



HAL
open science

Méthodes de séparation aveugle de sources non linéaires, étude du modèle quadratique 2*2

Chahinez Chaouchi

► **To cite this version:**

Chahinez Chaouchi. Méthodes de séparation aveugle de sources non linéaires, étude du modèle quadratique 2*2. Traitement du signal et de l'image [eess.SP]. Université Paul Sabatier - Toulouse III, 2011. Français. NNT: . tel-00683204

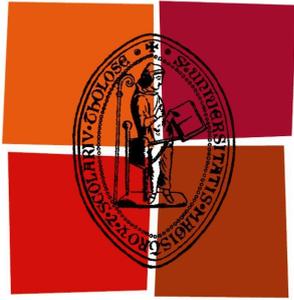
HAL Id: tel-00683204

<https://theses.hal.science/tel-00683204>

Submitted on 28 Mar 2012

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

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



Université
de Toulouse

THÈSE

En vue de l'obtention du

DOCTORAT DE L'UNIVERSITÉ DE TOULOUSE

Délivré par : *l'Université Toulouse III - Paul Sabatier*

Discipline ou spécialité : *Traitement du Signal*

Présentée et soutenue par *Chahinez CHAOUCHI*

Le *Juin 2011*

Titre : *Méthodes de séparation aveugle de sources non linéaires, étude du modèle quadratique*

Ecole doctorale : *Mathématique, Informatique et Télécommunications*

Jury

M. KARIM ABED-MERAIM	Professeur ENST, Paris	(Rapporteur)
M. XAVIER BRIOTTET	Directeur de recherche ONERA, Toulouse	(Examinateur)
M. YANNICK DEVILLE	Professeur, IRAP-Université de Toulouse	(Directeur)
M. SHAHRAM HOSSEINI	MCF, IRAP-Université de Toulouse	(Directeur)
M. CHRISTIAN JUTTEN	Professeur, Gipsa-UJF, Grenoble	(Rapporteur)
M. ERIC MOREAU	Professeur, ISITV, Toulon	(Président)

Unité de recherche : *Institut de Recherche en Astrophysique et Planétologie, IRAP*

Directeur(s) de Thèse : *Yannick Deville & Shahram Hosseini*

Table des matières

Table des figures	vii
Liste des tableaux	ix
Notations et Abréviations	xi
Introduction générale	1

Chapitre 1 État de l'art

1.1	Introduction	6
1.2	Mélanges linéaires	6
1.2.1	Différents types de mélanges	6
1.2.2	Familles de méthodes et critères pour la SAS	8
1.2.3	SAS et ACI : du cas linéaire au cas non-linéaire	12
1.3	Mélanges non-linéaires	13
1.3.1	Panorama des méthodes	13
1.3.2	Formulation mathématique du problème	17
1.3.3	Existence et Unicité des solutions fournies par l'ACI non-linéaire	19
1.3.4	Mélanges Post-Non-Linéaires (PNL)	21
1.4	Génèse de la thèse : du modèle linéaire-quadratique au modèle qua- dratique	23
1.4.1	Modèle linéaire-quadratique	24
1.4.2	Modèle quadratique	25
1.4.3	Objectifs de la thèse	26
1.5	Conclusion	26

Chapitre 2 Réseaux récurrents
--

2.1	Introduction	30
-----	------------------------	----

2.2	Travaux antérieurs	30
2.3	Modèle linéaire-quadratique	31
2.3.1	Généralités	31
2.3.2	Étude du réseau récurrent	34
2.4	Modèle quadratique	37
2.4.1	Réseau récurrent associé	38
2.4.2	Points d'équilibre du réseau	39
2.4.3	Stabilité locale des points d'équilibre du réseau	41
2.4.4	Limites de la stabilité locale du réseau étendu	42
2.5	Cas où les sorties sont permutées	44
2.6	Comparatif des travaux sur les deux modèles	45
2.7	Simulations	46
2.7.1	Réseau basique	47
2.7.2	Améliorations avec le réseau étendu : scénario n° 1	48
2.7.3	Améliorations avec le réseau étendu : scénario n° 2	49
2.8	Conclusion	51

<p>Chapitre 3</p> <p>Identification du mélange par une méthode fondée sur des cumulants</p>

3.1	Introduction	54
3.2	Travaux antérieurs	54
3.3	Approche proposée et expressions des cumulants des observations	55
3.3.1	Propriétés des cumulants et notations	56
3.3.2	Cumulants d'ordre 1	57
3.3.3	Cumulants d'ordre 2	57
3.3.4	Cumulants d'ordre 3	59
3.3.5	Cumulants d'ordre 4	63
3.4	Estimation des paramètres du mélange	64
3.4.1	Estimation des cumulants des observations	65
3.4.2	Résolution du système	65
3.5	Simulations	66
3.5.1	Premier scénario	66
3.5.2	Deuxième scénario	68
3.5.3	Troisième scénario	69
3.6	Conclusion	70

Chapitre 4**Approche par maximum de vraisemblance**

4.1	Introduction	74
4.2	Travaux antérieurs	74
4.2.1	Généralités	74
4.2.2	Modèles de densités des sources	75
4.3	Estimation des coefficients du mélange : approche par maximum de vraisemblance (MV)	76
4.3.1	Modèle de mélange	76
4.3.2	Critère d'estimation des coefficients du mélange : principe gé- néral	77
4.3.3	Détails des calculs du gradient de la log-vraisemblance	79
4.3.4	Algorithme d'optimisation	81
4.4	Structure récurrente associée	81
4.5	Simulations	83
4.6	Conclusions	90

Conclusions et perspectives 91

Annexes**Annexe A****Stabilité du réseau basique : cas d'existence d'un seul point d'équi-
libre**

A.1	Etude théorique de la condition d'existence du point non séparant dans \mathbb{R}	95
A.1.1	Etude du signe de δ_1'	96
A.1.2	Conséquence sur la condition (A.1)	96
A.1.3	Informations supplémentaires : signe des racines	97
A.2	Etude de la stabilité au point d'équilibre	97
A.2.1	Cas où les sources sont de même signe	98
A.2.2	Cas où les sources sont de signes contraires	102
A.3	Conclusion	102

Annexe B**Détails des calculs pour les « cumulants sources »**

B.1	Cumulants associés aux sources d'ordre 4 non nuls	105
B.2	Cumulants associés aux sources d'ordre 4 nuls	108

Annexe C

Méthodes de Gauss-Newton et Levenberg-Marquardt

C.1	Position du problème	113
C.2	Méthode de Gauss-Newton (GN)	114
C.3	Méthode de Levenberg-Marquardt (LM)	116
C.4	Rappel sur la méthode de Newton locale	117

Annexe D

Fonctions MATLAB

D.1	Estimation des cumulants	119
D.2	Résolution du système non-linéaire	123

Bibliographie	125
----------------------	------------

Publications scientifiques de l'auteur	135
---	------------

Résumé	137
---------------	------------

Table des figures

1.1	Structures de mélange et de séparation pour le modèle PNL	21
1.2	Récapitulatif des différents travaux effectués sur les modèles linéaire- quadratique et quadratique.	27
2.1	Structure récurrente basique pour le modèle linéaire-quadratique . . .	32
2.2	Structure récurrente étendue, avec termes d'auto-feedback, pour le modèle linéaire-quadratique.	33
2.3	Modèle de mélange et principe de construction du réseau récurrent pour le modèle linéaire-quadratique.	33
2.4	Structure récurrente proposée pour le modèle quadratique	38
2.5	Réseau basique : Convergence locale vers le point séparant (15, 3). $s_1 = 15, s_2 = 3, L_{12} = -0.5, L_{21} = -0.2, Q_1 = 1, Q_2 = 0.5,$ $M = 100, r = 0.1, K = 6.$	47
2.6	Réseau basique : Divergence locale du point non-séparant (14.7087, 2.3534). $s_1 = 15, s_2 = 3, L_{12} = -0.5, L_{21} = -0.2, Q_1 = 1, Q_2 = 0.5, M = 100,$ $r = 0.1, K = 6.$	48
2.7	Réseau étendu : Stabilisation locale du point séparant $s_1 = s_2 = 4,$ $L_{12} = L_{21} = L = -0.75, Q_1 = Q_2 = Q = -0.5, k_{i,i \in \{1,2\}} = \sqrt{2},$ $M = 100, r = 0.1, K = 6.$	49
2.8	Réseau étendu : Stabilisation locale du point séparant $s_1 = 1, s_2 = 1,$ $L_{12} = -0.2, L_{21} = -0.1, Q_1 = -0.9, Q_2 = -2.1, k_1 = 3, k_2 = 2,$ $M = 40, r = 0.1, K = 6.$	50
3.1	Convergence vers "S" pour les paramètres $\{L_{12}, L_{21}, Q_1, Q_2\}$	66
3.2	Phénomène d'inversion de signe pour les paramètres Q_1 et Q_2	67

3.3	Valeur moyenne pour 100 réalisations des sources de l'EQM en fonction de Q pour les 12 paramètres.	69
3.4	Valeur moyenne de l'EQM pour 100 réalisations des sources en fonction de Q calculée pour les paramètres $\{L_{12}, L_{21}, Q_1, Q_2\}$	70
4.1	Structure récurrente associée (Figure 2.4 avec $l_{11} = l_{22} = 0$)	82
4.2	Tracé de la source 1 :uniforme entre [5.5, 6.5]	86
4.3	Tracé de la source 2 :uniforme entre [2.5, 3.5]	87
4.4	Tracé de l'observation 1 paramètres du mélange fixés à $[L_{12}, L_{21}, Q_1, Q_2] = [-1; -0.2; 1; -0.5]$	87
4.5	Tracé de l'observation 2 paramètres du mélange fixés à $[L_{12}, L_{21}, Q_1, Q_2] = [-1; -0.2; 1; -0.5]$	88
4.6	Tracé de l'observation 1 en fonction de l'observation 2 paramètres du mélange fixés à $[L_{12}, L_{21}, Q_1, Q_2] = [-1; -0.2; 1; -0.5]$	88
4.7	Tracé de la source 1 en fonction de la source 2	89
4.8	Tracé de l'estimée 1 en fonction de l'estimée 2	89
A.1	Tracés des fonctions $Z_1 = 2y(1+x) - 2y(x-1)\sqrt{1-\frac{1}{y}}$ et $Z_2 = \frac{(1+x)^2}{x}$	101
A.2	Tracés de la fonction $Z_3 = 2y(1+x) - 2y(x-1)\sqrt{1-\frac{1}{y} - \frac{(1+x)^2}{x}}$	101

Liste des tableaux

3.1	Nombre d'équations et d'inconnues pour notre système en fonction de l'ordre des cumulants.	65
3.2	Nombre d'occurrence de la convergence vers le point d'intérêt S et les points parasites pour les algorithmes de Levenberg-Marquardt (LM) et Gauss-Newton (GN).	68
4.1	Moyenne, Ecart-type, Maximum, Minimum du SIR et moyenne de SIR_0 (dB) sur les sorties pour différentes statistiques des sources. . .	85
4.2	Moyenne, Ecart-type, Maximum, Minimum du SIR et moyenne de SIR_0 (dB) sur les sorties pour différentes statistiques des sources et différentes réalisations des sources.	86
A.1	Étude des signes pour la vérification de l'inéquation (A.23)	99

Notations et Abréviations

Notations

- Les vecteurs sont notés en minuscule et en **gras**.
- \mathbf{u}^T désigne la transposée du vecteur \mathbf{u} .
- \hat{u} est l'estimée de u .
- $s_i(t)$: un échantillon du signal unidimensionnel s_i (signal temporel sauf indication contraire).
- $\mathbf{s}(t)$: vecteur colonne des signaux sources $\mathbf{s}(t) = [s_1(t), \dots, s_k(t)]^T$.
- $\mathbf{x}(t)$: vecteur colonne des échantillons $x_i(t)$ des observations.
- \mathbf{A} : matrice de mélange.
- \mathbf{B} : matrice de séparation.
- $E[.]$ est l'espérance mathématique.

Sigles et abréviations

ACI (ou ICA)	:	Analyse en Composantes Indépendantes (<i>Independent Component Analysis</i>)
ACI NL (ou NL ICA)	:	Analyse en Composantes Indépendantes Non-Linéaire (<i>Nonlinear Independent Component Analysis</i>)
AMUSE	:	<i>Algorithm for Multiple Unknown Signal Extraction</i>
ATM	:	<i>Additive-Target Mixtures</i>
BPSK	:	<i>Binary Phase Shift Keying</i>
CCA	:	Composantes Cibles Additives
CCF	:	Capteurs à Courants de Foucault
COM	:	<i>Contrast Maximization</i>
EQM (ou MSE)	:	Erreur Quadratique Moyenne (<i>Mean Square Error</i>)
FMN (ou NMF)	:	Factorisation en Matrices Non-négatives (<i>Non-negative Matrix Factorization</i>)
GN	:	algorithme de Gauss-Newton
i.i.d.	:	Indépendants et Identiquement Distribués
IM	:	Information Mutuelle
JADE	:	<i>Joint Approximative Diagonalization of Eigen-matrices</i>
LI-TEPCORR	:	<i>Linear-Instantaneous Temporal correlation</i>
LQ-TEPCORR	:	<i>Linear-Quadratic Temporal correlation</i>
LM	:	algorithme de Levenberg-Marquardt
MC	:	Moindres Carrés (<i>Least Mean Square</i>)
MV (ou ML)	:	Maximum de Vraisemblance (<i>Maximum Likelihood</i>)
pdf	:	densité de probabilité (<i>Probability Density Function</i>)
PNL	:	Post-Non-Linéaire
SAS	:	Séparation Aveugle de Sources
SAS NL	:	Séparation Aveugle de Sources Non-Linéaire
SCA	:	Analyse en Composantes Parcimonieuses (<i>Sparse Component Analysis</i>)
RSI (ou SIR)	:	Rapport Signal à Interférences (<i>Signal to Interference Ratio</i>)
SOBI	:	<i>Second Order Blind Identification</i>

Introduction générale

La séparation de sources consiste à restaurer un ensemble de signaux sources à partir d'observations qui correspondent à des *mélanges* de ces sources. La Séparation de Sources est dite Aveugle (SAS) lorsqu'un minimum d'information a priori est disponible. Le problème de SAS date du début des années 1980, il a été formalisé par les travaux de Ans, Hérault et Jutten dans le cadre d'études neurobiologiques et peut-être indépendamment par Bar-Ness dans le domaine des communications [11, 48]. Ceci a suscité depuis de nombreuses publications et l'engouement de plusieurs communautés scientifiques. Les domaines d'application de la SAS sont variés : l'acoustique, l'audio, la parole, le domaine biomédical, l'astrophysique, la classification d'images, les télécommunications, les capteurs chimiques...

Au delà de l'intérêt porté au modèle de mélange linéaire, son utilisation peut s'avérer limitée pour des applications réelles, d'où la nécessité d'avoir recours à des modèles non linéaires, plus complexes certes, mais permettant de modéliser au mieux la réalité. Le cas du mélange linéaire a été largement abordé en SAS contrairement au cas non linéaire, plus délicat et moins étudié. Il n'existe pas de méthode standard pour aborder ce dernier type de modèle, l'étude se faisant pratiquement au cas par cas. Malgré ces difficultés, certains auteurs se sont tout de même penchés sur le problème des mélanges non linéaires qui restent plus difficiles à aborder. Un historique complet sur la séparation aveugle de sources suivi d'un développement sur les travaux effectués sur les mélanges non-linéaires peut être trouvé dans [11, 23, 47, 48].

Cette thèse s'inscrit dans le cadre de la SAS pour des modèles non linéaires. Elle fait suite aux travaux effectués par Y. Deville et S. Hosseini dans le cas particulier du mélange linéaire-quadratique (termes linéaires et termes croisés d'ordre 2) [28, 40, 41]. Dans ce manuscrit nous considérerons un autre cas particulier de mélange non linéaire, le mélange quadratique introduit dans [28] et qui sera défini dans

les prochains chapitres. Le choix d'un tel modèle s'explique d'abord par un souci de simplification. En effet, le modèle initialement envisagé dans cette thèse devait également comporter des termes linéaires, ce qui ferait du modèle linéaire-quadratique précédent un sous-ensemble de ce modèle, mais la contrepartie aurait été l'introduction de difficultés supplémentaires dans les développements effectués au cours de cette thèse. Rien n'empêche cependant de considérer le modèle quadratique indépendamment des termes linéaires. Ce type de modèle pourrait à l'avenir trouver son application dans le cas de capteurs fournissant des combinaisons linéaires des puissances propres à chaque source, et de la puissance d'interaction de ces sources. De plus, l'une des perspectives de cette étude est d'étendre les méthodes développées à des modèles polynômiaux d'ordre plus élevé et d'étendre le champ d'application [60].

Cette étude se fait sous l'hypothèse d'indépendance statistique des sources ce qui nous lie naturellement au concept d'Analyse en Composantes Indépendantes (ACI).

Dans ce document les premiers résultats présentés seront des « développements théoriques » puis leur validation se fera sur des signaux artificiels, les méthodes proposées n'ayant pas été appliquées à des mélanges réels. Dans un premier temps nous présenterons le système de séparation proposé, celui-ci sera suivi par une méthode dédiée à l'estimation des coefficients du mélange, mise en amont de cette structure de séparation. Puis, une seconde méthode, basée sur le Maximum de Vraisemblance, sera présentée. Les problèmes d'identifiabilité et de séparabilité du modèle ne seront cependant pas abordés.

Plan du document

Le document est organisé comme suit : le premier chapitre établit l'état de l'art. Après avoir abordé brièvement le cas des mélanges linéaires en SAS, le cas des mélanges non linéaires est décrit de manière plus détaillée, plus particulièrement dans le cas de l'ACI Non Linéaire (ACI NL), étant donné que l'on travaille sous l'hypothèse d'indépendance statistique des sources. La dernière partie de ce chapitre est consacrée à la genèse et aux objectifs de la thèse.

Le chapitre 2 est dédié à la présentation et à l'étude de la structure de séparation proposée, de type réseau récurrent. Ainsi, les propriétés du réseau récurrent sont étudiées (points d'équilibre du réseau, stabilité locale), et ce, en contexte non aveugle,

et à un instant donné. Un parallèle est fait entre les travaux effectués dans le cadre du modèle linéaire-quadratique et le cas du modèle quadratique.

Dans le chapitre 3 nous proposons une méthode basée sur les cumulants afin d'estimer les coefficients du mélange. Cette partie vient compléter le chapitre précédent. Une deuxième méthode de séparation aveugle de sources est développée pour notre modèle dans le chapitre 4, celle-ci cette fois étant basée sur le principe de Maximum de Vraisemblance (MV), toujours en ayant recours au réseau récurrent, mais sous sa version dite basique. Chacun des chapitres est complété par des simulations numériques sur des signaux artificiels. Enfin, nous terminons ce manuscrit par les conclusions générales et les perspectives relatives à cette thèse.

Chapitre 1

État de l'art

Sommaire

1.1	Introduction	6
1.2	Mélanges linéaires	6
1.2.1	Différents types de mélanges	6
1.2.2	Familles de méthodes et critères pour la SAS	8
1.2.3	SAS et ACI : du cas linéaire au cas non-linéaire	12
1.3	Mélanges non-linéaires	13
1.3.1	Panorama des méthodes	13
1.3.2	Formulation mathématique du problème	17
1.3.3	Existence et Unicité des solutions fournies par l'ACI non-linéaire	19
1.3.4	Mélanges Post-Non-Linéaires (PNL)	21
1.4	Génèse de la thèse : du modèle linéaire-quadratique au modèle quadratique	23
1.4.1	Modèle linéaire-quadratique	24
1.4.2	Modèle quadratique	25
1.4.3	Objectifs de la thèse	26
1.5	Conclusion	26

1.1 Introduction

Ce chapitre est dédié à l'état de l'art. Après avoir présenté de façon sommaire les mélanges linéaires, ainsi que les différentes méthodes et critères utilisés en SAS, nous entamerons alors de manière plus détaillée les mélanges non-linéaires. La transition entre ces deux parties est assurée par la comparaison faite entre la SAS et l'ACI dans les cas linéaire et non-linéaire. Dans la section consacrée aux mélanges non-linéaires, nous commencerons par dresser un panorama des travaux effectués dans ce domaine. Une formulation mathématique du problème sera ensuite donnée et suivie par la question d'existence et d'unicité des solutions en ACI NL. Le cas particulier des mélanges post-non-linéaires sera également abordé. Enfin, nous terminerons ce chapitre par la genèse et les objectifs de cette thèse.

1.2 Mélanges linéaires

Les mélanges linéaires font partie des mélanges les plus étudiés en SAS. Ce type de mélange ne sera abordé que brièvement. Nous effectuerons d'abord un rapide passage en revue des différents types de mélanges linéaires rencontrés en SAS, plus exactement, nous nous attacherons à présenter la formulation mathématique des mélanges. Puis nous présenterons les méthodes et les critères utilisés. Nous parlerons également des indéterminations liées au modèle de mélange linéaire. Des panoramas plus complets peuvent être trouvés dans [11, 12, 15, 23, 39, 42, 74, 80].

1.2.1 Différents types de mélanges

Une première classification des mélanges peut se faire suivant la *nature* du mélange, liant les observations aux sources. Pour aller du modèle le plus simple au plus complexe, nous commencerons donc par le modèle de mélange *linéaire instantané*. Sous forme matricielle le modèle liant P observations à N sources est formulé de la manière suivante :

$$\mathbf{x}(t) = A\mathbf{s}(t) \tag{1.1}$$

où $\mathbf{x}(t)$ désigne le vecteur d'observations $[x_1(t) \dots x_P(t)]^T$, A la matrice des coefficients constants du mélange et $\mathbf{s}(t)$ le vecteur sources $[s_1(t) \dots s_N(t)]^T$.

D'un point de vue mathématique le modèle liant une des observations aux N sources

peut alors s'écrire :

$$x_i(t) = \sum_{j=1}^N a_{ij} s_j(t), \quad i = 1 \dots P \quad (1.2)$$

avec $x_i(t)$ représentant l'observation i , $s_j(t)$, la j^{e} source et le coefficient a_{ij} réel, constant, pouvant représenter la contribution de la j^{e} source sur le i^{e} capteur (ou atténuation), et ce, à un instant t .

Les mélanges à *atténuations et retards* sont un peu plus délicats à traiter que le mélange linéaire instantané. Ces mélanges peuvent être décrits par l'équation :

$$x_i(t) = \sum_{j=1}^N a_{ij} s_j(t - t_{ij}), \quad i = 1 \dots P \quad (1.3)$$

De la même manière, $x_i(t)$ représente l'observation i à l'instant t , s_j , la j^{e} source, les coefficients a_{ij} et t_{ij} , sont réels et constants et correspondent respectivement à l'amplitude de la contribution et au retard de la source j sur le capteur i .

Aux deux types de mélanges présentés précédemment vient s'ajouter la classe plus complexe et plus générale des mélanges *convolutifs*. Ces derniers peuvent être formulés comme suit :

$$x_i(t) = \sum_{j=1}^N \sum_{k=-\infty}^{-\infty} a_{ij,k} s_j(t - t_{ij,k}), \quad i = 1 \dots P \quad (1.4)$$

Une autre formulation du problème, pour le même type de mélange que ci-dessus, est obtenue en modélisant la contribution globale d'une source sur un capteur. Ceci peut se faire en ayant recours à la réponse impulsionnelle d'un filtre, ainsi nous avons :

$$x_i(t) = \sum_{j=1}^N a_{ij}(t) * s_j(t), \quad i = 1 \dots P \quad (1.5)$$

où $a_{ij}(t) * s_j(t)$ correspond à la contribution globale de la source j sur le capteur i et $*$ désigne l'opérateur de convolution.

Dans un cas plus général, un bruit additif peut s'ajouter aux modèles précédents. En ce qui concerne les mélanges *non-linéaires*, ces derniers seront abordés ultérieurement en Section 1.3.

Une autre classification des mélanges peut s'opérer suivant le nombre P d'observations considéré et le nombre N de sources : on parle de mélange *sous-déterminé* lorsque le nombre de sources est supérieur au nombre d'observations, de mélange *surdéterminé* lorsque P est supérieur à N , et de mélange *déterminé* lorsque les nombres de sources et d'observations sont égaux.

Après avoir décrit rapidement les différents types de mélanges linéaires, nous passons maintenant aux méthodes et critères utilisés en SAS.

1.2.2 Familles de méthodes et critères pour la SAS

Comme nous l'avons mentionné précédemment, la SAS consiste à restaurer un ensemble de signaux sources à partir d'observations qui constituent des mélanges de ces signaux sources. Il existe différentes méthodes qui permettent d'atteindre ce but. Parmi les méthodes utilisées en SAS on peut citer : *l'ACI*, *la Factorisation en Matrices Non-négatives* et *l'Analyse en Composantes Parcimonieuses*.

L'*ACI* constitue l'une des méthodes les plus utilisées en SAS. P. Comon [22] a généralisé en 1994 le principe de l'*ACI* permettant d'obtenir des signaux de sortie, statistiquement indépendants, égaux aux signaux sources dans le cadre du mélange linéaire instantané à certaines indéterminations près (facteur d'échelle, permutations). Pour une analyse plus détaillée nous renvoyons le lecteur à [11, 23, 42], ouvrages dédiés à l'Analyse en Composantes Indépendantes et à la séparation aveugle de sources. Différents critères peuvent être utilisés pour réaliser l'*ACI*, parmi eux nous citons :

- la *maximisation de la non gaussianité* qui est l'un des critères les plus utilisés en *ACI*. Ceci est motivé par le lien qui existe avec le théorème central limite qui dit que la distribution d'une somme de variables aléatoires tend à être gaussienne sous certaines conditions. Si l'on considère le mélange instantané de N sources aléatoires, non gaussiennes et de même distribution, ce mélange est alors plus gaussien que la distribution de chacune des sources. Ceci suggère inversement, de démixer les signaux sources en forçant chaque signal de sortie à devenir le plus non-gaussien possible. En pratique on maximise la non-gaussianité des sorties du réseau. L'une des mesures de non-gaussianité les plus utilisées est le *kurtosis*.
- la *maximisation de la vraisemblance*. L'une des interprétations possibles du principe de maximum de vraisemblance (MV) en *ACI* est d'estimer les valeurs des paramètres du mélange qui maximisent la probabilité de réalisation des observations. Le lien entre le principe de MV et celui de l'*INFOMAX* a été établi par J.-F. Cardoso dans [14].

- la minimisation de l'Information Mutuelle (IM). Il s'agit d'une approche importante en ACI inspirée par la théorie de l'information. Il s'agit alors d'estimer les coefficients de la matrice de séparation afin de minimiser l'information mutuelle des sorties. Cette approche peut donc être développée en utilisant l'IM qui est un critère de mesure de dépendance statistique.

On rappelle que si \mathbf{x} est un vecteur aléatoire de densité conjointe $f_{\mathbf{x}}$, et de densités marginales f_{X_i} , l'IM du vecteur \mathbf{x} , notée $I(\mathbf{x})$ est définie par :

$$I(\mathbf{x}) = \delta(f_{\mathbf{x}}(\mathbf{x}) | \prod_{i=1}^N f_{X_i}(x_i)) \quad (1.6)$$

$$= \int_{-\infty}^{\infty} f_{\mathbf{x}}(\mathbf{x}) \log \frac{f_{\mathbf{x}}(\mathbf{x})}{\prod_{i=1}^N f_{X_i}(x_i)} dx_1 dx_2 \dots dx_N \quad (1.7)$$

La notation $\delta(f_X|f_Y)$ est celle de la divergence de Kullback-Leibler entre les densités de probabilité f_X et f_Y , définie par :

$$\delta(f_X|f_Y) = \int_{-\infty}^{\infty} f_X(x) \log \frac{f_X(x)}{f_Y(x)} dx \quad (1.8)$$

$$(1.9)$$

L'information mutuelle est par définition une grandeur positive qui s'annule si et seulement si \mathbf{x} a des composantes indépendantes.

Dans le cadre du mélange linéaire instantané,

$$\mathbf{x}(n) = A\mathbf{s}(n) \quad (1.10)$$

l'IM est utilisée pour estimer la matrice B telle que :

$$\mathbf{y}(n) = B\mathbf{x}(n) \quad (1.11)$$

ait des composantes indépendantes. Le problème de SAS peut alors être ramené à la minimisation de l'IM, I , des sources estimées :

$$I(y_1, \dots, y_N) = \sum_{i=1}^N H(y_i) - H(\mathbf{y}) \quad (1.12)$$

où $H(\mathbf{x})$ désigne l'entropie de Shannon qui pour un vecteur aléatoire \mathbf{x} de distribution $f_{\mathbf{x}}(\mathbf{x})$ est la quantité :

$$H(\mathbf{x}) = - \int_{-\infty}^{\infty} f_{\mathbf{x}}(\mathbf{x}) \log f_{\mathbf{x}}(\mathbf{x}) d\mathbf{x} \quad (1.13)$$

$$= -E[\log f_{\mathbf{x}}(\mathbf{x})] \quad (1.14)$$

avec $E[.]$ désignant l'opérateur d'espérance mathématique. Aucun modèle de distribution des sources n'est supposé. Dans la pratique, l'utilisation de cette méthode est très complexe puisque le calcul de l'IM nécessite le calcul de la densité conjointe $f_{\mathbf{Y}}$ des signaux. En général on préfère exploiter pour le modèle linéaire la relation :

$$I(y_1, \dots, y_N) = \sum_{i=1}^N H(y_i) - H(\mathbf{x}) - \log |\det B| \quad (1.15)$$

et en remarquant que $H(\mathbf{x})$ est constant, on préfère minimiser alors l'expression :

$$J(B) = \sum_{i=1}^N H(y_i) - \log |\det B| \quad (1.16)$$

- nous citerons également les *approches tensorielles*. En général, la restriction à l'ordre 2 n'est pas suffisante pour séparer les sources. D'un point de vue théorique, l'indépendance entre signaux correspond à la nullité des cumulants croisés entre les versions décalées ou non de ces signaux et à tous les ordres. Il est cependant difficile de tester cette condition avec un coût calculatoire fini. En pratique, les méthodes de SAS annulent donc les cumulants jusqu'à un certain ordre fixé en général à 4. La généralisation des méthodes basées sur les cumulants se fait par une approche tensorielle qui consiste en l'utilisation de tenseurs de cumulants d'ordre supérieur. Ces tenseurs de cumulants sont des généralisations de la matrice de covariance qui est un tenseur de cumulants du second ordre.

