum_{k=1}^M \left(\hat{\mathbf{X}}(\theta_i)^T \mathbf{S}^{(k)} \right)^2 \right) \left(\frac{1}{M} \sum_{k=1}^M \left(\hat{\mathbf{X}}(\theta_j)^T \mathbf{S}^{(k)} \right)^2 \right),$$

où $\mathbf{S}^{(k)}$ désigne la k -ième réalisation de \mathbf{S} . Nous gardons le couple de vecteurs $\hat{\mathbf{X}}(\theta_i), \hat{\mathbf{X}}(\theta_j)$ tel que $|\hat{C}_{i,j}|$ est minimal. Appelons $\hat{\mathbf{X}}_1$ et $\hat{\mathbf{X}}_2$ les deux vecteurs retenus.

(4) Estimer $\hat{\mathbf{V}}_1$ et $\hat{\mathbf{V}}_2$ comme étant les vecteurs de \mathbb{R}^2 orthogonaux respectivement à $\hat{\mathbf{X}}_2 = \begin{pmatrix} \cos(\hat{\theta}_2) \\ \sin(\hat{\theta}_2) \end{pmatrix}$ et $\hat{\mathbf{X}}_1 = \begin{pmatrix} \cos(\hat{\theta}_1) \\ \sin(\hat{\theta}_1) \end{pmatrix}$.

(5) Calculer $\hat{\tilde{\mathbf{V}}}_1$ et $\hat{\tilde{\mathbf{V}}}_2$ avec $\hat{\tilde{\mathbf{V}}}_i = \mathbf{F}^T \hat{\mathbf{V}}_i, i = 1, 2$.

2.3 Expériences numériques

On considère les signaux $\tilde{\mathbf{V}}_1$ et $\tilde{\mathbf{V}}_2$ dans $\mathbb{R}^N (N = 30)$, tels que :

$$\tilde{\mathbf{V}}_i = \frac{\mathbf{W}_i}{\|\mathbf{W}_i\|}, i = 1, 2, \text{ avec :}$$

$$W_1(i) = i, i = 0, 1, 2, \dots, N - 1$$

$$W_2(i) = i^2, i = 0, 1, 2, \dots, N - 1$$

On génère M signaux ($M = 30$) de la forme

$$\tilde{\mathbf{S}} = a_1 \tilde{\mathbf{V}}_1 + a_2 \tilde{\mathbf{V}}_2,$$

où a_1 et a_2 sont deux variables aléatoires uniformes et indépendantes telles que :
 $a_i \sim U[-2.5, 2.5], i = 1, 2$.

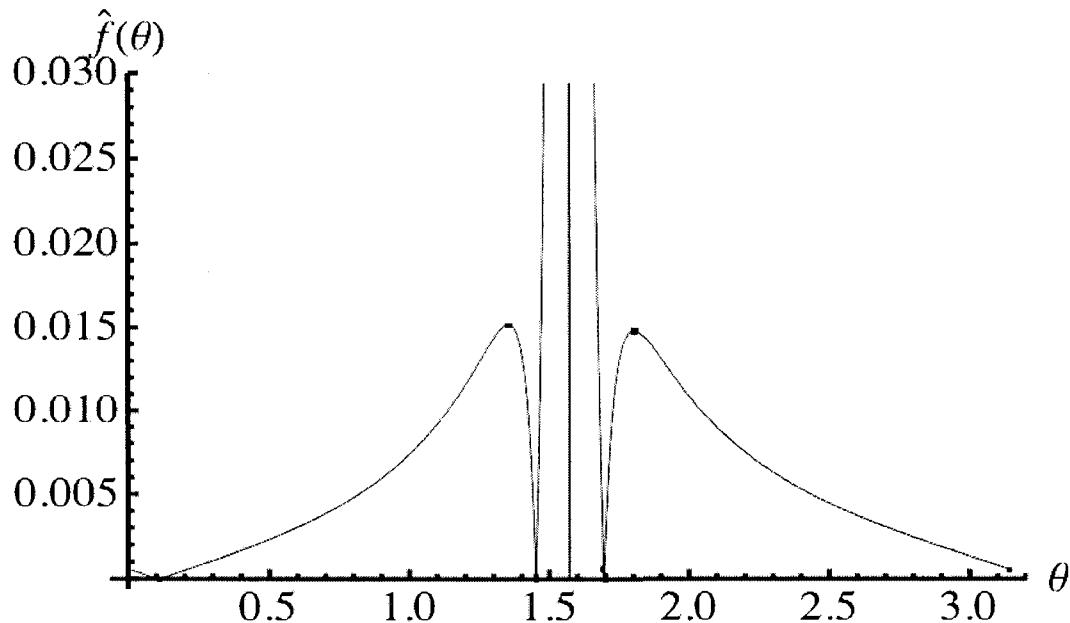


Figure 2.1 Graphe de $\hat{f}(\theta) = \|\hat{\mathbf{z}}_3(\mathbf{X}(\theta)) - \hat{\mathbf{z}}_1(\mathbf{X}(\theta))\|$.

Les cinq premières valeurs propres de $\hat{\mathbf{C}}$ sont :

$$\{3.04615, 0.0367635, -1.7698 \times 10^{-16}, 1.57924 \times 10^{-16}, 1.55897 \times 10^{-16}\}$$

On constate que seulement les deux premières valeurs propres sont significativement différentes de zéro, ce à quoi nous nous attendions.

En traçant la courbe de $\hat{f}(\theta)$ en fonction de θ pour $\theta \in [0, \pi]$ (Figure 2.1), on peut voir clairement les quatre angles $\theta_1 = 0, 107413, \theta_2 = 1, 45173, \theta_3 = 1, 57212, \theta_4 = 1, 69588$ où \hat{f} s'annule.

Pour estimer les quatre valeurs de θ , nous avons fixé un pas de $\Delta\theta = 0.1$ et nous avons échantillonné f dans $[0, \pi]$ puis détecté les minimums apparents de la fonction échantillonnée, qui sont les valeurs θ_0 de θ telles que $f(\theta_0) < f(\theta_0 - \Delta\theta)$ et $f(\theta_0) < f(\theta_0 + \Delta\theta)$. Pour chaque θ_0 , on cherche θ_1 , le minimum local de f dans

l'intervalle $[\theta_0 - \Delta\theta, \theta_0 + \Delta\theta]$. θ_1 est une racine si $f(\theta_1) = 0$. Puisque f admet un maximum de quatre racines dans $[0, \pi]$, si on détecte les quatre racines, l'algorithme s'arrête. Sinon, on divise le pas par cinq et on répète l'algorithme. L'algorithme s'arrête si on trouve les quatre racines, ou bien si $\Delta\theta < 10^{-4}$. Cet algorithme est rapide pour chercher des racines qui sont distantes l'une de l'autre. Il est aussi capable de trouver les racines même si elles sont très proches l'une de l'autre, ce qui est le cas quand f a des pics en forme d'aiguilles (Figure 2.8). Dans ce cas, le calcul est plus long car l'algorithme fera plusieurs itérations avant de converger.

Plus précisément, l'algorithme est le suivant :

```

pas = 0.1;
test = 0;
tant que [ (test == 0) et (pas > 0.0001) ]
    listex={};
    listef={};
    x0=0.0;
    tant que [ x0 <= Pi ]
        si [ ( f(x0) < f(x0-pas) ) et ( f(x0) < f(x0+pas) ) ]
            res=FindMinimum[f(x) , {x , x0-pas , x0+pas}];
            f(x1) = res[[1]];
            x1 = x/.res[[2]];
            si [ f(x1) <= 10^{-10} ]
                listex = Join [ listex, {x1}];
                listef = Join [ listef, {f(x1)}];
                x0=x0+pas;
                nbsol = Length[ listex ];
                si [ nbsol == 4 ]
                    test=1;

```

```
pas = pas/5;
```

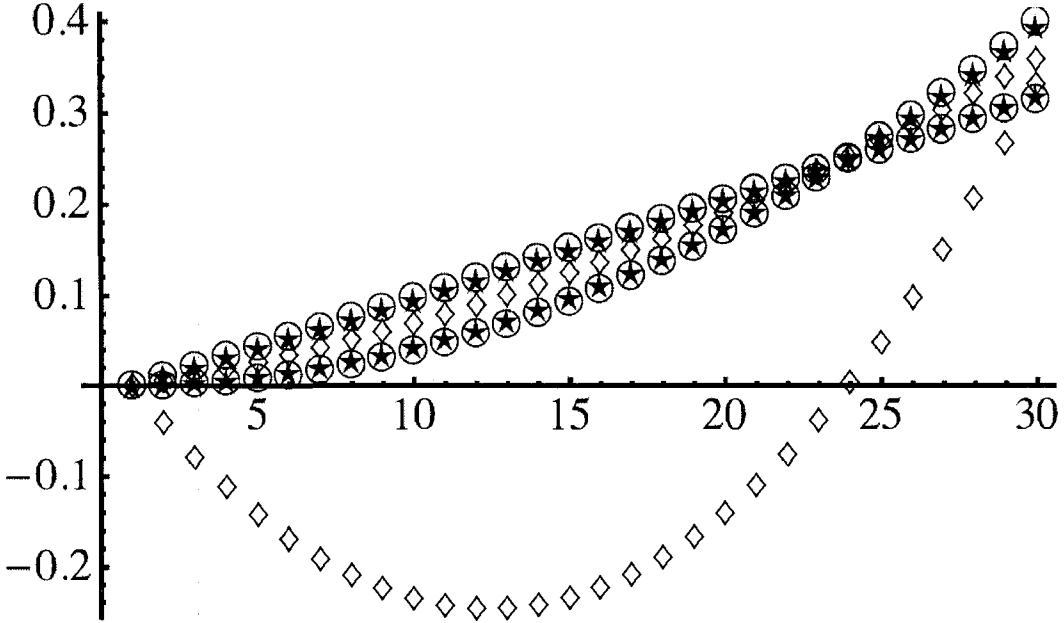


Figure 2.2 Les composantes estimées par notre méthode et les composantes indépendantes.

La figure 2.2 nous montre les composantes indépendantes $\tilde{\mathbf{V}}_1$ et $\tilde{\mathbf{V}}_2$ ainsi que les quatre vecteurs $\hat{\mathbf{V}}_1$, $\hat{\mathbf{V}}_2$, $\hat{\mathbf{V}}_3$ et $\hat{\mathbf{V}}_4$ estimés par notre méthode. Deux de ces vecteurs sont éliminés au moyen des $|\hat{\mathcal{C}}_{i,j}|$. $\hat{\mathbf{V}}_1$ et $\hat{\mathbf{V}}_2$ sont représentés par des étoiles, tandis que $\tilde{\mathbf{V}}_1$ et $\tilde{\mathbf{V}}_2$ sont représentés par des cercles. Les vecteurs $\hat{\mathbf{V}}_3$ et $\hat{\mathbf{V}}_4$, qui sont représentés par des losanges, sont les vecteurs rejetés par notre méthode.

Nous avons voulu tester les limitations pratiques de notre algorithme. Nous avons effectué des tests sur deux paramètres. Le premier paramètre ε contrôle la distance entre les deux composantes. Le deuxième paramètre n permet de faire tendre les lois de a_1 et a_2 vers une loi Gaussienne centrée réduite.

Pour faire varier la distance entre les deux composantes, on a utilisé les signaux

$$\tilde{W}_1(i) = i \text{ et } \tilde{W}_2(i) = i + \varepsilon, i = 0, \dots, N - 1.$$

Les variables aléatoires a_i sont définies par

$$a_i = \frac{1}{\sqrt{n}} \sum_{j=1}^n U_j,$$

où les U_j sont des variables aléatoires i.i.d. telles que $U_j \sim \mathbf{U}[-\sqrt{3}, \sqrt{3}] \forall j$. Comme $E\{U_j\} = 0$ et $\text{Var}(U_j) = 1$, alors $E\{a_i\} = 0$ et $\text{Var}(a_i) = 1$ et la loi de probabilité de a_i tend vers la loi normale centrée réduite quand $n \rightarrow \infty$.

Les figures 2.3, 2.4 et 2.5, nous montrent la variation de $f(\theta)$ en fonction de n et les figures 2.6, 2.7 et 2.8 nous montrent la variation de la courbe $f(\theta)$ en fonction de ε .

Dans les figures 2.3, 2.4 et 2.5, on remarque que lorsque n augmente, l'amplitude de la fonction $f(\theta)$ diminue (i.e la fonction tend vers 0 pour tout θ) et certains pics deviennent plus fins. Avec l'algorithme proposé, nous avons réussi à trouver les racines même dans le cas où $n = 16$. La figure 2.10 nous montre l'agrandissement de la courbe de $f(\theta)$ représentée en figure 2.5, on peut voir clairement les quatre racines de $f(\theta)$.

Dans les figures 2.6, 2.7 et 2.8, on constate que plus ε est petit, plus les pics de la fonction $f(\theta)$ sont rapprochés les uns des autres. Avec l'algorithme proposé, nous avons réussi à trouver les racines même dans le cas où $\varepsilon = 1$ et $n = 1$ (Figure 2.8), mais le calcul est un peu long, parce qu'on fait plusieurs itérations pour que l'algorithme converge. La figure 2.9 représente un agrandissement de la courbe de la fonction $f(\theta)$ représentée dans la figure 2.8, où on peut voir clairement les quatre racines.

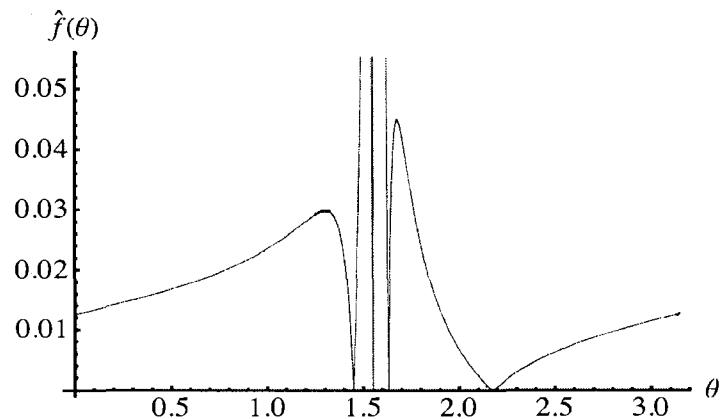


Figure 2.3 La courbe de $\hat{f}(\theta)$ avec $\epsilon = 10$ et $n = 1$.

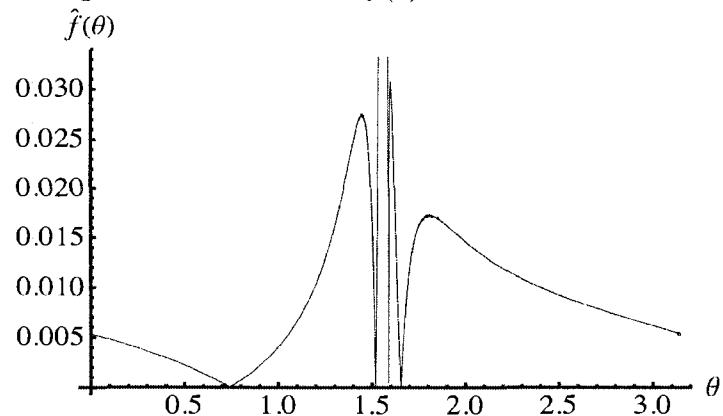


Figure 2.4 La courbe de $\hat{f}(\theta)$ avec $\epsilon = 10$ et $n = 4$.

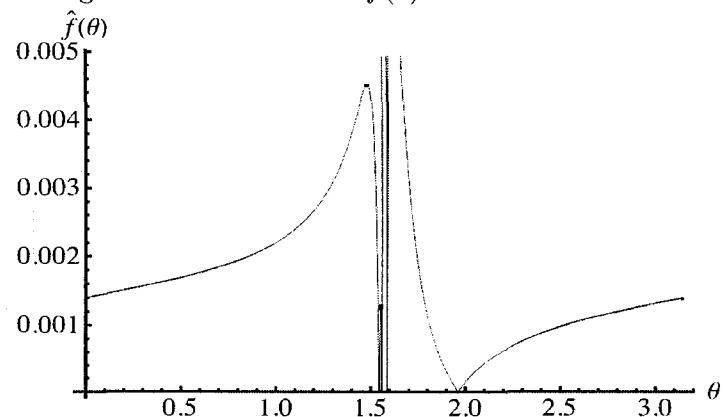
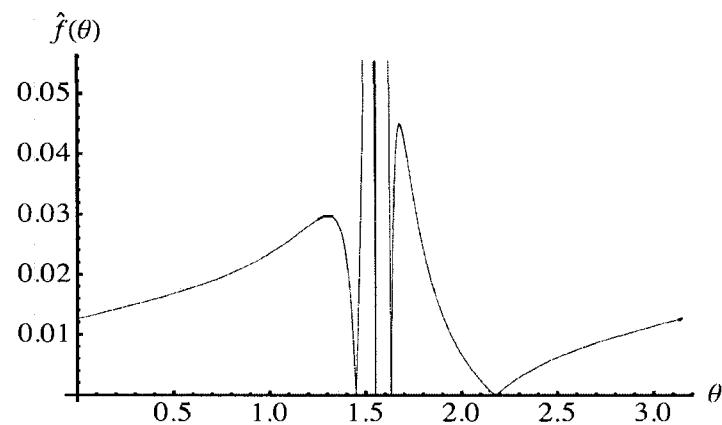
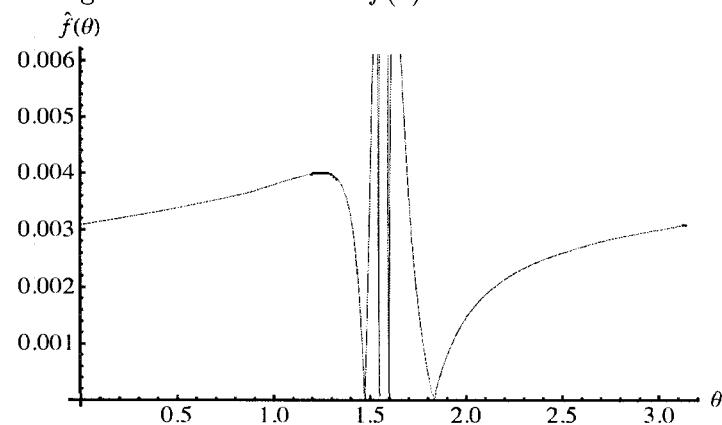
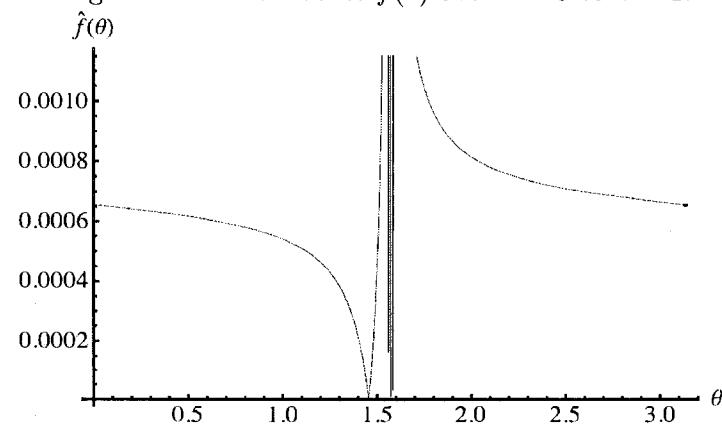


Figure 2.5 La courbe de $\hat{f}(\theta)$ avec $\epsilon = 10$ et $n = 16$.

Figure 2.6 La courbe de $\hat{f}(\theta)$ avec $\epsilon = 10$ et $n = 1$.Figure 2.7 La courbe de $\hat{f}(\theta)$ avec $\epsilon = 5$ et $n = 1$.Figure 2.8 La courbe de $\hat{f}(\theta)$ avec $\epsilon = 1$ et $n = 1$.

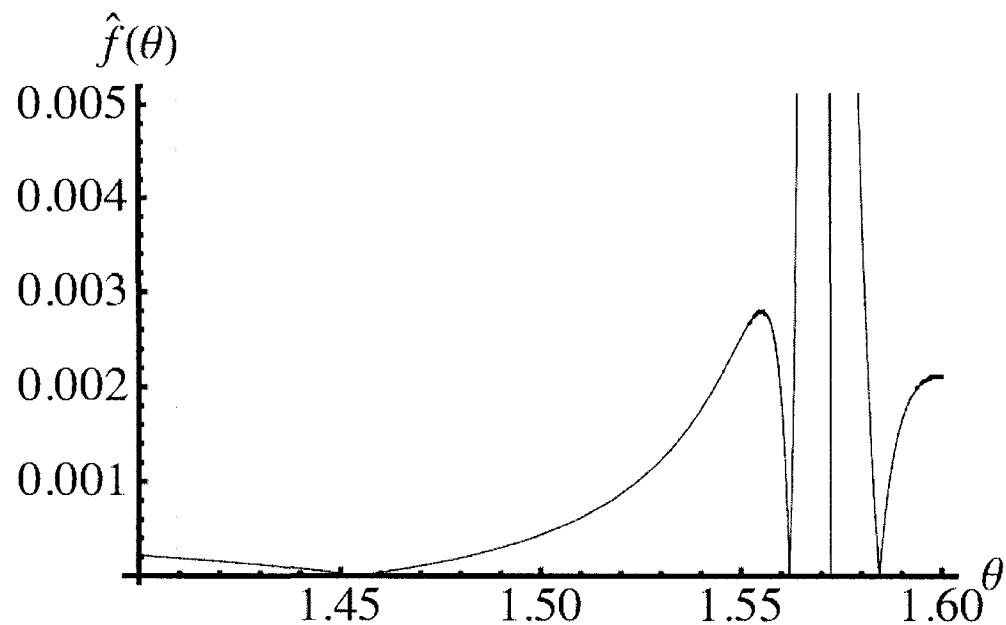


Figure 2.9 Agrandissement de la figure 2.8.