On définit par exemple un tenseur de cumulants F comme un opérateur linéaire défini par les cumulants croisés d'ordre 4 de ces données. C'est une transformation linéaire de l'espace $N \times N$ tel que, le i, j^e élément de la transformation noté $F_{ij}(M)$ est défini par :

$$F_{ij}(M) = \sum_{kl} m_{kl} cum_4(y_i(t), y_j(t), y_k(t), y_l(t)) \quad (1.17)$$

où m_{kl} est un élément de la matrice M qui est transformée.

Comme tout opérateur linéaire symétrique, le tenseur F a une décomposition en valeurs propres. Une matrice propre vérifie :

$$F(M) = \lambda M \quad (1.18)$$

les matrices propres seront notées M_i .

Cette décomposition en valeurs propres correspond à une diagonalisation conjointe des matrices $F(M_i)$. La méthode *JADE* [16], qui utilise des cumulants d'ordre 4, est basée sur un tel critère qui peut être réécrit suivant :

$$J_{JADE} = \sum_{(ijkl) \neq (iikl)} cum_4(y_i(t), y_j(t), y_k(t), y_l(t))^2 \quad (1.19)$$

E. Moreau [58] a proposé une généralisation de l'algorithme *JADE* en étendant le critère à des cumulants d'ordre quelconque supérieur ou égal à trois.

Méthodes à l'ordre 2. La possibilité d'effectuer la séparation au second ordre est apparue très tôt à condition de supposer par exemple les sources colorées et de spectres distincts. Il est vrai qu'à son origine, le problème de séparation de sources se place dans le cadre du mélange linéaire instantané, $\mathbf{x}(n) = A\mathbf{s}(n)$, et repose sur l'indépendance des sources. On cherche ainsi à reconstituer des sources par transformation inverse $\mathbf{y}(n) = B\mathbf{x}(n)$ dont les composantes soient les plus indépendantes possible. Il n'est pas nécessaire *a priori* de faire des hypothèses sur le comportement temporel de \mathbf{s} , par exemple de *stationnarité*. Cependant d'un point de vue statistique il s'agit de se placer idéalement dans un contexte de sources i.i.d et donc temporellement blanches. La décorrélation de \mathbf{x} ne permet pas d'identifier la matrice A . Ainsi l'analyse en composantes principales *laisse la place* à l'analyse en composantes indépendantes, ce qui induit également le développement et l'usage des statistiques d'ordre supérieur. Les premières méthodes ACI de second ordre exploitant l'autocorrélation ou la non-stationnarité des signaux reposent sur le constat que les matrices d'autocovariance $R_x(\nu)$, considérées à différents retards ν , sont simultanément diagonalisables puisque $R_x(\nu) = AR_s(\nu)A^T$ et que les matrices $R_s(\nu)$ sont diagonales dès que les sources sont spatialement non corrélées. La méthode *AMUSE* [82] est définie par la diagonalisation conjointe exacte des estimateurs $R_x(0)$ et $R_x(\nu_0)$ pour un certain retard ν_0 . La méthode *SOBI* [13] est obtenue par une diagonalisation conjointe approchée de $R_x(0)$ et de plusieurs matrices $R_x(\nu_i)$, $i = 1, \dots, p$.

Ensuite viennent les méthodes par *Factorisation en Matrices Non-négatives* (FMN)¹ [19]. Cette famille de méthodes fait l'hypothèse de sources et de coefficients de la

¹ou NMF pour Non-negative Matrix Factorization

matrice de mélange non-négatifs. Parmi les signaux qui vérifient ces propriétés figurent, par exemple, les signaux de spectroscopie.

Enfin, nous passons aux méthodes d'*Analyse en Composantes Parcimonieuses*². Cette classe de méthodes est actuellement très populaire en SAS. En général, les observations sont ramenées à une représentation qui améliore le caractère parcimonieux des sources, i.e. leur nullité dans certaines zones de la représentation. Il peut s'agir du domaine temps-fréquence ou temps-échelle par exemple. L'Analyse en Composantes Parcimonieuses permet surtout l'estimation des sources gaussiennes et/ou corrélées, ainsi que la séparation des mélanges sous-déterminés.

Il doit être mentionné qu'en SAS de mélanges linéaires, en utilisant l'ACI, les sources peuvent être retrouvées, au mieux à un facteur d'échelle (nous incluons également le signe) et à une permutation près. Ceci constitue les indéterminations liées au problème de SAS linéaire instantanée dans le cas de l'ACI.

1.2.3 SAS et ACI : du cas linéaire au cas non-linéaire

La comparaison entre les problèmes de SAS et d'ACI pour les mélanges linéaire et non-linéaire a été abordée dans les travaux de C. Jutten et J. Karhunen [47]. En première partie de l'article, les auteurs posent les problèmes de la SAS et de l'ACI pour un modèle de mélange très élémentaire, celui du mélange linéaire instantané déterminé. En SAS, il s'agit de retrouver *les* sources qui ont généré le mélange, tandis que l'ACI recherche les composantes les plus indépendantes possibles qui auraient pu générer le mélange. En supposant que la matrice de mélange soit de rang plein et qu'il existe au plus une source gaussienne (et en tenant compte également des indéterminations d'échelle et de permutation) il a été prouvé que ces deux problèmes sont *équivalents*. Il devient alors important de souligner que ceci n'est plus valable, dès lors que l'on considère des mélanges non-linéaires généraux. Les mélanges non-linéaires sont plus complexes et plus difficiles à étudier, c'est ce que nous nous proposons de présenter dans la section qui suit.

²ou SCA pour Sparse Component Analysis

1.3 Mélanges non-linéaires

1.3.1 Panorama des méthodes

Le mélange linéaire (1.2) est trop simple pour décrire toutes les situations réelles. Il devient alors naturel de considérer une extension de ce modèle. Sans avoir la prétention de présenter une liste exhaustive des travaux effectués dans le cadre des mélanges non-linéaires, ni d'entrer dans les détails des méthodes proposées (on peut trouver des ouvrages d'ordre général ou des travaux tels que [11, 12, 42, 48]), nous proposons un tour d'horizon des travaux effectués en séparation de sources non-linéaire, d'abord hors ACI, puis en nous focalisant sur les méthodes liées à l'ACI NL. Certains modèles présentés seront décrits dans la section 1.3.2.

À notre connaissance, en SAS NL (méthodes autres que ACI) nous pouvons citer l'article [29] abordant le problème d'identification aveugle de mélange et de SAS. Dans cet article, les auteurs considèrent que les sources sont non-stationnaires, ou plus spécifiquement que l'hypothèse de parcimonie temporelle des sources est faite. Le mélange étudié est de type linéaire-quadratique, il en découle la méthode LQ-TEPCORR, extension de leur précédente méthode LI-TEPCORR.

Pour ce tour d'horizon nous allons maintenant mettre l'accent sur les méthodes faisant appel à l'ACI NL. Dans le cadre non-linéaire les premiers travaux sont attribués à C. Jutten qui a utilisé des mélanges faiblement non-linéaires pour tester la robustesse et les performances de l'algorithme d'Hérault-Jutten introduit pour le problème de SAS linéaire [44]. Cependant Burel a probablement été le premier à introduire un algorithme spécifiquement pour l'ACI NL. Sa méthode est basée sur un réseau de neurones à rétropropagation pour des non-linéarités paramétriques mais souffre de complexité calculatoire et de problèmes de minima locaux [9].

Parra et al. [66] proposent une technique pour trouver une transformation non-linéaire qui produit une représentation statistiquement indépendante des distributions. Cette extraction de caractéristiques en termes d'information mutuelle minimale généralise le critère de décorrélation de l'ACP linéaire à des ordres supérieurs.

Lee et al. [51] s'intéressent au problème de SAS NL, ils présentent un ensemble de règles d'apprentissage basées sur le critère de maximisation de l'information. Le modèle est divisé en une partie linéaire et une partie représentant la fonction de transfert non-linéaire du canal de transmission. Le modèle proposé se focalise

sur les non-linéarités paramétriques sigmoïdales et les polynômes d'ordre supérieur. Les algorithmes proposés nécessitent cependant une étude complémentaire sur la stabilité des algorithmes et la convergence de l'algorithme. Les algorithmes proposés devraient être applicables à des données biomédicales de type IRM.

Dans [64], P. Pajunen *et al.* introduisent une nouvelle approche neuronale pour la séparation de sources non-linéaire où les cartes auto-organisatrices de Kohonen³ ont été utilisées comme algorithme de séparation.

P. Pajunen et J. Karhunen [65] ont proposé un algorithme utilisant les cartes topographiques génératives⁴ visant à remédier aux problèmes rencontrés avec la méthode précédente [64], et permettant de la généraliser, puisqu'elle s'applique à des sources dont la statistique est connue mais arbitraire. En effet pour pouvoir être appliquée la méthode précédente nécessitait que les densités des sorties soient uniformément distribuées ce qui posait problème dans le cas où les sources originales n'étaient pas uniformes. L'algorithme présenté en [65] utilise une approche par maximum de vraisemblance, et la transformation de mélange est supposée suffisamment régulière. Les auteurs montrent alors expérimentalement qu'en trouvant une transformation séparante, suffisamment régulière et qui maximise la vraisemblance, les sources peuvent être séparées.

Yang et al [84] ont envisagé un perceptron à deux couches comme structure séparante pour des mélanges non-linéaires avec des algorithmes d'apprentissage dérivant de deux approches, l'une basée sur le maximum d'entropie, l'autre sur la minimisation de l'information mutuelle par le biais de la descente en gradient naturel.

Dans [3], L. Almeida fait une synthèse de ces travaux antérieurs et conduisant à la méthode MISEP, extension de la méthode INFOMAX. Cette méthode sépare les sources à la fois dans les mélanges linéaires et les mélanges non-linéaires, et utilise le critère d'information mutuelle des sorties. Dans [4], l'auteur applique la méthode MISEP au problème non-linéaire des pages recto et verso d'images acquises par un scanner par exemple. Les résultats obtenus améliorent la qualité d'une séparation linéaire.

H. Valpola *et al* [83] ont proposé une approche d'ACI NL basée sur techniques bayésiennes. L'algorithme utilise un perceptron multicouches pour modéliser le mé-

³SOM en anglais pour Self Organizing Maps

⁴GTM en anglais pour Generative Topographic Maps

lange non-linéaire et l'ensemble d'apprentissage pour estimer les distributions *a posteriori* des variables inconnues du modèle, i.e. les paramètres du réseau multicouches.

Les méthodes présentées jusqu'à présent concernent les mélanges non-linéaires en tant que mélanges très généraux et, par conséquent, très difficiles à étudier. Pour simplifier ce modèle, des configurations particulières ont généralement été envisagées dans la littérature et nécessitent l'ajout de contraintes fortes. Les mélanges dits post-non-linéaires (PNL) correspondent à un type particulier de mélanges non-linéaires. Ils ont été introduits par A. Taleb et C. Jutten et sont des plus étudiés, nous leur consacrerons la sous-section 1.3.4. Il existe également d'autres approches qui traitent d'autres types de mélanges non-linéaires particuliers. Dans ce qui suit, nous en citerons quelques exemples, la plupart des approches citées correspondant à des modèles polynômiaux ou à des modèles qui du moins, s'y apparentent.

T.-W. Lee [50] présente des méthodes qui traitent du problème des modèles NL dans le cadre de l'ACI, d'abord dans le cas d'un « *modèle NL simple (PNL)*, » puis dans celui d'un modèle plus général que le premier, où *les non linéarités sont approximées par un développement en série de Taylor*.

M. Castella [17] traite de systèmes de mélanges MIMO instantanés NL, plus exactement *polynômiaux*. Il aborde le problème d'inversibilité du modèle (l'inverse appartient à la classe des polynômes). Son approche est basée sur l'utilisation des bases de Groebner, pour d'abord tester l'existence de cet inverse (solution exacte), et ensuite éventuellement le déterminer. Il s'intéresse également au cas plus général des modèles NL (interpolation sous certaines conditions), ce qui apporte une généralisation des résultats obtenus précédemment pour donner une condition d'existence de cet inverse et le calculer sous une forme polynômiale. Enfin, le problème de SAS sous la contrainte supplémentaire de sources à alphabet fini est abordé. Nous reviendrons sur ce modèle en Section 1.3.2.

J. Eriksson et V. Koivunen dans [33,34] s'intéressent respectivement à l'identifiabilité aveugle et à la séparation aveugle d'une certaine *classe de mélanges satisfaisant le théorème d'addition*. Ils proposent alors un algorithme de séparation pour de tels mélanges.

[79] est un article qui traite de non-linéarités polynômiales, avec un intérêt particulier pour *les monômes du second ordre et les formes quadratiques*. F. Theis et W. Nakamura examinent le modèle de séparation quadratique, et définissent le

modèle générant les données et le processus de séparation, en utilisant l'analyse faite à partir de l'ACI linéaire dans le cas surdéterminé. Une discussion sur l'identifiabilité du modèle de « démélange » quadratique est également présentée. Leur méthode est alors appliquée sur des données simulées et des images réelles.

M. Krob et M. Benidir [49] proposent une identification aveugle du *modèle linéaire-quadratique* en utilisant une décomposition en valeurs singulières des matrices de cumulants d'ordre 3 sur les sorties. Les sources sont supposées indépendantes, circulaires et centrées. L'identification se fait uniquement à partir des sorties du système.

Dans [1], K. Abed-Meraim *et al.* s'intéressent à l'identification aveugle d'un *mélange linéaire-quadratique* instantané bruité, modèle qui sera simplifié par la suite. Des hypothèses supplémentaires sur les sources sont faites, celles-ci sont supposées corrélées temporellement, circulaires complexes et mutuellement indépendantes. Ils proposent alors une nouvelle technique d'identification aveugle de ce mélange, basée sur la diagonalisation conjointe d'un ensemble de matrices de corrélation. Leur méthode fait également appel à l'utilisation des moments d'ordre 3. Un exemple d'application sur le couplage de phase quadratique est donné. Ce modèle sera présenté en Section 1.3.2.

Y. Deville et S. Hosseini se sont penchés sur le cas du *mélange linéaire-quadratique* (termes linéaires et termes croisés), en configuration aveugle [40,41], ou non [28] (coefficients du mélange supposés connus). Dans le premier volet [40,41], ils développent une structure de séparation de type réseau récurrent et les algorithmes d'adaptation font appel, soit aux statistiques d'ordre supérieur, soit au maximum de vraisemblance. Dans le deuxième volet [28], ils reviennent à une configuration non aveugle, car leur intérêt porte d'une part sur les problèmes de stabilité du réseau rencontrés dans les articles précédents, et d'autre part sur la généralisation à d'autres classes de mélanges. Nous reviendrons sur ces articles de façon plus approfondie dans la Section 1.4 et dans les chapitres qui suivront. Nous pouvons également préciser que leur modèle a été utilisé dans [35] dans le cas d'images scannées.

Dans [31,32] L. Tomazeli Duarte et C. Jutten s'intéressent au problème de SAS pour *une classe particulière de mélanges non-linéaires* qui peut *s'apparenter dans certains cas aux mélanges polynômiaux*. Ce type de modèle trouve des applications dans les capteurs chimiques. Dans ces deux articles, les auteurs ont recours égale-

ment à une structure récurrente comme structure de séparation, et les algorithmes d'apprentissage sont liés respectivement aux statistiques d'ordre supérieur et à la minimisation de l'information mutuelle des sorties du réseau. Une analyse de la stabilité locale du réseau est également établie. Nous présenterons leur modèle en Section 1.3.2.

Pour terminer ce panorama des méthodes nous citerons les travaux de Y. Naudet. Dans [60], Y. Naudet étudie les possibilités d'application de la séparation de sources au traitement de la réponse d'un capteur à courants de Foucault (CCF), en vue d'atténuer, ou mieux de supprimer, les effets néfastes dus à la température dans une cible lors d'une mesure de la distance la séparant du capteur. L'un des mélanges étudié dans cette thèse est de type linéaire-quadratique, incluant les termes linéaires mais également les auto-termes et inter-termes quadratiques .

1.3.2 Formulation mathématique du problème

Comme mentionné précédemment, les problèmes de SAS et d'ACI diffèrent dans le cas de mélanges non-linéaires. D'un point de vue mathématique, dans le cas de mélanges instantanés pour P observations et N sources, le problème général de l'ACI non-linéaire peut être formulé comme suit :

$$\mathbf{x} = F(\mathbf{s}) = \begin{cases} f_1(\mathbf{s}) \\ \cdot \\ \cdot \\ f_P(\mathbf{s}) \end{cases} \quad (1.20)$$

où \mathbf{x} est le vecteur d'observations $[x_1(t) \dots x_P(t)]^T$, F est un ensemble de fonctions $[f_1(t) \dots f_P(t)]^T$ de mélange non-linéaires réelles inconnues et \mathbf{s} est le vecteur sources $[s_1(t) \dots s_N(t)]^T$ dont les composantes sont indépendantes. Il s'agit alors de trouver la fonction H qui nous donne le vecteur \mathbf{y} tel que :

$$\mathbf{y} = H(\mathbf{x}) \quad (1.21)$$

et dont les composantes sont statistiquement indépendantes. Une caractéristique fondamentale de ce problème dans le cas général est que des solutions existent toujours et ne sont pas uniques. Pour la SAS il s'agirait de retrouver le vecteur source

s à l'origine du mélange. Cependant, en ACI NL le problème est plus délicat : en effet, si l'on considère deux variables aléatoires u et v statistiquement indépendantes, alors pour des fonctions f et g non-linéaires inversibles choisies arbitrairement, les variables aléatoires $f(u)$ et $g(v)$ restent indépendantes. Dans le cas général des mélanges non-linéaires, en ne considérant que l'hypothèse d'indépendance statistique, les sources ne peuvent être donc retrouvées au mieux, qu'à une fonction inconnue des sources près. De telles distorsions ne sont malheureusement pas acceptables. Une autre difficulté qui se présente dans le cas non-linéaire est le fait que le mélange de deux variables aléatoires indépendantes puisse donner à son tour des variables qui restent statistiquement indépendantes, tout en restant des mélanges de ces variables. Nous reviendrons plus en détail sur ce sujet dans la Section 1.3.3.

Le modèle (1.20) a été posé d'un point de vue général, nous nous proposons maintenant de donner des exemples de modèles contraints évoqués en Section 1.3.1.

M. Castella [17] a considéré, entre autres, le cas du mélange non-linéaire de 2 sources binaires (i.e. $\forall i \in \{1, 2\}, s_i \in \{-1; +1\}$). Plus précisément, il se place dans un contexte lié aux télécommunications et où les deux sources représentent deux signaux BPSK⁵. Les sources sont également supposées centrées et mutuellement indépendantes. Le modèle pour une observation x et 2 sources s_1 et s_2 est le suivant :

$$x = as_1s_2 + bs_1 + cs_2 + d \quad (1.22)$$

où a, b, c, d représentent des constantes.

La généralisation de ce modèle à Q observations peut s'écrire :

$$\mathbf{x} = \mathbf{A} \begin{pmatrix} s_1 \\ s_2 \\ s_1s_2 \end{pmatrix} + \mathbf{B} \quad (1.23)$$

avec \mathbf{A} une matrice $Q \times 3$, \mathbf{B} un vecteur colonne $Q \times 1$ correspondant à la contribution d'un monôme d'ordre 1 constant. Ce modèle peut être à nouveau simplifié en tenant compte du fait que les sources sont centrées.

Le modèle simplifié, considéré par K. Abed-Meraim *et al* dans [1] pour n observations et m sources, s'écrit :

$$\mathbf{x}(t) = \sum_{1 \leq i \leq m} s_i(t) \mathbf{l}_i + \sum_{1 \leq i \leq j \leq m} s_i(t) s_j(t) \mathbf{q}_{ij} + \mathbf{b}(t) \quad (1.24)$$

⁵pour Binary Phase Shift Keying

avec $\mathbf{b}(t)$ correspondant à un bruit blanc additif, les sources sont supposées mutuellement indépendantes, circulaires et corrélées temporellement. Les vecteurs de coefficients déterministes \mathbf{l}_i et \mathbf{q}_{ij} correspondent respectivement aux parties linéaire et quadratique. De manière plus condensée (1.24) se réécrit :

$$\mathbf{x}(t) = H\mathbf{s} + \mathbf{b} \quad (1.25)$$

où \mathbf{s} est le vecteur source de dimension $K = m + \frac{m(m+1)}{2}$, $\mathbf{s} = [\mathbf{s}_1^T, \mathbf{s}_2^T]^T$ avec $\mathbf{s}_1^T = [s_1 \dots s_m]$ et $\mathbf{s}_2^T = [s_1^2, s_1 s_2, \dots, s_1 s_m, s_2^2, \dots, s_m^2]$. H est la matrice de mélange de dimension $n \times K$. \mathbf{b} est le bruit blanc additif indépendant des signaux sources.

Un autre exemple de mélange contraint est celui considéré par L. Tomazeli Duarte et C. Jutten dans [31, 32], il s'agit du modèle simplifié suivant, mettant en jeu 2 sources et 2 observations, et lié à des capteurs chimiques :

$$\begin{cases} x_1 = s_1 + a_{12}s_2^k \\ x_2 = s_2 + a_{21}s_1^{\frac{1}{k}} \end{cases} \quad (1.26)$$

où x_i et s_j représentent respectivement le i^{e} mélange et la j^{e} source. Les sources sont positives (elles expriment des concentrations) et sont statistiquement indépendantes. Le coefficient $k \in \mathbb{N}$ est supposé connu et correspond à une constante liée au rapport des valences des 2 ions considérés.

1.3.3 Existence et Unicité des solutions fournies par l'ACI non-linéaire

A. Hyvarinen et P. Pajunen [43] ont été les premiers à poser le problème d'existence et d'unicité des solutions de l'ACI NL dans un cadre général. Ces questions restent cependant fondamentales dans la construction d'algorithmes en ACI NL.

Dans cet article, en conservant les notations utilisées dans la Section 1.3.2, équations (1.20,1.21), les auteurs indiquent de manière explicite comment construire une fonction H de manière à ce que les composantes du vecteur $\mathbf{y} = H(\mathbf{x})$ soient statistiquement indépendantes. Ils montrent que cette construction n'est pas unique en général. Ainsi dans une première partie, les auteurs démontrent que le problème *admet au moins une solution*. La construction de cette solution se fait par le biais d'une procédure qui généralise la procédure d'orthogonalisation de Gram-Schmidt, c'est

alors, de manière récursive, à partir des observations, qu'ils obtiennent des variables aléatoires uniformes, statistiquement indépendantes. Dans une seconde partie, ils abordent le problème de l'unicité de la solution : en appliquant une simple transformation linéaire aux observations, et en répétant la procédure définie précédemment, ils montrent qu'ils peuvent à nouveau construire une nouvelle solution, ce qui implique clairement la non-unicité des solutions. L'unicité des solutions peut cependant être atteinte si l'on rajoute des contraintes, à certaines indéterminations près.

Pour en revenir au problème des solutions obtenues en ACI NL, et qui restent toujours des mélanges des sources initiales, nous citerons l'exemple simple donné dans [47].

Supposons que l'on dispose de deux sources s_1 et s_2 statistiquement indépendantes, la première ayant une distribution de Rayleigh soit $p_{s_1} = s_1 \exp\left(-\frac{s_1^2}{2}\right)$ et la seconde, une distribution uniforme sur $[0, 2\pi)$. Si l'on considère la transformation non-linéaire H :

$$[y_1, y_2] = H(s_1, s_2) \quad (1.27)$$

$$= [s_1 \cos(s_2), s_1 \sin(s_2)] \quad (1.28)$$

la densité conjointe des variables y_1 et y_2 s'écrit alors :

$$p_{y_1, y_2}(y_1, y_2) = \frac{p_{s_1, s_2}(s_1, s_2)}{|J|}, \quad (1.29)$$

$$= \frac{1}{2\pi} \exp\left(\frac{-y_1^2 - y_2^2}{2}\right) \quad (1.30)$$

$$= \left[\frac{1}{\sqrt{2\pi}} \exp\left(\frac{-y_1^2}{2}\right) \right] \left[\frac{1}{\sqrt{2\pi}} \exp\left(\frac{-y_2^2}{2}\right) \right] \quad (1.31)$$

$$\text{avec } J = \begin{pmatrix} \cos(s_2) & -s_1 \sin(s_2) \\ \sin(s_2) & s_1 \cos(s_2) \end{pmatrix} \quad (1.32)$$

(1.31) montre clairement que les 2 variables y_1 et y_2 restent indépendantes.

Dans le registre particulier des mélanges post-non-linéaires, A. Taleb et C. Jutten ont également abordé le problème d'existence et d'unicité des solutions, c'est ce que nous étudierons dans la prochaine sous-section.

1.3.4 Mélanges Post-Non-Linéaires (PNL)

Comme nous l'avons mentionné précédemment, il s'agit d'un mélange non-linéaire particulier. Dans ce cas, le mélange peut se décomposer en une partie linéaire où les sources sont mélangées de manière linéaire instantanée, suivie d'une non-linéarité sur chaque observation. Ce modèle de mélange présente l'intérêt de représenter une certaine réalité physique, en effet, il peut être assimilé à un modèle où les signaux sont mélangés linéairement et instantanément par le canal, et les capteurs et leurs instrumentations (amplificateurs...) seraient responsables de la distorsion non-linéaire. Ce type de mélange a été énormément étudié. A. Taleb et C. Jutten ont proposé plusieurs approches pour résoudre le problème, en passant par le maximum de vraisemblance [77] ou l'information mutuelle [10, 78].

Le modèle PNL peut s'exprimer pour N sources et N capteurs comme :

$$x_i = f_i\left(\sum_{j=1}^N a_{ij}s_j\right), \quad i = 1 \dots N \quad (1.33)$$

où les x_i représentent les observations, les f_i sont des fonctions non-linéaires inversibles inconnues, les a_{ij} représentent les coefficients de mélange réels de la partie linéaire instantanée du modèle et les s_j les sources supposées statistiquement indépendantes.

En utilisant la notation matricielle nous avons :

$$\mathbf{x} = F(\mathbf{A}\mathbf{s}), \quad (1.34)$$

où \mathbf{x} désigne le vecteur observations, $F = [f_1 \dots f_N]^T$ les fonctions non-linéaires appliquées composante par composante, A la matrice de mélange linéaire instantané et \mathbf{s} le vecteur sources.

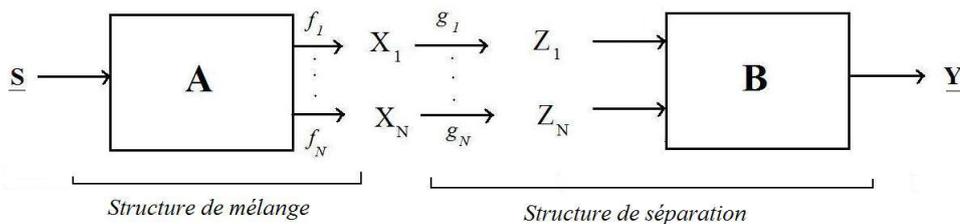


FIGURE 1.1 – Structures de mélange et de séparation pour le modèle PNL

La structure de séparation est également composée d'une partie linéaire et d'une partie non-linéaire. Le premier étage est composé de fonctions non-linéaires g_i qui inversent chacune des fonctions non-linéaires f_i , tandis que le second, représenté par une matrice B , est censé inverser la matrice de mélange A . La figure 1.1 illustre les structures de mélange et de séparation du modèle PNL.

A. Taleb et C. Jutten se sont également intéressés aux propriétés de séparabilité d'un tel mélange. Ils montrent que sous l'hypothèse de faibles contraintes sur la matrice de mélange A et sur les sources, ces dernières peuvent être retrouvées avec pratiquement les mêmes indéterminations que dans le cas du mélange linéaire instantané. Le vecteur sorties \mathbf{y} du système de séparation peut s'écrire :

$$\mathbf{y} = P\Lambda\mathbf{s} + \mathbf{T} \quad (1.35)$$

où P désigne une matrice de permutation, Λ une matrice diagonale et \mathbf{T} un vecteur de translation constant.

Il en découle le lemme de séparabilité suivant [78] :

Lemme : Si (A, f) désigne une structure PNL et (g, B) une structure séparante PNL où nous avons :

- la matrice A est régulière et qui a au moins 2 entrées non nulles par ligne et par colonne,
- les fonctions f_i , $i = 1 \dots N$ sont inversibles et différentiables,
- la matrice B est une matrice régulière,
- et $h_i = g_i \circ f_i$ satisfait $\forall u \in \mathbb{R}$, $h'_i(u) \neq 0$, $\forall i = 1 \dots N$,

et si l'on suppose que chaque source s_i a une densité qui s'annule au moins sur un compact, alors les sorties \mathbf{y} de la structure de séparation sont des composantes mutuellement indépendantes si et seulement si les composantes h_i sont linéaires, et B est une matrice séparante.

Ils montrent également que certaines hypothèses (sur les densités des sources) peuvent être relâchées, preuve donnée dans [2].

Les mélanges PNL ont été étudiés par d'autres auteurs (voir les références citées dans [12,42]). Dans [53], M. Babaie-Zadeh étend le cas des mélanges PNL à celui des

mélanges CPNL (mélanges post-non-linéaires convolutifs) et présente 2 algorithmes de séparation.

Après avoir abordé le cas particulier des mélanges PNL nous passons à d'autres mélanges non-linéaires particuliers qui font l'objet de cette thèse.

1.4 Génèse de la thèse : du modèle linéaire-quadratique au modèle quadratique

Les modèles linéaire-quadratique et quadratique considérés comprennent respectivement des termes linéaires et des termes croisés pour le premier et des auto-termes et des termes croisés pour le second. Cette thèse prend sa source dans les travaux de Y. Deville et S. Hosseini [28] qui s'inscrivent dans la continuité des développements effectués dans le cas spécifique du modèle de mélange linéaire-quadratique [26, 40, 41]. Pour « résumer », dans [28] un système qui implémente, en *contexte non aveugle* (coefficients du mélange connus), l'inverse de la fonction de mélange (1.20) est défini, et ce, pour une large classe de mélanges non-linéaires⁶. Nous ne citerons pas les différentes classes de mélanges proposées, seule l'une d'entre elle, qui s'applique à notre modèle de mélange, sera présentée.

Une particularité de l'approche consiste en l'**association** d'une structure récurrente aux classes de mélanges présentées, la structure récurrente proposée pouvant être vue comme une extension des structures rencontrées dans le cas de mélanges linéaires [45, 52, 59]. La version étendue de la structure présente *des degrés de liberté supplémentaires*, solution proposée afin de pallier les problèmes de stabilité rencontrés dans la version basique du réseau utilisée dans [40, 41]. L'avantage des structures présentées dans [28, 40, 41] est que la forme explicite du modèle inverse n'est pas nécessaire, d'où la possibilité d'une généralisation aisée pour des modèles de mélanges plus complexes, bien que l'utilisation de structures directes reste possible lorsque la forme explicite du modèle inverse est disponible.

L'article [28] définit différentes classes de mélanges dont les mélanges non-linéaires

⁶La classe générale dite ETM pour Extractable-Target Mixtures ne sera pas présentée dans ce manuscrit.

dits à « Composantes Cibles Additives » (CCA)⁷, qui sont de la forme :

$$x_i(n) = T_i[s_i(n)] - I_i[s(n)] \quad (1.36)$$

où les termes $T_i[s_i(n)]$ et $I_i[s(n)]$ désignent respectivement les termes d'intérêt et les termes d'interférence du modèle considéré. La structure associée au modèle conserve alors les premières composantes et supprime les secondes, de sorte à récupérer à l'équilibre⁸ les sources en sortie du réseau, deux exemples concrets (mélanges linéaire-quadratique et quadratique) seront donnés dans ce qui suit. Nous commencerons donc par présenter les travaux effectués dans le cadre du modèle linéaire-quadratique, avant de passer au cas du modèle quadratique et des objectifs de cette thèse.

1.4.1 Modèle linéaire-quadratique

Le mélange linéaire-quadratique a donné lieu à plusieurs travaux [26, 28, 40, 41]. Dans [40, 41], les auteurs s'y intéressent dans un contexte aveugle. Dans ces deux articles, l'accent est mis sur la structure récurrente utilisée dans sa version basique, tout en rappelant que l'utilisation de structures directes reste possible. Les méthodes de séparation proposées s'appuient respectivement sur les statistiques croisées d'ordre supérieur et le principe de maximum de vraisemblance. Dans [28], nous nous situons en contexte non aveugle, le modèle linéaire-quadratique est l'un des exemples cités comme application aux classes de mélanges présentées. Différentes structures de séparation sont proposées pour un tel modèle, mais il ne s'agit pas de l'unique mélange présenté, puisque c'est dans cet article que le modèle quadratique est introduit. Le mélange considéré dans [28, 40, 41] est un mélange instantané défini pour 2 observations et 2 sources par les équations :

$$\begin{cases} x_1(n) = a_{11}s'_1(n) + a_{12}s'_2(n) + b_1s'_1(n)s'_2(n) \\ x_2(n) = a_{21}s'_1(n) + a_{22}s'_2(n) + b_2s'_1(n)s'_2(n) \end{cases} \quad (1.37)$$

où $(x_1(n), x_2(n))$ désignent les observations et $(s'_1(n), s'_2(n))$ les sources, et supposées statistiquement indépendantes dans [40, 41]. Ce modèle a été plus précisément considéré dans sa version *normalisée*, i.e. en considérant que a_{11} et a_{22} sont non nuls,

⁷ou "ATM" pour Additive-Target Mixtures

⁸cette notion sera définie dans le chapitre suivant

nous obtenons les expressions :

$$\begin{cases} x_1(n) = s_1(n) - L_{12}s_2(n) - Q_1s_1(n)s_2(n) \\ x_2(n) = s_2(n) - L_{21}s_1(n) - Q_2s_1(n)s_2(n) \end{cases} \quad (1.38)$$

$$\text{avec } \begin{cases} s_i(n) = a_{ii}s'_i(n), \quad i \in \{1, 2\}, \\ L_{ij} = -\frac{a_{ij}}{a_{jj}} \\ Q_i = -\frac{b_i}{a_{ii}a_{jj}}, \quad (i, j) \in \{(1, 2), (2, 1)\} \end{cases} \quad (1.39)$$

Dans ce cas, $(x_1(n), x_2(n))$ désignent toujours les observations et (s_1, s_2) les versions normalisées des sources. Nous noterons que les coefficients du mélange ont également changé suite à cette normalisation. Les observations $(x_1(n), x_2(n))$ et les paramètres du modèle $(L_{12}, L_{21}, Q_1, Q_2)$ sont connus [28] ou inconnus [40, 41], et les paramètres du mélange comme les sources sont des réels.

Ce mélange est un mélange à CCA. En reprenant les notations précédentes nous avons dans ce cas précis les correspondances :

$$T_i[s_i(n)] = s_i(n) \quad \text{et} \quad I_i[s(n)] = -L_{ij}s_j(n) - Q_i s_i(n)s_j(n), \quad i \neq j. \quad (1.40)$$

Les différentes structures associées à ce modèle de mélange, ainsi que l'analyse de leurs propriétés seront présentées dans le chapitre 2.

1.4.2 Modèle quadratique

Ce modèle a été introduit dans [28] et nous le définirons conformément à cet article. Ainsi, le modèle quadratique est présenté comme un mélange appartenant également à la classe dite à CCA. Différentes structures de séparation pour un tel modèle sont présentées dans [28], mais les propriétés des structures ne sont pas étudiées. Comme précédemment, il s'agit d'un mélange instantané, quadratique (auto-termes et termes croisés), différent du modèle de mélange précédent, défini pour 2 observations et 2 sources dans sa version initiale par les équations :

$$\begin{cases} x_1(n) = [a_{11}s'_1(n)]^2 + [a_{12}s'_2(n)]^2 + b_1s'_1(n)s'_2(n) \\ x_2(n) = [a_{21}s'_1(n)]^2 + [a_{22}s'_2(n)]^2 + b_2s'_1(n)s'_2(n) \end{cases} \quad (1.41)$$

où $(x_1(n), x_2(n))$ désignent les observations, et $(s'_1(n), s'_2(n))$ les sources.

Pour la version normalisée le mélange s'écrit :

$$\begin{cases} x_1(n) = s_1^2(n) - L_{12}s_2^2(n) - Q_1s_1(n)s_2(n) \\ x_2(n) = s_2^2(n) - L_{21}s_1^2(n) - Q_2s_1(n)s_2(n) \end{cases} \quad (1.42)$$

$$\text{avec} \quad \begin{cases} s_i(n) = a_{ii}s'_i(n), \quad i \in \{1, 2\}; \\ L_{ij} = -\frac{a_{ij}^2}{a_{jj}^2}, \quad L_{ij} < 0, \\ Q_i = -\frac{b_i}{a_{ii}a_{jj}}, \quad (i, j) \in \{(1, 2), (2, 1)\} \end{cases} \quad (1.43)$$

Les observations et les coefficients du mélange sont supposés connus dans [28], ces derniers, ainsi que les sources, sont supposés réels. Ce mélange étant un mélange à CCA nous avons les correspondances suivantes :

$$T_i[s_i(n)] = s_i^2(n) \quad \text{et} \quad I_i[s(n)] = -L_{ij}s_j^2(n) - Q_i s_i(n)s_j(n), \quad i \neq j. \quad (1.44)$$

La présentation des structures de séparation proposées se fera dans le chapitre suivant.

Un récapitulatif des différents travaux effectués par Y. Deville et S. Hosseini sur les modèles linéaire-quadratique et quadratique est disponible à la figure 1.2.

1.4.3 Objectifs de la thèse

Comme nous l'avons mentionné précédemment, dans le cas du modèle quadratique, différentes structures de séparation ont été proposées dans [28], cependant les propriétés des réseaux n'ont pas été étudiées. Dans le prochain chapitre de cette thèse, nous nous intéresserons donc aux points d'équilibre de la structure, ainsi qu'à l'analyse de la stabilité locale aux points d'équilibre de cette structure. Ceci se fera à un instant donné, dans un contexte non aveugle.

Un autre objectif de cette thèse est l'évolution vers un contexte de SAS . Dans les chapitres qui suivront, deux méthodes seront développées permettant d'atteindre cet objectif. Les simulations se feront sur des signaux de synthèse.

1.5 Conclusion

Dans ce chapitre, nous avons commencé par présenter rapidement les mélanges linéaires et les différentes méthodes de SAS. Puis nous sommes passés au cas plus

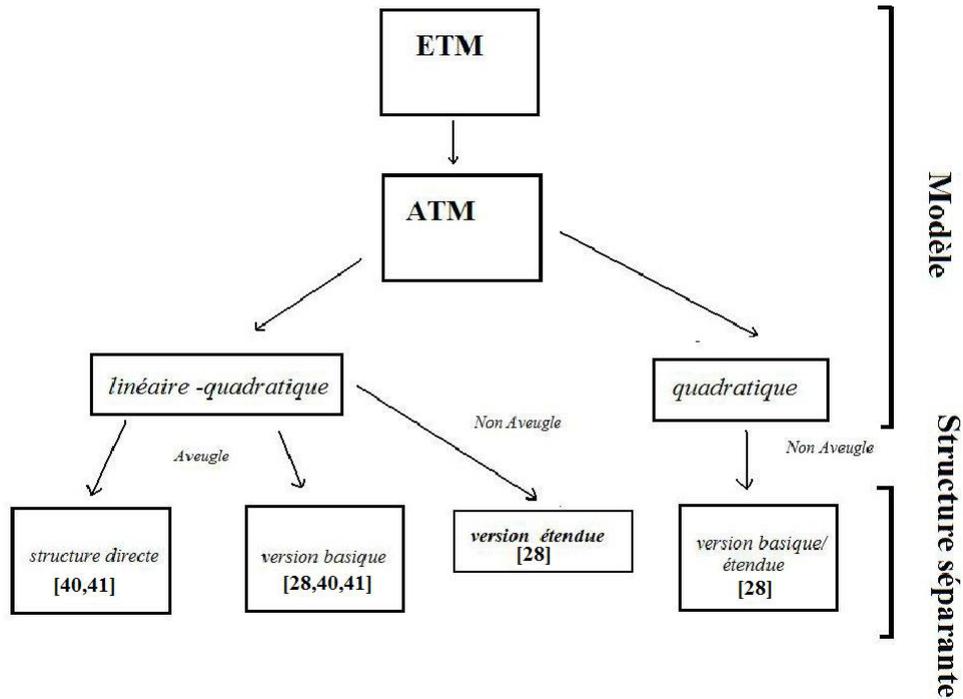


FIGURE 1.2 – Récapitulatif des différents travaux effectués sur les modèles linéaire-quadratique et quadratique.

complexe des mélanges non-linéaires, un panorama non exhaustif des principaux modèles et méthodes a été exposé, avant d’aborder le problème d’existence et d’unicité des solutions en ACI NL. Le cas particulier des mélanges PNL a ensuite été considéré.

Nous avons terminé ce chapitre par l’exposé de la genèse et de la problématique de cette thèse. Ainsi, après avoir présenté le modèle de mélange étudié, nous allons passer aux structures de séparation associées qui ont été proposées, et qui seront définies dans le prochain chapitre. Dans le chapitre qui va suivre, nous allons présenter et étudier la structure de séparation de type réseau récurrent qui a été utilisée dans cette thèse.

Chapitre 2

Réseaux récurrents

Sommaire

2.1	Introduction	30
2.2	Travaux antérieurs	30
2.3	Modèle linéaire-quadratique	31
2.3.1	Généralités	31
2.3.2	Étude du réseau récurrent	34
2.4	Modèle quadratique	37
2.4.1	Réseau récurrent associé	38
2.4.2	Points d'équilibre du réseau	39
2.4.3	Stabilité locale des points d'équilibre du réseau	41
2.4.4	Limites de la stabilité locale du réseau étendu	42
2.5	Cas où les sorties sont permutées	44
2.6	Comparatif des travaux sur les deux modèles	45
2.7	Simulations	46
2.7.1	Réseau basique	47
2.7.2	Améliorations avec le réseau étendu : scénario n° 1	48
2.7.3	Améliorations avec le réseau étendu : scénario n° 2	49
2.8	Conclusion	51

2.1 Introduction

Dans le premier chapitre, nous avons introduit les modèles de mélanges linéaire-quadratique et quadratique, dans leurs versions originales et normalisées. Dans le présent chapitre, nous présentons les structures de séparation, de type réseau récurrent, proposées pour ces mélanges non-linéaires particuliers. Dans ce chapitre, comme dans les suivants, seules les versions normalisées des modèles seront considérées. Nous commencerons avec les travaux antérieurs qui ont été faits en SAS dans le cadre des réseaux récurrents pour le modèle de mélange linéaire. Puis, nous exposerons les travaux effectués sur le modèle linéaire-quadratique. Nous présenterons ensuite nos contributions pour le mélange quadratique. Ceci nous permettra d'effectuer un parallèle entre les deux modèles. Les résultats présentés concerneront essentiellement les propriétés des réseaux étudiés, nous déterminerons les points d'équilibre des réseaux et étudierons leur stabilité locale. Il est important de souligner que nous nous situerons *en contexte non aveugle*, et à un instant donné, en d'autres termes nous déterminerons un vecteur de sorties pour un vecteur d'observations donné, les coefficients de mélange étant connus. Des simulations seront fournies en fin de chapitre.

2.2 Travaux antérieurs

Les réseaux récurrents introduits dans ce chapitre constituent un prolongement des travaux effectués en SAS avec les réseaux récurrents, dans le cadre du modèle linéaire instantané. Le premier réseau récurrent introduit, à l'origine de la SAS, est le réseau de Héroult-Jutten, bien connu. Ce réseau a été énormément étudié. Une synthèse sur leur algorithme itératif peut être trouvée dans [45, 46]. L'analyse de la stabilité a été effectuée par E. Sorouchyari dans la même série d'articles [75]. J.-C. Fort a également effectué une analyse de la stabilité de ce réseau dans [36]. Il existe cependant une différence entre les analyses de stabilité du réseau récurrent citées précédemment et l'analyse de stabilité que nous proposons, car elle s'effectuera *en contexte non aveugle* et à un instant donné, i.e. nous proposerons un algorithme de calcul d'un vecteur de sortie, pour un vecteur d'entrée, à poids fixés, tandis que les travaux cités précédemment traitent d'algorithmes d'adaptation des poids du réseau.

On peut également citer les travaux d'O. Macchi et E. Moreau qui ont également proposé des réseaux directs et hybrides pour le modèle de mélange linéaire. Après avoir analysé la convergence d'une structure directe contrôlée par l'algorithme de Héroult-Jutten dans [52], ils comparent plusieurs structures, directe, récurrente et hybride, dans [59]. Des structures étendues ont également été introduites dans le cas linéaire instantané, notamment par Cichocki *et al* [18]. Des compléments sur cette littérature peuvent être trouvés dans [28].

2.3 Modèle linéaire-quadratique

2.3.1 Généralités

Le modèle linéaire-quadratique appartient à la classe de mélanges dits à CCA, introduite dans [28]. Nous étudions dans cette section les différentes structures qui ont été proposées et étudiées dans [28, 40, 41], pour le modèle linéaire-quadratique. L'approche proposée pour construire la structure de séparation est simple à comprendre : il suffit de considérer que le modèle de mélange étudié est composé de *termes d'interférence* et de **termes d'interférence**, il s'agit alors de construire un réseau récurrent, qui *conserve* les premiers et **supprime** les seconds.

Nous rappelons que le modèle linéaire-quadratique, pour 2 sources et 2 observations, est le suivant :

$$\begin{cases} x_1(n) = s_1(n) - L_{12}s_2(n) - Q_1s_1(n)s_2(n) \\ x_2(n) = s_2(n) - L_{21}s_1(n) - Q_2s_1(n)s_2(n) \end{cases} \quad (2.1)$$

avec (s_1, s_2) , les sources normalisées et (x_1, x_2) les observations. Les termes L_{12} , L_{21} , Q_1 , Q_2 représentent les coefficients du mélange normalisés, constants et réels.

Pour ce type de modèle une structure de séparation, de type réseau récurrent, avait été considérée dans [40, 41], en contexte aveugle, dans une version du réseau dite basique. Les équations de récurrence sur m qui la définissent à chaque instant n sont :

$$\begin{cases} y_1(m+1) = x_1(n) + l_{12}y_2(m) + q_1y_1(m)y_2(m) \\ y_2(m+1) = x_2(n) + l_{21}y_1(m) + q_2y_1(m)y_2(m) \end{cases} \quad (2.2)$$

et la Figure 2.1 représente cette structure.

La différence entre les travaux présentés dans [40] et [41] réside dans le critère utilisé afin d'estimer les coefficients du mélange. Dans le premier article, les auteurs ont recours aux statistiques croisées d'ordre supérieur, tandis que dans le second, une approche par maximum de vraisemblance est utilisée. Dans [28], la structure

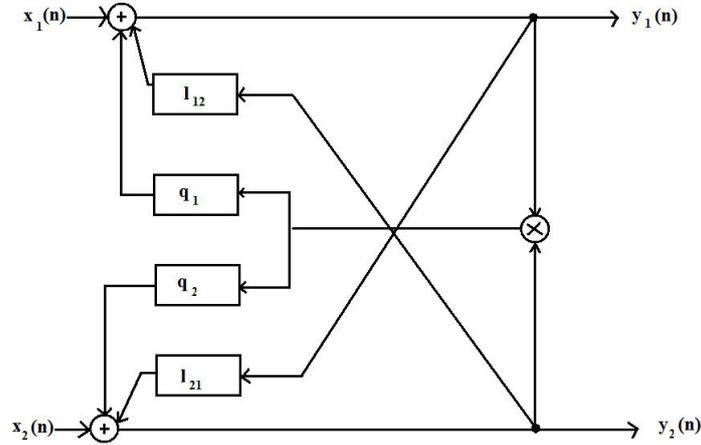


FIGURE 2.1 – Structure récurrente basique pour le modèle linéaire-quadratique

basique décrite précédemment est vue comme un cas particulier d'une structure dite étendue, comprenant des degrés de liberté supplémentaires. Cette structure est alors étudiée en *contexte non aveugle*. Elle est définie à chaque instant n par les équations de récurrence sur m suivantes :

$$\begin{cases} y_1(m+1) = x_1(n) + l_{11}y_1(m) + l_{12}y_2(m) + q_1y_1(m)y_2(m) \\ y_2(m+1) = x_2(n) + l_{22}y_2(m) + l_{21}y_1(n) + q_2y_1(m)y_2(m) \end{cases} \quad (2.3)$$

La Figure 2.2 illustre ce type de réseau.

D'autre part, la Figure 2.3 montre le principe de *réseau associé* dans ce cas au modèle linéaire-quadratique, où l'on considère que le modèle de mélange étudié est composé de *termes d'intérêt* et de **termes d'interférence**, et qu'il s'agit alors de construire un réseau récurrent, qui *conserve* les premiers et **supprime** les seconds. Les termes $l_{11} y_1$ et $l_{22} y_2$ sont les termes d'auto-feedback, et les facteurs l_{11} et l_{22} les degrés de liberté supplémentaires. Le lien entre les facteurs l_{ii} , l_{ij} , et q_i et les paramètres du mélange seront définis dans la sous-section suivante. La possibilité d'utiliser des structures directes avait également été évoquée dans [41]. En effet, la résolution du système (2.1) conduit à l'expression des solutions en fonction des

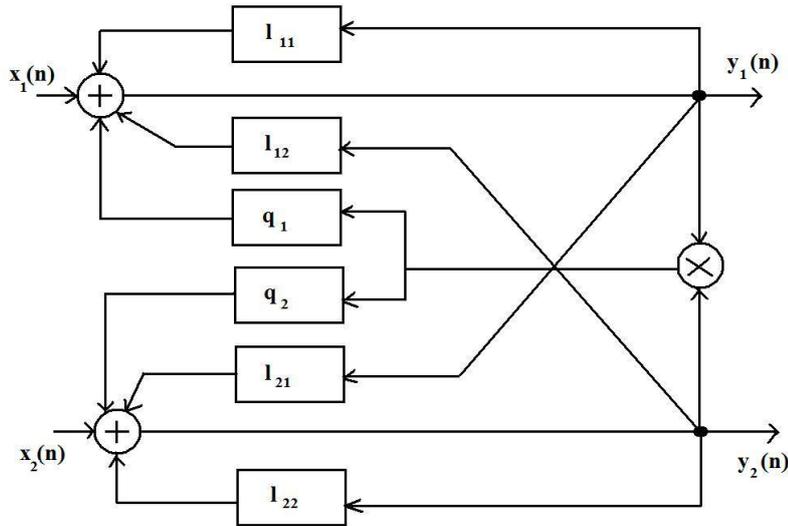
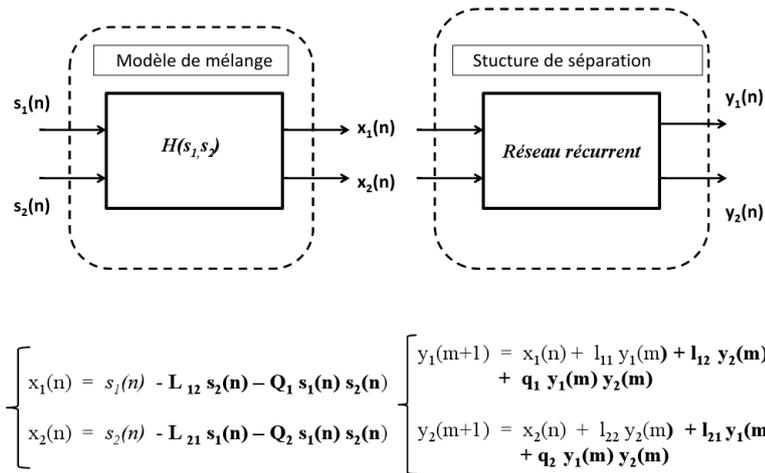


FIGURE 2.2 – Structure récurrente étendue, avec termes d'auto-feedback, pour le modèle linéaire-quadratique.



$$\begin{cases} x_1(n) = s_1(n) - l_{12} s_2(n) - q_1 s_1(n) s_2(n) \\ x_2(n) = s_2(n) - l_{21} s_1(n) - q_2 s_1(n) s_2(n) \end{cases} \quad \begin{cases} y_1(m+1) = x_1(n) + l_{11} y_1(m) + l_{12} y_2(m) \\ \quad + q_1 y_1(m) y_2(m) \\ y_2(m+1) = x_2(n) + l_{22} y_2(m) + l_{21} y_1(m) \\ \quad + q_2 y_1(m) y_2(m) \end{cases}$$

FIGURE 2.3 – Modèle de mélange et principe de construction du réseau récurrent pour le modèle linéaire-quadratique.

observations et des coefficients du mélange. Cette expression avait été donnée dans [41], et permet de déduire facilement les structures directes correspondantes. Mais l'intérêt d'un tel réseau récurrent, dans ses deux versions, réside dans sa simplicité de généralisation à des modèles plus complexes, pour lesquels la version explicite de l'inverse du modèle n'est pas disponible. Le réseau étendu a quant à lui été introduit comme solution aux problèmes de stabilité, rencontrés avec le réseau basique [40].

Après avoir défini de telles structures, il est intéressant d'étudier leurs propriétés, c'est ce que nous aborderons dans la sous-section qui va suivre.

2.3.2 Étude du réseau récurrent

Dans cette sous-section, nous évoquons les principaux résultats obtenus pour l'analyse du réseau récurrent linéaire-quadratique. Après avoir rappelé les limitations du réseau récurrent basique, nous passerons à l'analyse de la stabilité du réseau étendu, en contexte non aveugle. Le calcul des points d'équilibre du réseau ne sera pas effectué.

Limitations du réseau basique

L'étude de la stabilité locale du réseau avait été entreprise dans [40]. Les principales limitations du réseau récurrent dans sa version basique sont d'une part, que même pour les vraies valeurs des paramètres, le réseau peut être instable, et d'autre part qu'au courant de la procédure d'adaptation, les coefficients du réseau peuvent être amenés à appartenir à une région où le réseau récurrent est instable. L'analyse locale aux points d'équilibre⁹ peut se faire par un développement en série de Taylor de premier ordre autour de ce point. Pour un système avec 2 sources et 2 observations nous avons :

$$\mathbf{y}(m+1) \simeq \mathbf{f}(\mathbf{s}) + \mathbf{J}(\mathbf{y}(m) - \mathbf{s}) \quad (2.4)$$

$\mathbf{y} = [y_1, y_2]^T$ étant le vecteur de sorties, $\mathbf{s} = [s_1, s_2]^T = [y_{1E}, y_{2E}]^T$ le vecteur sources (point d'équilibre) autour duquel se fait le développement, et \mathbf{J} est la matrice jacobienne de $\mathbf{f} = [f_1, f_2]$, fonctions non-linéaires, définie par $\mathbf{J} = \frac{\partial \mathbf{f}(\mathbf{y}(m))}{\partial \mathbf{y}(m)} \Big|_{(y_{1E}, y_{2E})}$.

La stabilité locale au point d'équilibre $\mathbf{y}_E = (y_{1E}, y_{2E})$ dépend des valeurs propres de la matrice \mathbf{J} dont le module doit être inférieur à 1 pour assurer la stabilité. Dans [40],

⁹Cette notion sera définie explicitement dans la prochaine section.

après avoir donné l'expression des valeurs propres de \mathbf{J} , deux cas sont envisagés suivant que ces valeurs sont complexes ou réelles. Il en découle les conditions suivantes : si les valeurs sont complexes, la récurrence (2.2) est localement stable si et seulement si :

$$(Q_1 s_2 + Q_2 s_1)^2 + |(Q_1 s_2 + Q_2 s_1)^2 + 4(L_{12}L_{21} + L_{12}Q_2 s_2 + L_{21}Q_1 s_1)| < 4. \quad (2.5)$$

Si les valeurs propres sont réelles alors le réseau est localement stable si et seulement si :

$$|Q_1 s_2 + Q_2 s_1 \pm \sqrt{(Q_1 s_2 + Q_2 s_1)^2 + 4(L_{12}L_{21} + L_{12}Q_2 s_2 + L_{21}Q_1 s_1)}| < 2 \quad (2.6)$$

Ces résultats montrent que pour des coefficients de mélange fixés, l'une de ces inégalités ne peut être vérifiée que sur un certain domaine d'appartenance des sources, ce qui limite l'application de ce réseau basique.

Amélioration de la stabilité locale via le réseau étendu

Les points d'équilibre du réseau recherchés sont ceux de la forme¹⁰ [28] :

$$y_{1E} = k_1 s_1(n) \text{ et } y_{2E} = k_2 s_2(n) \quad (2.7)$$

c'est-à-dire que nous recherchons les sources normalisées à un facteur d'échelle près. Ceci nous amène à choisir les paramètres l_{ij} , q_i et k_i de telle sorte que les équations suivantes soient vérifiées :

$$\left\{ \begin{array}{l} l_{11} = -\frac{1}{k_1} + 1 \\ l_{12} = \frac{L_{12}}{k_2} = L_{12}l'_{22} \\ q_1 = \frac{Q_1}{k_1 k_2} = Q_1 l'_{11} l'_{22} \\ l_{22} = -\frac{1}{k_2} + 1 \\ l_{21} = \frac{L_{21}}{k_1} = L_{21}l'_{11} \\ q_2 = \frac{Q_2}{k_1 k_2} = Q_2 l'_{11} l'_{22} \end{array} \right. \quad (2.8)$$

avec :

$$l'_{11} = 1 - l_{11} \text{ et } l'_{22} = 1 - l_{22} \quad (2.9)$$

¹⁰le mode opératoire de calcul des points d'équilibre sera détaillé à la section suivante. Celui-ci étant le même pour le réseau linéaire-quadratique que pour le réseau quadratique, nous renvoyons donc le lecteur à la prochaine section pour le détail des calculs.

La procédure est alors la suivante : sélectionner une valeur pour les paramètres l_{11} et l_{22} , et ensuite fixer les autres paramètres à l'aide de (2.8) et (2.9). Par construction le réseau a donc au moins un point fixe. Il s'agit alors de déterminer s'il existe d'autres points fixes pour les valeurs des paramètres sélectionnées comme indiqué précédemment. Ceci nous conduit aux 2 solutions [28] :

$$\begin{cases} y_{1E} = \frac{1}{l'_{11}} s_1 \text{ et } y_{2E} = \frac{1}{l'_{22}} s_2 \\ y_{1E} = \frac{1}{l'_{11}} \left[\frac{\beta}{\alpha} s_2 + \frac{\gamma}{\alpha} \right] \text{ et } y_{2E} = \frac{1}{l'_{22}} \left[\frac{\alpha}{\beta} s_1 - \frac{\gamma}{\beta} \right] \end{cases} \quad (2.10)$$

$$\text{avec } \begin{cases} \alpha = Q_2 + Q_1 L_{21} \\ \beta = -(Q_1 + Q_2 L_{12}) \\ \gamma = 1 - L_{21} L_{12} \end{cases} \quad (2.11)$$

Ces solutions sont toujours des solutions séparantes, mais la deuxième est permutée avec une constante additive.

Si l'on reconsidère la matrice jacobienne \mathbf{J} , précédemment définie, la condition de stabilité aux points d'équilibre peut également s'écrire en fonction de la trace T et du déterminant D de cette matrice, à l'aide du système suivant [81] :

$$\begin{cases} T + D + 1 > 0, \\ T - D - 1 < 0, \\ D - 1 < 0. \end{cases} \quad (2.12)$$

Après quelques calculs ces inégalités deviennent :

$$\begin{cases} |l'_{11} l'_{22}| \sqrt{\delta_{y_1}} - 2A l'_{11} - 2B l'_{22} + 4 > 0 \\ |l'_{11} l'_{22}| \sqrt{\delta_{y_1}} - A l'_{11} - B l'_{22} < 0 \end{cases} \quad (2.13)$$

$$\text{avec } \begin{cases} \delta_{y_1} = [Q_2 x_1(n) - Q_1 x_2(n) + \gamma]^2 - 4\alpha [x_1(n) + x_2(n) L_{12}] \\ A = \frac{Q_1}{2\beta} [-Q_2 x_1(n) + Q_2 x_2(n) + \gamma - \text{sign}(l'_{11} l'_{22}) \sqrt{\delta_{y_1}}] + 1 \\ B = \frac{Q_2}{2\alpha} [-Q_2 x_1(n) + Q_2 x_2(n) - \gamma + \text{sign}(l'_{11} l'_{22}) \sqrt{\delta_{y_1}}] + 1 \end{cases} \quad (2.14)$$

Ces inéquations définissent complètement la stabilité d'un point d'équilibre. Ce qui est proposé, toujours dans [28] est de poser $l'_{22} = \lambda l'_{11}$, où λ est un paramètre. Sans entrer dans le détail des calculs, on en arrive à la conclusion de devoir choisir le

paramètre l'_{11} , pour un λ donné, qui garantit les conditions de stabilité, et défini par :

$$l'_{11} = \frac{\mu_{max} A + B\lambda}{2 |\lambda| \sqrt{\delta y_1}} \quad (2.15)$$

avec :

$$\mu_{max} = \begin{cases} 1 & \text{si } C(\lambda) \leq 4 \\ 1 - \sqrt{1 - \frac{4}{C(\lambda)}} & \text{sinon} \end{cases} \quad (2.16)$$

avec $C(\lambda) = \frac{(A+B\lambda)^2}{|\lambda| \sqrt{\delta y_1}}$

Ces développements nous conduisent à la procédure suivante, s'appliquant indépendamment à chaque instant n , et qui garantit d'avoir un système qui nous conduit vers une solution stable, i.e. que l'on converge vers l'un des deux points d'équilibre :

1. Sélectionner λ comme mentionné dans [28].
2. Sélectionner l'_{11} suivant l'équation (2.15) en tenant compte du fait que $sign(l'_{11}l'_{22}) = sign(\lambda)$.
3. Choisir $l'_{22} = \lambda l'_{11}$
4. Choisir les valeurs des autres paramètres suivant (2.8) et (2.9)

Un début d'évolution du réseau vers un contexte aveugle peut être trouvé dans [27].