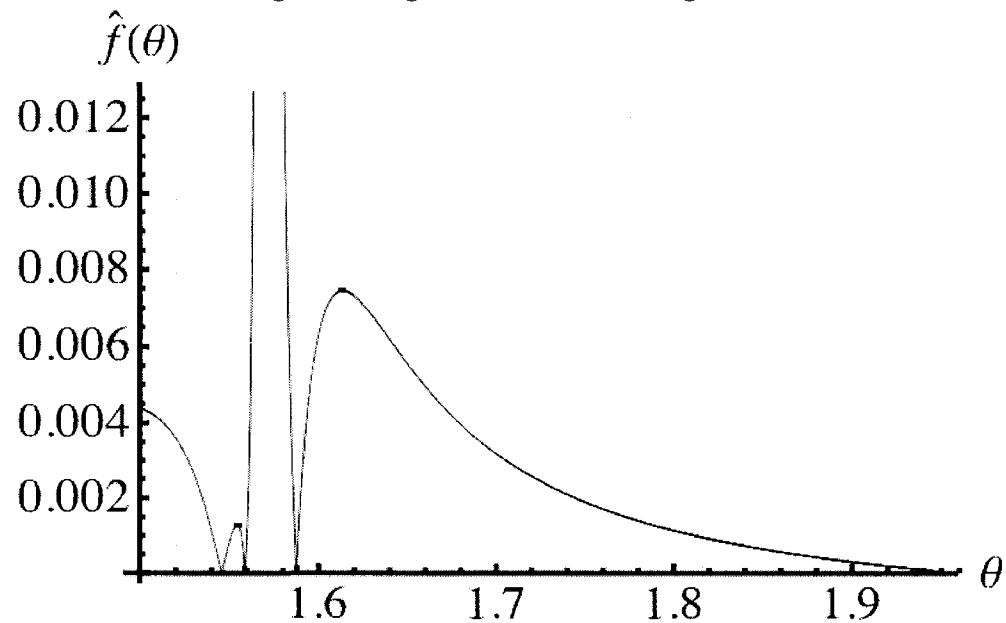
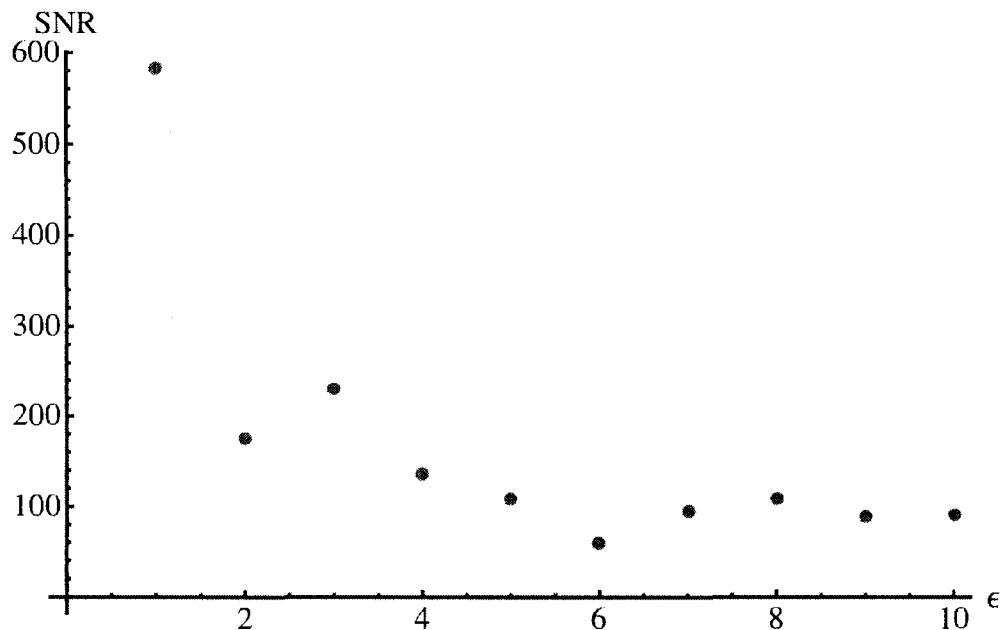
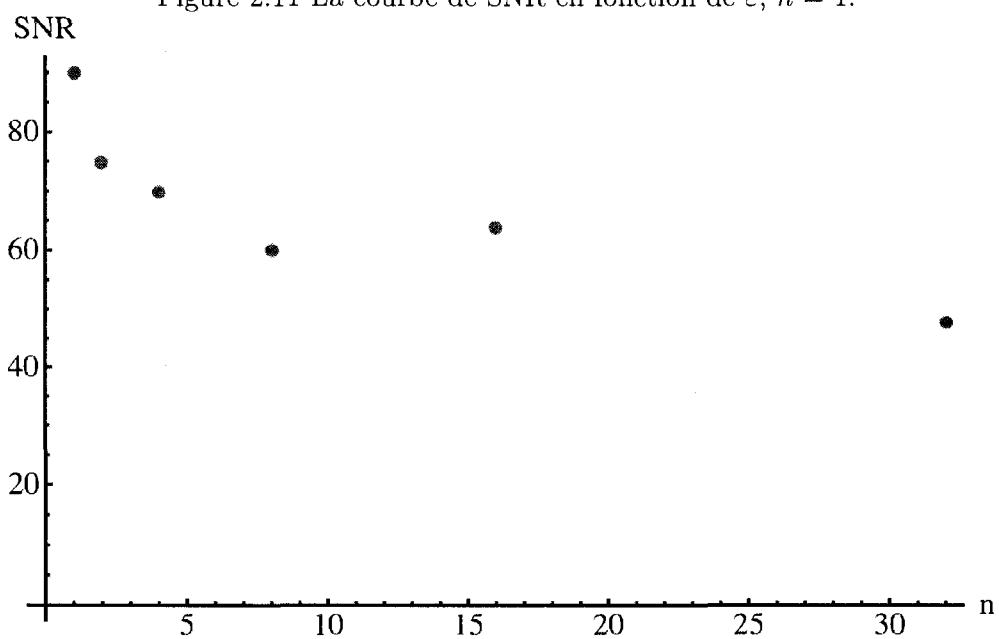


Figure 2.10 Agrandissement de la figure 2.5.

Figure 2.11 La courbe de SNR en fonction de ϵ , $n = 1$.Figure 2.12 La courbe de SNR en fonction de n , $\epsilon = 10$.

On définit le rapport signal-bruit (SNR) de nos vecteurs estimés $\hat{\mathbf{V}}_1$ et $\hat{\mathbf{V}}_2$ par

$$\text{SNR} = \frac{\text{SNR}_1 + \text{SNR}_2}{2},$$

$$\begin{aligned} \text{où, } \text{SNR}_1 &= \frac{\|\hat{\mathbf{V}}_1\|}{\|\hat{\mathbf{V}}_1 - \tilde{\mathbf{V}}_1\|} \\ &= \frac{1}{\|\hat{\mathbf{V}}_1 - \tilde{\mathbf{V}}_1\|} \\ \text{SNR}_2 &= \frac{\|\hat{\mathbf{V}}_2\|}{\|\hat{\mathbf{V}}_2 - \tilde{\mathbf{V}}_2\|} \\ &= \frac{1}{\|\hat{\mathbf{V}}_2 - \tilde{\mathbf{V}}_2\|}. \end{aligned}$$

Nous avons calculé le rapport signal bruit pour les différentes valeurs de ε et de n . Dans la figure 2.11, nous avons fixé $n = 1$ et nous avons utilisé $\varepsilon = 1, 2, \dots, 10$. Pour chaque valeur de ε , nous avons fait 20 essais en générant les M signaux \mathbf{S} ($M = 30$) pour chaque essai et nous avons calculé la médiane des SNR obtenus, qui est représentée dans la figure 2.11. Dans la figure 2.12, nous avons fixé $\varepsilon = 10$ et nous avons utilisé $n = 1, 2, 4, 8, 16, 32$. Pour chaque valeur de n , nous avons fait 20 essais en générant les M signaux \mathbf{S} pour chaque essai et nous avons calculé la médiane des SNR obtenus, qui est représentée dans la figure 2.12.

Dans la figure 2.11 nous remarquons que le SNR est peu sensible à la variation de ε . Nous avons obtenus des valeurs de SNR qui sont généralement de l'ordre de 100, sauf pour $\varepsilon = 1, 2, 3$, où la précision accrue des estimés obtenus est possiblement liée à la grande finesse des pics de la fonction $f(\theta)$. En général, les résultats obtenus sont satisfaisants, car nous avons effectué les essais avec une taille d'échantillon $M = 30$. Avec une taille d'échantillon plus grande on obtient des SNR plus élevés. La figure 2.13 représente les vecteurs obtenus en utilisant $\varepsilon = 10$ et $n = 1$, le SNR correspondant est $\text{SNR} = 83$ avec $\text{SNR}_1 = 110.24$ et $\text{SNR}_2 = 55.68$. On peut voir clairement la superposition des vecteurs estimés avec les composantes indépendantes.

Dans la figure 2.12, nous remarquons que lorsque n augmente, le SNR diminue. Nous remarquons aussi pour n assez grand (≥ 32), $f(\theta)$ sera presque nulle $\forall \theta \in [0, \pi]$ car

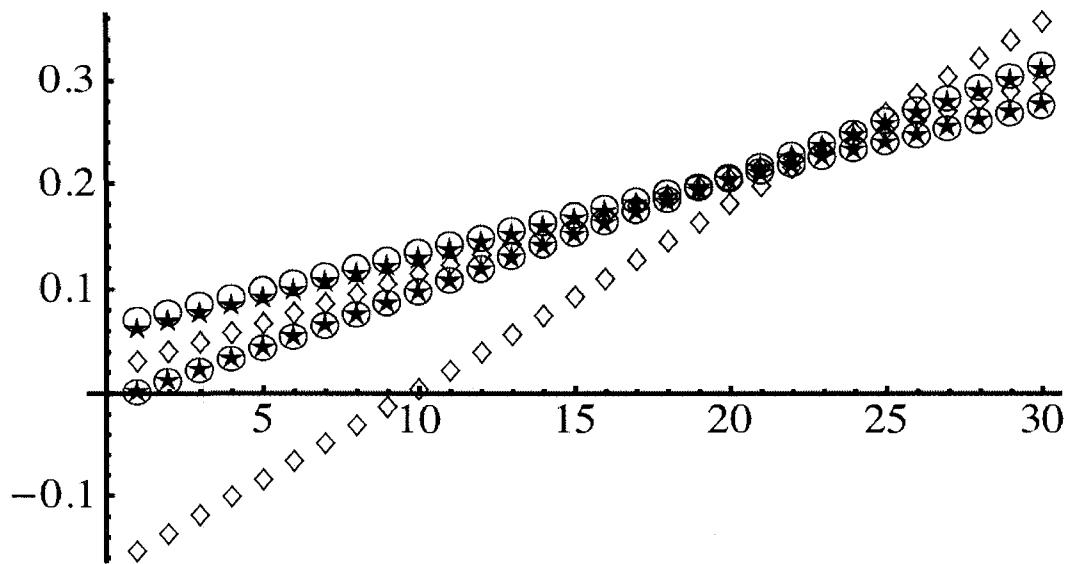


Figure 2.13 Les composantes obtenus avec $\epsilon = 10$ et $n = 1$, SNR=83.

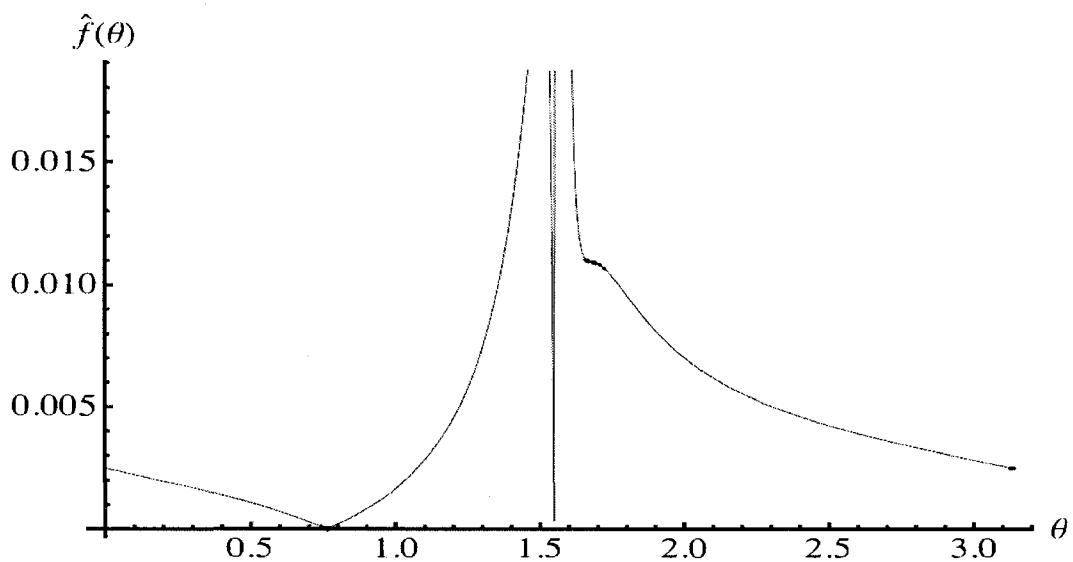


Figure 2.14 La courbe de $\hat{f}(\theta)$ avec $\epsilon = 10$ et $n = 32$.

les $a_i, i = 1, 2$, sont presque gaussiennes (si les $a_i, i = 1, 2$ sont gaussiennes, alors $f(\theta) = 0, \forall \theta$). Si n est assez élevé, alors les variables $a_i, i = 1, 2$ tendent vers la loi gaussienne, et par suite les $k_i = E\{a_i^4\} - 3E\{a_i^2\}^2, i = 1, 2$ seront presque nulles. Dans certains cas, \hat{k}_1 et \hat{k}_2 sont de signes différents et donc la fonction $\hat{f}(\theta)$ a seulement deux racines (figure 2.14).

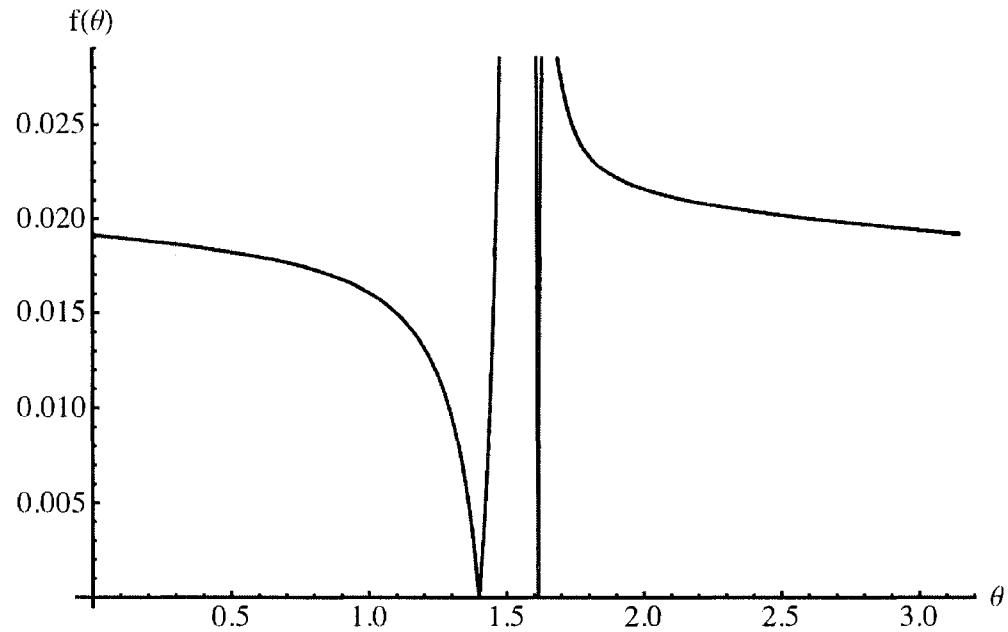
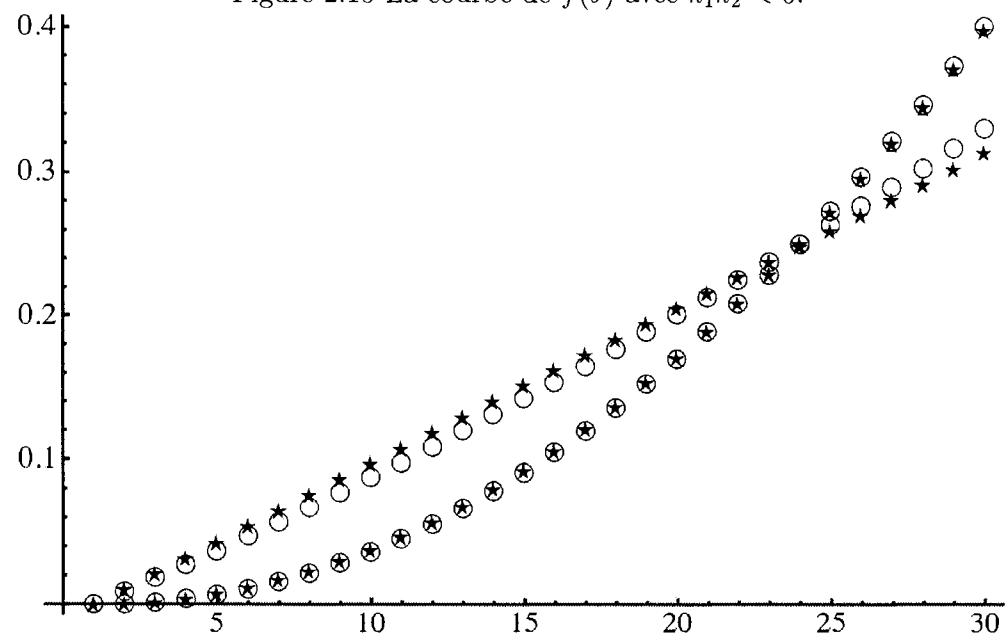
Généralement, si nous avons une taille d'échantillon assez élevée i.e. $M > 30$, notre méthode estime les composantes indépendantes avec un $\text{SNR} > 100$ dans plus que 90% des cas.

Rappelons que toutes les expériences faites étaient simulées en utilisant des variables aléatoires $a_i, i = 1, 2$ de distribution uniforme, donc $k_1 k_2 > 0$. Pour le deuxième cas où $k_1 k_2 < 0$, nous avons généré M signaux ($M = 100$) de la forme

$$\tilde{\mathbf{S}} = a_1 \tilde{\mathbf{V}}_1 + a_2 \tilde{\mathbf{V}}_2,$$

où a_1 et a_2 sont des variables aléatoires indépendantes telles que $a_1 \sim U[-1, 1]$ et $a_2 \sim T$, où T est la loi de Student avec 5 degrés de liberté. Donc $k_1 < 0$ et $k_2 > 0$.

En traçant la courbe de $\hat{f}(\theta)$ en fonction de θ pour $\theta \in [0, \pi]$ (Figure 2.15), on peut voir clairement les deux angles θ_1, θ_2 où la courbe passe par zéro. Nous avons estimé les deux valeurs de θ en utilisant l'algorithme décrit précédemment. Enfin nous avons estimé $\hat{\tilde{\mathbf{V}}}_1$ et $\hat{\tilde{\mathbf{V}}}_2$ et on les a représenté sur le même graphique avec $\tilde{\mathbf{V}}_1$ et $\tilde{\mathbf{V}}_2$ (Figure 2.16), où $\hat{\tilde{\mathbf{V}}}_1$ et $\hat{\tilde{\mathbf{V}}}_2$ sont représentés par des étoiles, tandis que $\tilde{\mathbf{V}}_1$ et $\tilde{\mathbf{V}}_2$ sont représentés par des cercles.

Figure 2.15 La courbe de $\hat{f}(\theta)$ avec $k_1 k_2 < 0$.Figure 2.16 La courbe de $\hat{V}_1, \hat{V}_2, \tilde{V}_1, \tilde{V}_2$ avec $k_1 k_2 < 0$.

CHAPITRE 3

LE CAS 2D : DÉTERMINATION DES COMPOSANTES INDÉPENDANTES PAR LA MINIMISATION D'UNE FONCTION OBJECTIVE

En deux dimensions, nous avons vu dans la première méthode que l'estimation des composantes indépendantes exploitant la colinéarité des $\mathbf{Z}_n(\mathbf{X})$ n'était pas toujours facile. D'une part, les quatre racines de la fonction $f(\theta) := \|\hat{\mathbf{z}}_3(\mathbf{X}(\theta)) - \hat{\mathbf{z}}_1(\mathbf{X}(\theta))\|$ peuvent être difficiles à localiser dans le cas où le graphe de la fonction f contient de fines aiguilles. D'autre part, nous avons vu que seules deux de ces racines sont utiles (on les détermine à l'aide du coefficient de corrélation $\mathcal{C}_{1,2}$), de sorte que les calculs effectués pour déterminer deux des quatre racines sont faits inutilement.

Pour éviter ces inconvénients, nous introduisons dans cette section une seconde méthode d'estimation des composantes indépendantes. Dans cette méthode, l'estimateur des C.I. est défini comme étant le minimum d'une fonction qui tient compte à la fois des conditions d'orthogonalité $\mathbf{X}_i^T \mathbf{Z}_n(\mathbf{X}_j) = 0, i \neq j$, et des conditions de corrélation nulle $\mathcal{C}_{1,2} = 0$. Nous verrons que les deux composantes indépendantes sont obtenues à partir des deux minima d'une fonction d'une seule variable, et que ces minima sont plus faciles à localiser qu'avec la première méthode.

La seconde méthode consiste à trouver les racines de la fonction $g(\mathbf{X}_1, \mathbf{X}_2) \geq 0$ qui est définie par

$$g(\mathbf{X}_1, \mathbf{X}_2) := (\mathbf{X}_1^T \mathbf{z}_1(\mathbf{X}_2))^2 + (\mathbf{X}_1^T \mathbf{z}_3(\mathbf{X}_2))^2 + (\mathbf{X}_2^T \mathbf{z}_3(\mathbf{X}_1))^2 + \gamma (\mathcal{C}_{1,2})^2, \quad (3.1)$$

où $\gamma > 0$ est un coefficient de pondération. Nous remarquons que si les facteurs $\mathbf{X}_1^T \mathbf{S}$ et $\mathbf{X}_2^T \mathbf{S}$ sont indépendants, alors $g(\mathbf{X}_1, \mathbf{X}_2) = 0$. Remarquons aussi que nous n'utilisons

pas le terme $\mathbf{X}_2^T \mathbf{z}_1(\mathbf{X}_1)$, parce que $\mathbf{X}_2^T \mathbf{Z}_1(\mathbf{X}_1) = E\{(\mathbf{X}_2^T \mathbf{S})(\mathbf{X}_1^T \mathbf{S})\} = \mathbf{X}_1^T \mathbf{Z}_1(\mathbf{X}_2)$

3.1 Définition de la méthode II

La deuxième méthode consiste à chercher les racines de la fonction $g(\mathbf{X}_1, \mathbf{X}_2)$. Or $g(\mathbf{X}_1, \mathbf{X}_2) = 0$ implique

$$\begin{cases} \mathbf{X}_1^T \mathbf{Z}_1(\mathbf{X}_2) = 0 \\ \mathbf{X}_1^T \mathbf{Z}_3(\mathbf{X}_2) = 0 \\ \mathbf{X}_2^T \mathbf{Z}_3(\mathbf{X}_1) = 0 \\ \mathcal{C}_{1,2} = 0. \end{cases} \quad (3.2)$$

Nous dénoterons par \mathbf{X}^\perp un vecteur normé orthogonal à \mathbf{X} défini par

$$\mathbf{X}^\perp = \epsilon \frac{R(\mathbf{X})}{\|R(\mathbf{X})\|},$$

où R représente une rotation de $\frac{\pi}{2}$ dans le plan (Φ_1, Φ_2) et le signe de $\epsilon = \pm 1$ est ajusté pour que \mathbf{X}^\perp satisfasse à la condition (1.2).

Nous remarquons que la première ligne de (3.2) implique $\mathbf{X}_1 := (\mathbf{Z}_1(\mathbf{X}_2))^\perp$, or $\mathbf{Z}_1(\mathbf{X}_2) = \mathbf{C}\mathbf{X}_2$. Si on définit la fonction $g_1(\mathbf{X})$ par

$$g_1(\mathbf{X}) := (\mathbf{C}\mathbf{X})^\perp \quad (3.3)$$

et la fonction $g_2(\mathbf{X})$ par

$$\begin{aligned} g_2(\mathbf{X}) &:= g(g_1(\mathbf{X}), \mathbf{X}) \\ &= (\mathbf{X}^T \mathbf{Z}_3(g_1(\mathbf{X}))^2 + (g_1(\mathbf{X})^T \mathbf{Z}_3(\mathbf{X}))^2 + \gamma (\mathcal{C}(\mathbf{X}, g_1(\mathbf{X})))^2), \end{aligned} \quad (3.4)$$

où $\mathcal{C}(\mathbf{X}, \mathbf{Y}) := E\{(\mathbf{X}^T \mathbf{S})^2 (\mathbf{Y}^T \mathbf{S})^2\} - E\{(\mathbf{X}^T \mathbf{S})^2\} E\{(\mathbf{Y}^T \mathbf{S})^2\}$, alors notre méthode se

réduit à chercher les racines de $g_2(\mathbf{X}) = 0$. Comme \mathbf{X} est complètement spécifié par un seul angle θ , i.e. $\mathbf{X} = \mathbf{X}(\theta)$, il suffit maintenant de trouver les zéros de $g_2(\mathbf{X}(\theta))$, qui est une fonction d'une seule variable.

L'algorithme d'estimation des $\bar{\mathbf{X}}_i$ se résume comme suit :

(1) Estimer la matrice de covariance $\hat{\mathbf{C}}$ du signal comme décrit dans la première méthode et calculer les valeurs et les vecteurs propres de $\hat{\mathbf{C}}$. Posons $\hat{E} = \text{span}\{\hat{\Phi}_1, \hat{\Phi}_2\}$.

(2) Rappelons que tout vecteur $\mathbf{X} \in \hat{E}$ s'écrit sous la forme

$$\mathbf{X}(\theta) = \cos(\theta)\hat{\Phi}_1 + \sin(\theta)\hat{\Phi}_2.$$

On estime $g_1(\mathbf{X}(\theta))$ par

$$\hat{g}_1(\mathbf{X}(\theta)) = (\hat{\mathbf{C}}\mathbf{X}(\theta))^\perp.$$

On définit

$$\begin{aligned}\hat{f}(\theta) &= \hat{g}_2(\mathbf{X}(\theta)) \\ &= \hat{g}(\hat{g}_1(\mathbf{X}(\theta)), \mathbf{X}(\theta)) \\ &= (\mathbf{X}(\theta)^T \hat{\mathbf{z}}_3(\hat{g}_1(\mathbf{X}(\theta)))^2 + (\hat{g}_1(\mathbf{X}(\theta))^T \hat{\mathbf{z}}_3(\mathbf{X}(\theta)))^2 + \gamma (\hat{\mathcal{C}}_{1,2}(\mathbf{X}, \hat{g}_1(\mathbf{X})))^2),\end{aligned}$$

où $\hat{\mathbf{z}}_n(\mathbf{X})$ et $\hat{\mathcal{C}}$ sont estimées comme dans la méthode précédente.

(3) Rechercher les zéros de la fonction $\hat{f}(\theta)$ sur $[0, \pi]$. On cherche deux zéros, disons θ_1 et θ_2 . L'existence de ces deux solutions sera justifiée à la section suivante. Posons $\hat{\mathbf{X}}(\theta_i) = \cos(\theta_i)\hat{\Phi}_1 + \sin(\theta_i)\hat{\Phi}_2, i = 1, 2$. Appelons $\hat{\mathbf{X}}_1 = \hat{\mathbf{X}}(\theta_1)$ et $\hat{\mathbf{X}}_2 = \hat{\mathbf{X}}(\theta_2)$.