2.4 Modèle quadratique

Nous abordons maintenant le cas du modèle quadratique. Comme précisé auparavant, le modèle et la structure de séparation avaient été introduits dans [28], cependant, les propriétés du réseau (existence et unicité des points d'équilibre, analyse de la stabilité aux points d'équilibre) n'avaient pas été étudiées. C'est ce que nous proposons d'examiner dans cette section. Contrairement au modèle linéaire-quadratique, les calculs seront détaillés, compte tenu de leur originalité.

Nous rappelons que le modèle considéré est un modèle aussi à CCA, pour 2 sources et 2 observations, il s'écrit :

$$\begin{cases} x_1(n) = s_1^2(n) - L_{12}s_2^2(n) - Q_1s_1(n)s_2(n) \\ x_2(n) = s_2^2(n) - L_{21}s_1^2(n) - Q_2s_1(n)s_2(n) \end{cases} \quad (2.17)$$

où (s_1, s_2) représentent les sources normalisées et (x_1, x_2) les observations. Les coefficients L_{ij} , $(i, j) \in \{(1, 2), (2, 1)\}$ et Q_i , $i \in \{1, 2\}$ représentent les coefficients du mélange normalisés, constants et réels.

Comme pour le modèle linéaire-quadratique, l'approche proposée consiste à considérer que le mélange étudié présente des composantes d'intérêt et des composantes d'interférence respectivement $[s_i^2(n)]$ et $[-L_{ij}s_j^2(n) - Q_i s_i(n)s_j(n)]$, $(i, j) \in \{(1, 2), (2, 1)\}$ pour (2.17). Nous lui associons alors un réseau récurrent, qui conserve les premières composantes et supprime les secondes, de sorte à récupérer à l'équilibre les sources en sortie du réseau, aux indéterminations de signe et d'échelle près. En d'autres termes, il s'agit d'obtenir à l'équilibre des sorties sous la forme :

$$y_i(m+1) = y_i(m) = y_{iE} = k_i s_i(n), \quad k_i \in \mathbb{R}^*, \quad i \in \{1, 2\} \quad (2.18)$$

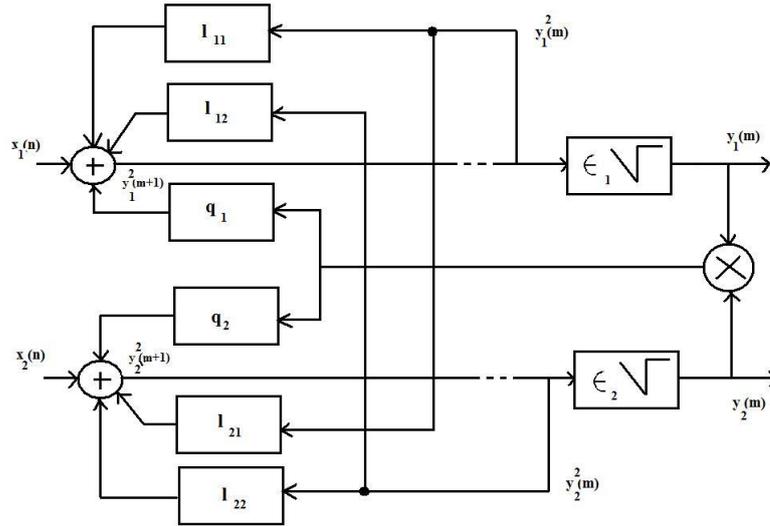


FIGURE 2.4 – Structure récurrente proposée pour le modèle quadratique

2.4.1 Réseau récurrent associé

Le réseau récurrent que nous utilisons est de type version étendue. En effet, il suffit de mettre les termes $l_{ii, i \in \{1, 2\}}$ à zéro pour revenir au réseau basique. A chaque instant n , comme l'illustre la figure 2.4, nous mettons en œuvre les récurrences sur m suivantes :

$$\begin{cases} y_1(m+1) = \epsilon_1 \sqrt{x_1(n) + l_{12}y_2^2(m) + q_1y_1(m)y_2(m) + l_{11}y_1^2(m)}, \\ y_2(m+1) = \epsilon_2 \sqrt{x_2(n) + l_{21}y_1^2(m) + q_2y_1(m)y_2(m) + l_{22}y_2^2(m)}, \end{cases} \quad (2.19)$$

où les $\epsilon_{i,i \in \{1,2\}}$ valent ± 1 et sont fixés à chaque instant n suivant le signe des sources. Lorsque les $k_{i,i \in \{1,2\}}$ sont déterminés, les paramètres $(l_{12}, l_{21}, l_{11}, l_{22}, q_1, q_2)$ sont ensuite fixés de manière à assurer la compatibilité des systèmes (2.17) et (2.19) à l'équilibre, avec (2.18) comme point d'équilibre. On peut montrer simplement qu'ils vérifient alors les relations suivantes :

$$\begin{cases} l_{11} = 1 - \frac{1}{k_1^2} \\ l_{22} = 1 - \frac{1}{k_2^2} \\ l_{12} = \frac{L_{12}}{k_2^2} \\ l_{21} = \frac{L_{21}}{k_1^2} \\ q_1 = \frac{Q_1}{k_1 k_2} \\ q_2 = \frac{Q_2}{k_1 k_2} \end{cases} \quad (2.20)$$

On pose

$$l'_{11} = 1 - l_{11} = \frac{1}{k_1^2} \text{ et } l'_{22} = 1 - l_{22} = \frac{1}{k_2^2}. \quad (2.21)$$

En pratique on commence d'abord par fixer les facteurs l_{11} et l_{22} , puis on en déduit les autres coefficients à l'aide de (2.20).

2.4.2 Points d'équilibre du réseau

L'étude de la structure définie par (2.19) nécessite la détermination de ses points d'équilibre et l'analyse de leur stabilité. Si l'on considère le système à l'équilibre (y_{1E}, y_{2E}) , on obtient en élevant (2.19) au carré et en utilisant (2.18) et (2.21) :

$$\begin{cases} x_1(n) + l_{12}y_{2E}^2 + q_1y_{1E}y_{2E} - l'_{11}y_{1E}^2 = 0 \\ x_2(n) + l_{21}y_{1E}^2 + q_2y_{1E}y_{2E} - l'_{22}y_{2E}^2 = 0 \end{cases} \quad (2.22)$$

qui mène à l'expression des termes croisés en supposant les $q_{i,i \in \{1,2\}}$ non nuls¹¹ :

$$y_{1E}y_{2E} = \frac{l'_{11}y_{1E}^2 - l_{12}y_{2E}^2 - x_1}{q_1} = \frac{l'_{22}y_{2E}^2 - l_{21}y_{1E}^2 - x_2}{q_2}; \quad (2.23)$$

ce qui conduit à :

$$y_{2E}^2 (q_1 l'_{22} + q_2 l_{12}) = y_{1E}^2 (l'_{11} q_2 + l_{21} q_1) + (x_2 q_1 - x_1 q_2) \quad (2.24)$$

¹¹Dans la suite des calculs l'indice n des termes x_1 et x_2 sera ignoré.

$$\text{et } y_{2E} = \epsilon_{2E} \sqrt{y_{1E}^2 \frac{\beta}{\alpha} + \frac{\gamma}{\alpha}}, \quad \epsilon_{2E} = \pm 1 \quad (2.25)$$

$$\text{avec } \begin{cases} q_1 l'_{22} + q_2 l_{12} = (Q_1 + Q_2 L_{12}) \frac{l'_{22}}{k_1 k_2} = \alpha \\ l'_{11} q_2 + l_{21} q_1 = (Q_2 + Q_1 L_{21}) \frac{l'_{11}}{k_1 k_2} = \beta \\ x_2 q_1 - x_1 q_2 = \gamma \end{cases} \quad (2.26)$$

En injectant (2.25) dans la première équation du système (2.22), on obtient :

$$q_1 \epsilon_{2E} y_{1E} \sqrt{y_{1E}^2 \frac{\beta}{\alpha} + \frac{\gamma}{\alpha}} = \left(l'_{11} - \frac{\beta}{\alpha} l_{12} \right) y_{1E}^2 - x_1 - l_{12} \frac{\gamma}{\alpha} \quad (2.27)$$

En élevant au carré (2.27), on obtient finalement l'équation du second degré en Y_1 :

$$aY_1^2 + bY_1 + c = 0, \quad Y_1 = y_{1E}^2. \quad (2.28)$$

$$\text{avec } \begin{cases} a = \left(l'_{11} - \frac{\beta}{\alpha} l_{12} \right)^2 - q_1^2 \frac{\beta}{\alpha}, \\ b = -q_1^2 \frac{\gamma}{\alpha} - 2 \left(l'_{11} - \frac{\beta}{\alpha} l_{12} \right) \left(x_1 + l_{12} \frac{\gamma}{\alpha} \right), \\ c = \left(x_1 + l_{12} \frac{\gamma}{\alpha} \right)^2 \end{cases} \quad (2.29)$$

Le discriminant de (2.28), $\delta = b^2 - 4ac$ s'écrit après simplification :

$$\begin{aligned} \delta &= \frac{q_1^4}{\alpha^2} \left(l'_{11} l'_{22} \right) \left[(Q_2 + Q_1 L_{21}) s_1^2(n) \right. \\ &\quad \left. + (Q_1 + Q_2 L_{12}) s_2^2(n) - 2(1 - L_{21} L_{12}) s_1(n) s_2(n) \right]^2 \end{aligned} \quad (2.30)$$

On remarque que $\delta \geq 0$. Par conséquent, (2.28) admet 2 racines réelles, correspondant à la première sortie au carré de la structure à l'équilibre :

$$\begin{cases} Y_1^{(1)} = \frac{s_1^2(n)}{l'_{11}} = k_1^2 s_1^2(n), \\ Y_1^{(2)} = k_1^2 \frac{[(1-L_{21}L_{12})s_1(n) - (Q_1+Q_2L_{12})s_2(n)]^2}{d} \\ \text{avec } d = (1-L_{21}L_{12})^2 - (Q_1+Q_2L_{12})(Q_2+Q_1L_{21}) \end{cases} \quad (2.31)$$

En utilisant (2.25), on en déduit les secondes sorties au carré associées ($Y_2 = y_{2E}^2$) :

$$\begin{cases} Y_2^{(1)} = \frac{s_2^2(n)}{l'_{22}} = k_2^2 s_2^2(n), \\ Y_2^{(2)} = k_2^2 \frac{[(Q_2+Q_1L_{21})s_1(n) - (1-L_{21}L_{12})s_2(n)]^2}{d} \end{cases} \quad (2.32)$$

Les sorties de la structure étant réelles, seules les solutions (2.31, 2.32) positives sont conservées : la première solution (2.31) existe toujours tandis que la seconde existe uniquement si $d > 0$. Nous obtenons alors, en prenant respectivement les racines

carrées des équations de (2.31) et (2.32), les expressions suivantes pour les points d'équilibre :

$$\begin{cases} (y_{1E}, y_{2E})^{(1)} = (\epsilon k_1 s_1(n), \epsilon k_2 s_2(n)) \\ (y_{1E}, y_{2E})^{(2)} = (\epsilon' k_1 \frac{[(1-L_{21}L_{12})s_1(n) - (Q_1 + Q_2 L_{12})s_2(n)]}{\sqrt{d}}, \\ \epsilon' k_2 \frac{[(Q_2 + Q_1 L_{21})s_1(n) - (1-L_{21}L_{12})s_2(n)]}{\sqrt{d}}), \text{ si } d > 0 \end{cases} \quad (2.33)$$

avec $d = (1 - L_{21}L_{12})^2 - (Q_1 + Q_2 L_{12})(Q_2 + Q_1 L_{21})$ et $\epsilon, \epsilon' = \pm 1$.

La première paire $(y_{1E}, y_{2E})^{(1)}$ correspond à une solution séparante, contrairement à la seconde $(y_{1E}, y_{2E})^{(2)}$ qui demeure un mélange linéaire des sources. Ces points d'équilibre doivent vérifier la relation :

$$\text{sign}(y_{iE}) = \text{sign}(\epsilon_i), \quad i \in \{1, 2\} \quad (2.34)$$

Pour les 2 équations de (2.33), (2.34) s'écrit respectivement :

$$\text{sign}(k_1)\text{sign}(k_2) = \text{sign}(\epsilon_1)\text{sign}(\epsilon_2)\text{sign}(s_1(n))\text{sign}(s_2(n)) \quad (2.35)$$

$$\text{et } \text{sign}(k_1)\text{sign}(k_2) = \text{sign}(\epsilon_1)\text{sign}(\epsilon_2)\text{sign}(r_1)\text{sign}(r_2) \quad (2.36)$$

$$\text{avec } \begin{cases} r_1 = (1 - L_{21}L_{12})s_1(n) - (Q_1 + Q_2 L_{12})s_2(n), \\ r_2 = (Q_2 + Q_1 L_{21})s_1(n) - (1 - L_{21}L_{12})s_2(n) \end{cases} \quad (2.37)$$

Si l'on suppose que (2.35) est vérifiée, la structure présente *au moins* un point d'équilibre, le point séparant défini par la première équation de (2.33). Le second point d'équilibre, deuxième équation de (2.33), non séparant, *existe* si d est positif, ce qui assure son existence dans \mathbb{R} , *et* si la condition (2.36) est vérifiée.

2.4.3 Stabilité locale des points d'équilibre du réseau

Dans cette partie on suppose que les paramètres $(l_{12}, l_{21}, l_{11}, l_{22}, q_1, q_2)$ vérifient les conditions (2.20) et que la condition (2.35) est valide. Pour étudier la stabilité locale¹² d'un point d'équilibre (y_{1E}, y_{2E}) , nous utilisons l'approximation au premier ordre de (2.19) au voisinage de celui-ci, définie par :

$$\begin{cases} y(m+1) \approx f(y_E) + J(y(m) - y_E), \quad y = \begin{pmatrix} y_1 \\ y_2 \end{pmatrix}^T \\ f(y) = \begin{pmatrix} \epsilon_1 \sqrt{x_1(n) + l_{12}y_2^2 + q_1 y_1 y_2 + l_{11}y_1^2} \\ \epsilon_2 \sqrt{x_2(n) + l_{21}y_1^2 + q_2 y_1 y_2 + l_{22}y_2^2} \end{pmatrix}, \end{cases} \quad (2.38)$$

¹²Nous n'étudierons la stabilité locale que pour le point séparant

où J désigne la matrice jacobienne au point d'équilibre (y_{1E}, y_{2E}) définie par :

$$J = \left(\frac{\partial f(y(m))}{\partial y(m)} \right)_{|(y_{1E}, y_{2E})} \quad (2.39)$$

$$J = \begin{pmatrix} \frac{q_1 y_{2E} + 2l_{11} y_{1E}}{2y_{1E}} & \frac{2l_{12} y_{2E} + q_1 y_{1E}}{2y_{1E}} \\ \frac{2l_{21} y_{1E} + q_2 y_{2E}}{2y_{2E}} & \frac{q_2 y_{1E} + 2l_{22} y_{2E}}{2y_{2E}} \end{pmatrix} \quad (2.40)$$

Une condition nécessaire et suffisante pour qu'un point d'équilibre soit localement stable est que les valeurs propres de J , $\lambda_{i,i \in \{1,2\}}$ vérifient la relation $|\lambda_{i,i \in \{1,2\}}| < 1$. Comme mentionné précédemment, le point d'équilibre vérifie alors le système [81] :

$$\begin{cases} T + D + 1 > 0, \\ T - D - 1 < 0, \\ D - 1 < 0. \end{cases} \quad (2.41)$$

où T et D désignent respectivement la trace et le déterminant de J .

Pour le point séparant, première équation dans (2.33), le système (2.41) s'écrit :

$$\begin{cases} \frac{s_1^2((2k_1^2-1)Q_2 - L_{21}Q_1) + s_2^2((2k_2^2-1)Q_1 - L_{12}Q_2) + 2s_1s_2(2(2k_1^2k_2^2 - k_1^2 - k_2^2) + (1 - L_{12}L_{21}))}{2(k_1k_2)^2s_1s_2} > 0 \\ \frac{(Q_2 + Q_1L_{21})s_1^2 + (Q_1 + Q_2L_{12})s_2^2 - 2s_1s_2(1 - L_{12}L_{21})}{2(k_1k_2)^2s_1s_2} < 0, \\ \frac{s_1^2((k_1^2-1)Q_2 - L_{21}Q_1) + s_2^2((k_2^2-1)Q_1 - L_{12}Q_2) + 2s_1s_2(-(k_1^2 + k_2^2) + (1 - L_{12}L_{21}))}{2(k_1k_2)^2s_1s_2} < 0 \end{cases} \quad (2.42)$$

Pour le point non séparant, lorsqu'il existe, on peut montrer que la deuxième équation de (2.41) s'écrit :

$$\frac{-d[(Q_2 + Q_1L_{21})s_1^2 + (Q_1 + Q_2L_{12})s_2^2 - 2(1 - L_{12}L_{21})s_1s_2]}{2(k_1k_2)^2r_1r_2} < 0, \quad (2.43)$$

La deuxième équation de (2.42) et l'équation (2.43) montrent, d'une part, que leur validité ne dépend pas des $k_{i,i \in \{1,2\}}$, et d'autre part, que si les 2 points d'équilibre existent, seul l'un d'entre eux peut être stable. La démonstration sera donnée dans la sous-section suivante.

2.4.4 Limites de la stabilité locale du réseau étendu

Dans cette sous-section, nous allons montrer qu'il existe des cas où l'auto-feedback n'a pas d'influence sur la stabilité du réseau au point séparant.

Il s'agit de démontrer la remarque faite précédemment, et qui va établir que suivant les valeurs que prennent les sources et les coefficients, si du moins le point séparant est stable ou pas, alors forcément le point non séparant sera instable ou pas. La démonstration est basée sur la comparaison des expressions dérivées à partir de la deuxième équation de (2.41) calculée pour les points séparant et non séparant.

Nous rappelons que l'expression de la matrice jacobienne au point d'équilibre (y_{1E}, y_{2E}) s'écrit :

$$J = \frac{1}{2} \begin{pmatrix} \frac{q_1 y_{2E} + 2l_{11} y_{1E}}{y_{1E}} & \frac{2l_{12} y_{2E} + q_1 y_{1E}}{y_{1E}} \\ \frac{2l_{21} y_{1E} + q_2 y_{2E}}{y_{2E}} & \frac{q_2 y_{1E} + 2l_{22} y_{2E}}{y_{2E}} \end{pmatrix} \quad (2.44)$$

Le point (y_{1E}, y_{2E}) est stable si (2.41) est vérifiée. Nous allons maintenant dériver l'équation 2 de (2.41) pour les points séparant et non séparant.

L'expression de $(T - D - 1)$ en fonction des sorties (y_1, y_2) à l'équilibre est :

$$T - D - 1 = \frac{y_{1E}^2 \beta' + y_{2E}^2 \alpha' - 2y_{1E} y_{2E} \frac{1 - L_{12} L_{21}}{k_1^2 k_2^2}}{2y_{1E} y_{2E}} \quad (2.45)$$

avec, $\beta' = q_2 - q_2 l_{11} + q_1 l_{21}$ et $\alpha' = q_1 - q_1 l_{22} + q_2 l_{12}$.

Expression au point séparant

Au point séparant, $y_{1E} = \epsilon k_1 s_1$ et $y_{2E} = \epsilon k_2 s_2$, nous avons obtenu l'expression :

$$(T - D - 1)_{sep} = \frac{s_1^2(Q_2 + Q_1 L_{21}) + s_2^2(Q_1 + Q_2 L_{12}) - 2s_1 s_2(1 - L_{12} L_{21})}{2k_1^2 k_2^2 s_1 s_2} \quad (2.46)$$

D'après l'équation (2.46), le signe de $(T - D - 1)$ ne dépend pas des paramètres k_1 et k_2 , il ne dépend que des paramètres du mélange et des sources.

Expression au point non séparant

Après quelques calculs et simplifications on aboutit à l'équation :

$$(T - D - 1)_{non-sep} = \frac{-d(s_1^2(Q_2 + Q_1 L_{21}) + s_2^2(Q_1 + Q_2 L_{12}) - 2s_1 s_2(1 - L_{12} L_{21}))}{2k_1^2 k_2^2 A_1 A_2} \quad (2.47)$$

$$\text{avec} \begin{cases} d = (1 - L_{12} L_{21})^2 - (Q_1 + Q_2 L_{12})(Q_2 + Q_1 L_{21}) \\ A_1 = (1 - L_{12} L_{21})s_1 - (Q_1 + Q_2 L_{12})s_2 \\ A_2 = (Q_2 + Q_1 L_{21})s_1 - (1 - L_{12} L_{21})s_2 \end{cases} \quad (2.48)$$

On a donc finalement,

$$(T - D - 1)_{sep} = -\frac{(T - D - 1)_{non-sep} A_1 A_2}{d s_1 s_2} \quad (2.49)$$

$$\text{or} \begin{cases} d > 0 \\ \text{sign}(A_1 A_2) = \text{sign}(s_1 s_2) \end{cases} \quad (2.50)$$

La deuxième égalité dans (2.50) est obtenue à partir du système (2.33), en effet, si les 2 points d'équilibre *existent en même temps*, alors d'après (2.33) :

$$\begin{cases} \text{sign}(\epsilon s_1) = \text{sign}(\epsilon' \frac{A_1}{\sqrt{d}}) \\ \text{sign}(\epsilon s_2) = \text{sign}(\epsilon' \frac{A_2}{\sqrt{d}}) \end{cases} \quad (2.51)$$

sachant que :

$$\text{sign}(y_{1E}^{(1)}) = \text{sign}(y_{1E}^{(2)}), \text{ et } \text{sign}(y_{2E}^{(1)}) = \text{sign}(y_{2E}^{(2)}) \quad (2.52)$$

nous avons :

$$\text{sign}(s_1 s_2) = \text{sign}(A_1 A_2). \quad (2.53)$$

On en conclut d'après (2.49) que si l'un des points d'équilibre est stable, l'autre ne l'est forcément pas car $(T - D - 1)_{sep}$ et $(T - D - 1)_{non-sep}$ sont de signes contraires. De plus, on remarque que les termes d'auto-feedback n'ont aucune influence sur ces inéquations.

Malgré ce type de comportement, l'intérêt de l'utilisation du réseau étendu n'est pas remis en cause, car les 2 autres inéquations de stabilité dans (2.41) dépendent des facteurs k_1 et k_2 , comme nous le montrerons à travers des exemples dans la section « Simulations ». De plus, en annexe (Annexe A), nous démontrerons que dans le cas du réseau basique, si le point « parasite » n'existe pas, alors le point séparant est forcément instable.

2.5 Cas où les sorties sont permutées

Jusqu'à présent, nous avons considéré des sorties obtenues à l'équilibre, à un facteur d'échelle près. Les développements théoriques correspondant ne tiennent pas compte d'une éventuelle permutation des sources, en d'autres termes, nous n'avons pas encore cherché des points d'équilibre du type :

$$y_{1E} = \epsilon k_1 s_2(n), \text{ et } y_{2E} = \epsilon' k_2 s_1(n), \quad (2.54)$$

Afin d'assurer la compatibilité des systèmes (2.17) et (2.19) à l'équilibre avec (2.54) comme point d'équilibre, nous obtenons dans ce cas les nouveaux poids du réseau :

$$\left\{ \begin{array}{l} l_{11} = \frac{L_{12}}{k_1^2} + 1 \\ l_{22} = \frac{L_{21}}{k_2^2} + 1 \\ l_{12} = -\frac{1}{k_2^2} \\ l_{21} = -\frac{1}{k_1^2} \\ q_1 = \frac{Q_1}{k_1 k_2} \\ q_2 = \frac{Q_2}{k_1 k_2} \end{array} \right. \quad (2.55)$$

De la même manière que précédemment on obtient les solutions au carré suivantes (il s'agit de tous les points d'équilibre résultant de (2.55)) :

$$\left\{ \begin{array}{l} Y_1^{(1)} = k_1^2 s_2^2(n), \\ Y_1^{(2)} = k_1^2 \frac{[(Q_2 + Q_1 L_{21})s_1(n) + (1 - L_{21} L_{12})s_2(n)]^2}{d}, \quad \text{si } d > 0 \\ Y_2^{(1)} = k_2^2 s_1^2(n), \\ Y_2^{(2)} = k_2^2 \frac{(1 - L_{21} L_{12})s_1(n) + (Q_1 + Q_2 L_{12})s_2(n)}{d}, \quad \text{si } d > 0 \end{array} \right. \quad (2.56)$$

avec $d = (1 - L_{21} L_{12})^2 - (Q_1 + Q_2 L_{12})(Q_2 + Q_1 L_{21})$,

soient les solutions :

$$\left\{ \begin{array}{l} (y_{1E}, y_{2E})^{(1)} = (\epsilon k_1 s_2(n), \epsilon k_2 s_1(n)) \\ (y_{1E}, y_{2E})^{(2)} = (\epsilon' k_1 \frac{[(Q_2 + Q_1 L_{21})s_1(n) + (1 - L_{21} L_{12})s_2(n)]}{\sqrt{d}}, \\ \epsilon' k_2 \frac{[(1 - L_{21} L_{12})s_1(n) + (Q_1 + Q_2 L_{12})s_2(n)]}{\sqrt{d}}), \quad \text{si } d > 0 \end{array} \right. \quad (2.57)$$

Nous obtenons donc les solutions permutées ainsi que des solutions qui constituent encore un mélange linéaire des sources. Les poids du réseau associés à ces solutions, sont évidemment différents des poids obtenus précédemment dans le cas de sources non permutées.

2.6 Comparatif des travaux sur les deux modèles

Dans un premier temps, nous avons exposé les résultats obtenus concernant les réseaux récurrents, et ce, pour le modèle linéaire-quadratique, avant de passer à nos contributions relatives au modèle quadratique. Le raisonnement utilisé pour

déterminer les points d'équilibre, les poids du réseau associés, ainsi que l'étude de la stabilité locale au point d'équilibre est le même pour les deux modèles. Cependant les calculs deviennent plus complexes pour le mélange quadratique. Contrairement au cas linéaire-quadratique, dans le cas quadratique il n'est pas possible de garantir la stabilité du point d'équilibre séparant. En effet, nous sommes confrontés à des problèmes au niveau de la stabilité locale du réseau dans le cas quadratique où nous avons montré que certaines inéquations étaient indépendantes des degrés de liberté introduits afin de stabiliser le réseau dans certains cas. Il serait alors peut-être nécessaire d'envisager un sous-cas du modèle quadratique, qui tiendrait compte du fait que ces inégalités, ne dépendent que des sources et des paramètres du mélange, et donc, ce sous-cas correspondrait alors à l'ajout d'hypothèses, menant à la validité de cette inéquation du système traduisant la stabilité du réseau au point séparant.

Les expressions obtenues pour les points d'équilibre sont plus complexes dans le cas quadratique. Pour ce dernier nous aboutissons à la solution « sources à un facteur d'échelle près » ainsi qu'à d'autres solutions qui restent des mélanges linéaires de ces sources, contrairement au cas linéaire-quadratique où nous aboutissons pour les deux points d'équilibre aux sources (à un facteur d'échelle ainsi qu'à une permutation et une constante additive près).

2.7 Simulations

Dans cette section, nous allons illustrer à l'aide de simulations les cas de stabilité/instabilité des réseaux basique et étendu dans le cas du modèle quadratique. Nous allons également valider les résultats théoriques obtenus ci-dessus. Nous montrons ainsi que l'introduction des $l_{ii, i \in \{1,2\}}$, et donc le recours à une structure étendue (comparativement à un réseau basique où ces derniers seraient mis à 0 soient les $k_{i, i \in \{1,2\}} = 1$), permet dans certains cas de stabiliser le réseau basique à un instant n .

Pour chaque simulation nous générons un voisinage de points d'initialisation autour du point d'équilibre considéré, suivant :

$$y_1(0) = y_{1E} + r \cos(\theta), \quad y_2(0) = y_{2E} + r \sin(\theta), \quad \text{avec } \theta = 2\pi \frac{k}{K} \quad (2.58)$$

où K représente le nombre de points d'initialisation générés. Dans ce qui va suivre,

M représente le nombre d'itérations de la récurrence. Les valeurs des sources normalisées, des coefficients du mélange (L_{12} , L_{21} , Q_1 , Q_2) ainsi que les valeurs des paramètres M , r et K seront mentionnées dans la légende des figures illustrant les simulations. Nous précisons également que les $\epsilon_{i,i \in \{1,2\}}$ sont fixés à 1 dans nos simulations.

2.7.1 Réseau basique

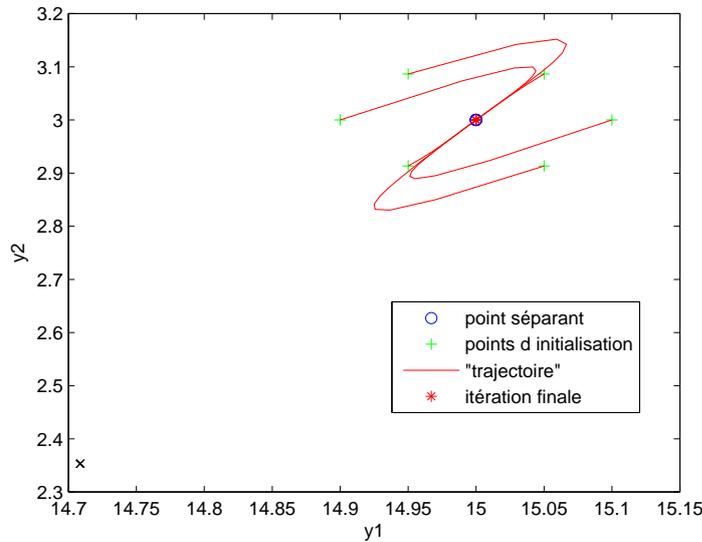


FIGURE 2.5 – Réseau basique : Convergence locale vers le point séparant (15, 3). $s_1 = 15$, $s_2 = 3$, $L_{12} = -0.5$, $L_{21} = -0.2$, $Q_1 = 1$, $Q_2 = 0.5$, $M = 100$, $r = 0.1$, $K = 6$.