(4) Estimer $\hat{\mathbf{V}}_1$ et $\hat{\mathbf{V}}_2$ comme étant les vecteurs de \hat{E} orthogonaux respectivement à $\hat{\mathbf{X}}_1$ et $\hat{\mathbf{X}}_2$.

(5) Calculer $\hat{\mathbf{V}}_1$ et $\hat{\mathbf{V}}_2$ avec $\hat{\mathbf{V}}_i = \mathbf{F}^T \hat{\mathbf{V}}_i, i = 1, 2$.

3.2 Analyse de la méthode II

Dans cette section, on utilise la fonction $f(\theta)$ dont les trois premiers termes sont les mêmes qu'on a utilisé à la première méthode, mais on a ajouté un quatrième terme $\mathcal{C}_{1,2}$ qu'on a déjà utilisé pour identifier les deux composantes indépendantes parmi les quatre trouvés en utilisant la colinéarité de l'ensemble $Z(\mathbf{X})$. Nous allons montrer que $f(\theta)$ ne possède que deux racines qui correspondent à $\bar{\mathbf{X}}_1$ et $\bar{\mathbf{X}}_2$.

Comme

$$\begin{cases} \mathbf{X}_1^T \mathbf{Z}_1(\mathbf{X}_2) = E\{(\mathbf{X}_1^T \mathbf{S})(\mathbf{X}_2^T \mathbf{S})\} \\ \mathbf{X}_1^T \mathbf{Z}_3(\mathbf{X}_2) = E\{(\mathbf{X}_1^T \mathbf{S})(\mathbf{X}_2^T \mathbf{S})^3\} \\ \mathbf{X}_2^T \mathbf{Z}_3(\mathbf{X}_1) = E\{(\mathbf{X}_2^T \mathbf{S})(\mathbf{X}_1^T \mathbf{S})^3\} \\ \mathcal{C}_{1,2} = E\{(\mathbf{X}_1^T \mathbf{S})^2(\mathbf{X}_2^T \mathbf{S})^2\} - E\{(\mathbf{X}_1^T \mathbf{S})^2\}E\{(\mathbf{X}_2^T \mathbf{S})^2\}, \end{cases}$$

alors (3.2) implique

$$\begin{cases} E\{(\mathbf{X}_1^T \mathbf{S})(\mathbf{X}_2^T \mathbf{S})\} = 0 \\ E\{(\mathbf{X}_1^T \mathbf{S})(\mathbf{X}_2^T \mathbf{S})^3\} = 0 \\ E\{(\mathbf{X}_2^T \mathbf{S})(\mathbf{X}_1^T \mathbf{S})^3\} = 0 \\ E\{(\mathbf{X}_1^T \mathbf{S})^2(\mathbf{X}_2^T \mathbf{S})^2\} - E\{(\mathbf{X}_1^T \mathbf{S})^2\}E\{(\mathbf{X}_2^T \mathbf{S})^2\} = 0. \end{cases} \quad (3.5)$$

Si on substitue (1.3) dans (3.5) et puisque a_1 et a_2 sont indépendants et que $E\{a_1\} = E\{a_2\} = 0$, alors (3.5) implique

$$\left\{ \begin{array}{l} E\{a_1^2\}(\mathbf{X}_1^T \mathbf{V}_1)(\mathbf{X}_2^T \mathbf{V}_1) + E\{a_2^2\}(\mathbf{X}_1^T \mathbf{V}_2)(\mathbf{X}_2^T \mathbf{V}_2) = 0 \\ \\ E\{a_1^4\}(\mathbf{X}_1^T \mathbf{V}_1)(\mathbf{X}_2^T \mathbf{V}_1)^3 + 3E\{a_1^2\}E\{a_2^2\}(\mathbf{X}_1^T \mathbf{V}_1)(\mathbf{X}_2^T \mathbf{V}_1)(\mathbf{X}_2^T \mathbf{V}_2)^2 \\ + E\{a_2^4\}(\mathbf{X}_1^T \mathbf{V}_2)(\mathbf{X}_2^T \mathbf{V}_2)^3 + 3E\{a_1^2\}E\{a_2^2\}(\mathbf{X}_1^T \mathbf{V}_2)(\mathbf{X}_2^T \mathbf{V}_2)(\mathbf{X}_2^T \mathbf{V}_1)^2 = 0 \\ \\ E\{a_1^4\}(\mathbf{X}_2^T \mathbf{V}_1)(\mathbf{X}_1^T \mathbf{V}_1)^3 + 3E\{a_1^2\}E\{a_2^2\}(\mathbf{X}_2^T \mathbf{V}_1)(\mathbf{X}_1^T \mathbf{V}_1)(\mathbf{X}_1^T \mathbf{V}_2)^2 \\ + E\{a_2^4\}(\mathbf{X}_2^T \mathbf{V}_2)(\mathbf{X}_1^T \mathbf{V}_2)^3 + 3E\{a_1^2\}E\{a_2^2\}(\mathbf{X}_2^T \mathbf{V}_2)(\mathbf{X}_1^T \mathbf{V}_2)(\mathbf{X}_1^T \mathbf{V}_1)^2 = 0 \\ \\ (\mathbf{X}_1^T \mathbf{V}_1)^2(\mathbf{X}_2^T \mathbf{V}_1)^2(E\{a_1^4\} - E\{a_1^2\}^2) \\ + (\mathbf{X}_1^T \mathbf{V}_2)^2(\mathbf{X}_2^T \mathbf{V}_2)^2(E\{a_2^4\} - E\{a_2^2\}^2) \\ + 4E\{a_1^2\}E\{a_2^2\}(\mathbf{X}_1^T \mathbf{V}_1)(\mathbf{X}_2^T \mathbf{V}_1)(\mathbf{X}_1^T \mathbf{V}_2)(\mathbf{X}_2^T \mathbf{V}_2) = 0. \end{array} \right. \quad (3.6)$$

On peut vérifier que le système (3.6) est satisfait si

$$\left\{ \begin{array}{l} \mathbf{X}_1^T \mathbf{V}_1 = 0 \\ \text{ou} \\ \mathbf{X}_2^T \mathbf{V}_1 = 0 \end{array} \right. \quad \text{et} \quad \left\{ \begin{array}{l} \mathbf{X}_1^T \mathbf{V}_2 = 0 \\ \text{ou} \\ \mathbf{X}_2^T \mathbf{V}_2 = 0 \end{array} \right. \quad (3.7)$$

Puisque \mathbf{V}_1 et \mathbf{V}_2 sont linéairement indépendants, alors les solutions non nulles de (3.7) sont

$$\left\{ \begin{array}{l} \mathbf{X}_1^T \mathbf{V}_1 = 0 \\ \mathbf{X}_2^T \mathbf{V}_2 = 0 \end{array} \right. \quad \text{ou} \quad \left\{ \begin{array}{l} \mathbf{X}_1^T \mathbf{V}_2 = 0 \\ \mathbf{X}_2^T \mathbf{V}_1 = 0, \end{array} \right. \quad (3.8)$$

qui correspondent à $\bar{\mathbf{X}}_1$ et $\bar{\mathbf{X}}_2$.

Si on définit les variables x, y, z et t par :

$$\begin{aligned} x &= \mathbf{X}_1^T \mathbf{V}_1 \\ y &= \mathbf{X}_1^T \mathbf{V}_2 \\ z &= \mathbf{X}_2^T \mathbf{V}_1 \\ t &= \mathbf{X}_2^T \mathbf{V}_2, \end{aligned} \quad (3.9)$$

alors le système (3.6) peut s'écrire sous la forme plus lisible

$$\left\{ \begin{array}{l} E\{a_1^2\}xz + E\{a_2^2\}yt = 0 \quad (1) \\ E\{a_1^4\}xz^3 + 3E\{a_1^2\}E\{a_2^2\}ytz^2 \quad (2) \\ + E\{a_2^4\}yt^3 + 3E\{a_1^2\}E\{a_2^2\}xzt^2 = 0 \quad (3) \\ E\{a_1^4\}zx^3 + 3E\{a_1^2\}E\{a_2^2\}tyx^2 \quad (4) \\ + E\{a_2^4\}ty^3 + 3E\{a_1^2\}E\{a_2^2\}zxy^2 = 0 \quad (5) \\ x^2z^2(E\{a_1^4\} - E\{a_1^2\}^2) \quad (6) \\ + y^2t^2(E\{a_2^4\} - E\{a_2^2\}^2) \quad (7) \\ + 4E\{a_1^2\}E\{a_2^2\}xyzt = 0. \quad (8) \end{array} \right. \quad (3.10)$$

Si $x \neq 0, y \neq 0, z \neq 0, t \neq 0$ et qu'on multiplie la première équation de (3.10) par $E\{a_1^2\}$ ou $E\{a_2^2\}$, on montre facilement que

$$E\{a_1^2\}E\{a_2^2\} = -E\{a_1^2\}^2 \frac{xz}{yt} \quad (3.11)$$

$$= -E\{a_2^2\}^2 \frac{yt}{xz}. \quad (3.12)$$

En sommant (3.11) et (3.12), on obtient aussi

$$E\{a_1^2\}E\{a_2^2\} = -\frac{1}{2}E\{a_1^2\}^2 \frac{xz}{yt} - \frac{1}{2}E\{a_2^2\}^2 \frac{yt}{xz}. \quad (3.13)$$

On peut faire apparaître k_1 et k_2 dans (3.10) grâce aux substitutions suivantes : on utilise (3.11) dans les lignes (2) et (4), (3.12) dans les lignes (3) et (5), et finalement (3.13) dans la ligne (8), ce qui conduit à

$$\begin{cases} \frac{E\{a_1^2\}}{E\{a_2^2\}} = -\frac{yt}{xz} \\ zx^3k_1E\{a_1^2\}^2 + ty^3k_2E\{a_2^2\}^2 = 0 \\ xz^3k_1E\{a_1^2\}^2 + yt^3k_2E\{a_2^2\}^2 = 0 \\ x^2z^2k_1E\{a_1^2\}^2 + y^2t^2k_2E\{a_2^2\}^2 = 0. \end{cases} \quad (3.14)$$

On remarque que si les $a_i, i = 1, 2$, ont un coefficient d'aplatissement égal à 3, ce qui est le cas de la loi gaussienne, alors $k_1 = 0$ et $k_2 = 0$ et les trois dernières équations du système (3.14) sont vérifiées pour tout couple $(\mathbf{X}_1, \mathbf{X}_2) \in \mathbb{R}^2 \times \mathbb{R}^2$. Le système (3.14) est donc vérifié pour tout couple $(\mathbf{X}_1, \mathbf{X}_2) \in \mathbb{R}^2 \times \mathbb{R}^2$ vérifiant $\mathbf{X}_1^T \mathbf{C} \mathbf{X}_2 = 0$. Si $k_1 k_2 > 0$, alors la quatrième équation de (3.14) n'a pas de solutions non-nulles. Les seules solutions de $g(\mathbf{X}_1, \mathbf{X}_2) = 0$ sont donc définies par (3.8) qui sont $\bar{\mathbf{X}}_1$ et $\bar{\mathbf{X}}_2$.

Si $k_1 k_2 < 0$, alors les trois dernières lignes de (3.14) peuvent s'écrire sous la forme

$$xz^3|k_1|E\{a_1^2\}^2 = yt^3|k_2|E\{a_2^2\}^2 \quad (3.15)$$

$$x^3z|k_1|E\{a_1^2\}^2 = y^3t|k_2|E\{a_2^2\}^2 \quad (3.16)$$

$$x^2z^2|k_1|E\{a_1^2\}^2 = y^2t^2|k_2|E\{a_2^2\}^2 \quad (3.17)$$

Les équations (3.15), (3.16) et (3.17) sont équivalentes à

$$\left| \frac{k_1}{k_2} \right| \left(\frac{E\{a_1^2\}}{E\{a_2^2\}} \right)^2 = \frac{yt^3}{xz^3} = \frac{y^3t}{x^3z} = \frac{y^2t^2}{x^2z^2},$$

d'où on tire

$$xt = yz. \quad (3.18)$$

De (3.18), on obtient $y = \frac{xt}{z}$ qu'on substitue dans la première équation de (3.14) pour obtenir

$$\left(\frac{t}{z} \right)^2 = -\frac{E\{a_1^2\}}{E\{a_2^2\}}$$

qui n'a pas de solutions non-nulles.

Il n'y a donc pas de solutions non-nulles si $k_1 k_2 < 0$.

Ces résultats peuvent se résumer comme suit :

- Si $k_1 k_2 = 0$, alors $g_2(\mathbf{X}) = 0$ admet une infinité de solutions.
- Si $k_1 k_2 \neq 0$, alors $g_2(\mathbf{X}) = 0$ admet deux racines qui correspondent à $\bar{\mathbf{X}}_1$ et $\bar{\mathbf{X}}_2$.

3.3 Expériences numériques

On considère les signaux $\tilde{\mathbf{V}}_1$ et $\tilde{\mathbf{V}}_2$ dans $\mathbb{R}^N (N = 30)$, tels que

$$\tilde{\mathbf{V}}_i = \frac{\mathbf{W}_i}{\|\mathbf{W}_i\|},$$

$i = 1, 2$, où $W_1(j) = j$, $W_2(j) = j^2$ et $j = 0, 1, 2, \dots, N - 1$.

On génère M signaux ($M = 30$) de la forme

$$\tilde{\mathbf{S}} = a_1 \tilde{\mathbf{V}}_1 + a_2 \tilde{\mathbf{V}}_2,$$

où a_1 et a_2 sont deux variables aléatoires uniformes et indépendantes telles que $a_i \sim U[-1, 1], i = 1, 2$. Comme a_1 et a_2 sont identiquement distribuées, alors $k_1 k_2 > 0$.

Les cinq premières valeurs propres de $\hat{\mathbf{C}}$ sont :

$$\{0.660424, 0.0113689, 7.10271 \times 10^{-17}, -4.60542 \times 10^{-17}, 2.54675 \times 10^{-17}\}$$

On constate que seules les deux premières valeurs propres sont significativement différentes de zéro, ce à quoi nous nous attendions.

Pour choisir la valeur de γ apparaissant dans la fonction g (formule 3.1), nous avons

comparé les valeurs des trois termes de la fonction \hat{f} . Nous avons donc effectué des essais avec des échantillons différents et des valeurs de γ différentes et nous avons comparé l'influence de chacun des termes dans la courbe de $\hat{f}(\theta), \theta \in [0, \pi]$. Nous avons remarqué que si on prend $\gamma < 10$, alors le terme $\mathcal{C}_{1,2}$ n'a pas d'influence et les autres termes dominent la courbe de la fonction \hat{f} , tandis que si on prend $\gamma > 30$, alors $\mathcal{C}_{1,2}$ domine la courbe de \hat{f} et les autres termes n'ont pas d'influence. Nous avons donc choisi $\gamma = 20$ pour que tous les termes aient une grandeur comparable.

En traçant la courbe de $\hat{f}(\theta)$ en fonction de θ pour $\theta \in [0, \pi]$ pour deux échantillons différents (Figures 3.1 et 3.2), on peut voir que ces courbes admettent quatre minimums locaux, mais seulement deux de ces minimums représentent des points où \hat{f} s'annule, soient $\hat{\theta}_1$ et $\hat{\theta}_2$.

Pour estimer $\hat{\theta}_1$ et $\hat{\theta}_2$, nous avons fixé un pas de $\Delta\theta = 0.01$ et nous avons échantillonné \hat{f} dans $[0, \pi]$ puis détecté les minimums apparents de la fonction échantillonnée, qui sont les valeurs θ_0 de θ telles que $\hat{f}(\theta_0) < \hat{f}(\theta_0 - \Delta\theta)$ et $\hat{f}(\theta_0) < \hat{f}(\theta_0 + \Delta\theta)$. Pour chaque θ_0 , on cherche θ_1 , le minimum local de \hat{f} dans l'intervalle $[\theta_0 - \Delta\theta, \theta_0 + \Delta\theta]$ en utilisant une méthode numérique. θ_1 est une racine si $\hat{f}'(\theta_1) = 0$. Puisque \hat{f} admet exactement deux racines dans $[0, \pi]$, on garde les deux valeurs de θ_1 qui ont les valeurs minimales de $\hat{f}(\theta)$.

Nous avons estimé $\hat{\mathbf{X}}_1$ et $\hat{\mathbf{X}}_2$ avec

$$\hat{\mathbf{X}}_i := \cos(\hat{\theta}_i) \hat{\Phi}_1 + \sin(\hat{\theta}_i) \hat{\Phi}_2, i = 1, 2.$$

Enfin $\hat{\mathbf{V}}_1 = \hat{\mathbf{X}}_2^\perp$ et $\hat{\mathbf{V}}_2 = \hat{\mathbf{X}}_1^\perp$ dans le sous-espace \hat{E} . La figure 3.3 nous montre la superposition des vecteurs estimés par la deuxième méthode avec les composantes indépendantes. $\hat{\mathbf{V}}_1$ et $\hat{\mathbf{V}}_2$ sont représentés par des cercles, tandis que $\tilde{\mathbf{V}}_1$ et $\tilde{\mathbf{V}}_2$ sont représentés par des étoiles.

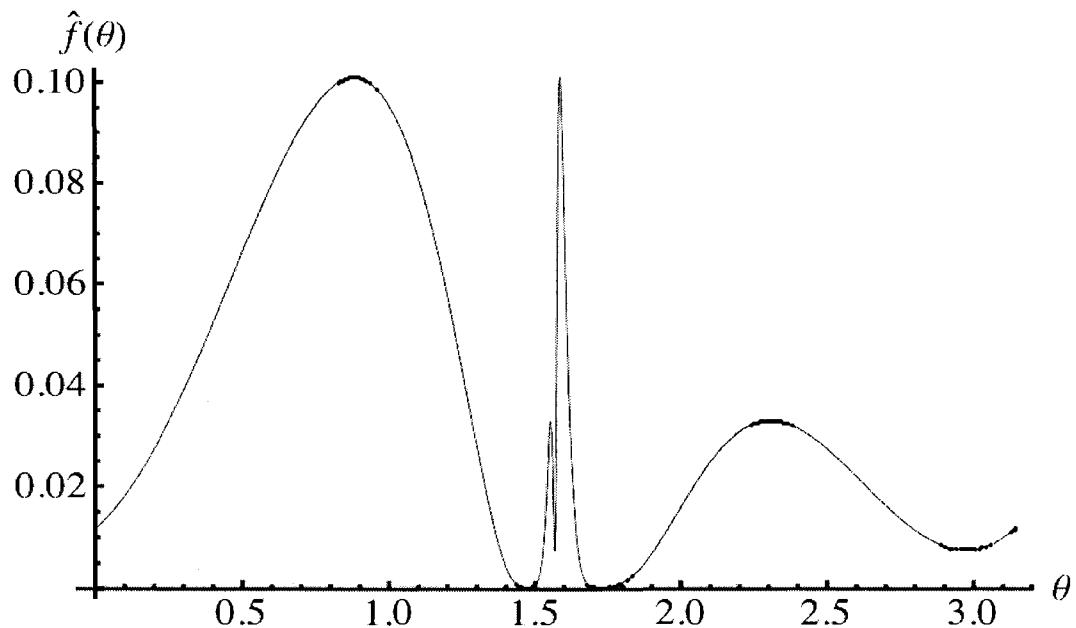


Figure 3.1 Graphe de $\hat{f}(\theta) = \hat{g}(\hat{g}_1(\mathbf{X}(\theta)), \mathbf{X}(\theta))$ dans un cas où $k_1 k_2 > 0$ pour l'échantillon 1.

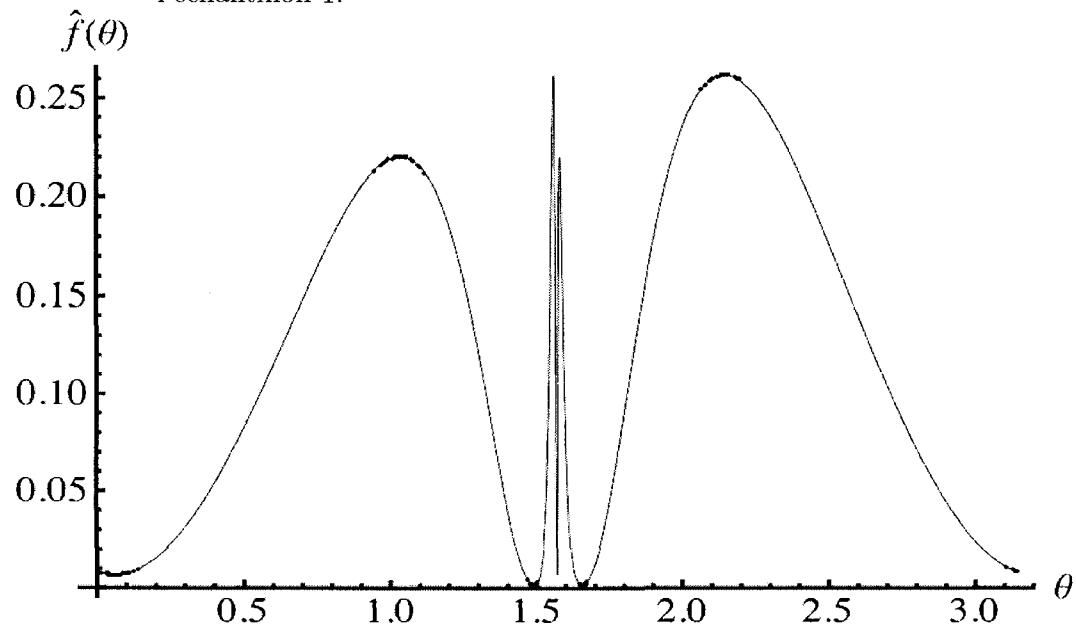


Figure 3.2 Graphe de $\hat{f}(\theta) = \hat{g}(\hat{g}_1(\mathbf{X}(\theta)), \mathbf{X}(\theta))$ dans un cas où $k_1 k_2 > 0$ pour l'échantillon 2.

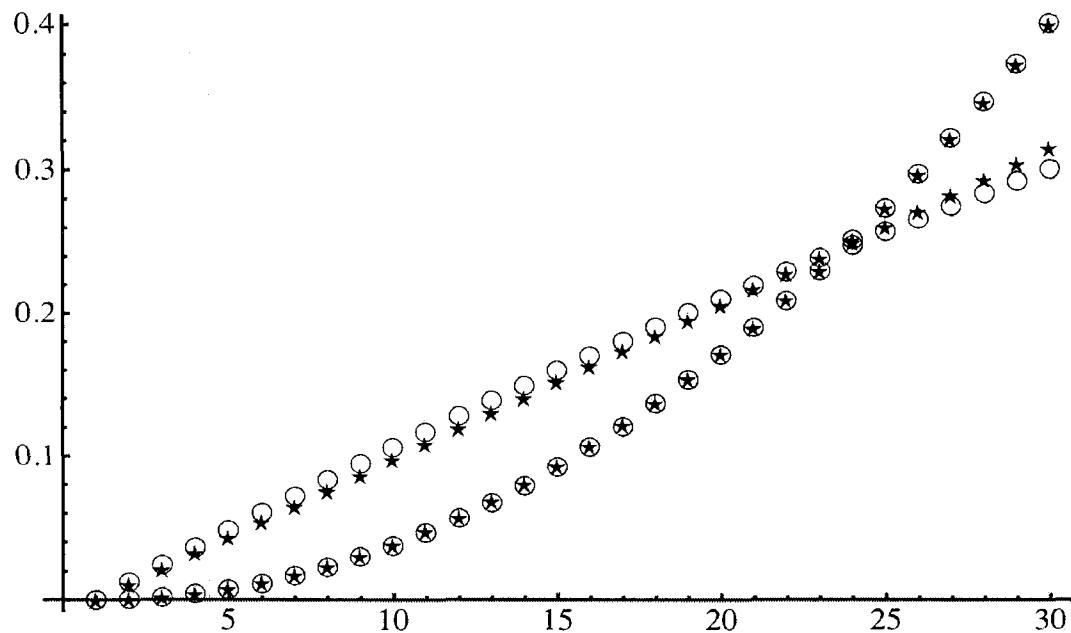


Figure 3.3 Les composantes estimées par la deuxième méthode (○) et les composantes indépendantes (★), dans un cas où $k_1 k_2 > 0$ pour l'échantillon 1.

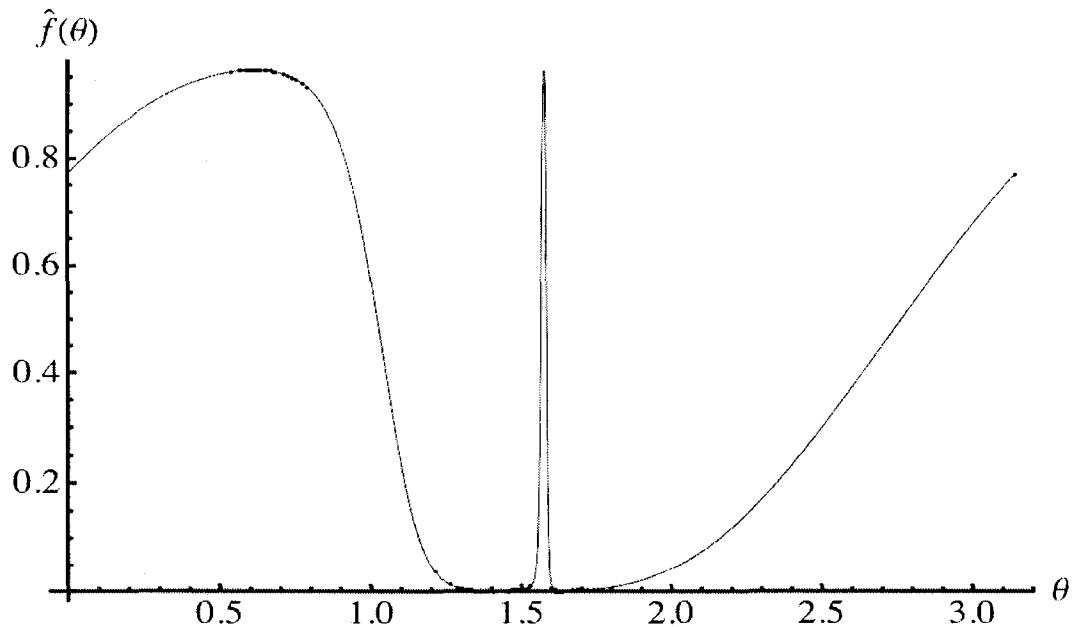


Figure 3.4 Graphe de $\hat{f}(\theta) = \hat{g}(\hat{g}_1(\mathbf{X}(\theta)), \mathbf{X}(\theta))$ pour un cas où $k_1 k_2 < 0$.

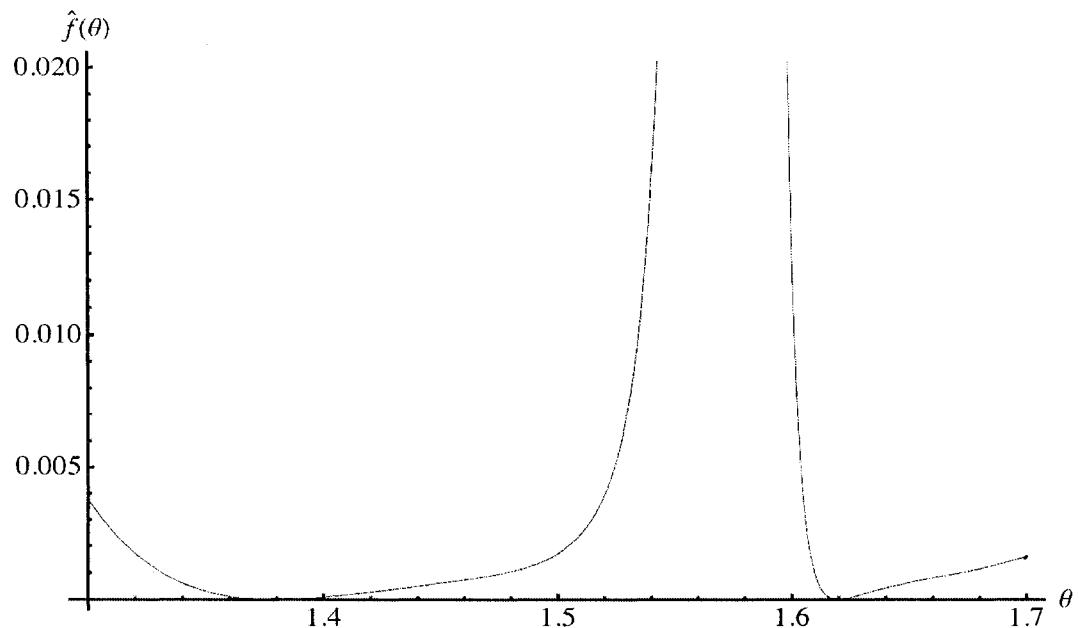


Figure 3.5 Agrandissement de la figure 3.4.

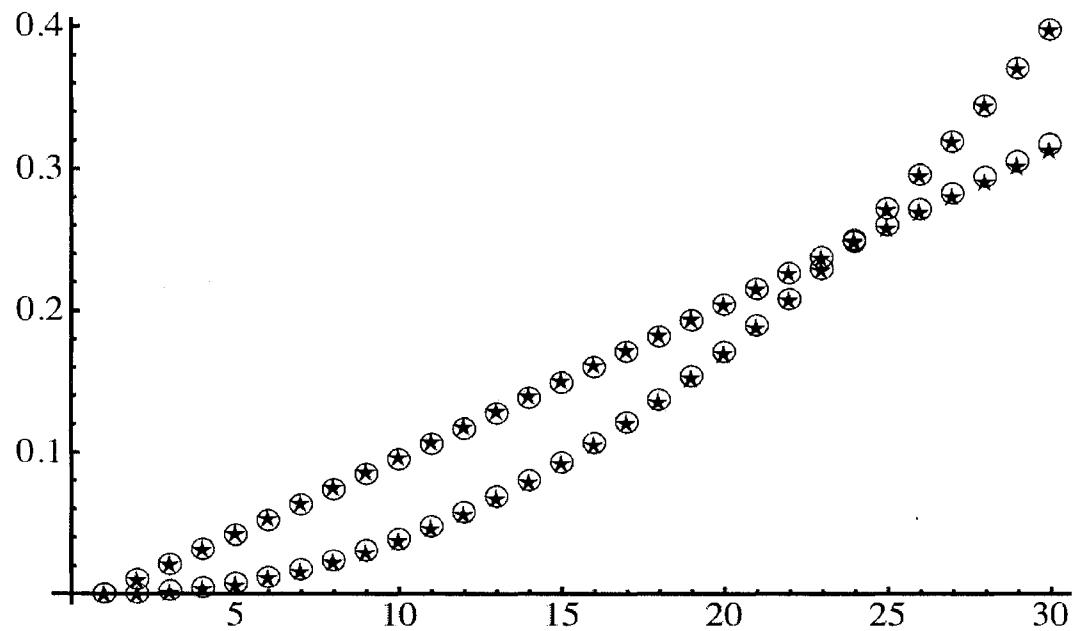


Figure 3.6 Les composantes estimées par la deuxième méthode (○) et les composantes indépendantes (★), $k_1 k_2 < 0$.

Pour le deuxième cas où $k_1 k_2 < 0$, nous avons générés M signaux \mathbf{S} ($M = 100$) de la forme

$$\tilde{\mathbf{S}} = a_1 \tilde{\mathbf{V}}_1 + a_2 \tilde{\mathbf{V}}_2,$$

où a_1 et a_2 sont des variables aléatoires indépendantes telles que $a_1 \sim U[-1, 1]$ et $a_2 \sim T$, où T est la loi de Student avec 5 degrés de liberté. Il s'ensuit que $k_1 < 0$ et $k_2 > 0$.

En traçant la courbe de $\hat{f}(\theta)$ en fonction de θ pour $\theta \in [0, \pi]$ (Figure 3.4), on peut voir clairement les deux angles θ_1 et θ_2 où la courbe passe par zéro (Figure 3.5). Nous avons estimé les deux valeurs de θ en utilisant l'algorithme décrit précédemment. Enfin nous avons estimé $\hat{\mathbf{V}}_1$ et $\hat{\mathbf{V}}_2$ et on les a représentés sur le même graphique avec $\tilde{\mathbf{V}}_1$ et $\tilde{\mathbf{V}}_2$ (Figure 3.6), où $\hat{\mathbf{V}}_1$ et $\hat{\mathbf{V}}_2$ sont représentés par des étoiles, tandis que $\tilde{\mathbf{V}}_1$ et $\tilde{\mathbf{V}}_2$ sont représentés par des cercles.

3.4 Comparaison des deux méthodes

Comme nous avons expliqué au début de cette section, avec la deuxième méthode les composantes indépendantes sont obtenues à partir des deux minima d'une fonction d'une variable, alors qu'avec la première méthode on obtient quatre racines et on distingue les composantes indépendantes à l'aide de $\mathcal{C}_{1,2}$. Aussi la courbe de la fonction de la première méthode dans certains cas contient de fines aiguilles de sorte que la localisation des quatre racines devient difficile, tandis que la courbe de la fonction de la deuxième méthode est généralement lisse autour des racines et même pour le cas où $\epsilon = 1$ (figure 3.7) les racines sont plus claires qu'avec la première méthode (figure 2.8).

Pour la rapidité des algorithmes, nous avons constaté que la première méthode est généralement plus rapide parce qu'on estime seulement $\mathbf{z}_i(\mathbf{X})$, $i = 1, 3$, alors que pour

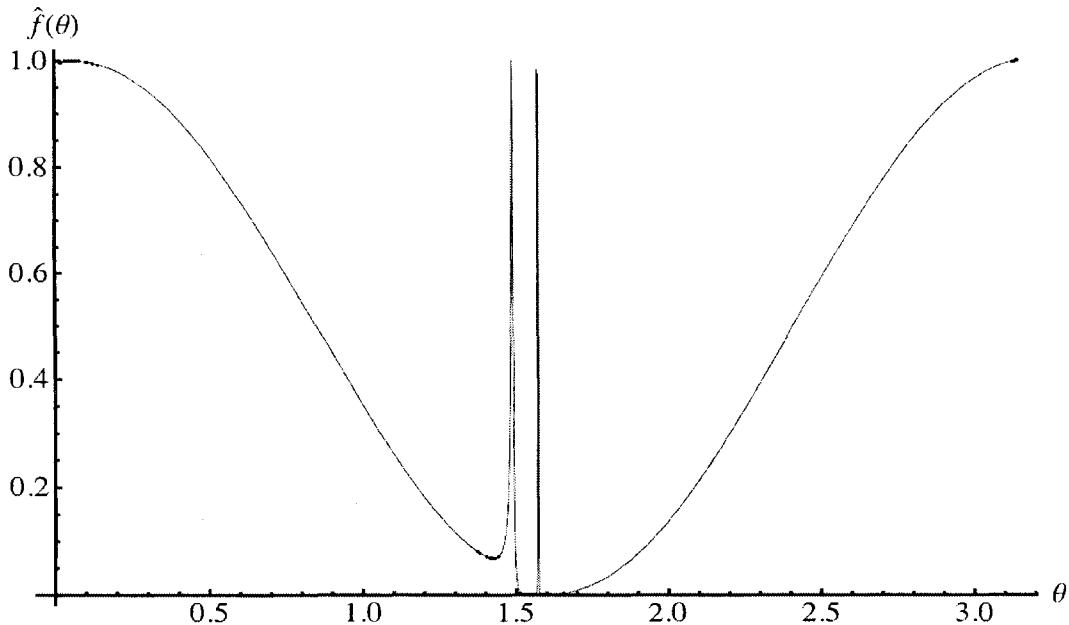


Figure 3.7 Graphe de $\hat{f}(\theta) = \hat{g}(\hat{g}_1(\mathbf{X}(\theta)), \mathbf{X}(\theta))$ pour un cas où $\epsilon = 1$.

la deuxième méthode on estime $\mathbf{X}_2 = (\mathbf{z}_1(\mathbf{X}_1))^\perp$, $\mathbf{z}_3(\mathbf{X}_1)$, $\mathbf{z}_3(\mathbf{X}_2)$ et $\mathcal{C}_{1,2}$ pour chaque \mathbf{X} . Nous pouvons donc conclure que le calcul est plus long avec la deuxième méthode, sauf pour le cas où nous avons de fines aiguilles qui nécessitent l'utilisation d'un pas très petit. En général le temps de calcul des deux méthodes est comparable parce qu'on utilise généralement un pas plus petit avec la première méthode.

Nous avons voulu comparer la précision des deux méthodes. Nous avons effectué des tests sur trois paramètres. Le premier paramètre ε contrôle la distance entre les deux composantes. Le deuxième paramètre n permet de faire tendre les lois de a_1 et a_2 vers une loi Gaussienne centrée réduite. Le troisième paramètre M est le nombre de signaux générés.

Pour faire varier la distance entre les deux composantes, on a utilisé les signaux

$$\tilde{V}_1(i) = i \text{ et } \tilde{V}_2(i) = i + \varepsilon, i = 0, \dots, N - 1,$$

avec $N = 30$. Les variables aléatoires a_i sont définies par

$$a_i = \frac{1}{\sqrt{n}} \sum_{j=1}^n U_j,$$

où les U_j sont des variables aléatoires i.i.d. telles que $U_j \sim \mathbf{U}[-\sqrt{3}, \sqrt{3}] \forall j$. Comme $E\{U_j\} = 0$ et $\text{Var}(U_j) = 1$, alors $E\{a_i\} = 0$ et $\text{Var}(a_i) = 1$ et la loi de probabilité de a_i tend vers la loi normale centrée réduite quand $n \rightarrow \infty$.

On définit la moyenne des rapports signal bruit SNR_{moy} par

$$\text{SNR}_{moy} = \frac{\text{SNR}_1 + \text{SNR}_2}{2},$$

et le minimum des rapports signal bruit par

$$\text{SNR}_{min} = \min\{\text{SNR}_1, \text{SNR}_2\},$$

$$\text{où } \text{SNR}_i = \frac{1}{\|\hat{\mathbf{V}}_i - \bar{\mathbf{V}}_i\|}, i = 1, 2.$$

Nous avons calculé SNR_{moy} et SNR_{min} pour les différentes valeurs de ε , M et n et pour chacune des deux méthodes.

Dans les figures 3.8 et 3.9, nous avons fixé $n = 1$, $M = 100$ et nous avons utilisé $\varepsilon = 1, 2, \dots, 10, 15, 20, 25$. Pour chaque valeur de ε , nous avons fait 20 essais en générant les M signaux \mathbf{S} pour chaque essai et nous avons calculé les médianes des SNR_{min} et SNR_{moy} obtenus avec chaque méthode, qui sont représentés respectivement dans les figures 3.8 et 3.9. Les SNR de la première méthode sont représentés par des cercles, tandis que les SNR de la deuxième méthode sont représentés par des étoiles.

Dans les figures 3.10 et 3.11, nous avons fixé $\varepsilon = 10$, $M = 100$ et nous avons utilisé $n = 1, 2, 4, 8, 16, 32$. Pour chaque valeur de n , nous avons fait 20 essais en générant les M signaux \mathbf{S} pour chaque essai et nous avons calculé les médianes des SNR_{min} et

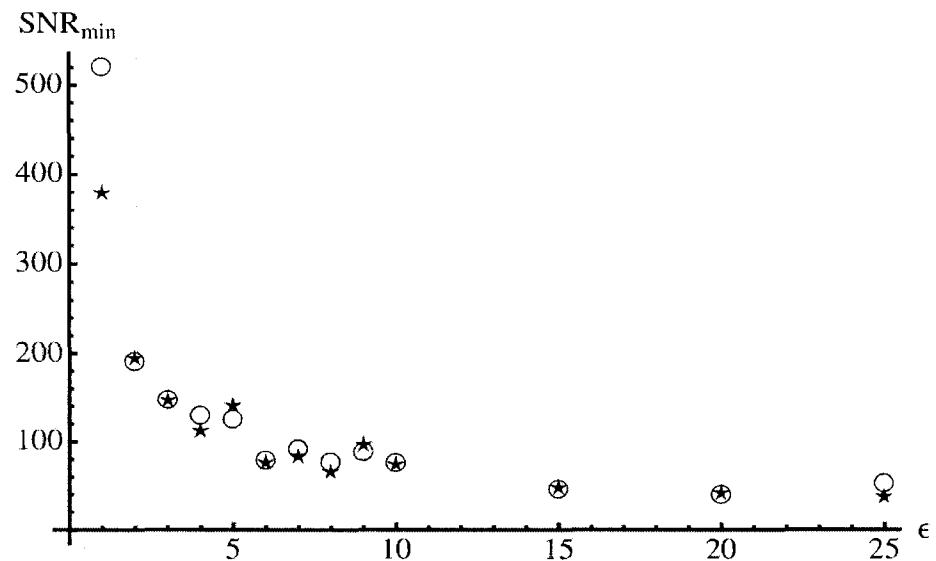


Figure 3.8 SNR_{\min} méthode I (○) VS. méthode II (★) en fonction de ϵ , $n = 1, M = 100$.

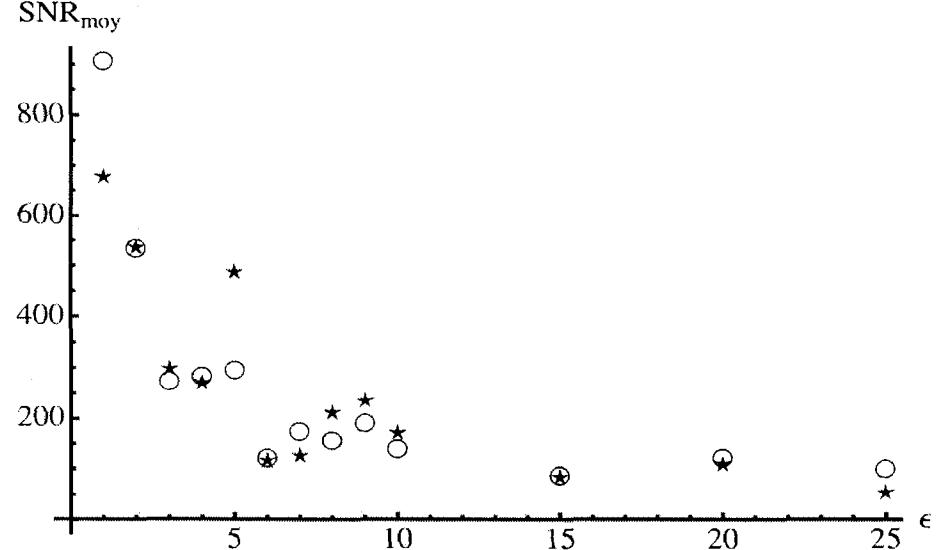


Figure 3.9 SNR_{moy} méthode I (○) VS. méthode II (★) en fonction de ϵ , $n = 1, M = 100$.

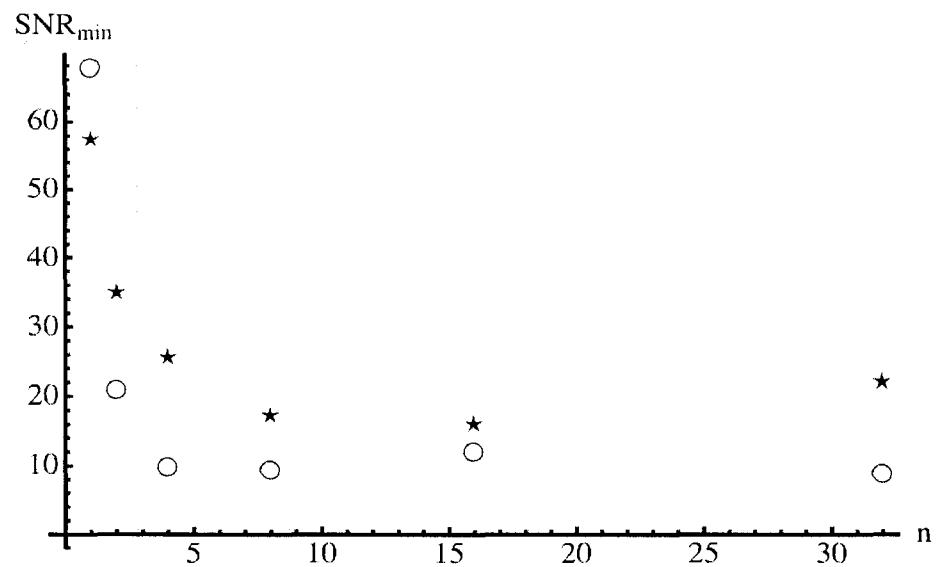


Figure 3.10 SNR_{\min} méthode I (○) VS. méthode II (★) en fonction de n , $\epsilon = 10$, $M = 100$.

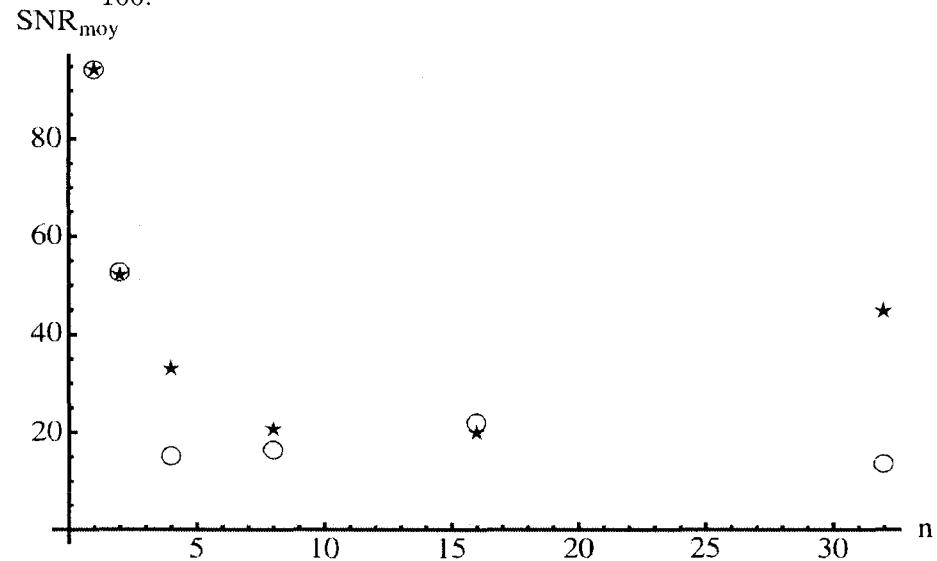


Figure 3.11 SNR_{moy} méthode I (○) VS. méthode II (★) en fonction de n , $\epsilon = 10$, $M = 100$.

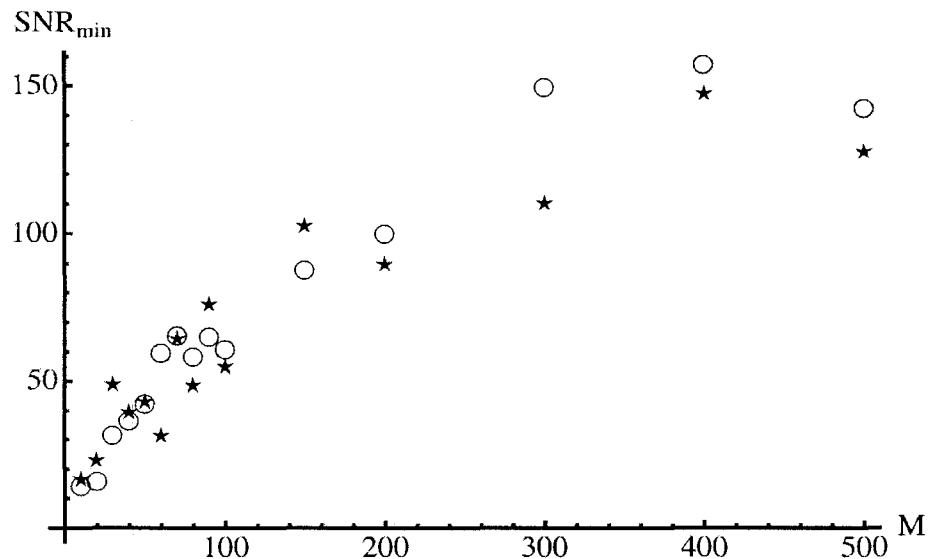


Figure 3.12 SNR_{\min} méthode I (\circ) VS. méthode II (\star) en fonction de M , $\epsilon = 10, n = 1$.