Dans un premier temps nous illustrons le comportement du réseau basique défini par les équations (2.19) avec $l_{11} = l_{22} = 0$, pour lequel le point séparant (15, 3) est localement stable (Figure 2.5), tandis que le point non-séparant (14.7087, 2.3534) est localement instable (Figure 2.6).

D'un point de vue théorique nous avons obtenu les résultats suivants :

- Pour le point séparant, nous avons obtenu les valeurs suivantes pour les membres de gauche du système (2.41) : $[20775, -0.075, -0.575]$, ce qui correspond aux modules des valeurs propres $|\lambda_1| = 0.85$ et $|\lambda_2| = 0.5$. Ces résultats confirment que le point séparant est localement stable car le système d'inéquations (2.41) est vérifié, tout comme le module des valeurs propres qui est inférieur à 1.

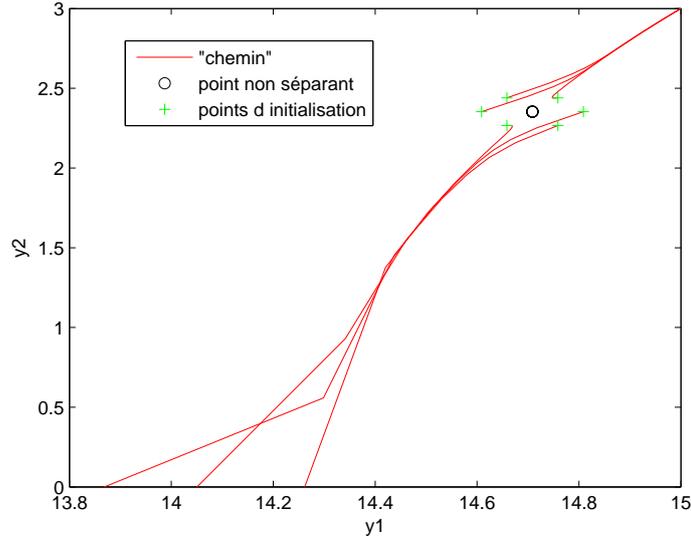


FIGURE 2.6 – Réseau basique : Divergence locale du point non-séparant $(14.7087, 2.3534)$. $s_1 = 15$, $s_2 = 3$, $L_{12} = -0.5$, $L_{21} = -0.2$, $Q_1 = 1$, $Q_2 = 0.5$, $M = 100$, $r = 0.1$, $K = 6$.

- Pour le point non-séparant, nous avons obtenu les valeurs suivantes pour les membres de gauche du système (2.41) : $[3.1875, 0.0975, -0.455]$, ce qui correspond aux modules des valeurs propres $|\lambda_3| = 1.181$ et $|\lambda_4| = 0.3152$. Ces résultats confirment la divergence locale du point non séparant. En effet, la seconde inéquation du système (2.41) n'est pas vérifiée tout comme la condition sur le module de la valeur propre $|\lambda_3|$ qui est supérieur à 1.

2.7.2 Améliorations avec le réseau étendu : scénario n° 1

Nous nous plaçons d'abord dans un cas de configuration symétrique du modèle (2.17), $L_{ij} = L = -0.75$, $Q_i = Q = -0.5$, $s_i(n) = s = 4$, $i, j \in \{1, 2\}$ pour lequel la structure définie par (2.19) avec les $k_{i,i \in \{1,2\}}$ fixés à 1 (soit un réseau basique avec des $l_{i,i \in \{1,2\}} = 0$) a été utilisée.

Ce réseau ne présente *qu'un seul point d'équilibre* ((2.36) non vérifiée), le point séparant $(4, 4)$ instable, en effet

- Les valeurs de gauche du système (2.41) obtenues sont les suivantes : $[-0.4375, -0.5625, -1.9375]$, la première inéquation n'est donc pas vérifiée.

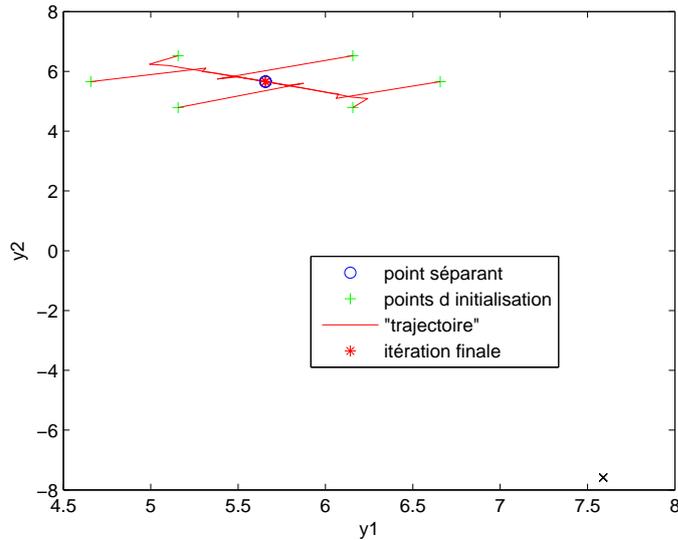


FIGURE 2.7 – Réseau étendu : Stabilisation locale du point séparant $s_1 = s_2 = 4$, $L_{12} = L_{21} = L = -0.75$, $Q_1 = Q_2 = Q = -0.5$, $k_{i,i \in \{1,2\}} = \sqrt{2}$, $M = 100$, $r = 0.1$, $K = 6$.

- Ce résultat est confirmé par le module des valeurs propres au point séparant $|\lambda_1| = 0.75$, $|\lambda_2| = 1.25 > 1$.

Grâce à la version étendue de notre réseau (i.e. $k_{i,i \in \{1,2\}} \neq 1$, $l_{i,i \in \{1,2\}} \neq 0$), en sélectionnant les $k_{i,i \in \{1,2\}} = \sqrt{2}$ de manière à assurer la validité de (2.42) puis en fixant $(l_{12}, l_{21}, l_{11}, l_{22}, q_1, q_2)$ suivant (2.20), le nouveau point séparant $(5.6569, 5.6569)$ devient localement stable :

- Les valeurs de gauche du système (2.41) obtenues sont les suivantes : $[1.6406, -0.1406, -1.1094]$, le système est donc vérifié,
- ce résultat est confirmé par le module des valeurs propres au point séparant $|\lambda_1| = 0.875$, $|\lambda_2| = 0.125$.

La Figure 2.7 montre que pour plusieurs initialisations générées au voisinage du nouveau point séparant, la structure converge vers ce dernier.

2.7.3 Améliorations avec le réseau étendu : scénario n° 2

On considère ensuite un mélange pour lequel $d < 0$ ($d = -0.0044$). Le point séparant $(1, 1)$, seul point d'équilibre de la structure est instable pour la structure basique

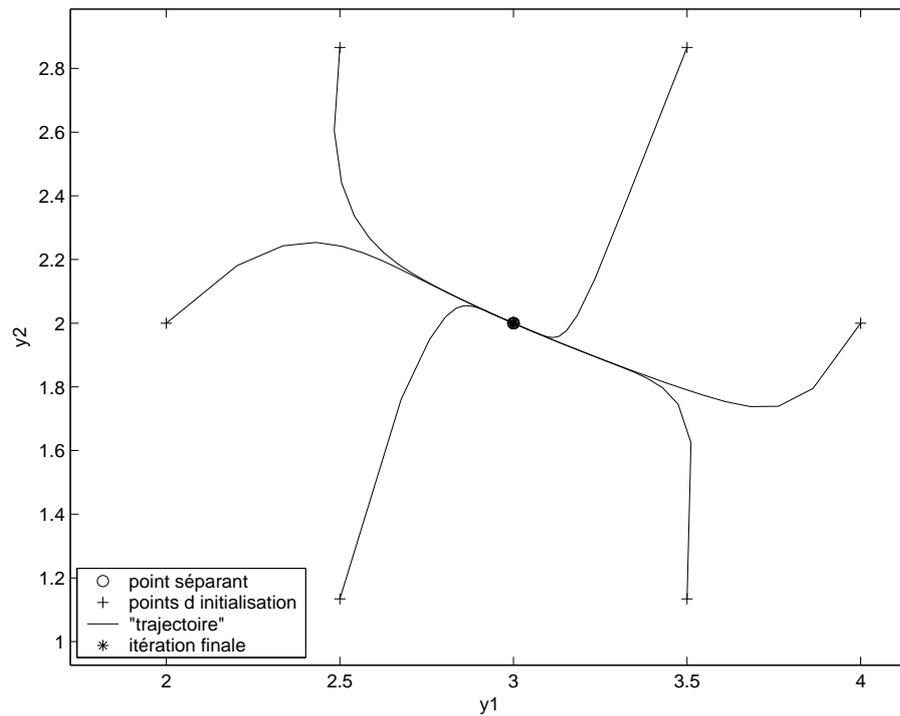


FIGURE 2.8 – Réseau étendu : Stabilisation locale du point séparant $s_1 = 1, s_2 = 1$, $L_{12} = -0.2, L_{21} = -0.1, Q_1 = -0.9, Q_2 = -2.1, k_1 = 3, k_2 = 2, M = 40, r = 0.1, K = 6$.

quelles que soient les valeurs des sources (voir Annexe A). Dans ce cas nous avons les résultats théoriques suivants :

- Les valeurs de gauche du système (2.41) obtenues sont les suivantes :
 $[-0.77, -2.2250, -1.275]$, il n'est donc pas vérifié pour la première inéquation,
- ce résultat est confirmé par le module des valeurs propres au point séparant
 $|\lambda_1| = 0.1652, |\lambda_2| = 1.6652 > 1$.

Pour la figure 2.8 les valeurs des sources ont été fixées à 1, la structure étendue ($k_1 = 3$ et $k_2 = 2$) converge vers le nouveau point séparant (3, 2) :

- Les valeurs de gauche du système (2.41) obtenues sont les suivantes :
 $[2.7163, -0.0635, -0.6101]$, le système est donc vérifié,
- ce résultat est confirmé par le module des valeurs propres au point séparant
 $|\lambda_1| = 0.8867, |\lambda_2| = 0.4397$.

Ces résultats confirment l'intérêt de la version étendue du réseau.

2.8 Conclusion

Nous avons considéré dans ce chapitre les modèles linéaire-quadratique et quadratique. Après avoir présenté à nouveau les 2 mélanges, nous avons détaillé les structures récurrentes associées. Pour le modèle quadratique nous avons défini les points d'équilibre du réseau. Il existe au moins un point d'équilibre correspondant aux sources à un facteur d'échelle près obtenu par construction. Nous avons ensuite, déterminé un autre point d'équilibre correspondant alors aux poids du réseau. Il pourrait exister d'autres points d'équilibre pour ce réseau (par exemple des polynômes d'un certain degré, dont la variable correspondrait à une source), mais nous ne les avons pas cherchés. Nous avons également étudié les limites de la stabilité locale des points d'équilibre, bien que l'apport de la structure étendue ne soit pas remis en cause, comme nous le montrent les tests effectués et exposés dans la partie « Simulations ». Auparavant, un comparatif entre les 2 modèles de mélange a été établi. Nous confirmons l'intérêt de la structure récurrente proposée pour le point d'équilibre séparant, via l'introduction de degrés de liberté supplémentaires. En effet, celle-ci permet dans certains cas de stabiliser localement le réseau à un instant

n . Il semblerait cependant que son application soit limitée à un certain domaine de valeurs de coefficients de mélanges.

Une procédure qui permettrait une sélection automatique des $k_{i,i \in \{1,2\}}$ peut-être considérée comme une perspective à apporter à ces travaux.

Nous allons maintenant passer au chapitre suivant où une extension au contexte aveugle, par le biais d'une méthode permettant l'estimation des coefficients du mélange en utilisant les cumulants, est proposée.

Chapitre 3

Identification du mélange par une méthode fondée sur des cumulants

Sommaire

3.1	Introduction	54
3.2	Travaux antérieurs	54
3.3	Approche proposée et expressions des cumulants des observations	55
3.3.1	Propriétés des cumulants et notations	56
3.3.2	Cumulants d'ordre 1	57
3.3.3	Cumulants d'ordre 2	57
3.3.4	Cumulants d'ordre 3	59
3.3.5	Cumulants d'ordre 4	63
3.4	Estimation des paramètres du mélange	64
3.4.1	Estimation des cumulants des observations	65
3.4.2	Résolution du système	65
3.5	Simulations	66
3.5.1	Premier scénario	66
3.5.2	Deuxième scénario	68
3.5.3	Troisième scénario	69
3.6	Conclusion	70

3.1 Introduction

Dans le chapitre 2 nous avons présenté la structure de séparation retenue pour cette étude, de type réseau récurrent, et nous avons analysé ses propriétés. Nous avons supposé que les valeurs des paramètres du mélange étaient connues. Dans ce chapitre, nous proposons une méthode qui permet justement l'estimation de ces coefficients à partir des observations. L'approche proposée est basée sur des cumulants. Nous proposons une estimation directe des coefficients de mélange dans le problème de SAS, en utilisant un système d'équations non-linéaires obtenu à partir des cumulants des observations. Ce type d'approche avait déjà été utilisé dans le cas du modèle linéaire, et nous montrerons que le passage au cas non-linéaire est plus complexe. Après avoir déduit les équations des cumulants des observations, nous les résoudrons numériquement, par le biais de différents algorithmes. Les coefficients ainsi obtenus peuvent être ensuite réinjectés dans des systèmes de séparation tels que les structures présentées précédemment. Des simulations numériques viendront compléter ce chapitre.

3.2 Travaux antérieurs

Pour estimer les coefficients du mélange, le recours aux *cumulants des sources et des observations* a été utilisé à diverses reprises, dans le cadre du modèle linéaire, et ce, par différents auteurs. Ainsi, on peut citer les travaux de A. Mansour [54, 56] ou de Y. Deville [24, 25]. Les développements effectués dans ce chapitre s'inscrivent dans la continuité de ces travaux, en effet, nous aurions pu déterminer notre système d'équations à partir des moments, mais nous avons préféré utiliser les cumulants du fait que ces derniers possèdent des propriétés intéressantes et que de nombreuses contributions existent dans ce domaine, vu leur utilisation très répandue.

D'autres travaux basés sur les cumulants ont été proposés et sont fondés sur les *cumulants des sorties du système de séparation*. Ainsi, on peut citer les algorithmes COM1, COM2 [21] et JADE [16]. Des extensions de l'algorithme JADE ont également été proposées, par exemple l'algorithme de A. Yeredor ACDC [85] qui a proposé une diagonalisation conjointe non orthogonale au sens des moindres carrés, et l'a appliqué à JADE. Nous pouvons également citer les extensions de JADE

(méthode ACDC) aux mélanges convolutifs [57].

3.3 Approche proposée et expressions des cumulants des observations

Nous rappelons que le modèle étudié est le suivant :

$$\begin{cases} x_1(n) = s_1^2(n) - L_{12}s_2^2(n) - Q_1s_1(n)s_2(n) \\ x_2(n) = s_2^2(n) - L_{21}s_1^2(n) - Q_2s_1(n)s_2(n) \end{cases} \quad (3.1)$$

où (s_1, s_2) représentent les sources normalisées et (x_1, x_2) les observations. Les coefficients $L_{ij, (i,j) \in \{(1,2), (2,1)\}}$ et $Q_{i, i \in \{1,2\}}$ représentent les coefficients du mélange, constants et réels.

Il s'agit d'estimer les paramètres du mélange de notre modèle. Pour ce, nous allons déterminer un système d'équations non-linéaires, obtenu à partir des expressions des cumulants des observations, et qui va dépendre des coefficients de mélange, et des moments des sources. Puis, à partir d'algorithmes numériques nous résolvons ce système en nous intéressant, plus particulièrement, aux paramètres du mélange. Les observations considérées dans (3.1) sont des combinaisons linéaires de fonctions puissance des sources. En utilisant alors les propriétés de multilinéarité des cumulants, les cumulants des observations peuvent donc être exprimés en fonction des cumulants des fonctions puissance des signaux sources, et des coefficients du mélange. Nous déduisons ces expressions, pour les cumulants des observations d'ordre un à quatre. De plus, certains cumulants associés aux sources sont nuls, car nous faisons les hypothèses que :

- les sources sont statistiquement indépendantes,
- les densités de probabilité des sources sont symétriques par rapport à 0, et elles sont donc centrées.

Afin de déterminer quels cumulants des sources considérés sont nuls, nous allons déduire leurs expressions en fonction des moments associés, en utilisant la formule classique liant cumulants et moments, et qui sera donnée ci-après. Ceci nous permet d'obtenir un système d'équations non-linéaires qui dépend des coefficients de mélange et des moments des sources.

3.3.1 Propriétés des cumulants et notations

Nous rappelons que le cumulants d'ordre r est lié aux moments d'ordre p , $p \leq r$, par la formule de Leonov et Shirayayev [63]

$$C[x_1, x_2, \dots, x_r] = \sum_{\pi} (-1)^{|\pi|-1} (|\pi| - 1)! \prod_{\beta \in \pi} E[\prod_{i \in \beta} x_i] \quad (3.2)$$

où π parcourt la liste de toutes les partitions de $\{1, \dots, r\}$, β parcourt la liste de tous les blocs de la partition π et $|\pi|$ est le nombre de parties dans la partition. $E[\cdot]$ est le symbole de l'espérance mathématique.

Nous allons également rappeler certaines propriétés des cumulants [61] :

PC1 : si λ_i , $i = 1, \dots, n$ sont des constantes, et x_i , $i = 1, \dots, n$ sont des variables aléatoires, alors

$$C(\lambda_1 x_1, \dots, \lambda_n x_n) = \prod_{i=1}^n \lambda_i C(x_1, \dots, x_n)$$

PC2 : Les cumulants sont symétriques vis à vis de leurs arguments, i.e.,

$$C(x_1, \dots, x_n) = C(x_{i_1}, \dots, x_{i_n})$$

où (i_1, \dots, i_n) est une permutation de $(1, \dots, n)$.

PC3 : Les cumulants sont additifs vis à vis de leurs arguments, c'est-à-dire que les cumulants de sommes sont égaux aux sommes des cumulants :

$$C(x_0 + y_0, z_1, \dots, z_n) = C(x_0, z_1, \dots, z_n) + C(y_0, z_1, \dots, z_n)$$

PC4 : Si un sous-ensemble de k variables aléatoires $\{x_i\}$ est indépendant du reste alors :

$$C(x_1, \dots, x_k, \dots, x_n) = 0$$

Les notations abrégées suivantes seront utilisées dans ce chapitre :

$C[(s_1^2)_2] = C[s_1^2, s_1^2]$ représente le cumulants d'ordre deux de $\{s_1^2\}$.

$C[(x_1)_2, x_2] = C[x_1, x_1, x_2]$ est le cumulants croisé d'ordre trois qui implique l'observation $\{x_1\}$ deux fois et l'observation $\{x_2\}$ une seule fois.

3.3.2 Cumulants d'ordre 1

Nous pouvons facilement déduire des équations (3.1), (3.2) et de l'hypothèse d'indépendance statistique des sources que le cumulants d'ordre un de la première observation est :

$$C[x_1] = E[x_1] \quad (3.3)$$

$$= E[s_1^2] - L_{12}E[s_2^2] - Q_1E[s_1]E[s_2]. \quad (3.4)$$

En utilisant l'hypothèse que les sources sont centrées, nous avons :

$$C[x_1] = E[s_1^2] - L_{12}E[s_2^2]. \quad (3.5)$$

De la même manière nous obtenons pour la deuxième observation :

$$C[x_2] = E[x_2] = E[s_2^2] - L_{21}E[s_1^2]. \quad (3.6)$$

3.3.3 Cumulants d'ordre 2

Cumulants des observations

Nous détaillons les calculs pour $C[(x_1)_2]$. En utilisant (3.1) et les propriétés des cumulants présentées ci-dessus nous avons :

$$\begin{aligned} C[(x_1)_2] &= C[x_1, x_1] \\ &= C[s_1^2 - L_{12}s_2^2 - Q_1s_1s_2, s_1^2 - L_{12}s_2^2 - Q_1s_1s_2], \\ &= C[s_1^2, s_1^2] - L_{12}C[s_1^2, s_2^2] - Q_1C[s_1^2, s_1s_2] \\ &\quad - L_{12}C[s_2^2, s_1^2] + L_{12}^2C[s_2^2, s_2^2] + L_{12}Q_1C[s_2^2, s_1s_2] \\ &\quad - Q_1C[s_1s_2, s_1^2] + L_{12}Q_1C[s_1s_2, s_2^2] + Q_1^2C[s_1s_2, s_1s_2]. \end{aligned}$$

Grâce aux propriétés des cumulants et en utilisant les équations (3.11) qui seront présentées ultérieurement, et correspondant à l'annulation de certains cumulants associés aux sources, nous obtenons finalement :

$$\begin{cases} C[(x_1)_2] = C[(s_1^2)_2] + L_{12}^2C[(s_2^2)_2] + Q_1^2C[(s_1s_2)_2], \\ C[(x_2)_2] = L_{21}^2C[(s_1^2)_2] + C[(s_2^2)_2] + Q_2^2C[(s_1s_2)_2], \\ C[x_1, x_2] = -L_{21}C[(s_1^2)_2] - L_{12}C[(s_2^2)_2] + Q_1Q_2C[(s_1s_2)_2]. \end{cases} \quad (3.7)$$

« **Cumulants associés aux sources** »

Nous commençons les calculs avec $C[(s_1^2)_2]$. À partir de l'équation (3.2), nous avons deux possibilités pour les partitions : $\{s_1^2, s_1^2\}$ et $\{\{s_1^2\}, \{s_1^2\}\}$, qui ont respectivement 1 et 2 éléments, donc en utilisant la formule (3.2) :

$$\begin{aligned} C[(s_1^2)_2] &= (-1)^{(1-1)}(1-1)!E[s_1^2 s_1^2] + \\ &\quad (-1)^{(2-1)}(2-1)!E[s_1^2]E[s_1^2] \\ &= E[s_1^4] - E[s_1^2]^2. \end{aligned} \quad (3.8)$$

Nous avons également

$$C[(s_2^2)_2] = E[s_2^4] - E[s_2^2]^2, \quad (3.9)$$

par des calculs similaires, en remplaçant uniquement s_1^2 par s_2^2 .

Pour $C[(s_1 s_2)_2]$, nous avons également 2 partitions, $\{s_1 s_2, s_1 s_2\}$ et $\{\{s_1 s_2\}, \{s_1 s_2\}\}$, donc :

$$\begin{aligned} C[(s_1 s_2)_2] &= (-1)^0 0! E[s_1 s_2 s_1 s_2] + (-1)^1 1! E[s_1 s_2] E[s_1 s_2] \\ &= E[s_1^2] E[s_2^2] - E[s_1]^2 E[s_2]^2 \\ &= E[s_1^2] E[s_2^2] \end{aligned} \quad (3.10)$$

où nous avons d'abord utilisé la propriété d'indépendance statistique des sources et ensuite, le fait que les sources soient centrées. Les autres cumulants associés aux sources sont nuls, c'est-à-dire :

$$C[s_1^2, s_2^2] = C[s_1^2, s_1 s_2] = C[s_2^2, s_1 s_2] = 0. \quad (3.11)$$

Ceci peut être démontré comme suit :

$$\begin{aligned} C[s_1^2, s_2^2] &= E[s_1^2 s_2^2] - E[s_1^2] E[s_2^2] \\ &= E[s_1^2] E[s_2^2] - E[s_1^2] E[s_2^2] = 0, \end{aligned} \quad (3.12)$$

en utilisant la propriété d'indépendance statistique, et

$$\begin{aligned} C[s_1^2, s_1 s_2] &= E[s_1^2 s_1 s_2] - E[s_1^2] E[s_1 s_2] \\ &= E[s_1^3] E[s_2] - E[s_1^2] E[s_1] E[s_2] = 0, \end{aligned} \quad (3.13)$$

où nous avons utilisé la propriété d'indépendance statistique et le fait que les sources présentent des densités symétriques. Un raisonnement similaire peut être effectué

pour $C[s_2^2, s_1 s_2]$, et nous obtenons $C[s_2^2, s_1 s_2] = 0$.

Les calculs qui vont suivre sont basés sur la même approche que celle présentée ci-dessus, mais ils sont plus complexes car ils nécessitent des statistiques d'ordre supérieur. Dans ce qui suit, nous adopterons le même schéma d'exposé : dans un premier temps nous donnons les expressions des cumulants des observations en fonction des cumulants associés aux sources, puis dans un deuxième temps nous donnons les expressions de ces cumulants associés aux sources en fonction des moments des sources.

3.3.4 Cumulants d'ordre 3

Cumulants des observations

Les cumulants des observations d'ordre trois sont :

$$\left\{ \begin{array}{l}
 C[(x_1)_3] = C[(s_1^2)_3] - L_{12}^3 C[(s_2^2)_3] \\
 \quad + 3Q_1^2 C[s_1^2, (s_1 s_2)_2] - 3L_{12} Q_1^2 C[s_2^2, (s_1 s_2)_2], \\
 C[(x_2)_3] = C[(s_2^2)_3] - L_{21}^3 C[(s_1^2)_3] \\
 \quad - 3L_{21} Q_2^2 C[s_1^2, (s_1 s_2)_2] + 3Q_2^2 C[s_2^2, (s_1 s_2)_2], \\
 C[(x_1)_2, x_2] = -L_{21} C[(s_1^2)_3] + L_{12}^2 C[(s_2^2)_3] \\
 \quad + (2Q_1 Q_2 - L_{21} Q_1^2) C[s_1^2, (s_1 s_2)_2] \\
 \quad + (Q_1^2 - 2L_{12} Q_1 Q_2) C[s_2^2, (s_1 s_2)_2], \\
 C[x_1, (x_2)_2] = L_{21}^2 C[(s_1^2)_3] - L_{12} C[(s_2^2)_3] \\
 \quad + (Q_2^2 - 2L_{21} Q_1 Q_2) C[s_1^2, (s_1 s_2)_2] \\
 \quad + (2Q_1 Q_2 - L_{12} Q_2^2) C[s_2^2, (s_1 s_2)_2].
 \end{array} \right. \quad (3.14)$$

Cumulants associés aux sources

Les cumulants associés aux sources peuvent être exprimés comme suit :

$$\begin{cases} C[(s_1^2)_3] = E[s_1^6] - 3E[s_1^4]E[s_1^2] + 2E[s_1^2]^3, \\ C[(s_2^2)_3] = E[s_2^6] - 3E[s_2^4]E[s_2^2] + 2E[s_2^2]^3, \\ C[s_1^2, (s_1s_2)_2] = E[s_2^2]E[s_1^4] - E[s_2^2]E[s_1^2]^2 \\ C[s_2^2, (s_1s_2)_2] = E[s_1^2]E[s_2^4] - E[s_1^2]E[s_2^2]^2. \end{cases} \quad (3.15)$$

Ces résultats ont été obtenus comme suit :

- pour $C[(s_1^2)_3]$, nous avons les partitions :

- * $\{s_1^2, s_1^2, s_1^2\}$ une fois,
- * $\{\{s_1^2\}, \{s_1^2\}, \{s_1^2\}\}$ une fois
- * et la partition $\{\{s_1^2\}, \{s_1^2, s_1^2\}\}$ trois fois.

Ces partitions contiennent respectivement 1, 3 et 2 éléments. Nous avons donc :

$$\begin{aligned} C[(s_1^2)_3] &= (-1)^0 0! E[s_1^2 s_1^2 s_1^2] + 3(-1)^1 1! E[s_1^2] E[s_1^2 s_1^2] + (-1)^2 (2)! E[s_1^2] E[s_1^2] E[s_1^2] \\ &= E[s_1^6] - 3E[s_1^2]E[s_1^4] + 2E[s_1^2]^3 \end{aligned} \quad (3.16)$$

- De la même manière on calcule $C[(s_2^2)_3]$, en remplaçant s_1^2 par s_2^2 .

- Pour $C[s_1^2, (s_1s_2)_2]$ nous avons les partitions :

- * $\{s_1^2, s_1s_2, s_1s_2\}$ une fois,
- * $\{\{s_1^2\}, \{s_1s_2\}, \{s_1s_2\}\}$ une fois,
- * la partition $\{\{s_1^2\}, \{s_1s_2, s_1s_2\}\}$ une fois,
- * et $\{\{s_1s_2\}, \{s_1^2, s_1s_2\}\}$ deux fois,

contenant respectivement 1, 3, 2 et 2 éléments. Nous obtenons alors :

$$\begin{aligned} C[s_1^2, (s_1s_2)_2] &= (-1)^0 0! E[s_1^2 s_1 s_2 s_1 s_2] + (-1)^2 2! E[s_1^2] E[s_1 s_2] E[s_1 s_2] \\ &\quad + (-1)^1 1! E[s_1^2] E[s_1 s_2 s_1 s_2] + 2(-1)^{2-1} (2-1)! E[s_1 s_2] E[s_1^2 s_1 s_2] \\ &= E[s_1^4 s_2^2] - E^2[s_1^2] E[s_2^2] \end{aligned} \quad (3.17)$$

car $E[s_1 s_2] = E[s_1]E[s_2] = 0$ puisque les sources sont statistiquement indépendantes et centrées, et on rappelle que $E[s_1 s_2 s_1 s_2] = E[s_1^2]E[s_2^2]$.

- Les calculs sont les mêmes pour $C[s_2^2, (s_1 s_2)_2]$, en remplaçant s_1^2 par s_2^2 dans les partitions à prendre en compte.

Les cumulants associés aux sources nuls pour cet ordre sont :

$$\begin{cases} C[(s_1 s_2)_3] = C[(s_1^2)_2, s_1 s_2] = C[(s_2^2)_2, s_1 s_2] = 0, \\ C[s_1^2, (s_2^2)_2] = C[s_2^2, (s_1^2)_2] = C[s_1^2, s_2^2, s_1 s_2] = 0. \end{cases} \quad (3.18)$$

Le détail des calculs est le suivant :

- Pour $C[(s_1 s_2)_3]$ nous reprenons les calculs de $C[(s_1^2)_3]$ en remplaçant s_1^2 par $s_1 s_2$. Nous avons alors :

$$\begin{aligned} C[(s_1 s_2)_3] &= (-1)^0 0! E[s_1 s_2 s_1 s_2 s_1 s_2] + 3(-1)^1 1! E[s_1 s_2] E[s_1 s_2 s_1 s_2] \\ &\quad + (-1)^2 2! E[s_1 s_2] E[s_1 s_2] E[s_1 s_2] \\ &= E[s_1^3 s_2^3] - 3E[s_1 s_2] E[s_1^2 s_2^2] + 2E^3[s_1 s_2] \\ &= 0 \end{aligned} \quad (3.19)$$

car $E[s_1 s_2] = E[s_1]E[s_2] = 0$, puisque les sources sont indépendantes et centrées, $E[s_1^3 s_2^3] = E[s_1^3]E[s_2^3]$, $E[s_1^3] = E[s_2^3] = 0$ car les sources sont statistiquement indépendantes et de densité de probabilité symétrique.

- Pour $C[(s_1^2)_2, s_1 s_2]$, nous avons les partitions

- * $\{s_1^2, s_1^2, s_1 s_2\}$ une fois,
- * $\{\{s_1^2\}, \{s_1^2\}, \{s_1 s_2\}\}$ une fois,
- * $\{\{s_1^2\}, \{s_1^2, s_1 s_2\}\}$ deux fois
- * et $\{\{s_1^2, s_1^2\}, \{s_1 s_2\}\}$ une fois.

Les nombres d'éléments de ces partitions sont respectivement 1, 3, 2 et 2. On en déduit :

$$\begin{aligned} C[(s_1^2)_2, s_1 s_2] &= (-1)^0 0! E[s_1^2 s_1^2 s_1 s_2] + (-1)^2 2! E[s_1^2] E[s_1^2] E[s_1 s_2] \\ &\quad + 2(-1)^1 1! E[s_1^2] E[s_1^2 s_1 s_2] + (-1)^1 1! E[s_1 s_2] E[s_1^2 s_1^2] \\ &= E[s_1^5 s_2] + 2E^2[s_1^2] E[s_1] E[s_2] + 2E[s_1^2] E[s_1^3] E[s_2] - E[s_1] E[s_2] E[s_1^4] \\ &= 0 \end{aligned} \quad (3.20)$$

car $E[s_1^5 s_2] = E[s_1^5]E[s_2] = 0$ car les sources sont indépendantes et centrées ($E[s_2] = 0$).

- Pour $C[(s_2^2)_2, s_1 s_2]$ nous raisonnons de la même manière que précédemment en permutant s_1 et s_2 .
- Le cumulante $C[s_1^2, (s_2^2)_2]$ a pour partitions
 - * $\{s_1^2, s_2^2, s_2^2\}$ une fois,
 - * $\{\{s_1^2\}, \{s_2^2\}, \{s_2^2\}\}$ une fois,
 - * $\{\{s_2^2\}, \{s_1^2, s_2^2\}\}$ deux fois
 - * et $\{\{s_2^2, s_2^2\}, \{s_1^2\}\}$ une fois.

Le nombre d'éléments de ces partitions sont respectivement 1, 3, 2 et 2. Nous avons donc :

$$\begin{aligned}
 C[s_1^2, (s_2^2)_2] &= (-1)^0 0! E[s_1^2 s_2^2 s_2^2] + (-1)^2 2! E[s_1^2] E[s_2^2] E[s_2^2] \\
 &\quad + 2(-1)^1 1! E[s_2^2] E[s_1^2 s_2^2] + (-1)^1 1! E[s_1^2] E[s_2^2 s_2^2] \\
 &= E[s_1^2 s_2^4] + 2E[s_1^2] E^2[s_2^2] - 2E^2[s_2^2] E[s_1^2] - E[s_1^2] E[s_2^4] \\
 &= 0
 \end{aligned} \tag{3.21}$$

car $E[s_1^2 s_2^4] = E[s_1^2] E[s_2^4]$ en tenant compte de l'hypothèse d'indépendance statistique des sources.

- Pour $C[s_2^2, (s_1^2)_2]$ on suit le même raisonnement que ci-dessus, en intervertissant s_1 et s_2 .
- Enfin pour $C[s_1^2, s_2^2, s_1 s_2]$, nous avons les partitions $\{s_1^2, s_2^2, s_1 s_2\}$, $\{\{s_1^2\}, \{s_2^2\}, \{s_1 s_2\}\}$, $\{\{s_1^2\}, \{s_2^2, s_1 s_2\}\}$, $\{\{s_2^2\}, \{s_1^2, s_1 s_2\}\}$ et $\{\{s_1^2, s_2^2\}, \{s_1 s_2\}\}$ une fois pour chacune d'entre elles. Elles ont respectivement 1, 3, 2, 2 et 2 éléments. Il en découle :

$$\begin{aligned}
 C[s_1^2, s_2^2, s_1 s_2] &= (-1)^0 0! E[s_1^2 s_2^2 s_1 s_2] + (-1)^2 2! E[s_1^2] E[s_2^2] E[s_1 s_2] \\
 &\quad - E[s_1^2] E[s_2^2 s_1 s_2] - E[s_2^2] E[s_1^2 s_1 s_2] - E[s_1 s_2] E[s_2^2 s_1^2] \\
 &= E[s_1^3 s_2^3] + 2E[s_1^2] E[s_2^2] E[s_1] E[s_2] - 3E[s_1^2] E[s_2^2] E[s_1] E[s_2] \\
 &= 0
 \end{aligned} \tag{3.22}$$

puisque $E[s_1^3 s_2^3] = E[s_1^3]E[s_2^3] = 0$ comme les sources sont statistiquement indépendantes et de densité de probabilité symétrique, donc centrées.

3.3.5 Cumulants d'ordre 4

Nous ne détaillerons pas ici les calculs pour l'ordre quatre. Ces derniers peuvent être trouvés en Annexe B, où les calculs des cumulants associés aux sources sont disponibles. Les cumulants des observations d'ordre quatre donnent le système d'équations non-linéaires suivant :

$$\left\{ \begin{array}{l}
 C[(x_1)_4] = C[(s_1^2)_4] + L_{12}^4 C[(s_2^2)_4] + Q_1^4 C[(s_1 s_2)_4] \\
 \quad + 6Q_1^2 C[(s_1^2)_2, (s_1 s_2)_2] + 6L_{12}^2 Q_1^2 C[(s_2^2)_2, (s_1 s_2)_2] \\
 \quad - 12L_{12} Q_1^2 C[s_1^2, s_2^2, (s_1 s_2)_2], \\
 C[(x_2)_4] = L_{21}^4 C[(s_1^2)_4] + C[(s_2^2)_4] + Q_2^4 C[(s_1 s_2)_4] \\
 \quad + 6L_{21}^2 Q_2^2 C[(s_1^2)_2, (s_1 s_2)_2] + 6Q_2^2 C[(s_2^2)_2, (s_1 s_2)_2] \\
 \quad - 12L_{21} Q_2^2 C[s_1^2, s_2^2, (s_1 s_2)_2], \\
 C[(x_1), (x_2)_3] = -L_{21}^3 C[(s_1^2)_4] - L_{12} C[(s_2^2)_4] + Q_1 Q_2^3 C[(s_1 s_2)_4] \\
 \quad + (-3Q_2^2 L_{21} + 3Q_1 Q_2 L_{21}^2) C[(s_1^2)_2, (s_1 s_2)_2] \\
 \quad + (-3L_{12} Q_2^2 + 3Q_1 Q_2) C[(s_2^2)_2, (s_1 s_2)_2] \\
 \quad + (3Q_2^2 - 6L_{21} Q_2 Q_1 + 3L_{12} L_{21} Q_2^2) C[s_1^2, s_2^2, (s_1 s_2)_2], \tag{3.23} \\
 C[(x_2), (x_1)_3] = -L_{21} C[(s_1^2)_4] - L_{12}^3 C[(s_2^2)_4] + Q_2 Q_1^3 C[(s_1 s_2)_4] \\
 \quad + (-3L_{21} Q_1^2 + 3Q_2 Q_1) C[(s_1^2)_2, (s_1 s_2)_2] \\
 \quad + (-3L_{12} Q_1^2 + 3Q_2 Q_1 L_{12}^2) C[(s_2^2)_2, (s_1 s_2)_2] \\
 \quad + (3Q_1^2 - 6L_{12} Q_2 Q_1 + 3L_{12} L_{21} Q_1^2) C[s_1^2, s_2^2, (s_1 s_2)_2], \\
 C[(x_1)_2, (x_2)_2] = L_{21}^2 C[(s_1^2)_4] + L_{12}^2 C[(s_2^2)_4] + Q_2^2 Q_1^2 C[(s_1 s_2)_4] \\
 \quad + (L_{21}^2 Q_1^2 + Q_2^2 - 4Q_1 Q_2 L_{21}) C[(s_1^2)_2, (s_1 s_2)_2] \\
 \quad + (L_{12}^2 Q_2^2 + Q_1^2 - 4L_{12} Q_1 Q_2) C[(s_2^2)_2, (s_1 s_2)_2] \\
 \quad + (4Q_1 Q_2 - 2L_{12} Q_2^2 - 2L_{21} Q_1^2 \\
 \quad \quad + 4Q_1 Q_2 L_{12} L_{21}) C[s_1^2, s_2^2, (s_1 s_2)_2].
 \end{array} \right.$$

et les cumulants associés aux sources d'ordre quatre s'écrivent :

$$\left\{ \begin{array}{l} C[(s_1^2)_4] = E[s_1^8] - 4E[s_1^6]E[s_1^2] - 6E[s_1^4]^2 - 3E[s_1^4]^2 + 12E[s_1^4]E[s_1^2]^2, \\ C[(s_2^2)_4] = E[s_2^8] - 4E[s_2^6]E[s_2^2] - 6E[s_2^4]^2 - 3E[s_2^4]^2 + 12E[s_2^4]E[s_2^2]^2, \\ C[(s_1s_2)_4] = E[s_1^4]E[s_2^4] - 3E[s_1^2]^2E[s_2^2]^2, \\ C[(s_1^2)_2, (s_1s_2)_2] = E[s_1^6]E[s_2^2] - E[s_1^2]E[s_2^2](3E[s_1^4] - 2E[s_1^2]^2), \\ C[(s_2^2)_2, (s_1s_2)_2] = E[s_2^6]E[s_1^2] - E[s_2^2]E[s_1^2](3E[s_2^4] - 2E[s_2^2]^2), \\ C[s_1^2, s_2^2, (s_1s_2)_2] = E[s_1^4]E[s_2^4] - E[s_1^2]^2E[s_2^4] - E[s_2^2]^2E[s_1^4] + E[s_1^2]^2E[s_2^2]^2. \end{array} \right. \quad (3.24)$$

$$\left\{ \begin{array}{l} C[(s_1^2)_3, s_2^2] = C[(s_2^2)_3, s_1^2] = C[(s_1s_2)_3, s_1^2] = 0. \\ C[(s_1s_2)_3, s_2^2] = C[(s_1^2)_3, s_1s_2] = C[(s_2^2)_3, s_1s_2] = 0. \\ C[(s_1^2), s_1s_2, (s_2^2)_2] = C[s_2^2, s_1s_2, (s_1^2)_2] = C[(s_1^2)_2, (s_2^2)_2] = 0. \end{array} \right. \quad (3.25)$$

Les systèmes (3.5), (3.6), (3.7), (3.14) et (3.23) forment ensemble le système final d'équations à résoudre, qui est fortement non-linéaire par rapport aux L_{ij} , Q_i et aux moments des sources. Nous obtenons alors un système de 14 équations, qui dépend de 12 inconnues ($L_{12}, L_{21}, Q_1, Q_2, E[s_1^2], E[s_1^4], E[s_1^6], E[s_1^8], E[s_2^2], E[s_2^4], E[s_2^6], E[s_2^8]$) mais nous ne serons finalement intéressés que par les coefficients de mélange $\{L_{12}, L_{21}, Q_1, Q_2\}$. Il faut remarquer que si nous n'utilisons pas les cumulants d'ordre quatre, nous obtenons un système de 9 équations avec 10 inconnues (voir le Tableau 4.1), qui ne peut être résolu sans contraintes supplémentaires sur les sources. Nous pouvons également avoir recours à des cumulants d'ordre supérieur à quatre ce qui conduit à avoir plus d'équations et ... plus d'inconnues.

Nous proposons alors de résoudre le système ci-dessus par le biais de méthodes d'optimisation numériques qui vont être définies dans ce qui suit.

3.4 Estimation des paramètres du mélange

Ayant obtenu le système d'équations défini ci-dessus, nous utilisons des algorithmes d'optimisation pour le résoudre (voir Annexe C), le but étant d'estimer les paramètres de mélange (L_{12}, L_{21}, Q_1, Q_2). Tout d'abord, il est nécessaire d'estimer les cumulants des observations aux différents ordres avant de résoudre le système en utilisant des méthodes numériques.

Ordre	Cumulants des observations (nombre)	Inconnues rajoutées
1	$C[x_1], C[x_2]$ (2)	$L_{12}, L_{21},$ $E[s_1^2], E[s_2^2]$
2	$C[(x_1)_2], C[(x_2)_2],$ $C[x_1, x_2]$ (3)	Q_1, Q_2 $E[s_1^4], E[s_2^4]$
3	$C[(x_1)_3], C[(x_2)_3],$ $C[(x_1)_2, x_2], C[x_1, (x_2)_2]$ (4)	$E[s_1^6], E[s_2^6]$
4	$C[(x_1)_4], C[(x_2)_4], C[(x_1), (x_2)_3],$ $C[(x_2), (x_1)_3], C[(x_1)_2, (x_2)_2]$ (5)	$E[s_1^8], E[s_2^8]$

TABLE 3.1 – Nombre d'équations et d'inconnues pour notre système en fonction de l'ordre des cumulants.

3.4.1 Estimation des cumulants des observations

Il existe différents estimateurs des auto-cumulants et des cumulants conjoints dans la littérature [55], [8]. Malheureusement, ces estimateurs ne s'appliquent qu'à des signaux centrés, ce qui n'est pas le cas des cumulants des observations (voir Équations (3.5) et (3.6)). Nous avons donc utilisé les fonctions proposées par la Toolbox HOSA [76], i.e. *cumest*, pour estimer les auto-cumulants ainsi que *cum2x*, *cum3x* et *cum4x* pour estimer les cumulants conjoints des observations (voir Annexe D).

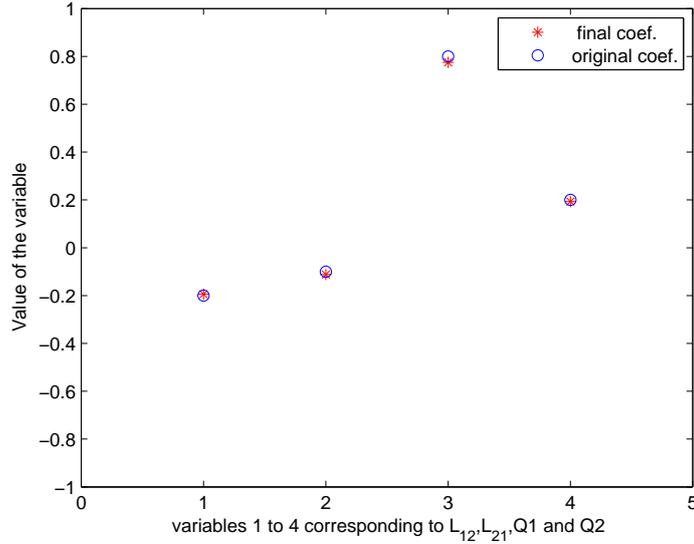
3.4.2 Résolution du système

Dans cette section, nous présentons les algorithmes d'optimisation utilisés afin de résoudre le système d'équations non-linéaires. Nous utilisons la fonction *fsolve* fournie par la toolbox MATLAB « Optimization toolbox » [20] (voir Annexe D), qui résout les systèmes d'équations non-linéaires de la forme :

$$F(v) = 0,$$

$$\text{avec } \begin{cases} v = [v_1, \dots, v_m]^T \\ F(v) = [f_1(v), \dots, f_n(v)]^T, \end{cases}$$

v_i , $i = 1..m$ sont les variables (i.e. les 12 inconnues dans notre cas) et f_i , $i = 1..n$ sont les fonctions non-linéaires. Donc pour obtenir cette forme avec un membre de


 FIGURE 3.1 – Convergence vers "S" pour les paramètres $\{L_{12}, L_{21}, Q_1, Q_2\}$

droite des équations égal à zéro, nous réarrangeons le système formé de $\{(3.5), (3.6), (3.7), (3.14), (3.23)\}$, en déplaçant tous les termes du membre de droite, à gauche. La fonction *fsolve* est utilisée suivant 2 configurations : la première fait appel à l'algorithme de Levenberg-Marquardt, tandis que la seconde utilise l'algorithme de Gauss-Newton (voir Annexe C).

3.5 Simulations

3.5.1 Premier scénario

Pour ce premier scénario, nous considérons l'ensemble des paramètres $\{L_{12}, L_{21}, Q_1, Q_2\} = \{-0.2, -0.1, 0.8, 0.2\}$, et les sources de longueur $N = 10000$ échantillons sont uniformément distribuées sur l'intervalle $[-1, 1]$. La fonction *fsolve* retourne en sortie un indicateur « EXITFLAG » qui décrit si l'algorithme converge ou non. Pour chaque configuration de la fonction *fsolve*, nous effectuons 450 tests. Pour chaque test, nous utilisons l'indicateur EXITFLAG pour réinitialiser la variable $v = [L_{12}, L_{21}, Q_1, Q_2, E[s_1^2], E[s_1^4], E[s_1^6], E[s_1^8], E[s_2^2], E[s_2^4], E[s_2^6], E[s_2^8]]$ de la fonction *fsolve* (les initialisations sont différentes et générées aléatoirement) si l'algorithme ne converge pas, chaque test étant effectué au plus 150 fois. Pour la configuration suivant Levenberg-Marquardt, l'algorithme converge dans 95% des tests

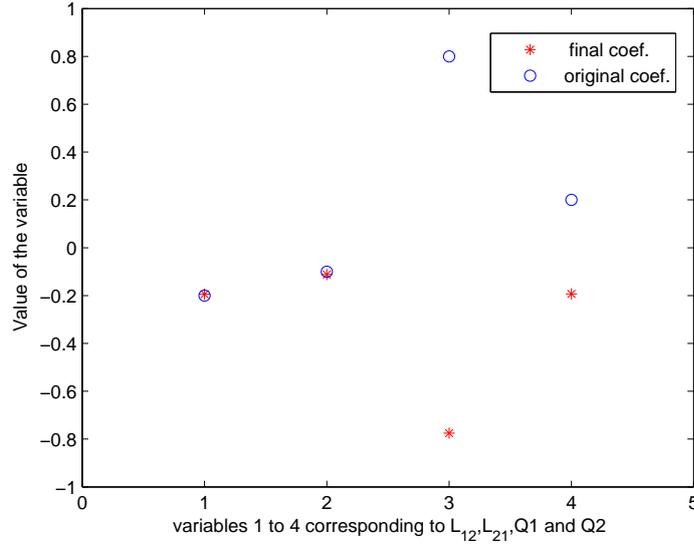


FIGURE 3.2 – Phénomène d’inversion de signe pour les paramètres Q_1 et Q_2

(dans le cas contraire, l’algorithme diverge), dont environ 35% (soit 161 tests) vers le point d’intérêt, appelé S , le reste correspondant à d’autres points solutions de notre système d’équations (voir Tableau 3.2).

Nous avons observé un phénomène d’inversion de signes pour les paramètres estimés Q_1 et Q_2 , comme l’illustre la figure 3.2. Ceci est dû à la manière dont ils apparaissent dans notre système, i.e. si (Q_1, Q_2) est solution du système, alors $(-Q_1, -Q_2)$ est aussi solution.

Pour la configuration de Gauss-Newton, nous obtenons 91% de cas de convergence en général et 37% (167 tests) de convergence vers le point S . De meilleurs résultats de convergence sont obtenus dans ce cas, car moins de points parasites sont rencontrés en nombre suffisant dans nos tests (voir le tableau 3.2). Dans cette configuration, nous retrouvons également le phénomène d’inversion cité plus haut. Ainsi, pour les deux configurations dans ce scénario, l’algorithme converge le plus souvent vers le point d’intérêt, comparativement aux autres points de convergence. Le tableau 3.2 montre que ces points parasites peuvent être éliminés en tenant compte du fait que $L_{12}, L_{21} < 0$, car cette condition n’est pas vérifiée pour les valeurs de L_{12} et L_{21} estimées en ces points.

Algo.	Point de convergence $\{L_{12}, L_{21}, Q_1, Q_2\}$		Nb.
L. M.	S	$-0.183, -0.089, -0.758, -0.19728$	161
	Pt 1	$1.684, -0.210, -0.217, -0.723$	62
	Pt 2	$0.162, 0.099, 1.373, 0.091$	53
	Pt 3	$1.760, -0.207, -0.228, 0.715$	47
	Pt 4	$-0.192, 0.092, -1.487, -0.076$	43
G. N.	S	$-0.1747, -0.110, 0.778, 0.209$	167
	Pt 5	$-0.318, 0.468, 0.881, 0.619$	58
	Pt 6	$-0.130, 0.070, -1.546, -0.0944$	49
	Pt 7	$6.362, -0.353, 0.131, 1.152$	25

TABLE 3.2 – Nombre d’occurrence de la convergence vers le point d’intérêt S et les points parasites pour les algorithmes de Levenberg-Marquardt (LM) et Gauss-Newton (GN).

3.5.2 Deuxième scénario

Pour ce scénario, nous considérons l’ensemble des paramètres $\{L_{12}, L_{21}, Q_1, Q_2\} = \{-0.9, -0.75, 0.2, 0.1\}$, et les sources de longueur $N = 10000$ échantillons suivent respectivement, la loi uniforme sur $[-1, 1]$ et la loi normale $\mathcal{N}(0, 0.2)$. Pour chaque configuration de la fonction f_{solve} , nous considérons 450 tests avec les mêmes conditions que celles mentionnées précédemment.

Pour la configuration de Levenberg-Marquardt, l’algorithme converge dans 100% des tests, dont 10.5% (43 tests) vers le point d’intérêt, appelé S' , le reste correspondant à d’autres solutions de notre système.

Pour la configuration de Gauss-Newton, nous obtenons également 100% de cas de convergence en général et approximativement 21% de convergence vers S' (94 tests).

Dans ce scénario, nous avons également observé le phénomène d’inversion des signes.

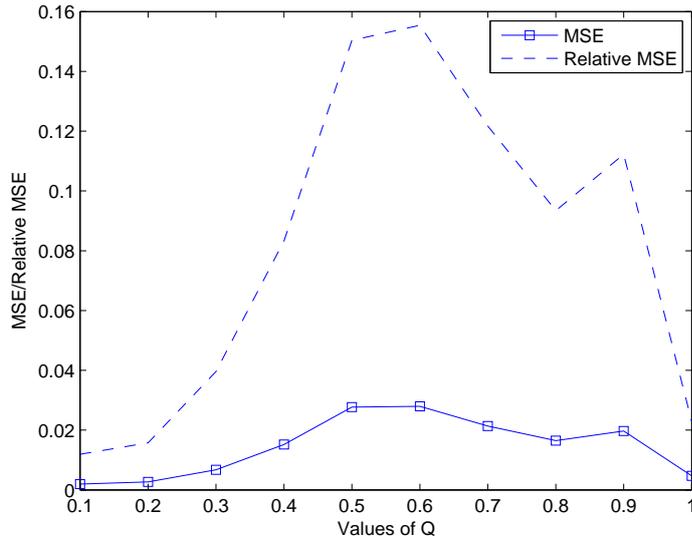


FIGURE 3.3 – Valeur moyenne pour 100 réalisations des sources de l’EQM en fonction de Q pour les 12 paramètres.

3.5.3 Troisième scénario

Nous avons généré des réalisations de Monte-Carlo, en utilisant 100 réalisations aléatoires des sources s_1 et s_2 de longueur $N = 10000$ et distribuées uniformément sur $[-1, 1]$. Les réalisations sont indépendantes. Les paramètres L_{12}, L_{21} , sont respectivement fixés à -0.8 et -0.9 , et $Q_1 = Q_2 = Q$ où Q varie entre 0.1 et 1 avec un pas de 0.1. Alors, pour chaque réalisation de nos mélanges correspondant à une valeur du paramètre Q , nous faisons les expériences pour toutes les sources en utilisant la procédure décrite dans le premier scénario et la fonction `fsolve` n’est utilisée qu’avec la configuration de l’algorithme de Levenberg-Marquardt.

À cause du phénomène d’inversion, l’erreur et l’erreur relative sont calculées comme suit :

$$\epsilon_i = |\hat{v}_i| - |v_i|, \quad i = 1..12 \quad (3.26)$$

$$\epsilon_{rel,i} = \frac{|\hat{v}_i| - |v_i|}{|v_i|}, \quad i = 1..12 \quad (3.27)$$

où i est l’indice des variables, \hat{v} contient les variables à la convergence et v est notre point d’intérêt.

L’*Erreur Quadratique Moyenne (EQM)* associée et *EQM relative* sont respecti-

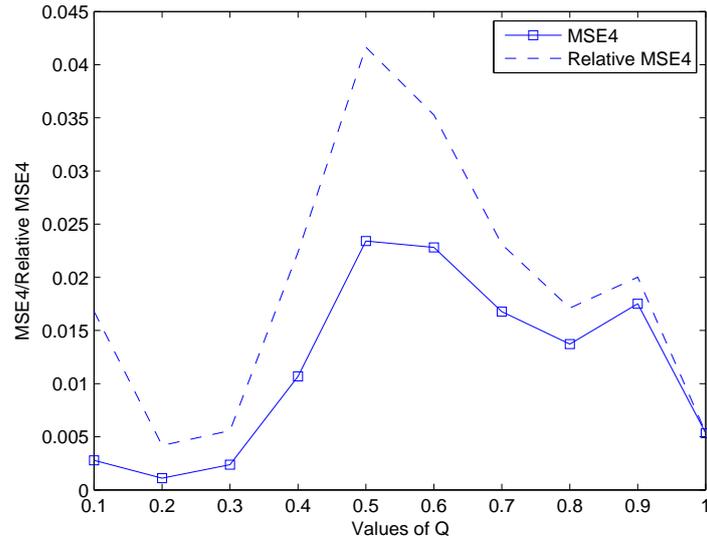


FIGURE 3.4 – Valeur moyenne de l’EQM pour 100 réalisations des sources en fonction de Q calculée pour les paramètres $\{L_{12}, L_{21}, Q_1, Q_2\}$.

vement :

$$EQM_n = \frac{1}{n} \sum_{i=1}^n (|\hat{v}_i| - |v_i|)^2 = \frac{1}{n} \sum_{i=1}^n \epsilon_i^2, \quad (3.28)$$

$$EQM_{rel,n} = \frac{1}{n} \sum_{i=1}^n \left(\frac{|\hat{v}_i| - |v_i|}{|v_i|} \right)^2 = \frac{1}{n} \sum_{i=1}^n \epsilon_{rel,i}^2, \quad (3.29)$$

où l’indice n est mis à 12 et 4 lorsque nous calculons ces critères respectivement pour l’ensemble des paramètres ou uniquement pour les 4 paramètres L_{12}, L_{21}, Q_1, Q_2 car les autres paramètres ne nous importent pas au final. Nous avons recours à des quantités relatives, car elles sont plus appropriées lorsque nous additionnons plusieurs variables.

Les figures 3.3 et 3.4 montrent les valeurs moyennes de EQM_n and $EQM_{rel,n}$ sur les 100 ensembles de sources pour les différentes valeurs du paramètre Q . Le graphe de la figure 3.4 montre que la moyenne de l’EQM relative est inférieure ou égale à 4%, lorsque nous considérons seulement les 4 paramètres d’intérêt, ce qui prouve l’efficacité de notre méthode.

3.6 Conclusion

Dans ce chapitre nous avons présenté une méthode afin d’estimer les paramètres du mélange. Cette méthode fait appel aux cumulants des observations, dont nous

avons déduit un système d'équations non-linéaires que nous résolvons à l'aide d'algorithmes d'optimisation numérique, les méthodes de Levenberg-Marquardt et Gauss-Newton. Il reste à noter que pour regrouper ce bloc placé en amont des travaux exposés dans le chapitre 2, ces équations doivent être modifiées car certaines de nos hypothèses ne sont plus valables. En effet, les sources ne doivent plus être supposées à densité symétriques, et donc elles seront non centrées (condition pour respecter l'inversibilité du modèle quadratique). Ceci alourdit considérablement nos développements (plus d'équations et d'inconnues à considérer), qui même dans le cas d'hypothèses simplificatrices restent lourds (se référer aux détails des calculs).

Chapitre 4

Approche par maximum de vraisemblance

Sommaire

4.1	Introduction	74
4.2	Travaux antérieurs	74
4.2.1	Généralités	74
4.2.2	Modèles de densités des sources	75
4.3	Estimation des coefficients du mélange : approche par maximum de vraisemblance (MV)	76
4.3.1	Modèle de mélange	76
4.3.2	Critère d'estimation des coefficients du mélange : principe général	77
4.3.3	Détails des calculs du gradient de la log-vraisemblance	79
4.3.4	Algorithme d'optimisation	81
4.4	Structure récurrente associée	81
4.5	Simulations	83
4.6	Conclusions	90

4.1 Introduction

Dans ce chapitre nous proposons une autre approche pour séparer les sources. Pour cette méthode, nous considérons la version basique du réseau récurrent associé à notre modèle quadratique introduite au Chapitre 2, et, une approche par Maximum de Vraisemblance (MV) est utilisée afin d'estimer les coefficients du mélange. Nous présentons un algorithme d'optimisation itératif qui nécessite l'estimation des fonctions score des signaux sources. Nous montrerons que l'algorithme final se décline suivant deux boucles imbriquées : la boucle principale permet d'estimer les coefficients du mélange, tandis que la boucle interne sert à l'estimation des signaux sources. Des tests et les résultats seront présentés en fin de chapitre.

4.2 Travaux antérieurs

4.2.1 Généralités

L'approche par MV est très populaire en SAS. Comme nous l'avions précisé au Chapitre 1, son but est de trouver les paramètres du mélange qui maximisent la probabilité d'occurrence des observations. On trouve de nombreuses contributions dans la littérature. Des approches basées sur le MV ont d'abord été proposées par M. Gaeta et J.-L. Lacoume [37, 38]. On peut également citer dans le cas linéaire les travaux de Pham et al. [72, 73], L. Parra et al. [67, 68] ou encore de A. Belouchrani et J.-F. Cardoso [6]. Le maximum de vraisemblance a été également utilisé pour les mélanges PNL par A. Taleb et C. Jutten [77] et pour des mélanges non-linéaires par Y. Deville et S. Hosseini [30, 41].

Le principal inconvénient de l'approche par MV est qu'elle met en jeu les densités de probabilité des sources (pdf), qui ne sont généralement pas connues. On peut alors supposer qu'elles sont connues, *a priori*, ou qu'elles appartiennent à une famille de densités donnée, les méthodes de MV font appel aux fonctions score des sources $g_i(\cdot) = -(\log(f_{S_i}(\cdot)))'$.