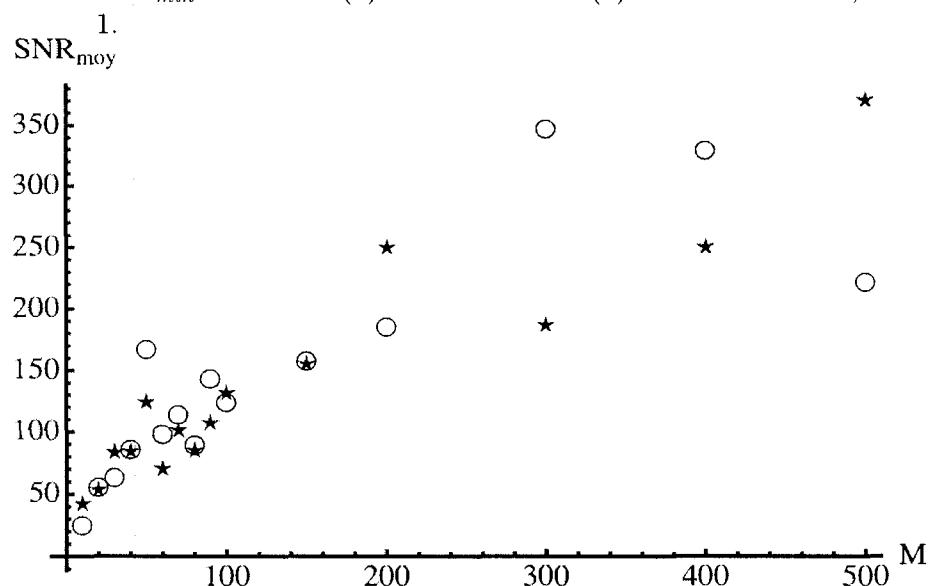


Figure 3.13 SNR_{moy} méthode I (\circ) VS. méthode II (\star) en fonction de M , $\epsilon = 10, n = 1$.

SNR_{moy} obtenus avec chaque méthode, qui sont représentés respectivement dans les figures 3.10 et 3.11. Les SNR de la première méthode sont représentés par des cercles, tandis que les SNR de la deuxième méthode sont représentés par des étoiles.

Dans les figures 3.12 et 3.13, nous avons fixé $\varepsilon = 10$, $n = 1$ et nous avons utilisé $M = 10, 20, \dots, 90, 100, 150, 200, 300, 400, 500$. Pour chaque valeur de M , nous avons fait 20 essais en générant les M signaux \mathbf{S} pour chaque essai et nous avons calculé les médianes des SNR_{min} et SNR_{moy} obtenus avec chaque méthode, qui sont représentés respectivement dans les figures 3.12 et 3.13. Les SNR de la première méthode sont représentés par des cercles, tandis que les SNR de la deuxième méthode sont représentés par des étoiles.

En examinant les figures 3.8, 3.9, 3.10, 3.11, 3.12 et 3.13, on constate que les deux méthodes ont une précision comparable.

3.5 Comparaison de la méthode I avec la méthode *FastICA*

Nous avons voulu comparer nos méthodes avec la méthode *FastICA* de (Hyvärinen et Oja (2000)).

Rappelons que le modèle utilisé pour la méthode *FastICA* est de la forme suivante

$$x^{(i)}(t) = A_1^{(i)} s_1(t) + A_2^{(i)} s_2(t), t = 1, \dots, N^*, i = 1, \dots, M^*, \quad (3.19)$$

où les $\mathbf{s}_k, i = 1, 2$, sont les signaux initiaux à estimer, tels que $s_1(t)$ et $s_2(t)$ sont statistiquement indépendants pour chaque instant $t \in \{1, \dots, N^*\}$ et les $\mathbf{x}^{(i)}, i = 1, \dots, N^*$, sont les signaux mesurés à chaque instant $t \in \{1, \dots, N^*\}$.

Si on introduit les notations

$$\mathbf{s} = \begin{pmatrix} s_1(1) & \dots & s_1(N^*) \\ s_2(1) & \dots & s_2(N^*) \end{pmatrix}, \quad (3.20)$$

$$\mathbf{A} = \begin{pmatrix} A_1^{(1)} & A_2^{(1)} \\ \dots & \dots \\ A_1^{(M^*)} & A_2^{(M^*)} \end{pmatrix}, \quad (3.21)$$

et

$$\mathbf{x} = \begin{pmatrix} x^{(1)}(1) & \dots & x^{(1)}(N^*) \\ \dots & \dots & \dots \\ x^{(M^*)}(1) & \dots & x^{(M^*)}(N^*) \end{pmatrix}, \quad (3.22)$$

alors (3.19) peut s'écrire sous la forme matricielle

$$\mathbf{x} = \mathbf{A} \mathbf{s}. \quad (3.23)$$

Notons que notre modèle peut s'écrire sous la forme

$$\tilde{S}^{(i)}(t) = a_1^{(i)} \tilde{V}_1(t) + a_2^{(i)} \tilde{V}_2(t), t = 1, \dots, N, i = 1, \dots, M. \quad (3.24)$$

Si on introduit les notations

$$\tilde{\mathbf{V}}^T = \begin{pmatrix} \tilde{V}_1(1) & \dots & \tilde{V}_1(N) \\ \tilde{V}_2(1) & \dots & \tilde{V}_2(N) \end{pmatrix}, \quad (3.25)$$

$$\mathbf{a} = \begin{pmatrix} a_1^{(1)} & a_2^{(1)} \\ \dots & \dots \\ a_1^{(M)} & a_2^{(M)} \end{pmatrix}, \quad (3.26)$$

et

$$\tilde{\mathbf{S}}_M^T = \begin{pmatrix} \tilde{S}^{(1)}(1) & \dots & \tilde{S}^{(1)}(N) \\ \dots & \dots & \dots \\ \tilde{S}^{(M)}(1) & \dots & \tilde{S}^{(M)}(N) \end{pmatrix}, \quad (3.27)$$

alors (3.24) peut s'écrire sous la forme matricielle

$$\tilde{\mathbf{S}}_M^T = \mathbf{a} \tilde{\mathbf{V}}^T. \quad (3.28)$$

Dans notre modèle, les \mathbf{a}_k sont mutuellement indépendants. Dans le modèle *FastICA*, ce sont les \mathbf{s}_k qui sont mutuellement indépendants. Il s'ensuit que \mathbf{s} correspond à \mathbf{a} dans notre modèle. Or

$$(3.28) \implies \tilde{\mathbf{S}}_M = \tilde{\mathbf{V}} \mathbf{a}^T. \quad (3.29)$$

Pour utiliser l'algorithme *FastICA* pour estimer les composantes indépendantes avec nos données, il faut donc utiliser les correspondances

$$\begin{cases} \mathbf{x} = \tilde{\mathbf{S}}_M \\ \mathbf{A} = \tilde{\mathbf{V}} \\ \mathbf{s} = \mathbf{a}^T, \end{cases} \quad (3.30)$$

et donc

$$\begin{cases} N^* = M \\ M^* = N. \end{cases}$$

Remarque : Dans le modèle *FastICA* chaque instant t correspond à une réalisation, alors que dans notre modèle, chaque $i, i = 1, \dots, M$, correspond à une réalisation.

Notons que l'algorithme de la méthode *FastICA* estime \mathbf{s}_1 et \mathbf{s}_2 , donc nous devons estimer $\mathbf{a}_1 := (a_1^{(1)}, \dots, a_1^{(M)})^T$ et $\mathbf{a}_2 := (a_2^{(1)}, \dots, a_2^{(M)})^T$. Puisque $\tilde{\mathbf{S}} = \mathbf{F}^T \mathbf{S}$,

$\tilde{\mathbf{V}} = \mathbf{F}^T \mathbf{V}$ et $\mathbf{F} \mathbf{F}^T = \mathbf{I}$, alors le résultat (3.29) implique successivement

$$\begin{aligned}\mathbf{F} \tilde{\mathbf{S}}_M &= \mathbf{F} \tilde{\mathbf{V}} \mathbf{a}^T \\ \implies \mathbf{S}_M &= \mathbf{V} \mathbf{a}^T.\end{aligned}$$

Notre méthode estime \mathbf{V} , on peut donc estimer \mathbf{a} à partir de \mathbf{S} et $\tilde{\mathbf{V}}$ avec

$$\hat{\mathbf{a}} = (\hat{\mathbf{V}}^{-1} \mathbf{S}_M)^T. \quad (3.31)$$

Puisque \mathbf{V}_1 et \mathbf{V}_2 sont des vecteurs non-nuls, linéairement indépendants de dimensions deux, alors \mathbf{V} est une matrice inversible de dimension 2×2 . On estime donc $\hat{\mathbf{a}}_1$ et $\hat{\mathbf{a}}_2$ avec

$$\begin{cases} \hat{\mathbf{a}}_1 \text{ est la première ligne de } \hat{\mathbf{a}} \\ \hat{\mathbf{a}}_2 \text{ est la deuxième ligne de } \hat{\mathbf{a}}. \end{cases} \quad (3.32)$$

Nous avons voulu comparer la précision de notre méthode et de la méthode "FastICA". On considère les signaux $\tilde{\mathbf{V}}_1$ et $\tilde{\mathbf{V}}_2$ dans $\mathbb{R}^N (N = 30)$, tels que

$$\tilde{\mathbf{V}}_i = \frac{\mathbf{U}_i}{\|\mathbf{U}_i\|},$$

$i = 1, 2$, où $U_1(k) = k, U_2(k) = k + 1, k = 0, 1, \dots, N - 1$. Nous avons généré M signaux ($M = 30$) de la forme

$$\tilde{\mathbf{S}}^{(i)} = a_1(i) \tilde{\mathbf{V}}_1 + a_2(i) \tilde{\mathbf{V}}_2,$$

où $a_i \sim U[-1, 1], i = 1, 2$.

Nous avons utilisé l'algorithme de *FastICA* pour estimer $\hat{\mathbf{a}}_{1FI}$ et $\hat{\mathbf{a}}_{2FI}$ et notre algorithme (méthode I) pour estimer $\hat{\mathbf{a}}_1$ et $\hat{\mathbf{a}}_2$. Pour comparer les signaux estimés avec \mathbf{a}_1 et \mathbf{a}_2 , nous avons utilisé les vecteurs normalisés $\frac{\mathbf{a}_i}{\|\mathbf{a}_i\|}$ et $\frac{\hat{\mathbf{a}}_i}{\|\hat{\mathbf{a}}_i\|}, i = 1, 2$ parce que ces vecteurs sont estimés à un facteur près.

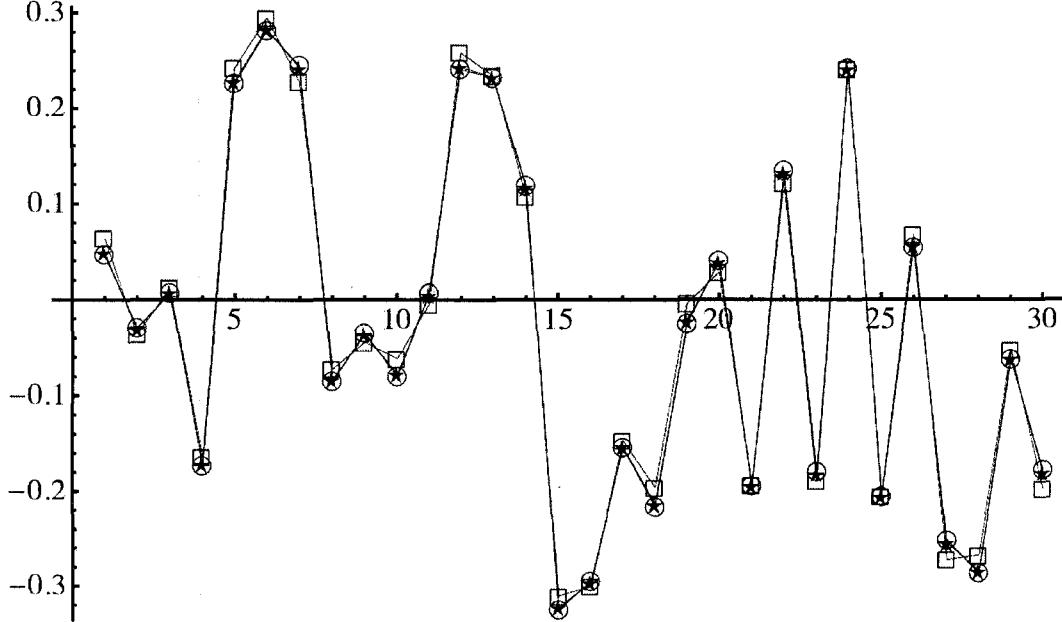


Figure 3.14 Graphe de $\frac{\mathbf{a}_1}{\|\mathbf{a}_1\|}(\star)$, $\frac{\hat{\mathbf{a}}_1}{\|\hat{\mathbf{a}}_1\|}(\circ)$ et $\frac{\hat{\mathbf{a}}_{1FI}}{\|\hat{\mathbf{a}}_{1FI}\|}(\square)$, dans le cas où $a_i \sim U[-1, 1]$, $i = 1, 2$.

La figure 3.14 nous montre les signaux $\frac{\mathbf{a}_1}{\|\mathbf{a}_1\|}$, $\frac{\hat{\mathbf{a}}_1}{\|\hat{\mathbf{a}}_1\|}$ et $\frac{\hat{\mathbf{a}}_{1FI}}{\|\hat{\mathbf{a}}_{1FI}\|}$ et la figure 3.15 nous montre les signaux $\frac{\mathbf{a}_2}{\|\mathbf{a}_2\|}$, $\frac{\hat{\mathbf{a}}_2}{\|\hat{\mathbf{a}}_2\|}$ et $\frac{\hat{\mathbf{a}}_{2FI}}{\|\hat{\mathbf{a}}_{2FI}\|}$. Les figures 3.16 et 3.17, nous montrent les résidus $\hat{a}_i - a_i$ pour les deux signaux en utilisant notre méthode et la méthode *FastICA*. Avec notre méthode, nous avons obtenu $\text{SNR}_1 = 97.58$ pour a_1 et $\text{SNR}_2 = 38.08$ pour a_2 . Avec la méthode "FastICA", nous avons obtenu $\text{SNR}_1 = 15.81$ pour a_1 et $\text{SNR}_2 = 13.54$ pour a_2 .

Nous avons fait un deuxième essai où nous avons considéré les signaux $\tilde{\mathbf{V}}_1$ et $\tilde{\mathbf{V}}_2$ dans \mathbb{R}^N ($N = 30$), tels que

$$\tilde{\mathbf{V}}_i = \frac{\mathbf{U}_i}{\|\mathbf{U}_i\|},$$

$i = 1, 2$, où $U_1(k) = k$, $U_2(k) = k + 1$, $k = 0, 1, \dots, N - 1$. Nous avons généré M signaux ($M = 30$) de la forme

$$\tilde{\mathbf{S}}^{(i)} = a_1(i)\tilde{\mathbf{V}}_1 + a_2(i)\tilde{\mathbf{V}}_2,$$

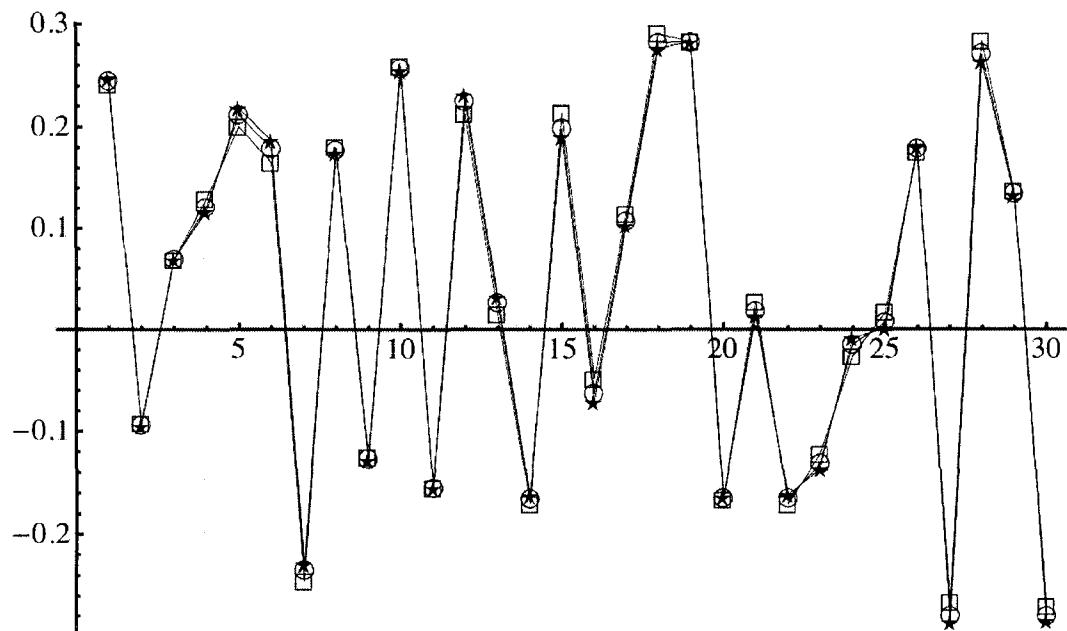


Figure 3.15 Graphe de $\frac{\mathbf{a}_2}{\|\mathbf{a}_2\|}(\star)$, $\frac{\hat{\mathbf{a}}_2}{\|\hat{\mathbf{a}}_2\|}(\circ)$ et $\frac{\hat{\mathbf{a}}_{2FL}}{\|\hat{\mathbf{a}}_{2FL}\|}(\square)$, dans le cas où $a_i \sim U[-1, 1], i = 1, 2$.

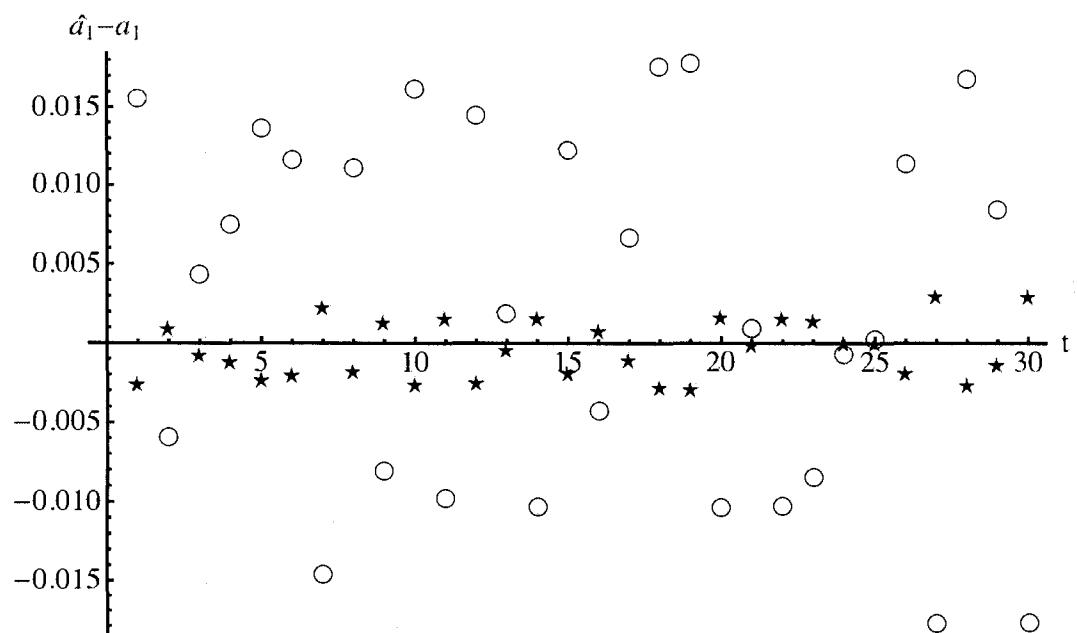


Figure 3.16 Graphe des résidus ($\frac{\hat{\mathbf{a}}_1}{\|\hat{\mathbf{a}}_1\|} - \frac{\mathbf{a}_1}{\|\mathbf{a}_1\|}$) obtenus avec notre méthode (\star) et la méthode *FastICA* (\circ), dans le cas où $a_i \sim U[-1, 1], i = 1, 2$.

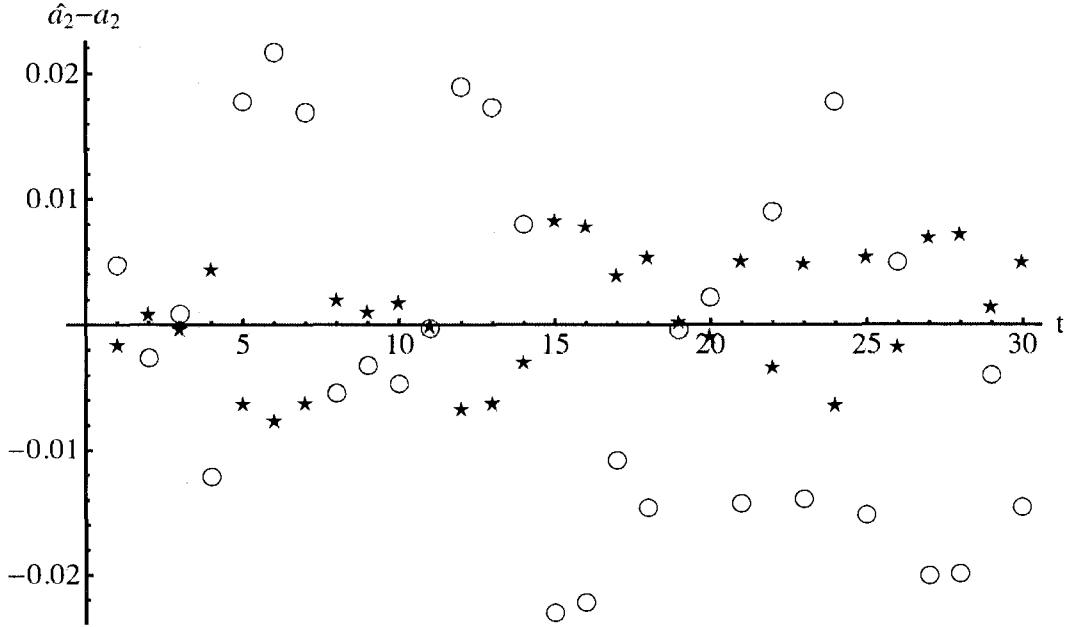


Figure 3.17 Graphe des résidus $(\frac{\hat{a}_2}{\|\hat{a}_2\|} - \frac{a_2}{\|a_2\|})$ obtenus avec notre méthode (\star) et la méthode *FastICA* (\circ), dans le cas où $a_i \sim U[-1, 1]$, $i = 1, 2$.

où $a_1(i) = \sin(\frac{i}{2})$ et $a_2(i) = \cos(\frac{i}{4})$, $i = 1, \dots, M$.

On voit que a_1 et a_2 ne sont pas des variables aléatoires indépendantes, puisqu'elles sont liées d'une façon déterministe. Cependant, a_1 et a_2 satisfont

$$\mathcal{C}(n, m) := \frac{1}{M} \sum_{i=1}^M (a_1(i))^n (a_2(i))^m - \frac{1}{M} \sum_{i=1}^M (a_1(i))^n * \frac{1}{M} \sum_{i=1}^M (a_2(i))^m \simeq 0 \quad (3.33)$$

pour $(n, m) \in \{(1, 1), (1, 3), (3, 1)\}$ et $M \simeq 4\pi k$, où $k \in \mathbb{N}$. Le résultat (3.33) peut s'établir en évaluant les sommes numériquement puis en posant $M = 4\pi k$ dans la formule résultante. Du point de vue des coefficients de corrélation $\mathcal{C}(n, m)$, a_1 et a_2 se comportent donc comme des variables aléatoires indépendantes.

Nous avons utilisé l'algorithme *FastICA* pour estimer \hat{a}_{1FI} et \hat{a}_{2FI} et notre algorithme (méthode I) pour estimer \hat{a}_1 et \hat{a}_2 . Pour comparer les signaux estimés avec a_1 et a_2 , nous avons utilisé les vecteurs normalisés $\frac{a_i}{\|a_i\|}$ et $\frac{\hat{a}_i}{\|\hat{a}_i\|}$, $i = 1, 2$.

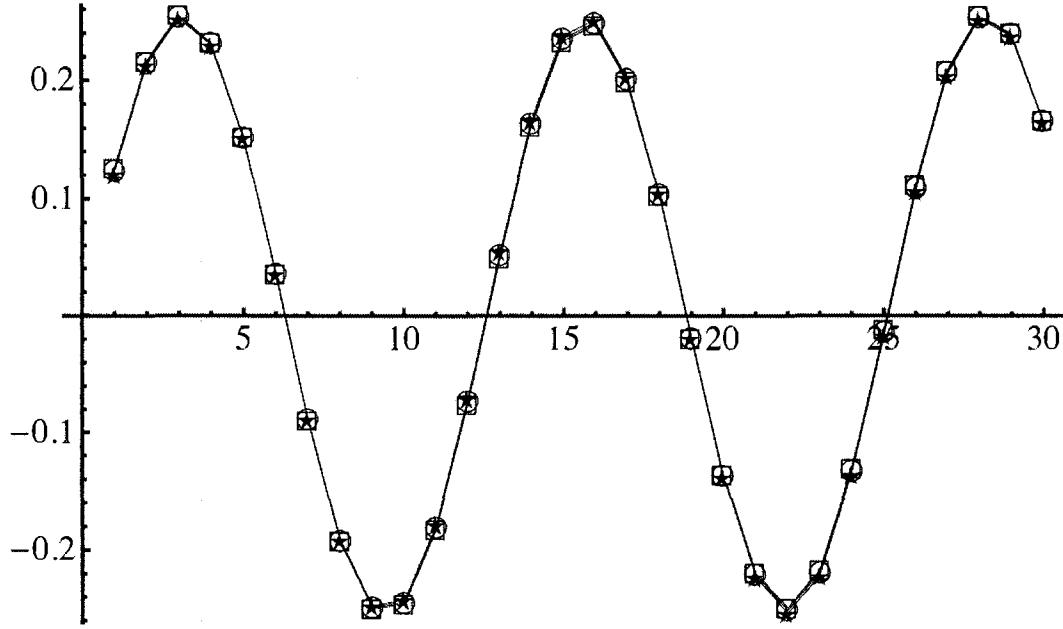


Figure 3.18 Graphe de $\frac{\mathbf{a}_1}{\|\mathbf{a}_1\|}(\star)$, $\frac{\hat{\mathbf{a}}_1}{\|\hat{\mathbf{a}}_1\|}(o)$ et $\frac{\hat{\mathbf{a}}_{1FI}}{\|\hat{\mathbf{a}}_{1FI}\|}(\square)$, dans le cas où $U_1(k) = k$ et $U_2(k) = k + 1$.