4.2.2 Modèles de densités des sources

Le vrai modèle des pdf des sources est généralement inconnu, nous avons alors recours à des modèles plus ou moins précis pour approximer les distributions des sources. Nous allons donc présenter certains de ces modèles utilisés pour représenter les distributions des sources.

Développement en série d'une densité

Il s'agit d'une méthode permettant d'approximer les densités en utilisant les cumulants d'ordre supérieur. Les principaux développements utilisés sont ceux en série de Gram-Charlier et d'Edgeworth. Ces deux développements ne diffèrent que par la façon d'ordonner les termes. Les formes tronquées de ces séries diffèrent alors à cause de ce changement d'ordre.

M. Gaeta et J.-L. Lacoume [37,38] ont utilisé le développement de Gram-Charlier, les premières fonctions caractéristiques des sources ont été exprimées par des séries tronquées à l'ordre 4. La distribution des sources s'écrit dans ce cas en fonction des cumulants d'ordre 3 et 4 suivant la relation [5] :

$$f_i(s_i) \approx \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{s_i^2}{2}\right) \left(1 + \frac{\kappa_3^i}{3!} H_3(s_i) + \frac{\kappa_4^i}{4!} H_4(s_i)\right) \quad (4.1)$$

où κ_3^i et κ_4^i désignent respectivement les cumulants normalisés d'ordre 3 et 4, et $H_k(s_i)$ les polynômes de Chebyshev-Hermite.

Le développement d'Edgeworth est très similaire, il a été utilisé par P. Comon [12] pour approximer la néguentropie en fonction des cumulants dont l'expression est :

$$J(y) \approx \frac{1}{12} E[y^3]^2 + \frac{1}{48} kurt(y)^2 \quad (4.2)$$

où y est une variable aléatoire centrée de variance unité et $kurt$ désigne le kurtosis.

À partir de ces développements, nous pouvons obtenir l'approximation des fonctions score, par exemple pour la fonction score obtenue par le développement d'Edgeworth de la distribution des sources au voisinage d'une distribution normale nous avons [15] :

$$\psi_i(s_i) \approx s_i - \frac{\kappa_4^i}{6} (s_i^3 - 3s_i) \quad (4.3)$$

Estimateurs à noyaux

Le modèle de distribution d'un estimateur à noyau peut s'exprimer suivant :

$$f_i(s_i(t)) = \frac{1}{N} \sum_{\tau=1}^N \frac{1}{h_N} \kappa \left(\frac{s_i(t) - s_i(\tau)}{h_N} \right) \quad (4.4)$$

où $\kappa(\cdot)$ est la fonction de noyau, h_N est la largeur d'intervalle, et N est le nombre d'échantillons du signal.

Un modèle à noyau gaussien a été proposé par D.-T. Pham dans [69]. Le noyau gaussien a été remplacé par des fonctions splines cardinales d'ordre 3 dans [70].

Familles de distributions paramétriques

L'exemple le plus utilisé est celui de la gaussienne généralisée qui s'écrit dans le cas univarié pour les sources centrées :

$$f_{\alpha,\beta}(s_i) = \frac{\beta}{2\alpha\Gamma(\frac{1}{\beta})} \exp \left(- \left(\frac{|s_i|}{\alpha} \right)^\beta \right) \quad (4.5)$$

où $\Gamma(y)$ est la fonction Gamma, $\alpha > 0$ est un facteur d'échelle qui régit l'étalement de la distribution et $\beta > 0$ est un paramètre qui définit la forme du pic.

On peut également citer les familles d'exponentielles généralisées [86], les distributions hyperboliques généralisées [71].

4.3 Estimation des coefficients du mélange : approche par maximum de vraisemblance (MV)

4.3.1 Modèle de mélange

Nous rappelons que le modèle considéré est un modèle aussi à CCA. Pour 2 sources et 2 observations, il s'écrit :

$$\begin{cases} x_1(n) = s_1^2(n) - L_{12}s_2^2(n) - Q_1s_1(n)s_2(n) \\ x_2(n) = s_2^2(n) - L_{21}s_1^2(n) - Q_2s_1(n)s_2(n) \end{cases} \quad (4.6)$$

où (s_1, s_2) représentent les sources normalisées, supposées statistiquement indépendantes, et (x_1, x_2) les observations. Les coefficients L_{ij} , $(i, j) \in \{(1, 2), (2, 1)\}$ et Q_i , $i \in \{1, 2\}$ représentent les coefficients du mélange, constants et réels.

Le modèle considéré sera noté :

$$\mathbf{x} = g(\mathbf{s}, \boldsymbol{\theta}_*) \quad (4.7)$$

où $\mathbf{x} = [x_1, x_2]^T$, $\mathbf{s} = [s_1, s_2]^T$ et $\boldsymbol{\theta}_* = [L_{12}, L_{21}, Q_1, Q_2]^T$.

4.3.2 Critère d'estimation des coefficients du mélange : principe général

L'algorithme de séparation nécessite l'estimation des coefficients du mélange. Dans cette section, nous présentons l'approche par MV [11, 23, 41, 73] utilisée pour estimer le vecteur des paramètres du mélange $\boldsymbol{\theta}_*$, inconnus, sous l'hypothèse que le modèle (4.6) est bijectif dans les domaines considérés des sources et des coefficients du mélange.

Le but de l'approche est d'identifier $\boldsymbol{\theta}_*$, en supposant que :

1. chaque source s_i est un signal i.i.d. Par conséquent, les signaux observés x_i sont également i.i.d.
2. Les sources sont statistiquement indépendantes, nous attribuons une densité hypothétique $f_{S_i}(s_i)$ à chacune des sources. Leur densité conjointe s'écrit alors :

$$f_{S_1, S_2}(s_1, s_2) = f_{S_1}(s_1)f_{S_2}(s_2). \quad (4.8)$$

Nous notons Θ l'ensemble des vecteurs de paramètres $\boldsymbol{\theta} = [l_{12}, l_{21}, q_1, q_2]^T$ tels que le modèle (4.7) soit bijectif sur le domaine de variation des sources. L'inverse de ce modèle pour chaque vecteur $\boldsymbol{\theta}$ sera noté par :

$$\hat{\mathbf{s}} = g^{-1}(\mathbf{x}, \boldsymbol{\theta}) \quad (4.9)$$

La distribution des sources étant fixée, la distribution de la transformation $g(\mathbf{s}, \boldsymbol{\theta})$ ne dépend que de $\boldsymbol{\theta}$. Nous utilisons cette famille de distributions comme un modèle paramétrique pour la pdf des observations et nous les notons par $P = \{p_{\boldsymbol{\theta}}(x_1(n), x_2(n)) | \boldsymbol{\theta} \in \Theta\}$.

Si nous considérons T échantillons des observations, la vraisemblance que ces échantillons soient tirés avec une pdf particulière $p_{\boldsymbol{\theta}}(\mathbf{x})$ est donnée par :

$$E = p_{\boldsymbol{\theta}}(x_1(1), x_2(1), \dots, x_1(T), x_2(T)) \quad (4.10)$$

$$= \prod_{n=1}^T p_{\boldsymbol{\theta}}(x_1(n), x_2(n)) \quad (4.11)$$

car les signaux sont supposés i.i.d.

Le mélange étant supposé bijectif, nous pouvons écrire :

$$E = \prod_{i=1}^T \frac{f_{S_1, S_2}(\hat{s}_1(i), \hat{s}_2(i))}{|J_g(\hat{s}_1, \hat{s}_2)|} \quad (4.12)$$

$$= \prod_{i=1}^T \frac{f_{S_1}(\hat{s}_1(i)) f_{S_2}(\hat{s}_2(i))}{|J_g(\hat{s}_1, \hat{s}_2)|} \quad (4.13)$$

où $[\hat{s}_1, \hat{s}_2] = g^{-1}(\mathbf{x}, \boldsymbol{\theta})$, et $J_g = \det \frac{\partial g}{\partial \mathbf{s}}$ est le jacobien du mélange évalué en $\hat{\mathbf{s}}$ et $\boldsymbol{\theta}$ qui vaut :

$$J_g = 4\hat{s}_1\hat{s}_2(1 - l_{12}l_{21}) - 2\hat{s}_1^2(q_2 + l_{21}q_1) - 2\hat{s}_2^2(q_1 + l_{12}q_2). \quad (4.14)$$

L'estimation par MV des paramètres $\boldsymbol{\theta}_*$ consiste à maximiser cette vraisemblance par rapport à $\boldsymbol{\theta}$. En passant au logarithme de E , nous avons :

$$\log E = \sum_{i=1}^T \log p_{\boldsymbol{\theta}}(x_1(i), x_2(i)). \quad (4.15)$$

L'expression de la fonction log-vraisemblance à maximiser devient alors :

$$L = \frac{1}{T} \log E = E_t[\log p_{\boldsymbol{\theta}}(x_1(t), x_2(t))], \quad (4.16)$$

où $E_t[\cdot]$ représente l'opérateur de moyennage temporel. En utilisant (4.13), nous obtenons :

$$L = E_t[\log f_{S_1}(\hat{s}_1)] + E_t[\log f_{S_2}(\hat{s}_2)] - E_t[\log |J_g(\hat{s}_1, \hat{s}_2)|], \quad (4.17)$$

et nous devons annuler son gradient $\frac{dL}{d\boldsymbol{\theta}}$, avec :

$$\frac{dL}{d\boldsymbol{\theta}} = -E_t[\Psi_{S_1}(\hat{s}_1) \frac{\partial \hat{s}_1}{\partial \boldsymbol{\theta}}] - E_t[\Psi_{S_2}(\hat{s}_2) \frac{\partial \hat{s}_2}{\partial \boldsymbol{\theta}}] - E_t[\frac{1}{J_g} \frac{dJ_g}{d\boldsymbol{\theta}}], \quad (4.18)$$

où $\Psi_{S_i}(u) = -\frac{\partial \log f_{S_i}(u)}{\partial u}$ représente la fonction score de la source s_i . En pratique les fonctions scores sont souvent estimées. Dans nos simulations nous utiliserons l'estimateur des fonctions score proposé dans [72].

Finalement, le gradient de la fonction log-vraisemblance s'écrit (voir le détail des calculs ci-après) :

$$\frac{dL}{d\boldsymbol{\theta}} = -E_t \left\{ \frac{1}{J_g} \left[\begin{array}{l} -2\hat{s}_2(2l_{21}\hat{s}_1 + \hat{s}_2q_2) + 4[\hat{s}_2^2(2\hat{s}_2 - q_2\hat{s}_1)(\hat{s}_2(1 - l_{12}l_{21}) - \hat{s}_1(q_2 + l_{21}q_1)) \\ + (2l_{21}\hat{s}_1 + q_2\hat{s}_2)\hat{s}_2^2(\hat{s}_1(1 - l_{12}l_{21}) - \hat{s}_2(q_1 + l_{12}q_2))]/J_g \\ + [\hat{s}_2^2(2\hat{s}_2 - q_2\hat{s}_1)\Psi_{S_1}(\hat{s}_1) + (2l_{21}\hat{s}_1 + q_2\hat{s}_2)\hat{s}_2^2\Psi_{S_2}(\hat{s}_2)], \\ -2\hat{s}_1(2\hat{s}_2l_{12} + q_1\hat{s}_1) + 4[\hat{s}_1^2(2l_{12}\hat{s}_2 + q_1\hat{s}_1)(\hat{s}_2(1 - l_{12}l_{21}) - \hat{s}_1(q_2 + l_{21}q_1)) \\ + \hat{s}_1^2(2\hat{s}_1 - q_1\hat{s}_2)(\hat{s}_1(1 - l_{12}l_{21}) - \hat{s}_2(q_1 + l_{12}q_2))]/J_g \\ + [\hat{s}_1^2(2l_{12}\hat{s}_2 + q_1\hat{s}_1)\Psi_{S_1}(\hat{s}_1) + (2\hat{s}_1 - q_1\hat{s}_2)\hat{s}_1^2\Psi_{S_2}(\hat{s}_2)], \\ -2(l_{21}\hat{s}_1^2 + \hat{s}_2^2) + 4[\hat{s}_1\hat{s}_2(2\hat{s}_2 - q_2\hat{s}_1)(\hat{s}_2(1 - l_{12}l_{21}) - \hat{s}_1(q_2 + l_{21}q_1)) \\ + \hat{s}_1\hat{s}_2(2l_{21}\hat{s}_1 + q_2\hat{s}_2)(\hat{s}_1(1 - l_{12}l_{21})) - \hat{s}_2(q_1 + l_{12}q_2)]/J_g \\ + [\hat{s}_1\hat{s}_2(2\hat{s}_2 - q_2\hat{s}_1)\Psi_{S_1}(\hat{s}_1) + \hat{s}_1\hat{s}_2(2l_{21}\hat{s}_1 + q_2\hat{s}_2)\Psi_{S_2}(\hat{s}_2)], \\ -2(l_{12}\hat{s}_2^2 + \hat{s}_1^2) + 4[\hat{s}_1\hat{s}_2(2l_{12}\hat{s}_2 + q_1\hat{s}_1)(\hat{s}_2(1 - l_{12}l_{21}) - \hat{s}_1(q_2 + l_{21}q_1)) \\ + \hat{s}_1\hat{s}_2(2\hat{s}_1 - q_1\hat{s}_2)(\hat{s}_1(1 - l_{12}l_{21}) - \hat{s}_2(q_1 + l_{12}q_2))]/J_g \\ + [\hat{s}_1\hat{s}_2(2l_{12}\hat{s}_2 + q_1\hat{s}_1)\Psi_{S_1}(\hat{s}_1) + \hat{s}_1\hat{s}_2(2\hat{s}_1 - q_1\hat{s}_2)\Psi_{S_2}(\hat{s}_2)] \end{array} \right] \right\} \quad (4.19)$$

4.3.3 Détails des calculs du gradient de la log-vraisemblance

Nous allons détailler les calculs de la fonction log-vraisemblance. L'équation (4.18) peut se mettre sous la forme :

$$\frac{dL}{d\boldsymbol{\theta}} = -E_t \left[\frac{d\hat{\mathbf{s}}}{d\boldsymbol{\theta}} \Psi_S(\hat{\mathbf{s}}) + \frac{1}{J_g} \frac{dJ_g}{d\boldsymbol{\theta}} \right], \quad (4.20)$$

où $\psi_S(\hat{\mathbf{s}}) = [\psi_{S_1}(\hat{s}_1), \psi_{S_2}(\hat{s}_2)]^T$ et

$$\frac{dJ_g}{d\boldsymbol{\theta}} = \frac{\partial J_g}{\partial \boldsymbol{\theta}} + \frac{d\hat{\mathbf{s}}}{d\boldsymbol{\theta}} \frac{\partial J_g}{\partial \hat{\mathbf{s}}} \quad (4.21)$$

Nous calculons en premier lieu les termes de l'équation (4.21). En procédant étape par étape, nous avons successivement :

$$\frac{\partial J_g}{\partial \boldsymbol{\theta}} = -2 \begin{bmatrix} \hat{s}_2(2l_{21}\hat{s}_1 + \hat{s}_2q_2) \\ \hat{s}_1(2l_{12}\hat{s}_2 + \hat{s}_1q_1) \\ l_{21}\hat{s}_1^2 + \hat{s}_2^2 \\ \hat{s}_1^2 + l_{12}\hat{s}_2^2 \end{bmatrix} \quad (4.22)$$

$$\frac{\partial J_g}{\partial \hat{\mathbf{s}}} = 4 \begin{bmatrix} \hat{s}_2(1 - l_{12}l_{21}) - \hat{s}_1(q_2 + l_{21}q_1) \\ \hat{s}_1(1 - l_{12}l_{21}) - \hat{s}_2(q_1 + l_{12}q_2) \end{bmatrix}. \quad (4.23)$$

En utilisant la relation :

$$\frac{d\hat{\mathbf{s}}}{d\boldsymbol{\theta}} = -\frac{\partial g}{\partial \boldsymbol{\theta}} \left(\frac{\partial g}{\partial \hat{\mathbf{s}}} \right)^{-1} \quad (4.24)$$

nous avons successivement :

$$\frac{\partial g}{\partial \boldsymbol{\theta}} = \begin{bmatrix} -\hat{s}_2^2 & 0 \\ 0 & -\hat{s}_1^2 \\ -\hat{s}_1 \hat{s}_2 & 0 \\ 0 & -\hat{s}_1 \hat{s}_2 \end{bmatrix} \quad (4.25)$$

$$\frac{\partial g}{\partial \hat{\mathbf{s}}} = \begin{bmatrix} 2\hat{s}_1 - q_1 \hat{s}_2 & -2l_{21} \hat{s}_1 - q_2 \hat{s}_2 \\ -2l_{12} \hat{s}_2 - q_1 \hat{s}_1 & 2\hat{s}_2 - q_2 \hat{s}_1 \end{bmatrix} \quad (4.26)$$

et donc :

$$\left(\frac{\partial g}{\partial \hat{\mathbf{s}}} \right)^{-1} = \frac{1}{J_g} \begin{bmatrix} 2\hat{s}_2 - q_2 \hat{s}_1 & 2l_{21} \hat{s}_1 + q_2 \hat{s}_2 \\ 2l_{12} \hat{s}_2 + q_1 \hat{s}_1 & 2\hat{s}_1 - q_1 \hat{s}_2 \end{bmatrix}. \quad (4.27)$$

En introduisant (4.25), (4.27) dans (4.24), nous obtenons :

$$\frac{d\hat{\mathbf{s}}}{d\boldsymbol{\theta}} = \frac{1}{J_g} \begin{bmatrix} (2\hat{s}_2 - q_2 \hat{s}_1) \hat{s}_2^2 & (2l_{21} \hat{s}_1 + q_2 \hat{s}_2) \hat{s}_2^2 \\ (2l_{12} \hat{s}_2 + q_1 \hat{s}_1) \hat{s}_1^2 & (2\hat{s}_1 - q_1 \hat{s}_2) \hat{s}_1^2 \\ (2\hat{s}_2 - q_2 \hat{s}_1) \hat{s}_1 \hat{s}_2 & (2l_{21} \hat{s}_1 + q_2 \hat{s}_2) \hat{s}_1 \hat{s}_2 \\ (2l_{12} \hat{s}_2 + q_1 \hat{s}_1) \hat{s}_1 \hat{s}_2 & (2\hat{s}_1 - q_1 \hat{s}_2) \hat{s}_1 \hat{s}_2 \end{bmatrix} \quad (4.28)$$

$$\frac{d\hat{\mathbf{s}}}{d\boldsymbol{\theta}} \frac{\partial J_g}{\partial \hat{\mathbf{s}}} = \frac{4}{J_g} \begin{bmatrix} a_1 \\ a_2 \\ a_3 \\ a_4 \end{bmatrix} \quad (4.29)$$

où

$$\left\{ \begin{array}{l} a_1 = \hat{s}_2^2 (2\hat{s}_2 - q_2 \hat{s}_1) (\hat{s}_2 (1 - l_{12} l_{21}) - \hat{s}_1 (q_2 + l_{21} q_1)) \\ \quad + (2l_{21} \hat{s}_1 + q_2 \hat{s}_2) \hat{s}_2^2 (\hat{s}_1 (1 - l_{12} l_{21}) - \hat{s}_2 (q_1 + l_{12} q_2)) \\ a_2 = \hat{s}_1^2 (2l_{12} \hat{s}_2 + q_1 \hat{s}_1) (\hat{s}_2 (1 - l_{12} l_{21}) - \hat{s}_1 (q_2 + l_{21} q_1)) \\ \quad + \hat{s}_1^2 (2\hat{s}_1 - q_1 \hat{s}_2) (\hat{s}_1 (1 - l_{12} l_{21}) - \hat{s}_2 (q_1 + l_{12} q_2)) \\ a_3 = \hat{s}_1 \hat{s}_2 (2\hat{s}_2 - q_2 \hat{s}_1) (\hat{s}_2 (1 - l_{12} l_{21}) - \hat{s}_1 (q_2 + l_{21} q_1)) \\ \quad + \hat{s}_1 \hat{s}_2 (2l_{21} \hat{s}_1 + q_2 \hat{s}_2) (\hat{s}_1 (1 - l_{12} l_{21}) - \hat{s}_2 (q_1 + l_{12} q_2)) \\ a_4 = \hat{s}_1 \hat{s}_2 (2l_{12} \hat{s}_2 + q_1 \hat{s}_1) (\hat{s}_2 (1 - l_{12} l_{21}) - \hat{s}_1 (q_2 + l_{21} q_1)) \\ \quad + \hat{s}_1 \hat{s}_2 (2\hat{s}_1 - q_1 \hat{s}_2) (\hat{s}_1 (1 - l_{12} l_{21}) - \hat{s}_2 (q_1 + l_{12} q_2)) \end{array} \right. \quad (4.30)$$

Ensuite, nous calculons le premier terme du membre de droite de l'équation (4.20)

$$\frac{d\hat{\mathbf{s}}}{d\boldsymbol{\theta}} \Psi_S(\hat{\mathbf{s}}) = \frac{1}{J_g} \begin{bmatrix} \hat{s}_2^2(2\hat{s}_2 - q_2\hat{s}_1)\Psi_{S_1}(\hat{s}_1) + (2l_{21}\hat{s}_1 + q_2\hat{s}_2)\hat{s}_2^2\Psi_{S_2}(\hat{s}_2) \\ \hat{s}_1^2(2l_{12}\hat{s}_2 + q_1\hat{s}_1)\Psi_{S_1}(\hat{s}_1) + (2\hat{s}_1 - q_1\hat{s}_2)\hat{s}_1^2\Psi_{S_2}(\hat{s}_2) \\ \hat{s}_1\hat{s}_2(2\hat{s}_2 - q_2\hat{s}_1)\Psi_{S_1}(\hat{s}_1) + \hat{s}_1\hat{s}_2(2l_{21}\hat{s}_1 + q_2\hat{s}_2)\Psi_{S_2}(\hat{s}_2) \\ \hat{s}_1\hat{s}_2(2l_{12}\hat{s}_2 + q_1\hat{s}_1)\Psi_{S_1}(\hat{s}_1) + \hat{s}_1\hat{s}_2(2\hat{s}_1 - q_1\hat{s}_2)\Psi_{S_2}(\hat{s}_2) \end{bmatrix}. \quad (4.31)$$

Finalement, en insérant les équations (4.22), (4.29), (4.30) et (4.31) dans (4.20), nous obtenons l'expression (4.19).

Le développement de ces calculs a été mené comme exposé dans [30].

4.3.4 Algorithme d'optimisation

Une fois l'expression de $\frac{dL}{d\boldsymbol{\theta}}$ déterminée, les estimateurs des paramètres du mélange sont mises à jour suivant la règle d'apprentissage utilisant le gradient de la log-vraisemblance :

$$\boldsymbol{\theta}(k+1) \leftarrow \boldsymbol{\theta}(k) + \mu \frac{dL}{d\boldsymbol{\theta}} \quad (4.32)$$

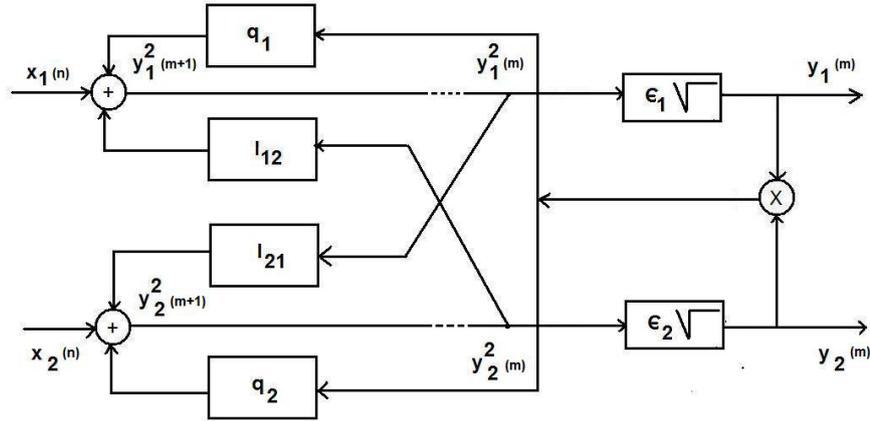
où μ correspond à un pas d'apprentissage positif.

Comme nous pouvons le constater à partir de (4.19), le calcul du gradient nécessite la connaissance des signaux $\hat{\mathbf{s}}_1$ and $\hat{\mathbf{s}}_2$ pour l'estimée courante de $\boldsymbol{\theta}$ à chaque itération de l'algorithme du gradient. Dans la section qui va suivre, nous proposons une structure récurrente pour calculer ces signaux à partir des observations. De plus, en pratique, les fonctions score des vraies sources sont généralement inconnues, et seront remplacées par les fonctions score estimées des signaux $\hat{\mathbf{s}}_1$ and $\hat{\mathbf{s}}_2$.

4.4 Structure récurrente associée

Afin d'estimer les signaux $\hat{\mathbf{s}}_1$ et $\hat{\mathbf{s}}_2$ à partir des observations et des valeurs courantes des estimées des paramètres du mélange, nous développons une structure récurrente, associée au modèle de mélange quadratique, de type réseau basique ($l_{ii} = 0$, $i = \{1, 2\}$). Cette récurrence s'écrit suivant (2.19) que nous rappelons ici :

$$\begin{cases} y_1(m+1) = \epsilon_1 \sqrt{x_1(n) + l_{12}y_2^2(m) + q_1y_1(m)y_2(m)}, \\ y_2(m+1) = \epsilon_2 \sqrt{x_2(n) + l_{21}y_1^2(m) + q_2y_1(m)y_2(m)}, \end{cases} \quad (4.33)$$


 FIGURE 4.1 – Structure récurrente associée (Figure 2.4 avec $l_{11} = l_{22} = 0$)

où les $\epsilon_i = \pm 1$, $i \in \{1, 2\}$ sont fixés à chaque instant n . Nous recherchons des sorties $y_{i,i \in \{1,2\}}(m)$ réelles $\forall m$. Les coefficients $q_{i,i \in \{1,2\}}$, l_{12} , et l_{21} , sont constants pendant la récurrence sur m . La figure (4.1) illustre le réseau récurrent. Cette structure nous donne les points d'équilibre suivants (que nous pouvons déterminer à partir des équations du Chapitre 2 (2.33) pour le cas particulier du réseau basique, i.e. avec $l_{11} = l_{22} = 0$) :

$$(y_{1E}; y_{2E})^{(1)} = (\epsilon s_1(n), \epsilon s_2(n)) \quad (4.34)$$

$$(y_{1E}; y_{2E})^{(2)} = \left(\epsilon' \frac{[(1-l_{21}l_{12})s_1(n) - (q_1+q_2l_{12})s_2(n)]}{\sqrt{d}}, \right. \\ \left. \epsilon' \frac{[(q_2+q_1l_{21})s_1(n) - (1-l_{21}l_{12})s_2(n)]}{\sqrt{d}} \right), \text{ si } d > 0 \quad (4.35)$$

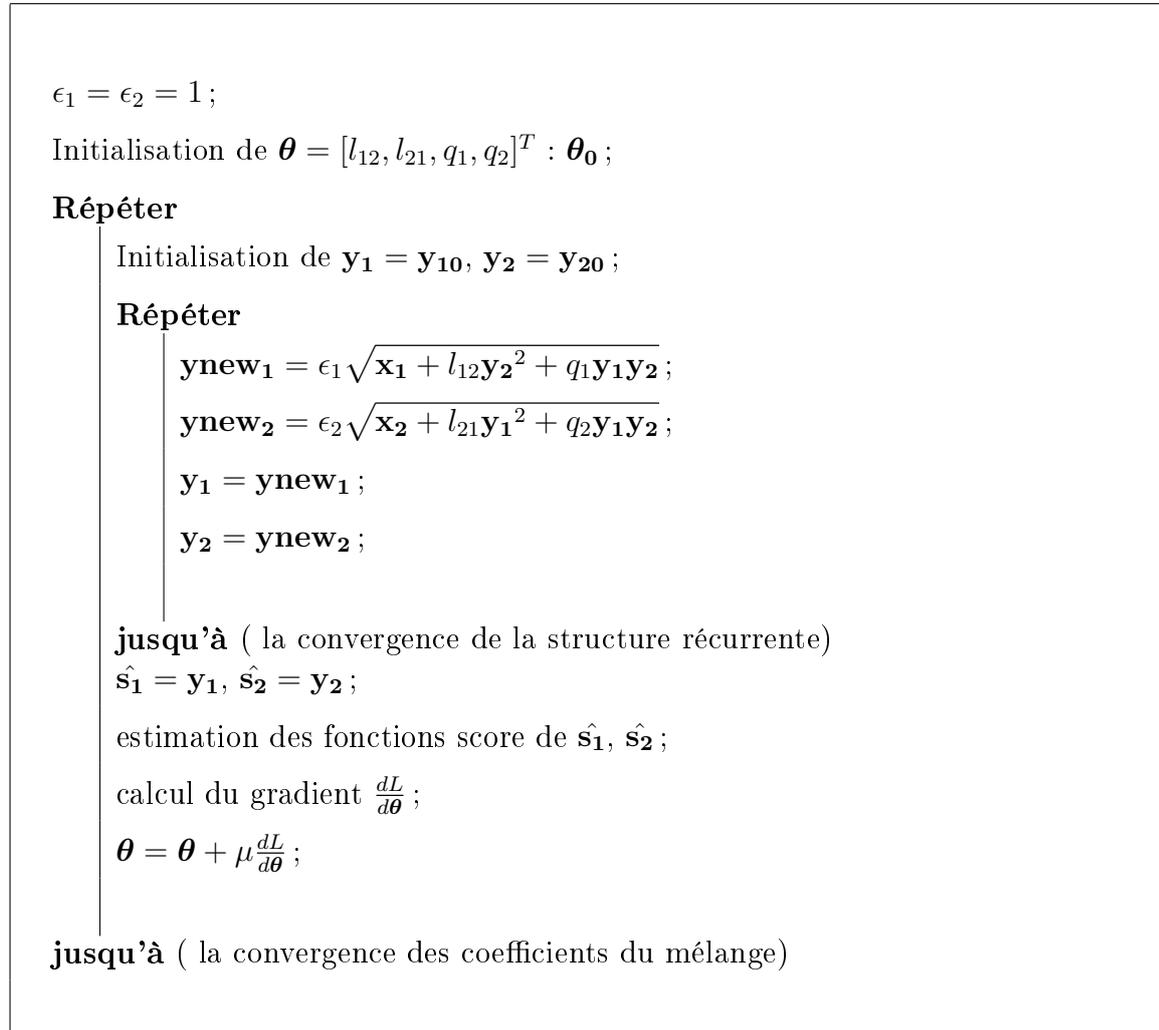
$$\text{avec } d = (1-l_{21}l_{12})^2 - (q_1+q_2l_{12})(q_2+q_1l_{21}) \quad (4.36)$$

avec $\epsilon, \epsilon' = \pm 1$. Pour $\theta = \theta_*$, la première paire de solution (4.34) correspond donc à un point séparant, qui extrait les sources, tandis que la seconde (4.35) correspond à des sources mélangées linéairement. De la même manière qu'au Chapitre 2, nous montrons que les conditions de stabilité locale du point d'équilibre séparant sont :

$$\begin{cases} \frac{(Q_2 - L_{21}Q_1)s_1^2 + (Q_1 - L_{12}Q_2)s_2^2 + 2(1 - L_{12}L_{21})s_1s_2}{2s_1s_2} > 0 \\ \frac{(Q_2 + Q_1L_{21})s_1^2 + (Q_1 + Q_2L_{12})s_2^2 - 2(1 - L_{12}L_{21})s_1s_2}{2s_1s_2} < 0 \\ \frac{-L_{21}Q_1s_1^2 - L_{12}Q_2s_2^2 - 2(1 + L_{12}L_{21})s_1s_2}{2s_1s_2} < 0 \end{cases} \quad (4.37)$$

Dans ce chapitre nous n'entrons pas dans les détails de l'étude de la stabilité du réseau, nous invitons le lecteur à se reporter au Chapitre 2 (section 2.4.3).