La figure 3.18 nous montre les signaux $\frac{\mathbf{a}_1}{\|\mathbf{a}_1\|}$, $\frac{\hat{\mathbf{a}}_1}{\|\hat{\mathbf{a}}_1\|}$ et $\frac{\hat{\mathbf{a}}_{1FI}}{\|\hat{\mathbf{a}}_{1FI}\|}$ et la figure 3.19 nous montre les signaux $\frac{\mathbf{a}_2}{\|\mathbf{a}_2\|}$, $\frac{\hat{\mathbf{a}}_2}{\|\hat{\mathbf{a}}_2\|}$ et $\frac{\hat{\mathbf{a}}_{2FI}}{\|\hat{\mathbf{a}}_{2FI}\|}$. Les figures 3.20 et 3.21, nous montrent les résidus des deux signaux en utilisant notre méthode et la méthode de *FastICA*. Avec notre méthode, nous avons obtenu $\text{SNR}_1 = 92.7$ pour a_1 et $\text{SNR}_2 = 30.64$ pour a_2 . Avec la méthode *FastICA*, nous avons obtenu $\text{SNR}_1 = 51.6$ pour a_1 et $\text{SNR}_2 = 6.05$ pour a_2 .

Nous avons fait un troisième essai où nous avons utilisé les signaux $\tilde{\mathbf{V}}_1$ et $\tilde{\mathbf{V}}_2$ dans $\mathbb{R}^N (N = 30)$, tels que

$$\tilde{\mathbf{V}}_i = \frac{\mathbf{U}_i}{\|\mathbf{U}_i\|},$$

$i = 1, 2$, où $U_1(k) = k$, $U_2(k) = k^2$, $k = 0, 1, \dots, N - 1$. Nous avons généré M signaux ($M = 30$) de la forme

$$\tilde{\mathbf{S}}^{(i)} = a_1(i)\tilde{\mathbf{V}}_1 + a_2(i)\tilde{\mathbf{V}}_2,$$

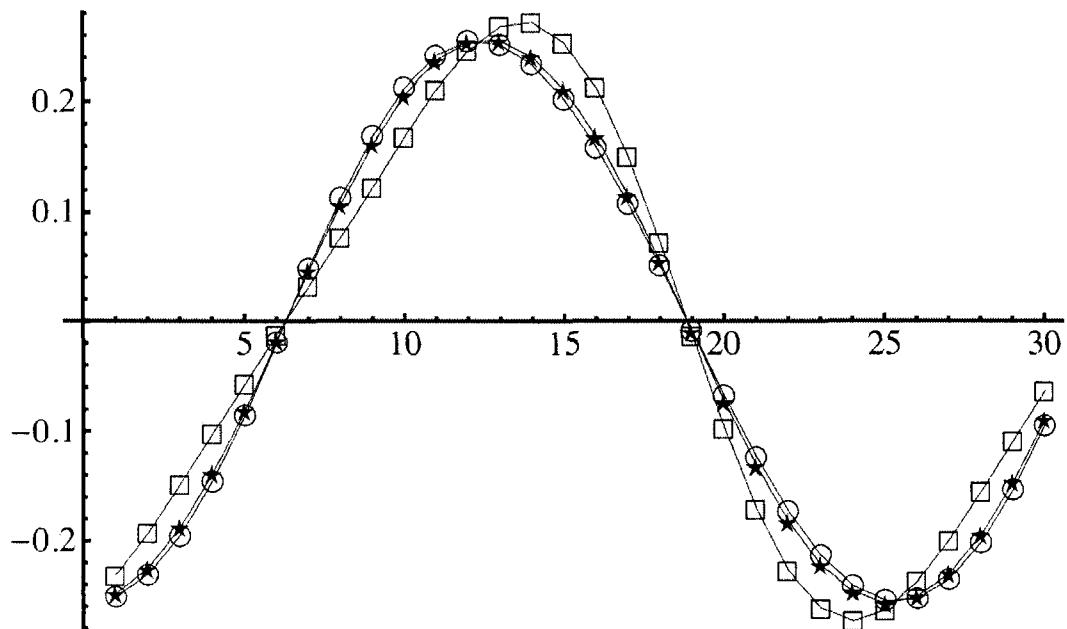


Figure 3.19 Graphe de $\frac{\mathbf{a}_2}{\|\mathbf{a}_2\|}(\star)$, $\frac{\hat{\mathbf{a}}_2}{\|\hat{\mathbf{a}}_2\|}(o)$ et $\frac{\hat{\mathbf{a}}_{2FI}}{\|\hat{\mathbf{a}}_{2FI}\|}(\square)$, dans le cas où $U_1(k) = k$ et $U_2(k) = k + 1$.

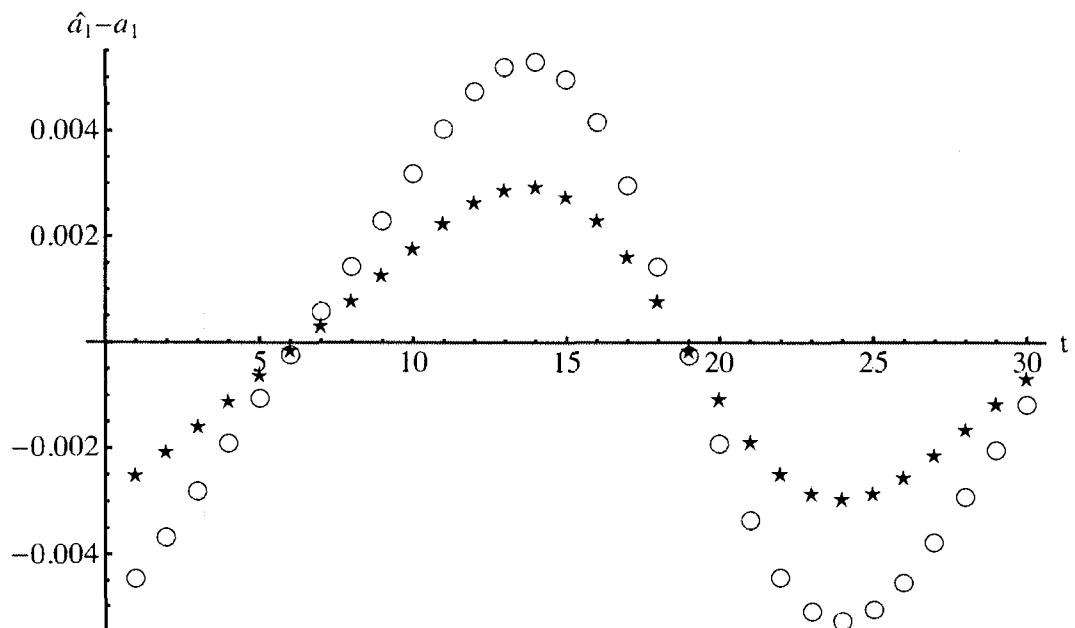


Figure 3.20 Graphe des résidus $(\frac{\hat{\mathbf{a}}_1}{\|\hat{\mathbf{a}}_1\|} - \frac{\mathbf{a}_1}{\|\mathbf{a}_1\|})$ obtenus avec notre méthode (\star) et la méthode *FastICA* (\circ), dans le cas où $U_1(k) = k$ et $U_2(k) = k + 1$.

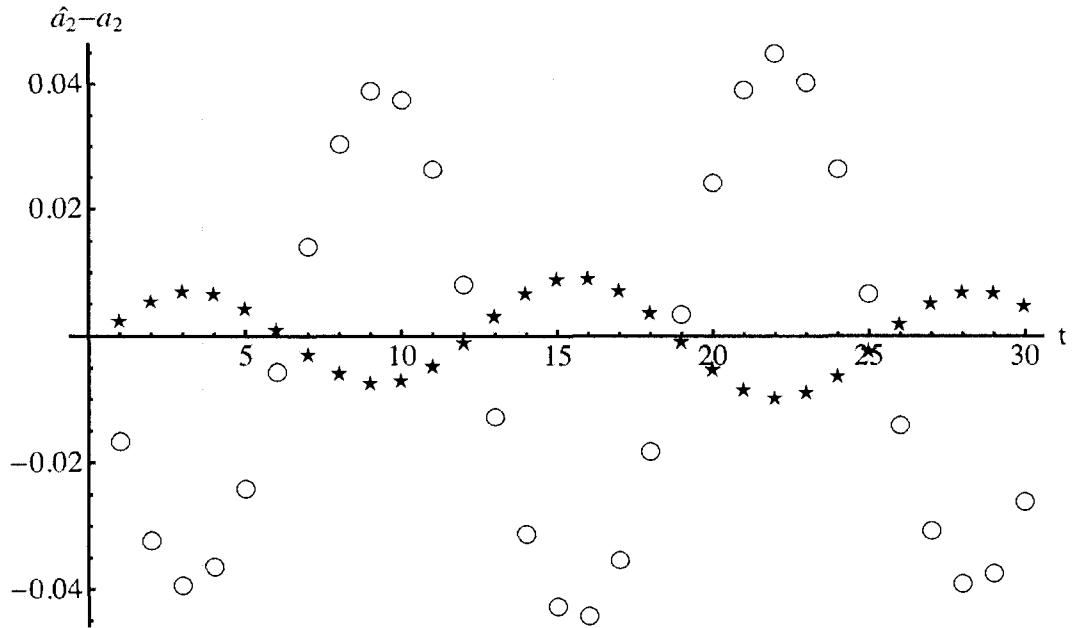


Figure 3.21 Graphe des résidus ($\frac{\hat{a}_2}{\|\hat{a}_2\|} - \frac{a_2}{\|a_2\|}$) obtenus avec notre méthode (\star) et la méthode *FastICA* (\circ), dans le cas où $U_1(k) = k$ et $U_2(k) = k + 1$.

où $a_1(i) = \sin(\frac{i}{2})$ et $a_2(i) = \cos(\frac{i}{4})$, $i = 1, \dots, M$.

Les résultats obtenus sont représentés par les figures 3.22, 3.23, 3.24 et 3.25. Avec notre méthode, nous avons obtenu $\text{SNR}_1 = 92.7$ pour a_1 et $\text{SNR}_2 = 30.64$ pour a_2 . Avec la méthode *FastICA*, nous avons obtenu $\text{SNR}_1 = 5.65$ pour a_1 et $\text{SNR}_2 = 31.25$ pour a_2 . On peut voir que notre méthode est plus précise.

Nous avons fait d'autres essais en changeant chaque fois les paramètres de génération des signaux, nous avons constaté que notre méthode estime généralement les composantes indépendantes avec une précision égale ou meilleure que la méthode *FastICA*.

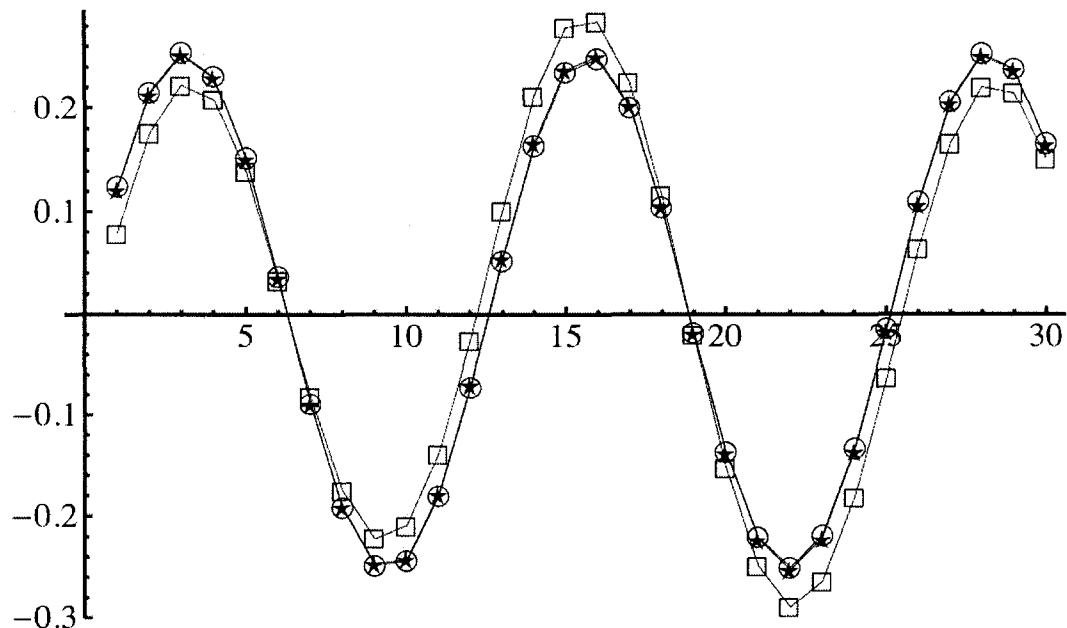


Figure 3.22 Graphe de $\frac{\mathbf{a}_1}{\|\mathbf{a}_1\|}(\star)$, $\frac{\hat{\mathbf{a}}_1}{\|\hat{\mathbf{a}}_1\|}(\circ)$ et $\frac{\hat{\mathbf{a}}_{1FL}}{\|\hat{\mathbf{a}}_{1FL}\|}(\square)$, dans le cas où $U_1(k) = k$ et $U_2(k) = k^2$.

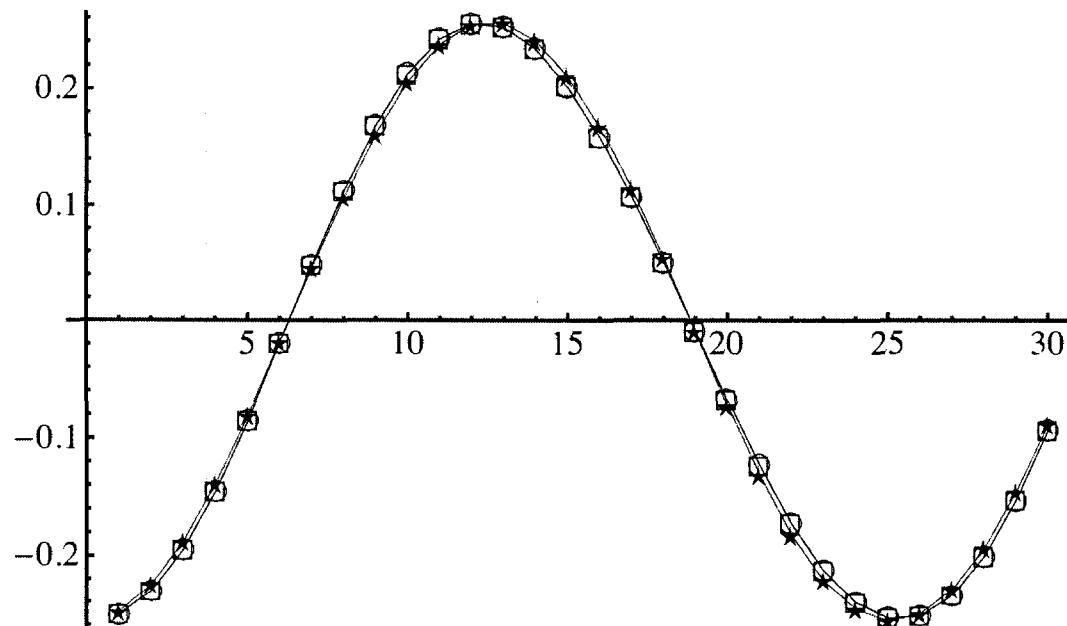


Figure 3.23 Graphe de $\frac{\mathbf{a}_2}{\|\mathbf{a}_2\|}(\star)$, $\frac{\hat{\mathbf{a}}_2}{\|\hat{\mathbf{a}}_2\|}(\circ)$ et $\frac{\hat{\mathbf{a}}_{2FL}}{\|\hat{\mathbf{a}}_{2FL}\|}(\square)$, dans le cas où $U_1(k) = k$ et $U_2(k) = k^2$.

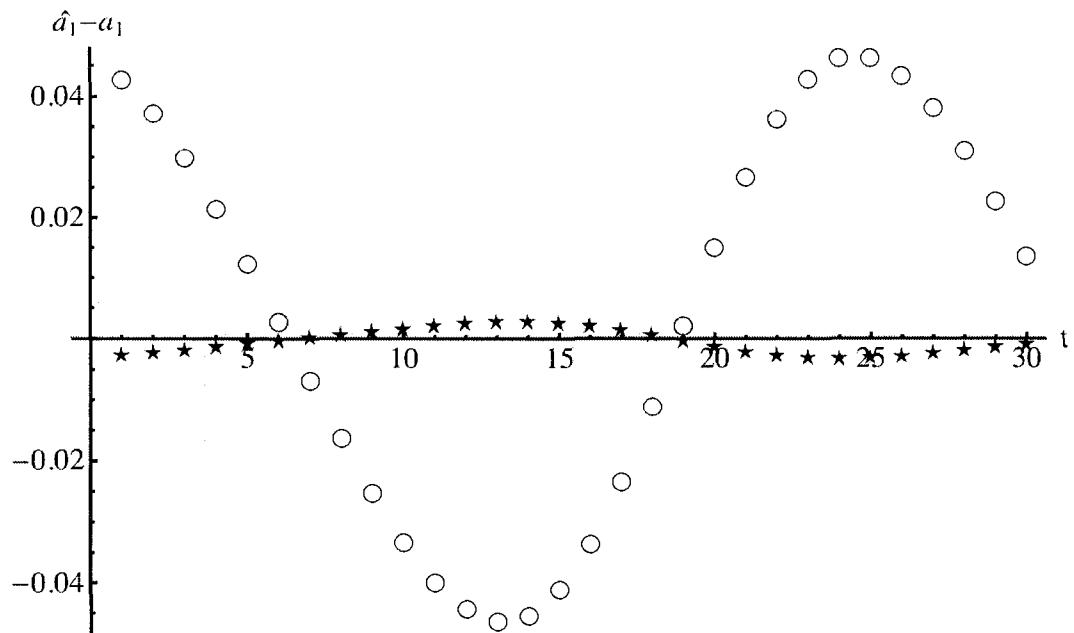


Figure 3.24 Graphe des résidus $(\frac{\hat{a}_1}{\|\hat{a}_1\|} - \frac{a_1}{\|a_1\|})$ obtenus avec notre méthode (\star) et la méthode *FastICA* (\circ), dans le cas où $U_1(k) = k$ et $U_2(k) = k^2$.

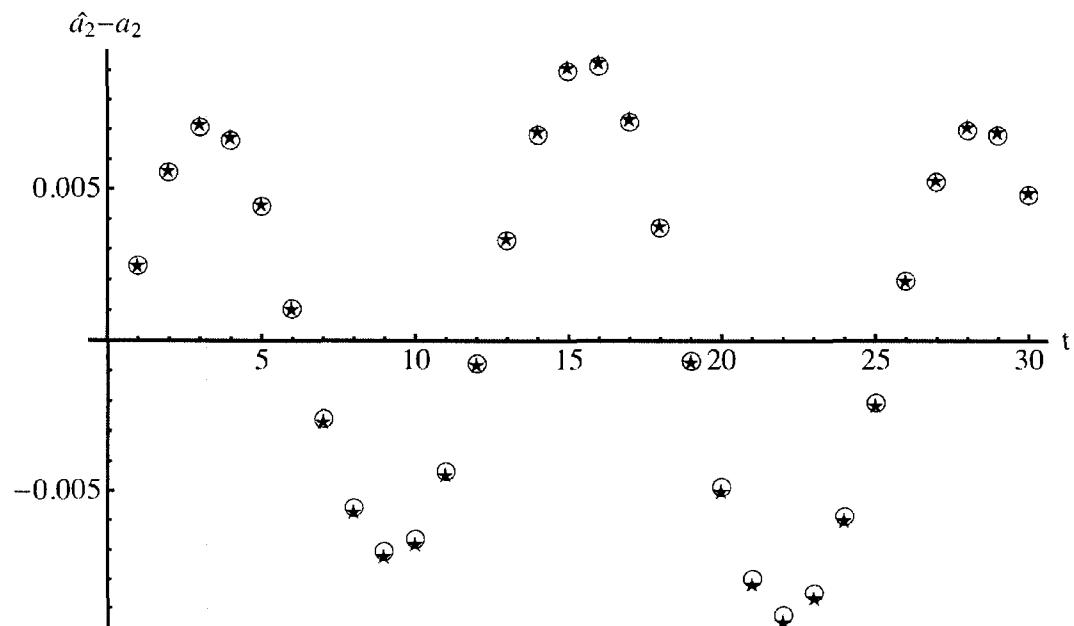


Figure 3.25 Graphe des résidus $(\frac{\hat{a}_2}{\|\hat{a}_2\|} - \frac{a_2}{\|a_2\|})$ obtenus avec notre méthode (\star) et la méthode *FastICA* (\circ), dans le cas où $U_1(k) = k$ et $U_2(k) = k^2$.

CHAPITRE 4

CAS DE TROIS COMPOSANTES ($p = 3$)

Ici $E = \text{span}\{\tilde{\mathbf{V}}_1, \tilde{\mathbf{V}}_2, \tilde{\mathbf{V}}_3\} = \text{span}\{\tilde{\Phi}_1, \tilde{\Phi}_2, \tilde{\Phi}_3\} = \mathbb{R}^3$. Puisque la dimension de l'espace E est trois, l'ensemble $Z(\mathbf{X})$ contient un maximum de trois éléments linéairement indépendants.

La méthode que nous allons utiliser est une extension de la méthode II du modèle à deux composantes. Les vecteurs $\bar{\mathbf{X}}_i, i = 1, 2, 3$, sont des racines de la fonction $f(\mathbf{X}_1, \mathbf{X}_2, \mathbf{X}_3) \geq 0$ définie par

$$\begin{aligned}
 f(\mathbf{X}_1, \mathbf{X}_2, \mathbf{X}_3) := & (\mathbf{X}_1^T \mathbf{z}_1(\mathbf{X}_2))^2 + (\mathbf{X}_1^T \mathbf{z}_1(\mathbf{X}_3))^2 + (\mathbf{X}_2^T \mathbf{z}_1(\mathbf{X}_3))^2 \\
 & + (\mathbf{X}_1^T \mathbf{z}_3(\mathbf{X}_2))^2 + (\mathbf{X}_1^T \mathbf{z}_3(\mathbf{X}_3))^2 + (\mathbf{X}_2^T \mathbf{z}_3(\mathbf{X}_1))^2 \\
 & + (\mathbf{X}_2^T \mathbf{z}_3(\mathbf{X}_3))^2 + (\mathbf{X}_3^T \mathbf{z}_3(\mathbf{X}_1))^2 + (\mathbf{X}_3^T \mathbf{z}_3(\mathbf{X}_2))^2 \\
 & + (\mathbf{X}_1^T \mathbf{z}_5(\mathbf{X}_2))^2 + (\mathbf{X}_1^T \mathbf{z}_5(\mathbf{X}_3))^2 + (\mathbf{X}_2^T \mathbf{z}_5(\mathbf{X}_1))^2 \\
 & + (\mathbf{X}_2^T \mathbf{z}_5(\mathbf{X}_3))^2 + (\mathbf{X}_3^T \mathbf{z}_5(\mathbf{X}_1))^2 + (\mathbf{X}_3^T \mathbf{z}_5(\mathbf{X}_2))^2 \\
 & + \gamma [(\mathcal{C}_{1,2})^2 + (\mathcal{C}_{1,3})^2 + (\mathcal{C}_{2,3})^2],
 \end{aligned} \tag{4.1}$$

où $\gamma > 0$ est un coefficient de pondération. On utilise les $\mathbf{z}_k(\mathbf{X}), k = 1, 3, 5$, qui sont de norme un, pour que les termes de la fonction f aient des grandeurs comparables. Remarquons que si les facteurs $\mathbf{X}_1^T \mathbf{S}$, $\mathbf{X}_2^T \mathbf{S}$ et $\mathbf{X}_3^T \mathbf{S}$ sont indépendants, alors $f(\mathbf{X}_1, \mathbf{X}_2, \mathbf{X}_3) = 0$. Soulignons aussi que la fonction $f(\mathbf{X}_1, \mathbf{X}_2, \mathbf{X}_3)$ reste invariante si on permute l'ordre des vecteurs $\mathbf{X}_1, \mathbf{X}_2, \mathbf{X}_3$.

4.1 Description de la méthode

Cette méthode consiste à chercher les racines de la fonction $f(\mathbf{X}_1, \mathbf{X}_2, \mathbf{X}_3)$. Or

$$f(\mathbf{X}_1, \mathbf{X}_2, \mathbf{X}_3) = 0 \iff \left\{ \begin{array}{l} \mathbf{X}_1^T \mathbf{z}_1(\mathbf{X}_2) = 0 \\ \mathbf{X}_1^T \mathbf{z}_1(\mathbf{X}_3) = 0 \\ \mathbf{X}_2^T \mathbf{z}_1(\mathbf{X}_3) = 0 \\ \mathbf{X}_1^T \mathbf{z}_3(\mathbf{X}_2) = 0 \\ \mathbf{X}_1^T \mathbf{z}_3(\mathbf{X}_3) = 0 \\ \mathbf{X}_2^T \mathbf{z}_3(\mathbf{X}_1) = 0 \\ \mathbf{X}_2^T \mathbf{z}_3(\mathbf{X}_3) = 0 \\ \mathbf{X}_3^T \mathbf{z}_3(\mathbf{X}_1) = 0 \\ \mathbf{X}_3^T \mathbf{z}_3(\mathbf{X}_2) = 0 \\ \mathbf{X}_1^T \mathbf{z}_5(\mathbf{X}_2) = 0 \\ \mathbf{X}_1^T \mathbf{z}_5(\mathbf{X}_3) = 0 \\ \mathbf{X}_2^T \mathbf{z}_5(\mathbf{X}_1) = 0 \\ \mathbf{X}_2^T \mathbf{z}_5(\mathbf{X}_3) = 0 \\ \mathbf{X}_3^T \mathbf{z}_5(\mathbf{X}_1) = 0 \\ \mathbf{X}_3^T \mathbf{z}_5(\mathbf{X}_2) = 0 \\ \mathcal{C}_{1,2} = 0 \\ \mathcal{C}_{1,2} = 0 \\ \mathcal{C}_{1,2} = 0. \end{array} \right. \quad (4.2)$$

Puisque $\mathbf{z}_1(\mathbf{X}) = \epsilon \frac{\mathbf{C} \mathbf{X}}{\|\mathbf{C} \mathbf{X}\|}$, où le signe de $\epsilon = \pm 1$ est ajusté pour que $\mathbf{z}_1(\mathbf{X})$ satisfasse la condition (1.2), alors les trois premières lignes de (4.2) impliquent

$$\begin{aligned}
\mathbf{X}_1 &\text{ est orthogonal à } \mathbf{C}\mathbf{X}_2 \text{ et } \mathbf{C}\mathbf{X}_3 \\
\mathbf{X}_2 &\text{ est orthogonal à } \mathbf{C}\mathbf{X}_1 \text{ et } \mathbf{C}\mathbf{X}_3 \\
\mathbf{X}_3 &\text{ est orthogonal à } \mathbf{C}\mathbf{X}_1 \text{ et } \mathbf{C}\mathbf{X}_2.
\end{aligned} \tag{4.3}$$

Remarquons que si un des trois vecteurs $\bar{\mathbf{X}}_i, i = 1, 2, 3$ est connu, alors l'estimation des deux autres vecteurs se réduit à un problème à deux dimensions et une inconnue. Par exemple, si $\bar{\mathbf{X}}_1$ est connu, alors $\bar{\mathbf{X}}_2$ et $\bar{\mathbf{X}}_3$ appartiennent au plan orthogonal à $\mathbf{C}\bar{\mathbf{X}}_1$ et ils satisfont $\bar{\mathbf{X}}_3\mathbf{C}\bar{\mathbf{X}}_2 = 0$ dans \mathbb{R}^3 . Si \mathbf{X}_1 et \mathbf{X}_2 sont donnés, alors

$$\mathbf{X}_3 = [\mathbf{Z}_1(\mathbf{X}_1), \mathbf{Z}_1(\mathbf{X}_2)]^\perp =: \mathbf{X}_3(\mathbf{X}_1, \mathbf{X}_2). \tag{4.4}$$

où $[\mathbf{Y}_1, \mathbf{Y}_2]^\perp$ dénote le vecteur normal au plan engendré par \mathbf{Y}_1 et \mathbf{Y}_2 et qui satisfait la condition (1.2). On définit

$$f_3(\mathbf{X}_1, \mathbf{X}_2) := f(\mathbf{X}_1, \mathbf{X}_2, \mathbf{X}_3(\mathbf{X}_1, \mathbf{X}_2)).$$

Par définition des composantes principales, la composante maximale des signaux $\tilde{\mathbf{S}}$ est celle de $\tilde{\Phi}_1$. Par conséquent, la coordonnée maximale des $\mathbf{V}_i, i = 1, 2, 3$, est aussi celle de $\tilde{\Phi}_1$.

Puisque $\bar{\mathbf{X}}_i$ est orthogonal à la fois à \mathbf{V}_j et \mathbf{V}_k , où $i, j, k = 1, 2, 3$ et $i \neq j, i \neq k$ et $j \neq k$, alors les $\bar{\mathbf{X}}_i, i = 1, 2, 3$, seront approximativement orthogonaux à $\tilde{\Phi}_1$. Il y aura donc au moins un vecteur $\bar{\mathbf{X}}_i, i = 1, 2, 3$, qui sera approximativement contenu dans le plan engendré par $\tilde{\Phi}_2$ et $\tilde{\Phi}_3$.

En première approximation, nous allons chercher \mathbf{X}_1 dans le plan engendré par $\tilde{\Phi}_2$ et $\tilde{\Phi}_3$. Puisque $\|\mathbf{X}_1\| = 1$ et $\tilde{\Phi}_2^T \tilde{\Phi}_3 = 0$, alors \mathbf{X}_1 peut s'écrire sous la forme

$$\mathbf{X}_1(\theta) = \cos(\theta)\tilde{\Phi}_2 + \sin(\theta)\tilde{\Phi}_3, \theta \in [0, \pi].$$

À chaque valeur de $\theta \in [0, \pi]$, nous associons un couple de vecteurs $(\mathbf{X}_1(\theta), \mathbf{X}_2(\theta))$ tels que

$$\begin{cases} \mathbf{X}_1(\theta) = \cos(\theta)\tilde{\Phi}_2 + \sin(\theta)\tilde{\Phi}_3 \\ \mathbf{X}_2(\theta) = [\mathbf{z}_1(\mathbf{X}_1(\theta)), \mathbf{z}_3(\mathbf{X}_1(\theta))]^\perp. \end{cases} \quad (4.5)$$

Le couple $(\mathbf{X}_1(\theta), \mathbf{X}_2(\theta))$ est complètement spécifié par un seul angle θ . Il suffit donc de trouver le minimum de $f_3(\mathbf{X}_1(\theta), \mathbf{X}_2(\theta))$, qui est une fonction d'une seule variable θ . Soit

$$\theta^{(1)} = \operatorname{argmin}_{\theta \in [0, \pi]} f_3(\mathbf{X}_1(\theta), \mathbf{X}_2(\theta)).$$

Le vecteur $\mathbf{X}_1^{(1)} := \mathbf{X}_1(\theta^{(1)})$ ainsi obtenu est une approximation d'un des vecteurs $\tilde{\mathbf{X}}_i$, $i = 1, 2, 3$.

On veut trouver une meilleure approximation de ces vecteurs en minimisant f_3 avec une méthode de descente qui s'appuie sur les conditions d'orthogonalité (4.3). Si $\mathbf{X}_1^{(1)}$ est une solution, alors le vecteur \mathbf{X}_2 appartient au plan $[\mathbf{C} \mathbf{X}_1^{(1)}]^\perp$, où $[\mathbf{Y}]^\perp$ dénote le plan dont la normale est le vecteur \mathbf{Y} . Dénotons par $\Psi_1^{(1)}$ et $\Psi_2^{(1)}$ deux vecteurs orthonormaux qui engendrent $[\mathbf{C} \mathbf{X}_1^{(1)}]^\perp$. On cherche donc le vecteur $\mathbf{X}_2(\theta)$, $\theta \in [0, \pi]$, sous la forme

$$\mathbf{X}_2(\theta) = \cos(\theta)\Psi_1^{(1)} + \sin(\theta)\Psi_2^{(1)}. \quad (4.6)$$

On cherchera l'angle $\theta \in [0, \pi]$ qui minimise $f_3(\mathbf{X}_1^{(1)}, \mathbf{X}_2(\theta))$, où $\mathbf{X}_2(\theta)$ est donné par (4.6). Soit

$$\theta^{(2)} = \operatorname{argmin}_{\theta \in [0, \pi]} f_3(\mathbf{X}_1^{(1)}, \mathbf{X}_2(\theta)).$$

Le vecteur $\mathbf{X}_2^{(2)} := \mathbf{X}_2(\theta^{(2)})$ ainsi obtenu est aussi une approximation d'un des vecteurs $\tilde{\mathbf{X}}_i$, $i = 1, 2, 3$. On cherchera donc un nouveau vecteur \mathbf{X}_1 dans le plan $[\mathbf{C} \mathbf{X}_2^{(2)}]^\perp$. Dénotons par $\Psi_1^{(2)}$ et $\Psi_2^{(2)}$ deux vecteurs orthonormaux qui engendrent $[\mathbf{C} \mathbf{X}_2^{(2)}]^\perp$. On

cherche donc le vecteur $\mathbf{X}_1(\theta), \theta \in [0, \pi]$, sous la forme

$$\mathbf{X}_1(\theta) = \cos(\theta)\Psi_1^{(2)} + \sin(\theta)\Psi_2^{(2)}. \quad (4.7)$$

On cherchera l'angle $\theta \in [0, \pi]$ qui minimise $f_3(\mathbf{X}_1(\theta), \mathbf{X}_2^{(2)})$, où $\mathbf{X}_1(\theta)$ est donné par (4.7). Soit

$$\theta^{(3)} = \operatorname{argmin}_{\theta \in [0, \pi]} f_3(\mathbf{X}_1(\theta), \mathbf{X}_2^{(2)}),$$

et

$$\mathbf{X}_1^{(2)} := \mathbf{X}_1(\theta^{(3)}).$$

Comme $\mathbf{X}_2^{(2)T} \mathbf{C} \mathbf{X}_1^{(1)} = \mathbf{X}_1^{(1)T} \mathbf{C} \mathbf{X}_2^{(2)} = 0$, alors $\mathbf{X}_1^{(1)}$ et $\mathbf{X}_1^{(2)}$ appartiennent tous les deux au plan $[\mathbf{C} \mathbf{X}_2^{(2)}]^\perp$. Or $f_3(\mathbf{X}_1^{(2)}, \mathbf{X}_2^{(2)}) \leq f_3(\mathbf{X}_1(\theta), \mathbf{X}_2^{(2)})$ pour tout $\theta \in [0, \pi]$, donc

$$f_3(\mathbf{X}_1^{(2)}, \mathbf{X}_2^{(2)}) \leq f_3(\mathbf{X}_1^{(1)}, \mathbf{X}_2^{(2)}).$$

On fixera ensuite $\mathbf{X}_1^{(2)}$ et on cherchera un nouveau vecteur $\mathbf{X}_2^{(3)}$, et on itère l'algorithme en fixant $\mathbf{X}_1^{(i)}$ et cherchant $\mathbf{X}_2^{(i+1)}$, ou en fixant $\mathbf{X}_2^{(i)}$ et cherchant $\mathbf{X}_1^{(i)}$. On s'arrête lorsque

$$f_3(\mathbf{X}_1^{(i)}, \mathbf{X}_2^{(i+1)}) \simeq f_3(\mathbf{X}_1^{(i+1)}, \mathbf{X}_2^{(i+1)}).$$

L'algorithme de descente qui conduit à l'estimation de $\bar{\mathbf{X}} := (\bar{\mathbf{X}}_1, \bar{\mathbf{X}}_2)$ se résume comme suit :

(1) Initialisation

On cherche \mathbf{X}_1 dans $\operatorname{span}(\tilde{\Phi}_2, \tilde{\Phi}_3)$.

On trouve une première estimation $\hat{\mathbf{X}}^{(1)} := (\hat{\mathbf{X}}_1^{(1)}, \hat{\mathbf{X}}_2^{(1)})$ où $\hat{\mathbf{X}}_2^{(1)}$ satisfait

$$\hat{\mathbf{X}}_2^{(1)T} \mathbf{Z}_k(\hat{\mathbf{X}}_1^{(1)}) = 0, k = 1, 3. \quad (4.8)$$

(2) Descente cyclique

On fixe $\mathbf{X}_1 = \hat{\mathbf{X}}_1^{(1)}$, puis on minimise $f_3(\hat{\mathbf{X}}_1^{(1)}, \mathbf{X}_2)$ en faisant faire une rotation à \mathbf{X}_2 dans le plan $[\mathbf{C}\hat{\mathbf{X}}_1^{(1)}]^\perp$, ce qui permet de respecter la condition d'orthogonalité

$$\mathbf{X}_2^T \mathbf{C} \hat{\mathbf{X}}_1^{(1)} = 0.$$

Le résultat de cette minimisation est un vecteur dénoté par $\hat{\mathbf{X}}_2^{(2)}$. Soulignons que d'après (4.8), le vecteur $\hat{\mathbf{X}}_2^{(1)}$ appartient aussi au plan $[\mathbf{C}\hat{\mathbf{X}}_1^{(1)}]^\perp$, et donc la minimisation conduit forcément à une nouvelle valeur de f_3 qui sera inférieure ou égale à la précédente. On trouve donc une deuxième estimation $\hat{\mathbf{X}}^{(2)} := (\hat{\mathbf{X}}_1^{(1)}, \hat{\mathbf{X}}_2^{(2)})$ tel que $f_3(\hat{\mathbf{X}}_1^{(1)}, \hat{\mathbf{X}}_2^{(2)}) \leq f_3(\hat{\mathbf{X}}_1^{(1)}, \hat{\mathbf{X}}_2^{(1)})$. Ces nouvelles estimations sont reliées aux précédentes par

$$\hat{\mathbf{X}}_2^{(2)T} \mathbf{C} \hat{\mathbf{X}}_1^{(1)} = 0. \quad (4.9)$$

Ici $f_3(\mathbf{X}_1^{(1)}, \mathbf{X}_2(\theta))$ aura deux minimums $\hat{\theta}_2^{(2)}$ et $\hat{\theta}_3^{(2)}$ tels que

$$f_3(\mathbf{X}_1^{(1)}, \mathbf{X}_2(\hat{\theta}^{(2)})) = f_3(\mathbf{X}_1^{(1)}, \mathbf{X}_2(\hat{\theta}^{(3)})).$$

En effet,

$$\begin{aligned} \beta_2 &:= f_3(\mathbf{X}_1^{(1)}, \mathbf{X}_2(\hat{\theta}^{(2)})) \\ &= f(\mathbf{X}_1^{(1)}, \mathbf{X}_2(\hat{\theta}^{(2)}), \mathbf{X}_3^{(2)}) \end{aligned}$$

où

$$\mathbf{X}_3^{(2)} := [\hat{\mathbf{C}}\mathbf{X}_1^{(1)}, \hat{\mathbf{C}}\mathbf{X}_2(\hat{\theta}_2^{(2)})]^\perp.$$

Comme f est invariante aux permutations de ses arguments, alors

$$\beta_2 = f(\mathbf{X}_1^{(1)}, \mathbf{X}_3^{(2)}, \mathbf{X}_2(\hat{\theta}^{(2)}))$$

or $\mathbf{X}_2(\hat{\theta}_2^{(2)}) \in [\hat{\mathbf{C}}\hat{\mathbf{X}}_1^{(1)}]^\perp$ et $\mathbf{X}_3^{(2)} \in [\hat{\mathbf{C}}\mathbf{X}_2(\hat{\theta}_2^{(2)})]^\perp \implies \mathbf{X}_2(\hat{\theta}_2^{(2)}) \in [\hat{\mathbf{C}}\mathbf{X}_3^{(2)}]^\perp$ car \mathbf{C} est symétrique, donc $\mathbf{X}_2(\hat{\theta}_2^{(2)}) = [\hat{\mathbf{C}}\hat{\mathbf{X}}_1^{(1)}, \hat{\mathbf{C}}\mathbf{X}_3^{(2)}]^\perp$, d'où

$$\beta_2 = f_3(\mathbf{X}_1^{(1)}, \mathbf{X}_3^{(2)}).$$

Or $\mathbf{X}_3^{(2)}$ et $\mathbf{X}_2(\hat{\theta}_2^{(2)})$ sont distincts car $\mathbf{X}_2(\hat{\theta}_2^{(2)})^T \hat{\mathbf{C}}\mathbf{X}_3^{(2)} = 0$, donc on a bien deux solutions distinctes qui donnent la même valeur de f_3 .

(3) On fixe $\mathbf{X}_2 = \hat{\mathbf{X}}_2^{(2)}$, puis on minimise $f_3(\mathbf{X}_1, \hat{\mathbf{X}}_2^{(2)})$ en faisant faire une rotation à \mathbf{X}_1 dans le plan $[\mathbf{C}\hat{\mathbf{X}}_2^{(2)}]^\perp$, ce qui permet de respecter la condition d'orthogonalité

$$\mathbf{X}_1^T \mathbf{C} \hat{\mathbf{X}}_2^{(2)} = 0.$$

Le résultat de cette minimisation est un vecteur dénoté par $\hat{\mathbf{X}}_1^{(2)}$. Soulignons que d'après (4.9) le vecteur $\hat{\mathbf{X}}_1^{(1)}$ appartient aussi au plan $[\mathbf{C}\hat{\mathbf{X}}_2^{(2)}]^\perp$, puisque $\hat{\mathbf{X}}_2^{(2)T} \mathbf{C} \hat{\mathbf{X}}_1^{(1)} = \hat{\mathbf{X}}_1^{(1)T} \mathbf{C} \hat{\mathbf{X}}_2^{(2)} = 0$, et donc la minimisation conduit forcément à une nouvelle valeur de f_3 qui sera inférieure ou égale à la précédente. On trouve donc un troisième estimé $\hat{\mathbf{X}}^{(3)} := (\hat{\mathbf{X}}_1^{(2)}, \hat{\mathbf{X}}_2^{(2)})$ tel que $f_3(\hat{\mathbf{X}}_1^{(2)}, \hat{\mathbf{X}}_2^{(2)}) \leq f_3(\hat{\mathbf{X}}_1^{(1)}, \hat{\mathbf{X}}_2^{(2)})$.

(4) On itère en fixant $\hat{\mathbf{X}}_1^{(i)}$ et en cherchant $\hat{\mathbf{X}}_2^{(i+1)}$, puis en fixant $\hat{\mathbf{X}}_2^{(i+1)}$ et en cherchant $\hat{\mathbf{X}}_1^{(i+1)}$. On s'arrête lorsque l'algorithme converge, i.e. lorsque $f_3(\hat{\mathbf{X}}^{(i+1)}) \simeq f_3(\hat{\mathbf{X}}^{(i)})$.

Remarquons que

$$f_3(\hat{\mathbf{X}}^{(i)}) \geq 0$$

et que

$$f_3(\hat{\mathbf{X}}^{(i+1)}) \leq f_3(\hat{\mathbf{X}}^{(i)})$$

pour tout i , et donc la limite $\lim_{i \rightarrow \infty} f_3(\hat{\mathbf{X}}^{(i)})$ existe et est positive ou nulle.

Cette descente cyclique assure donc la convergence de $f_3(\hat{\mathbf{X}}_1^{(i)}, \hat{\mathbf{X}}_2^{(i-1)})$ vers une valeur minimale.

Soulignons que rien n'assure que le minimum global de f_3 soit obtenu par cette descente cyclique.

4.2 Expériences numériques

On considère les signaux $\tilde{\mathbf{V}}_1$, $\tilde{\mathbf{V}}_2$ et $\tilde{\mathbf{V}}_3$ dans $\mathbb{R}^N (N = 30)$, tels que

$$\tilde{\mathbf{V}}_i = \frac{\mathbf{W}_i}{\|\mathbf{W}_i\|},$$

$i=1,2,3$, où $W_1(j) = j$, $W_2(j) = j^2$, $W_3(j) = j^3$ et $j = 0, 1, 2, \dots, N - 1$. On génère M signaux ($M = 500$) de la forme

$$\tilde{\mathbf{S}} = a_1 \tilde{\mathbf{V}}_1 + a_2 \tilde{\mathbf{V}}_2 + a_3 \tilde{\mathbf{V}}_3,$$

où a_1 , a_2 et a_3 sont des variables aléatoires uniformes et mutuellement indépendantes telles que $a_i \sim U[-\sqrt{3}, \sqrt{3}]$, $i = 1, 2, 3$.

Les cinq premières valeurs propres de $\tilde{\mathbf{C}}$ sont :

$$\{3.01798, 0.081469, 0.00110198, -1.17468 \times 10^{-16}, 6.3345 \times 10^{-17}\}$$

On constate que seules les trois premières valeurs propres sont significativement différentes de zéro, ce à quoi nous nous attendions.

Pour choisir la valeur de γ , nous avons comparé les valeurs des termes de la fonction \hat{f} . Nous avons donc effectué des essais avec des échantillons différents et des valeurs de γ différentes et nous avons comparé l'influence de chacun des termes dans la courbe de

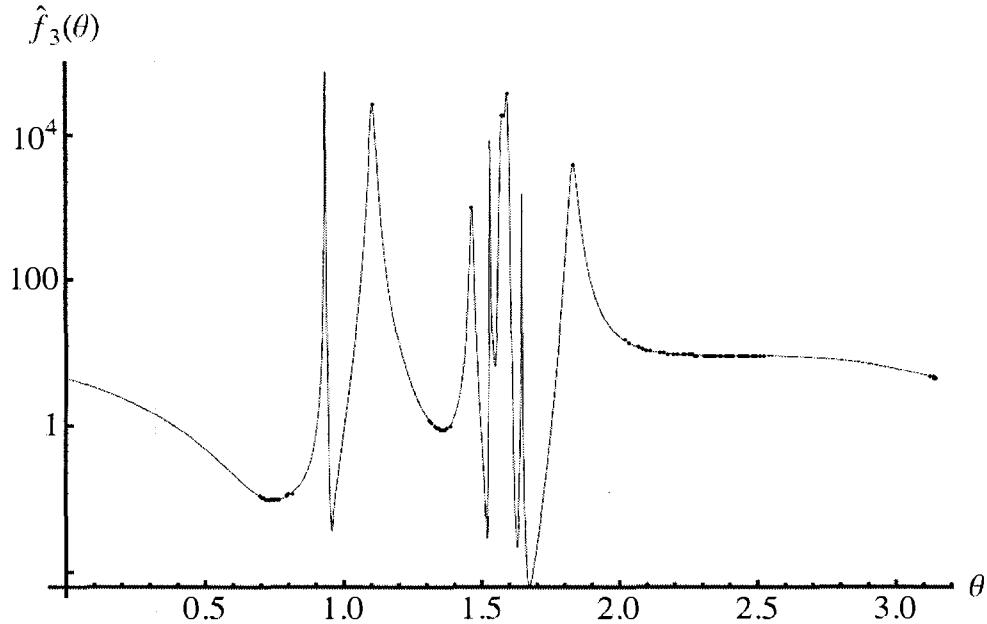


Figure 4.1 Graphe de $\tilde{f}_3(\theta)$, $\theta \in [0, \pi]$, dans $\text{span}(\tilde{\Phi}_2, \tilde{\Phi}_3)$.

$\hat{f}(\theta)$, $\theta \in [0, \pi]$. Nous avons remarqué que si on prend $\gamma < 5 \times 10^3$, alors le terme $\mathcal{C}_{1,2}$ n'a pas d'influence et les autres termes dominent la courbe de la fonction \hat{f} , tandis que si on prend $\gamma > 5 \times 10^4$, alors $\mathcal{C}_{1,2}$ domine la courbe de \hat{f} et les autres termes n'ont pas d'influence. Nous avons donc choisi $\gamma = 10^4$ pour que tous les termes aient une grandeur comparable.

Nous avons tracé la courbe de $\hat{f}_3(\mathbf{X}_1(\theta), \mathbf{X}_2(\theta))$ en fonction de θ pour $\theta \in [0, \pi]$ (figure 4.1), où $\mathbf{X}_1 \in \text{span}(\tilde{\Phi}_2, \tilde{\Phi}_3)$. Les vecteurs $\mathbf{X}_1(\theta)$ et $\mathbf{X}_2(\theta)$ sont obtenus avec (4.5).

Nous avons estimé $\hat{\theta}$ qui minimise $\hat{f}_3(\mathbf{X}_1(\theta), \mathbf{X}_2(\theta))$, qui conduit à la première estimation $\hat{X}^{(1)} := (\hat{\mathbf{X}}_1^{(1)}, \hat{\mathbf{X}}_2^{(1)})$.

Nous avons fixé $\mathbf{X}_1 := \hat{\mathbf{X}}_1^{(1)}$, puis nous avons minimisé $f_3(\hat{\mathbf{X}}_1^{(1)}, \mathbf{X}_2)$ en faisant faire une rotation à \mathbf{X}_2 dans le plan $[\hat{\mathbf{C}}\hat{\mathbf{X}}_1^{(1)}]^\perp$. Soient Ψ_1 et Ψ_2 deux vecteurs orthonormaux du plan $[\hat{\mathbf{C}}\hat{\mathbf{X}}_1^{(1)}]^\perp$.

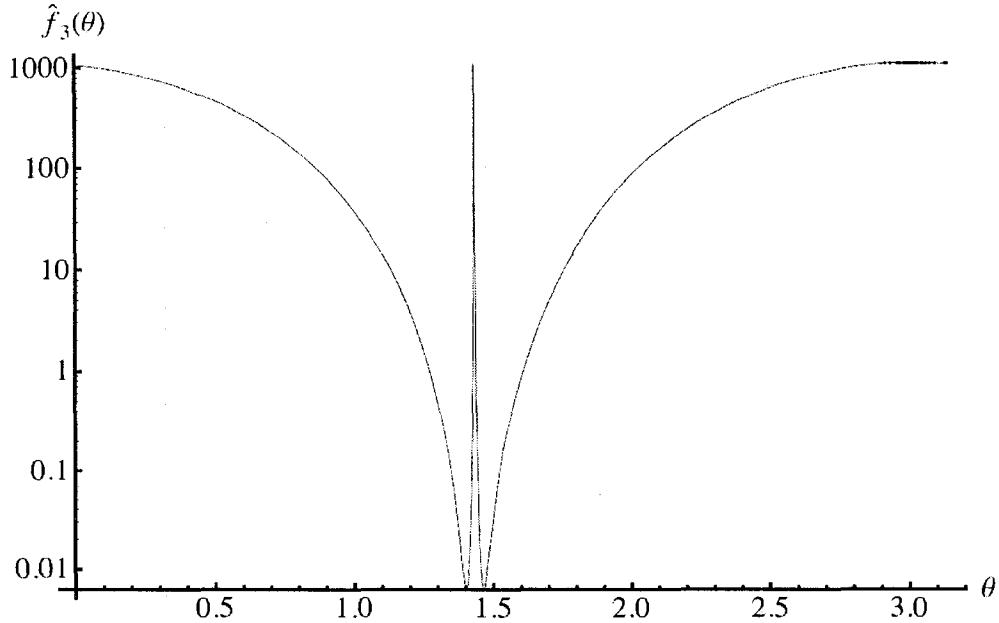


Figure 4.2 Graphe de $\tilde{f}_3(\theta)$, $\theta \in [0, \pi]$, dans $\text{span}([\hat{\mathbf{C}}\hat{\mathbf{X}}_1^{(1)}]^\perp)$.

Nous avons tracé la courbe de $f_3(\mathbf{X}_1^{(1)}, \mathbf{X}_2(\theta))$ en fonction de θ , $\theta \in [0, \pi]$ (figure 4.2), où

$$\mathbf{X}_2(\theta) := \cos(\theta)\mathbf{\Psi}_1 + \sin(\theta)\mathbf{\Psi}_2.$$

Comme $f_3(\mathbf{X}_1^{(1)}, \mathbf{X}_2(\theta))$ aura deux minimums $\hat{\theta}_2^{(2)}$ et $\hat{\theta}_3^{(2)}$ tels que

$$f_3(\mathbf{X}_1^{(1)}, \mathbf{X}_2(\hat{\theta}^{(2)})) = f_3(\mathbf{X}_1^{(1)}, \mathbf{X}_2(\hat{\theta}^{(3)})),$$

on utilise alors $\hat{\theta}_2^{(2)}$ ou $\hat{\theta}_3^{(2)}$ pour obtenir la deuxième estimation de $\bar{\mathbf{X}}$, $\hat{\mathbf{X}}^{(2)} := (\hat{\mathbf{X}}_1^{(1)}, \hat{\mathbf{X}}_2^{(2)})$ tel que $f_3(\hat{\mathbf{X}}^{(2)}) \leq f_3(\hat{\mathbf{X}}^{(1)})$.

Nous avons ensuite fixé $\mathbf{X}_2 := \hat{\mathbf{X}}_2^{(2)}$ et nous avons minimisé $f_3(\mathbf{X}_1, \hat{\mathbf{X}}_2^{(2)})$ en faisant faire une rotation à \mathbf{X}_1 dans le plan $[\hat{\mathbf{C}}\hat{\mathbf{X}}_2^{(2)}]^\perp$. Nous avons utilisé la même procédure pour obtenir une troisième estimation $\hat{\mathbf{X}}^{(3)}$.

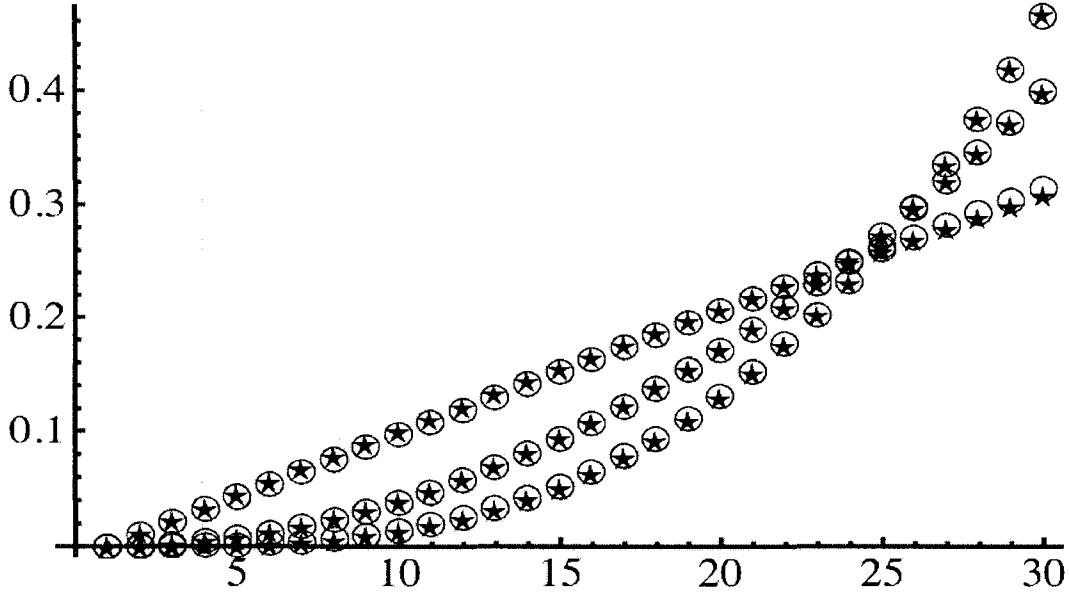


Figure 4.3 Graphe des composantes estimées avec notre méthode (*) et les composantes indépendantes (o), où $W_1(j) = j, W_2(j) = j^2, W_3(j) = j^3, j = 0, \dots, N - 1$ et $M = 500$.

Nous avons itéré cette procédure en fixant successivement $\hat{\mathbf{X}}_1^{(i)}$ et en cherchant $\hat{\mathbf{X}}_2^{(i+1)}$, et en fixant $\hat{\mathbf{X}}_2^{(i)}$ et en cherchant $\hat{\mathbf{X}}_1^{(i)}$. On s'arrête lorsque l'algorithme converge, i.e. lorsque $f_3(\hat{\mathbf{X}}^{(k)}) - f_3(\hat{\mathbf{X}}^{(k-1)}) \leq 10^{-8}$.

Nous avons estimé ensuite $\hat{\mathbf{V}}_1, \hat{\mathbf{V}}_2$ et $\hat{\mathbf{V}}_3$ en utilisant

$$\left\{ \begin{array}{l} \hat{\mathbf{V}}_1 = [\hat{\mathbf{X}}_2^{(k)}, \hat{\mathbf{X}}_3^{(k)}]^\perp \\ \hat{\mathbf{V}}_2 = [\hat{\mathbf{X}}_1^{(k)}, \hat{\mathbf{X}}_3^{(k)}]^\perp \\ \hat{\mathbf{V}}_3 = [\hat{\mathbf{X}}_1^{(k)}, \hat{\mathbf{X}}_2^{(k)}]^\perp. \end{array} \right.$$

La figure (4.3) nous montre la superposition des vecteurs estimés $\hat{\mathbf{V}}_1, \hat{\mathbf{V}}_2$ et $\hat{\mathbf{V}}_3$ qui sont représentés par des étoiles avec les composantes indépendantes $\tilde{\mathbf{V}}_1, \tilde{\mathbf{V}}_2$ et $\tilde{\mathbf{V}}_3$ qui sont représentés par des cercles.

Nous avons fait d'autres essais en utilisant chaque fois des composantes et des

distributions différentes, notre algorithme a estimé les composantes indépendantes avec précision dans au moins 90% des cas pour $M \geq 500$ et dans presque tous les cas pour $M \geq 5000$.

La figure (4.4) nous montre la superposition des vecteurs estimés $\hat{\mathbf{V}}_1$, $\hat{\mathbf{V}}_2$ et $\hat{\mathbf{V}}_3$ qui sont représentés par des étoiles avec les composantes indépendantes $\tilde{\mathbf{V}}_1$, $\tilde{\mathbf{V}}_2$ et $\tilde{\mathbf{V}}_3$ qui sont représentés par des cercles, où on a utilisé $W_1(j) = j$, $W_2(j) = j^2$, $W_3(j) = \log(j + 1)$, $j = 0, \dots, N - 1$ et $M = 5000$.

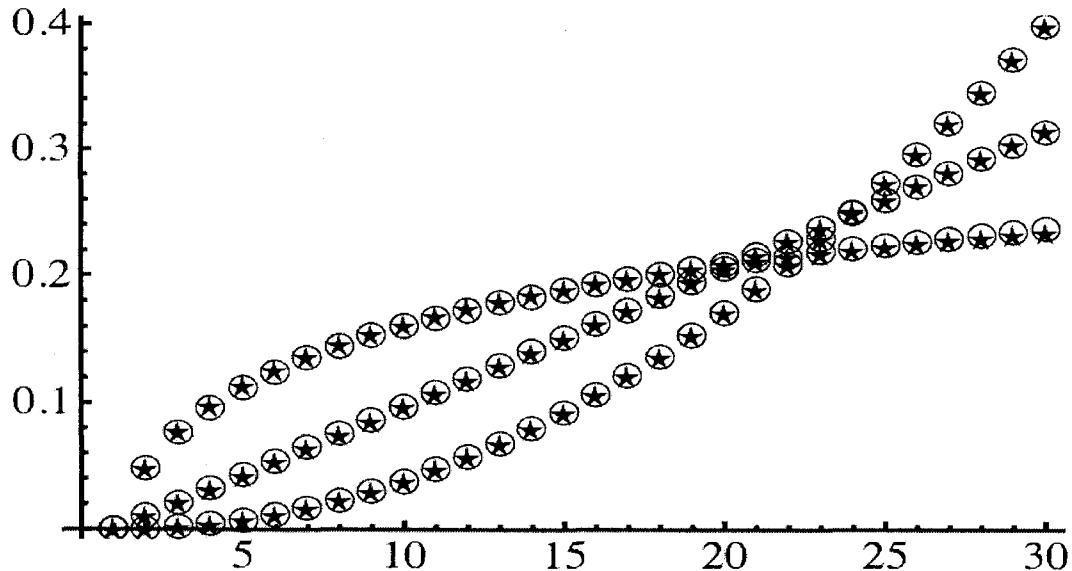


Figure 4.4 Graphe des composantes estimées avec notre méthode (*) et les composantes indépendantes (o), où $W_1(j) = j$, $W_2(j) = j^2$, $W_3(j) = \log(j + 1)$, $j = 0, \dots, N - 1$ et $M = 5000$.

Remarquons que pour les cas où $M < 500$, on obtient généralement une estimation approximative des $\hat{\mathbf{V}}_i$, $i = 1, 2, 3$, mais avec une précision inférieure à celle des cas où $M \geq 500$ (figure (4.5)). L'erreur d'estimation sur les \mathbf{V}_i , $i = 1, 2, 3$ a deux origines distinctes : l'estimation des $\mathbf{Z}_n(\mathbf{X})$ d'une part, et la minimisation de $f_3(\mathbf{X}_1, \mathbf{X}_2)$ d'autre part. Cette dernière ne conduit pas nécessairement au minimum global de f_3 . En effet, le point de convergence de la descente dépend de l'estimation initiale

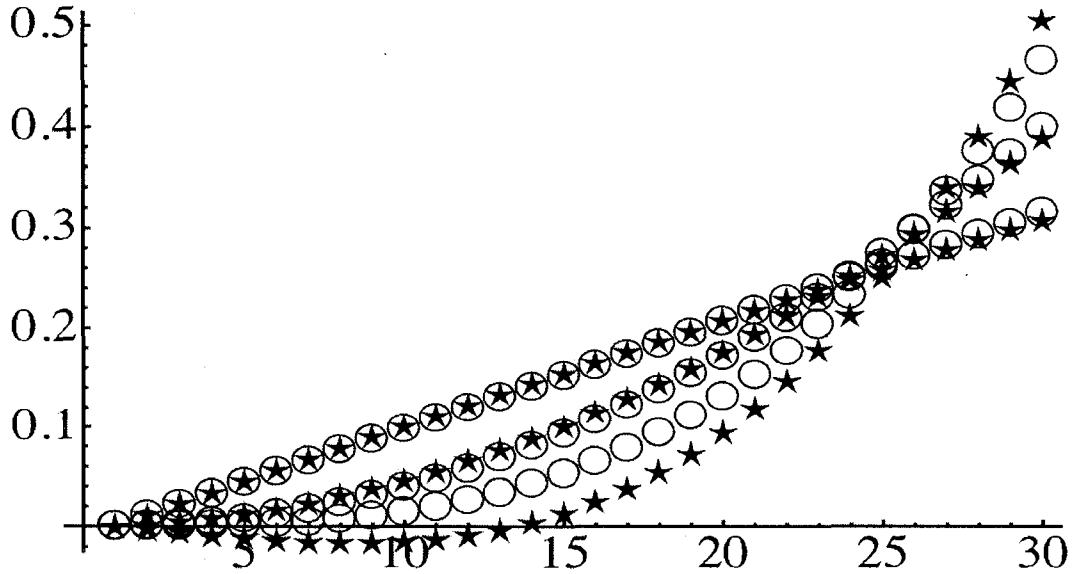


Figure 4.5 Graphe des composantes estimées avec notre méthode (\star) et les composantes indépendantes (\circ), où $W_1(j) = j, W_2(j) = j^2, W_3(j) = j^3, j = 0, \dots, N - 1$ et $M = 100$.

$\hat{\mathbf{X}}_1^{(1)}$ qui est imparfaite.

Nous avons utilisé un ordinateur *Apple MacBook Pro* dont les spécifications sont présentées dans la figure (4.6) et le logiciel *Mathematica* version 6.0.1.0 pour l'exécution de cet algorithme. Le temps d'exécution observé de cet algorithme dans le cas où $500 \leq M \leq 5000$ est de l'ordre de 15 secondes en moyenne. Ce temps dépend de la taille d'échantillon et du nombre d'itérations. En fait, le temps d'estimation de la matrice $\hat{\mathbf{C}}$ et les $\hat{\Phi}_i, i = 1, 2, 3$, est proportionnel à la taille d'échantillon, et le temps moyen d'une itération est d'environ 2 secondes. En pratique, le nombre d'itérations est généralement entre 3 et 12, en moyenne l'algorithme converge après 5 ou 6 itérations.



Figure 4.6 Spécifications de l'ordinateur utilisé pour l'exécution de nos algorithmes.

CONCLUSION

Dans ce travail nous avons développé deux méthodes pour l'estimation des composantes indépendantes dans le cas où $p = 2$. Nous avons démontré que ces deux méthodes estiment les composantes indépendantes avec une bonne précision dans la majorité des cas, même dans le cas où la taille d'échantillon est petite, i.e. $M = 30$.

La première méthode consiste à déterminer les vecteurs \mathbf{X} , tels que tous les vecteurs $\{\mathbf{Z}_n(\mathbf{X}), n \in \mathbb{N}\}$ soient colinéaires. Nous avons donc cherché les vecteurs \mathbf{X} dans \mathbb{R}^2 , tels que $\mathbf{Z}_1(\mathbf{X})$ et $\mathbf{Z}_3(\mathbf{X})$ soient colinéaires. Comme les vecteurs dans \mathbb{R}^2 sont déterminés à partir d'un seul paramètre qui est l'angle θ , notre problème consiste alors à chercher les angles θ où $\mathbf{Z}_1(\mathbf{X}(\theta))$ et $\mathbf{Z}_3(\mathbf{X}(\theta))$ sont colinéaires. Nous avons démontré que la résolution de ce problème nous donne quatre vecteurs et qu'on peut distinguer les estimations de \mathbf{X}_1 et \mathbf{X}_2 en utilisant la covariance $\mathcal{C}_{i,j}$ des vecteurs trouvés.

La deuxième méthode est basée sur la minimisation d'une fonction objectif. Nous avons démontré que cette fonction est positive et les solutions de cette dernière sont les composantes indépendantes. L'analyse des deux méthodes nous montre que les estimations sont précises. Nous avons aussi comparé nos méthodes avec la méthode *FastICA*, et nous avons montré que nos méthodes sont plus précises. Notons aussi que nos méthodes sont très rapides, le temps d'estimation est de l'ordre de fractions de seconde.

Nous avons développé aussi une méthode pour la détermination des composantes indépendantes dans le cas où $p = 3$. Cette méthode est une extension de la deuxième méthode, on minimise une fonction objectif à l'aide d'une procédure de descente cyclique. Nous avons démontré que cette descente converge vers une valeur minimale.

Notre objectif était de développer une méthode de détermination des composantes

indépendantes pour tout p . Nous avons pu développer les méthodes pour $p = 2$ et $p = 3$. Pour la suite, on peut peut-être développer une extension de la méthode pour les cas où $p \geq 4$.

RÉFÉRENCES

- AN, G., WU, J. ET RUAN, Q. (2007). A novel model for two-dimensional independent tv-based quotient image and its application to face recognition. Piscataway, NJ, USA, 345 – 8.
- BINGHAM, E. ET HYVÄRINEN, A. (2000). A fast fixed-point algorithm for independent component analysis of complex valued signals. *Int. J. Neural Syst. (Singapore)*, 10, 1 – 8.
- CARDOSO, J.-F. ET COMON, P. (1996). Independent component analysis, a survey of some algebraic methods. New York, NY, USA, vol. vol.2, 93 – 6.
- CHAABOUNI-CHOUAYAKH, H. ET DATCU, M. (2007). PCA vs. ICA decomposition of high resolution sar images : application to urban structures recognition. USA, vol. 6746, 674608 – 1.
- COMON, P. (1990a). Independent component analysis and blind identification. *Trait. Signal (France)*, 7, 435 – 50.
- COMON, P. (1990b). Statistical approach to the Jutten-Hérault algorithm. Berlin, Germany, 81 – 8.
- COMON, P. (1992). Independent component analysis [signal processing]. Amsterdam, Netherlands, 29 – 38.
- COMON, P. (1994). Independent component analysis, a new concept ? *Signal Process. (Netherlands)*, 36, 287 – 314.
- COMON, P., JUTTEN, C. ET HÉRAULT, J. (1991). Blind separation of sources. ii. problems statement. *Signal Process. (Netherlands)*, 24, 11 – 20.

- DE LATHAUWER, L., COMON, P., DE MOOR, B. ET VANDEWALLE, J. (1999). Ica algorithms for 3 sources and 2 sensors. Los Alamitos, CA, USA, 116 – 20.
- GRELLIER, O. ET COMON, P. (1998). Blind separation of discrete sources. *IEEE Signal Process. Lett. (USA)*, 5, 212 – 14.
- HÉRAULT, J., JUTTEN, C. ET ANS, B. (1985). Détection des grandeurs primitives dans un message composite par une architecture de calcul neuromimétique en apprentissage non supervisé. *Actes du Xième colloque GRETSI, Nice, France*. vol. 2, 1017–1022.
- HERRMANN, J. ET THEIS, F. (2007). Statistical analysis of sample-size effects in ICA. *Intelligent Data Engineering and Automated Learning - IDEAL 2007. Proceedings 8th International Conference. (Lecture Notes in Computer Science vol. 4881)*, 416 – 25.
- HYVÄRINEN, A. (1997). Independent component analysis : a neural network approach. *Acta Polytech. Scand. Math. Comput. Manag. Eng. Ser. (Finland)*, 1 – 60.
- HYVÄRINEN, A. (2001). Blind source separation by nonstationarity of variance : a cumulant-based approach. *IEEE Trans. Neural Netw. (USA)*, 12, 1471 – 4.
- HYVÄRINEN, A. (2005). A unifying model for blind separation of independent sources. *Signal Process. (Netherlands)*, 85, 1419 – 27. Blind separation source ;independent source ;unifying model ;autoregressive model ;maximum likelihood estimation ;independent component analysis ;.
- HYVÄRINEN, A. ET OJA, E. (1996). Simple neuron models for independent component analysis. *Int. J. Neural Syst. (Singapore)*, 7, 671 – 87.

- HYVARINEN, A. ET OJA, E. (1997). A fast fixed-point algorithm for independent component analysis. *Neural Comput. (USA)*, 9, 1483 – 92.
- HYVARINEN, A. ET OJA, E. (1998). Independent component analysis by general nonlinear hebbian-like learning rules. *Signal Process. (Netherlands)*, 64, 301 – 13.
- HYVARINEN, A. ET OJA, E. (2000). Independent component analysis : algorithms and applications. *Neural Netw. (UK)*, 13, 411 – 30.
- INKI, M. ET HYVARINEN, A. (2002). Two approaches to estimation of overcomplete independent component bases. Piscataway, NJ, USA, vol. vol.1, 454 – 9.
- JUTTEN, C. ET HÉRAULT, J. (1988). Independent components analysis (inca) versus principal components analysis. Amsterdam, Netherlands, 643 – 6.
- JUTTEN, C. ET HÉRAULT, J. (1991). Blind separation of sources i. an adaptive algorithm based on neuromimetic architecture. *Signal Process. (Netherlands)*, 24, 1 – 10.
- KOHNO, S., MIYAI, I., SEIYAMA, A., ODA, I., ISHIKAWA, A., TSUNEISHI, S., AMITA, T. ET SHIMIZU, K. (2007). Removal of the skin blood flow artifact in functional near-infrared spectroscopic imaging data through independent component analysis. *J. Biomed. Opt. (USA)*, 12, 1 – 9.
- LAI, D., RAO, N., HUI KUO, C., BHATT, S. ET DOGRA, V. (2007). Independent component analysis applied to ultrasound speckle texture analysis and tissue characterization. Piscataway, NJ, USA, 6523 – 6.
- LI, R. ET WANG, F. (2007). Adaptive semiblind signal separation approach using temporal structure of sources. Piscataway, NJ, USA, 315 – 18.

- NAIK, G., KUMAR, D. ET WEGHORN, H. (2007). Performance comparison of ica algorithms for isometric hand gesture identification using surface emg. Piscataway, NJ, USA, 613 – 18.
- OU, S., ZHAO, X. ET GAO, Y. (2007). Linear system identification employing independent component analysis. Piscataway, NJ, USA, 1336 – 40.
- SONG-YUN, X., ZHEN-ZHONG, Z., WEI-PING, Z. ET HAI-TAO, Z. (2007). Method and application of removing noise from EEG signals based on ICA method. *Chin. J. Med. Imag. Technol. (China)*, 23, 1562 – 5.
- YANG, S. ET YI, Z. (2007). Nonnegative matrix factorization for independent component analysis. Kokura, Japan, 769 – 71.
- ZEWAIL, R., ELSAFI, A. ET DURDLE, N. (2007). Wavelet-based independent component analysis for statistical shape modeling. Vancouver, BC, Canada, 1325 – 8.
- ZHANG, R., ZHANG, X.-P. ET GUAN, L. (2007). Wavelet-based texture retrieval using independent component analysis. Piscataway, NJ, USA, 341 – 4.