L'algorithme final, combinant le principe du maximum de vraisemblance et la structure de séparation de type réseau récurrent peut être résumé comme suit :



Algorithme 1: Algorithme général

où $\mathbf{y}_{\text{new}_1}, \mathbf{y}_{\text{new}_2}, \mathbf{y}_1, \mathbf{y}_2, \hat{\mathbf{s}}_1$ et $\hat{\mathbf{s}}_2$ sont des vecteurs ($1 \dots N, N = \text{longueur des signaux}$).

4.5 Simulations

Pour le premier scénario, nous générons 3 sortes de couples de sources aléatoires (les valeurs des sources sont choisies de manière à assurer la stabilité du réseau) :

- deux sources statistiquement indépendantes, distribuées uniformément sur [5.5, 6.5], avec $N = 1000$ échantillons,

- deux sources statistiquement indépendantes, de densité normale $\mathcal{N}(6, 0.2)$ avec $N = 1000$ échantillons,
- deux sources statistiquement indépendantes, l'une distribuée uniformément sur $[5.5, 6.5]$, la seconde de distribution normale $\mathcal{N}(6, 0.2)$ avec $N = 1000$ échantillons.

Les observations sont générées suivant (4.6) avec des paramètres de mélange $\theta_* = [L_{12}, L_{21}, Q_1, Q_2]$ fixés à $[-1; -0.2; 1; -0.5]$. Ces valeurs assurent la convergence du réseau récurrent vers le point séparant ((4.37) satisfaite). Le pas d'apprentissage est fixé pour toutes les simulations à $\mu = 0.0001$. L'expérience a été répétée 100 fois avec différentes initialisations des estimées des paramètres de mélange, et les valeurs initiales des sorties par rapport à m . Nous avons généré des initialisations au voisinage des vraies valeurs des paramètres de mélange et des sources, respectivement définies par les équations¹³ :

$$\theta_0 = \theta_* + 0.5 * (\text{rand}(4, 1) - 0.5) \quad (4.38)$$

$$y_{10} = s_1 + 1.2 * (\text{rand}(N, 1) - 0.5); \quad (4.39)$$

$$y_{20} = s_2 + 1.2 * (\text{rand}(N, 1) - 0.5). \quad (4.40)$$

Nous avons également utilisé l'estimateur des fonctions score proposé dans [72]. Pour chaque expérience, le SIR (Signal to Interference Ratio) sur les sorties est calculé suivant la formule :

$$SIR = 0.5 \sum_{i=1}^2 10 \log_{10} \frac{E[s_i^2]}{E[(\tilde{s}_i - s_i)^2]}, \quad (4.41)$$

où \tilde{s}_i sont les sources estimées finales.

Le tableau 4.1 résume les résultats obtenus. Nous remarquons que le SIR a été amélioré pour toutes les configurations. En effet, la moyenne de SIR_0 calculée pour les valeurs initiales du réseau est inférieure aux valeurs obtenues lorsque le réseau récurrent est mis en œuvre. Les résultats les plus satisfaisants concernent le cas d'une source uniforme et d'une source gaussienne car le SIR final est le plus élevé, cependant les résultats demeurent satisfaisants dans les autres cas.

¹³Nous utilisons la notation de la fonction *Matlab rand(N,1)* qui génère un vecteur aléatoire uniformément distribué sur l'intervalle $[0, 1]$ et de longueur N .

Nous signalons par ailleurs que le réseau est sensible aux initialisations des paramètres de mélange et des sorties du réseau récurrent, et le principal problème rencontré est que durant la procédure d'adaptation, les coefficients estimés peuvent se trouver dans une région où le réseau est instable.

Sources	moyenne(SIR)	écart-type(SIR)	Max/Min(SIR)	moyenne SIR_0
s_1 , <i>uniforme</i>	24.37	6.26	41.1/9.72	18.62
s_2 <i>uniforme</i>				
s_1 , <i>gaussienne</i>	23.56	5.42	44.89/13.13	18.60
s_2 <i>gaussienne</i>				
s_1 , <i>uniforme</i>	25.8	5.42	39.2/11.5	18.56
s_2 <i>gaussienne</i>				

TABLE 4.1 – Moyenne, Ecart-type, Maximum , Minimum du SIR et moyenne de SIR_0 (dB) sur les sorties pour différentes statistiques des sources.

Dans le second scénario, nous générons également, trois sortes de sources aléatoires comme précédemment. La différence réside dans le fait que dans cette expérience, nous générons 100 différentes réalisations des sources, tandis que les conditions d'initialisation sont constantes.

Le tableau 4.2 illustre les résultats obtenus à partir de ces simulations. Nous remarquons que les performances se sont dégradées pour les deux premières configurations (2 sources uniformes et 2 sources gaussiennes), tandis que pour le dernier cas (1 source uniforme et une source gaussienne) qui présente les meilleurs résultats, le SIR s'est amélioré.

Comparativement aux résultats du premier scénario, nous remarquons que les résultats obtenus sont moins performants, surtout pour les deux premières configurations de sources.

Nous proposons maintenant d'illustrer à travers différentes figures (voir Figures 4.2 à 4.8) les observations, les sources originales ainsi que les sorties de notre réseau (sources estimées) obtenues suivant le premier scénario proposé. Le RSB moyen est de 28.14 dB.

Sources	moyenne(SIR)	écart-type(SIR)	Max/Min(SIR)	moyenne SIR_0
s_1 , <i>uniforme</i> s_2 <i>uniforme</i>	16.97	5.2	17.37/16.89	18.6
s_1 , <i>gaussienne</i> s_2 <i>gaussienne</i>	16.97	5.2	25.4/10.5	18.6
s_1 , <i>uniforme</i> s_2 <i>gaussienne</i>	24.38	5.5	42.03/11.53	18.59

TABLE 4.2 – Moyenne, Ecart-type, Maximum, Minimum du SIR et moyenne de SIR_0 (dB) sur les sorties pour différentes statistiques des sources et différentes réalisations des sources.

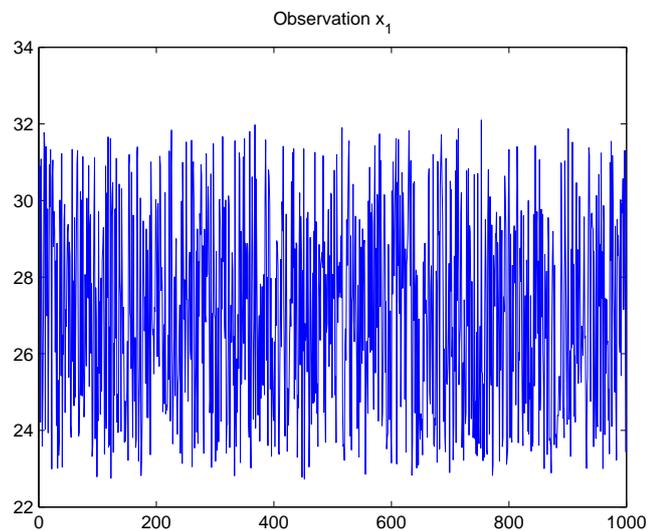
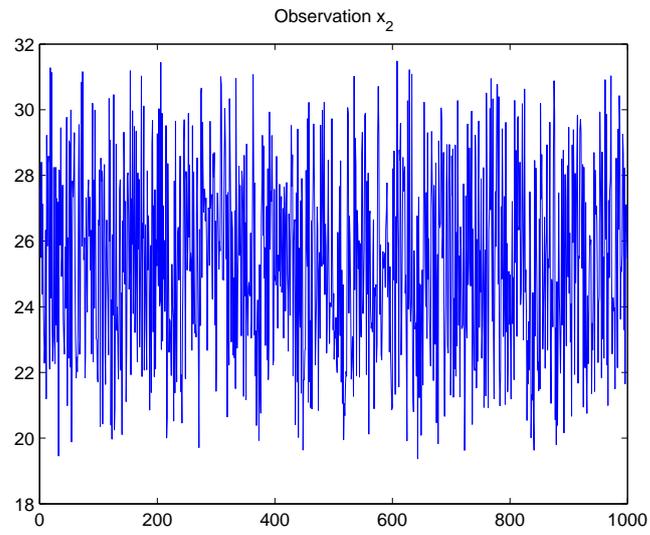
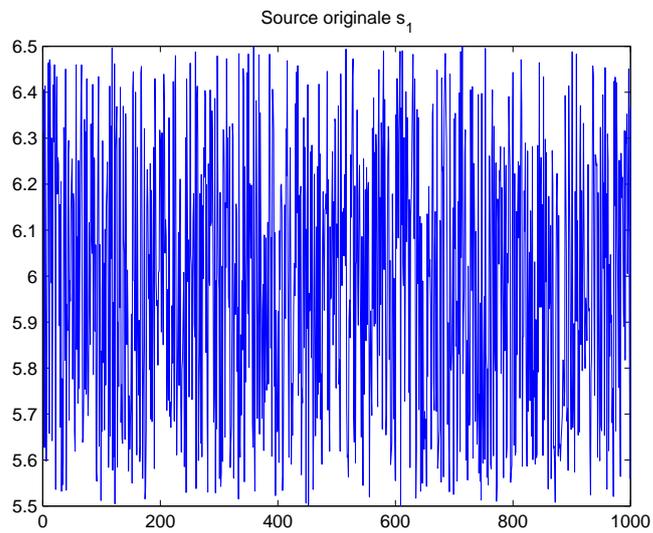


FIGURE 4.2 – Tracé de la source 1 :uniforme entre [5.5, 6.5]

FIGURE 4.3 – Tracé de la source 2 : uniforme entre $[2.5, 3.5]$ FIGURE 4.4 – Tracé de l'observation 1 paramètres du mélange fixés à $[L_{12}, L_{21}, Q_1, Q_2] = [-1; -0.2; 1; -0.5]$

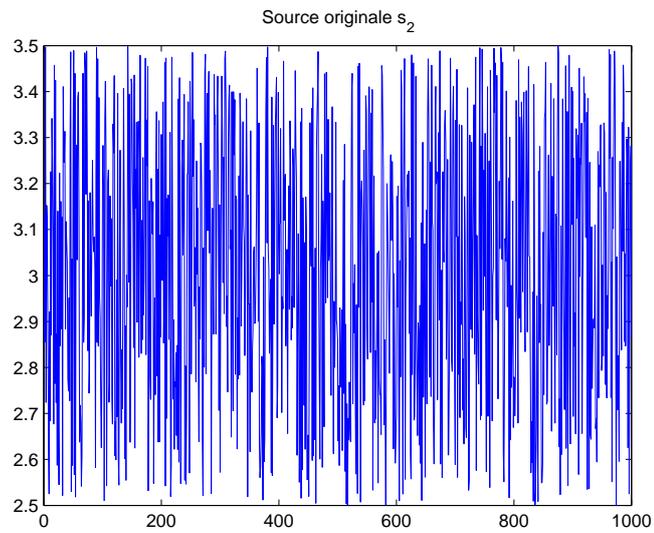


FIGURE 4.5 – Tracé de l'observation 2 paramètres du mélange fixés à $[L_{12}, L_{21}, Q_1, Q_2] = [-1; -0.2; 1; -0.5]$

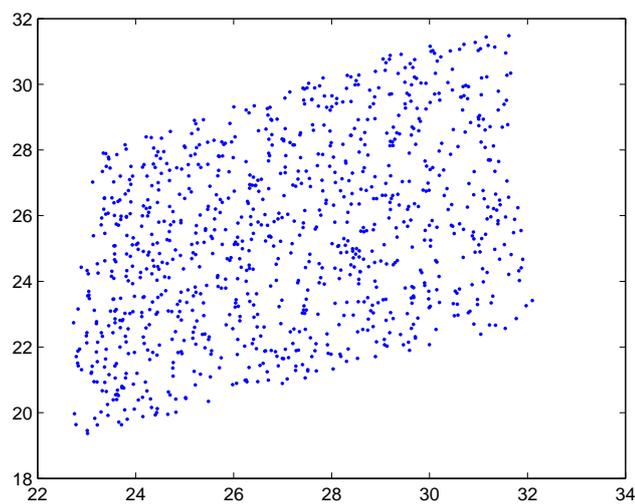


FIGURE 4.6 – Tracé de l'observation 1 en fonction de l'observation 2 paramètres du mélange fixés à $[L_{12}, L_{21}, Q_1, Q_2] = [-1; -0.2; 1; -0.5]$

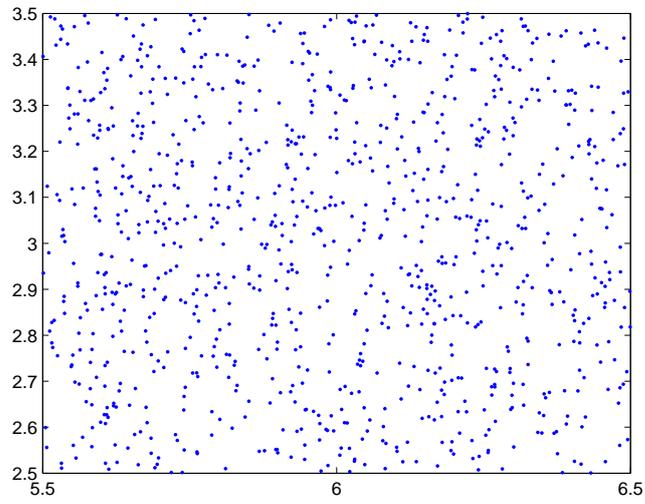


FIGURE 4.7 – Tracé de la source 1 en fonction de la source 2

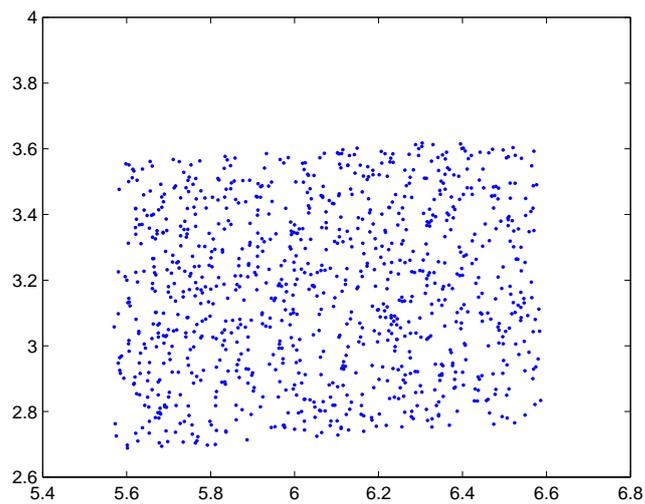


FIGURE 4.8 – Tracé de l'estimée 1 en fonction de l'estimée 2

4.6 Conclusions

Dans ce chapitre, nous avons présenté une nouvelle méthode de séparation aveugle de sources pour notre modèle quadratique. Une solution basée sur le maximum de vraisemblance a été proposée pour estimer les coefficients du mélange, à laquelle a été associé un réseau récurrent de type basique pour retrouver les sources. Les résultats obtenus restent satisfaisants, malgré le fait que ces derniers soient sensibles aux conditions initiales des simulations et à l'estimateur des fonctions score.

Conclusions et perspectives

Nous nous sommes intéressés au cours de cette thèse à un modèle non-linéaire particulier, le modèle quadratique composé d'auto-termes et de termes croisés. Nous avons alors proposé deux approches de séparation aveugle de sources (SAS), toutes deux basées sur des réseaux récurrents. Pour la première méthode présentée, un bloc situé en amont de la structure estime les paramètres du mélange via une approche basée sur les cumulants. La deuxième méthode repose sur une approche par maximum de vraisemblance (MV).

Après avoir fait un tour d'horizon sur la SAS dans les cas linéaire et non-linéaire, nous avons présenté au Chapitre 2 le réseau récurrent sur lequel sont basées nos approches. Ces dernières s'inscrivent dans la continuité des travaux développés par Y. Deville et S. Hosseini [26, 28, 40, 41]. Ainsi, nous avons présenté la structure récurrente utilisée, celle-ci se décline en une version étendue et une version plus simple dite basique. Nous avons alors étudié les propriétés de cette structure à travers les points d'équilibre et la stabilité locale du réseau en ces points à un instant donné en configuration non-aveugle. Un parallèle avec le cas du modèle linéaire-quadratique a aussi été introduit. Nous avons également étudié les limites de la stabilité locale des points d'équilibre dans le cas du modèle quadratique car contrairement au cas linéaire-quadratique, il est apparu que nous ne pouvions pas déduire une procédure qui permettrait une sélection automatique des degrés de liberté supplémentaires introduits dans le cas du modèle de mélange quadratique. À travers notre étude, nous avons vu que l'apport de la structure étendue n'est pas remis en cause pour autant, comme nous le montrent les tests effectués et exposés dans la partie « Simulations » du Chapitre 2. Nous confirmons donc l'intérêt de la structure récurrente étendue proposée pour le point d'équilibre séparant, via l'introduction de degrés de liberté supplémentaires car celle-ci permet dans certains cas de stabiliser localement le ré-

seau à un instant n . Il semblerait cependant que son application soit limitée à un certain domaine de valeurs de coefficients de mélange.

Puis nous sommes passés à la configuration aveugle. Nous avons introduit une méthode basée sur les cumulants qui permet d'estimer les coefficients du mélange. Nous avons montré que les développements étaient assez fastidieux, mais moyennant certaines hypothèses nous sommes parvenus à simplifier le problème. Sans l'ajout de ces hypothèses, le système d'équations non-linéaires que nous avons obtenu par le biais des cumulants des observations et qui a été ensuite résolu à l'aide d'algorithmes d'optimisation (Levenberg-Marquardt, Gauss-Newton), devra être remanié pour traiter directement le problème. Il apparaît pourtant que cette approche compliquerait davantage le problème mais reste tout de même envisageable.

Enfin, nous avons présenté une deuxième méthode de séparation basée cette fois sur le maximum de vraisemblance. Pour cette approche nous avons utilisé la version basique de la structure, car n'ayant pas totalement solutionné le problème de stabilité du réseau étendu, nous avons donc préféré utiliser la version simplifiée du modèle.

Concernant les perspectives de ce travail, beaucoup de choses restent à faire :

- Nous devons poursuivre les développements effectués pour la structure étendue en ce qui concerne l'analyse de la stabilité au point d'équilibre, même si cela doit se faire au prix d'hypothèses supplémentaires, afin de parvenir à automatiser le procédé de sélection des degrés de liberté supplémentaires introduits.
- L'analyse de la stabilité locale aux points d'équilibre, pour des sorties égales aux sources permutées, n'a pas été effectuée.
- Nous pouvons également chercher d'autres types de points d'équilibre pour notre réseau récurrent.
- Concernant l'estimation des coefficients du mélange par la méthode des cumulants, nous pourrions envisager de déterminer les équations du système des cumulants des observations sans avoir recours aux hypothèses supplémentaires même si cela doit se faire au prix de développements plus conséquents.

-
- Enfin, concernant la méthode basée sur le maximum de vraisemblance, si nous arrivons à stabiliser notre version étendue du réseau récurrent nous pourrions alors envisager de développer la structure étendue pour cette méthode.

À ces perspectives s'ajoutent les problèmes d'identifiabilité et de séparabilité du modèle qui n'ont pas du tout été abordés dans cette thèse.

Annexe A

Stabilité du réseau basique : cas d'existence d'un seul point d'équilibre

A.1 Etude théorique de la condition d'existence du point non séparant dans \mathbb{R}

Il s'agit de déterminer dans quels cas il n'existe qu'un seul point d'équilibre, le point séparant pour la structure basique du modèle quadratique, et d'étudier alors la stabilité locale de ce point. On rappelle que la condition vérifiée pour que le point non-séparant ne soit pas réel est $d < 0$ (2.33) que nous rappelons ici :

$$(1 - L_{21}L_{12})^2 - (Q_1 + Q_2L_{12})(Q_2 + Q_1L_{21}) < 0, \quad \text{avec } L_{12}, L_{21} < 0 \quad (\text{A.1})$$

On notera δ_1 le terme de gauche de cette expression, soit :

$$\delta_1 = (1 - L_{21}L_{12})^2 - (Q_1 + Q_2L_{12})(Q_2 + Q_1L_{21}).$$

En développant l'expression de gauche, on peut considérer le polynôme du second degré en Q_2 :

$$P(Q_2) = -L_{12}Q_2^2 - Q_1(1 + L_{21}L_{12})Q_2 + (1 - L_{21}L_{12})^2 - Q_1^2L_{21} \quad (\text{A.2})$$

qui est de la forme :

$$P(Q_2) = AQ_2^2 + BQ_2 + C \quad (\text{A.3})$$

avec

$$A = -L_{12} \quad (\text{A.4})$$

$$B = -Q_1(1 + L_{21}L_{12}) \quad (\text{A.5})$$

$$C = (1 - L_{21}L_{12})^2 - Q_1^2L_{21} \quad (\text{A.6})$$

On suppose que les conditions suivantes sont vérifiées¹⁴ :

$$A \neq 0, \quad (\text{A.7})$$

$$B \neq 0, \quad (\text{A.8})$$

$$C \neq 0, \quad (\text{A.9})$$

Le discriminant de $P(Q_2)$ s'écrit :

$$\delta'_1 = B^2 - 4AC \quad (\text{A.10})$$

$$= Q_1^2(1 + L_{21}L_{12})^2 + 4L_{12}((1 - L_{21}L_{12})^2 - Q_1^2L_{21})$$

$$= Q_1^2((1 + L_{21}L_{12})^2 - 4L_{21}L_{12}) + 4L_{12}(1 - L_{21}L_{12})^2$$

$$= Q_1^2(1 - L_{21}L_{12})^2 + 4L_{12}(1 - L_{21}L_{12})^2$$

$$= (1 - L_{21}L_{12})^2(Q_1^2 + 4L_{12}) \quad (\text{A.11})$$

A.1.1 Etude du signe de δ'_1

Nous commençons par étudier le signe de δ'_1 afin de déterminer le signe de $P(Q_2)$.

On rappelle que $L_{12}, L_{21} < 0$ d'après (1.43). Nous avons :

$$\delta'_1 < 0 \Rightarrow |Q_1| < 2\sqrt{-L_{12}} \quad (\text{A.12})$$

$$\delta'_1 = 0 \Rightarrow Q_1 = 2\sqrt{-L_{12}} \text{ ou } 1 - L_{21}L_{12} = 0 \quad (\text{A.13})$$

$$\delta'_1 > 0 \Rightarrow |Q_1| > 2\sqrt{-L_{12}} \quad (\text{A.14})$$

A.1.2 Conséquence sur la condition (A.1)

- Lorsque $\delta'_1 \leq 0$, $P(Q_2)$ est du même signe que $-L_{12}$ donc $P(Q_2) > 0$, et la relation (A.1) n'est pas vérifiée, il n'existe pas alors qu'un seul point d'équilibre dans ce cas.

¹⁴On ne traitera que le cas du polynôme du second degré général.

• Pour $\delta'_1 > 0$,

– on a $P(Q_2) < 0$ pour $Q_2 \in]R_1, R_2[$, où R_1 et R_2 désignent les racines de (A.3) et sont définies par :

$$R_1 = \frac{Q_1(1 + L_{21}L_{12}) - \sqrt{(1 - L_{21}L_{12})^2(Q_1^2 + 4L_{12})}}{-2L_{12}} \quad (\text{A.15})$$

$$R_2 = \frac{Q_1(1 + L_{21}L_{12}) + \sqrt{(1 - L_{21}L_{12})^2(Q_1^2 + 4L_{12})}}{-2L_{12}} \quad (\text{A.16})$$

– et $P(Q_2) \geq 0$ ailleurs.

Conclusion : Lorsque $\delta'_1 > 0$, si Q_2 est à l'extérieur des racines R_1 et R_2 la relation (A.1) n'est pas vérifiée et nous ne sommes pas dans le cas d'existence d'un seul point d'équilibre. Or le seul cas qui nous intéresse est celui de l'existence d'un seul point d'équilibre, le point séparant. Par contre à l'intérieur des racines, il n'existe qu'un seul point d'équilibre, c'est ce cas que nous allons approfondir par la suite.

A.1.3 Informations supplémentaires : signe des racines

On a d'après les relations (A.3) et (A.4)-(A.6) les relations suivantes :

$$\begin{cases} \frac{C}{A} = \frac{(1-L_{21}L_{12})^2 - Q_1^2 L_{21}}{-L_{12}} \Rightarrow \frac{C}{A} > 0 \text{ car } l_{ij} < 0 \\ \frac{-B}{A} = \frac{Q_1(1+L_{21}L_{12})}{-L_{12}} \Rightarrow \text{sgn}\left(\frac{-B}{A}\right) = \text{sgn}(Q_1) \end{cases} \quad (\text{A.17})$$

Pour $\delta'_1 > 0$, d'après (A.17), R_1 et R_2 ont le même signe que Q_1 .

Conclusion : Q_1 et Q_2 sont de même signe quand $P(Q_2) < 0$, i.e. lorsqu'il n'existe pas de point non séparant, $Q_2 \in]R_1, R_2[$ dans ce cas. Q_2 est donc de même signe que R_1 et R_2 .

A.2 Etude de la stabilité au point d'équilibre

Nous rappelons que nous nous plaçons dans le cas où $P(Q_2) < 0$, donc avec $\delta'_1 > 0$ et $Q_2 \in]R_1, R_2[$. Nous rappelons également que Q_1 et Q_2 sont de même signe. Nous allons maintenant effectuer un raisonnement par l'absurde : nous allons supposer que le point séparant est stable et montrer finalement que ce ne peut être le cas. Nous supposons que la condition (A.1) est vérifiée et nous étudions la stabilité au

point séparant, seul point d'équilibre de la structure parmi les deux points possibles du type considéré. Les conditions de stabilité locale au point d'équilibre (2.42, avec $k_1 = k_2 = 1$) s'écrivent d'après (2.42) :

$$\frac{a_i s_1^2 + b_i s_2^2 + c_i s_1 s_2}{s_1 s_2} > 0, \quad i \in \{1, 2, 3\} \quad (\text{A.18})$$

avec

$$a_1 = Q_2 - L_{21}Q_1; \quad b_1 = Q_1 - L_{12}Q_2; \quad c_1 = 2(1 - L_{21}L_{12}); \quad (\text{A.19})$$

$$a_2 = -(L_{21}Q_1 + Q_2); \quad b_2 = -(Q_1 + L_{12}Q_2); \quad c_2 = 2(1 - L_{21}L_{12}) \quad (\text{A.20})$$

$$a_3 = L_{21}Q_1; \quad b_3 = L_{12}Q_2; \quad c_3 = 2(1 + L_{21}L_{12}) \quad (\text{A.21})$$

Les discriminants des polynômes en s_1 de gauche (numérateurs) des inéquations précédentes ont pour expressions respectives :

$$\delta_{11} = 4s_2^2[(1 - L_{21}L_{12})^2 - (Q_2 - L_{21}Q_1)(Q_1 - L_{12}Q_2)] = 4s_2^2\delta_1^{cond} \quad (\text{A.22})$$

$$\delta_{22} = 4s_2^2[(1 - L_{21}L_{12})^2 - (Q_2 + L_{21}Q_1)(Q_1 + L_{12}Q_2)] = 4s_2^2\delta_1 \quad (\text{A.23})$$

$$\delta_{33} = 4s_2^2[(1 + L_{21}L_{12})^2 - L_{12}L_{21}Q_1Q_2] = 4s_2^2\delta_3 \quad (\text{A.24})$$

On a pour les 2 premiers discriminants :

$$\delta_{11} < 0, \quad (\text{A.25})$$

$$\text{car } \delta_1^{cond} = \delta_1 + 2(L_{21}Q_1^2 + L_{12}Q_2^2), \quad \delta_1 < 0,$$

$$\text{avec } L_{21}Q_1^2 + L_{12}Q_2^2 < 0 \quad (\text{A.26})$$

$$\delta_{22} < 0 \quad (\text{A.27})$$

A.2.1 Cas où les sources sont de même signe

Les relations (A.25,A.27) montrent que les polynômes des inéquations (A.18) $_{i \in \{1,2\}}$ sont du signe des $a_{i,i \in \{1,2\}}$ si s_1 et s_2 sont de même signe. Les conditions nécessaires et suffisantes pour assurer le respect de ces inéquations sont :

$$a_1 > 0 \Rightarrow Q_2 > L_{21}Q_1, \quad (\text{A.28})$$

$$a_2 > 0 \Rightarrow Q_2 < -L_{21}Q_1, \quad (\text{A.29})$$

$$\text{On note : } \alpha = L_{21}Q_1, \quad \beta = -L_{21}Q_1 = -\alpha \quad (\text{A.30})$$

Condition étudiée (CE)	signe(Q_1)	signe(Q_2)	signe(α)	signe(β)	Q_2 t.q. CE vérifiée	remarque
$Q_2 > \alpha$	+	+	-	+	$Q_2 > 0$	T. V
$Q_2 < \beta$	+	+	-	+	$0 < Q_2 < \beta$	possible
$Q_2 > \alpha$	-	-	+	-	$Q_2 > \alpha$	\emptyset
$Q_2 < \beta$	-	-	+	-	$Q_2 < \beta$	possible

TABLE A.1 – Étude des signes pour la vérification de l'inéquation (A.23)

avec T.V= toujours vérifiée.

On en conclut que l'inéquation (A.28) n'est jamais vérifiée pour $Q_1 < 0$, on en déduit que Q_1 et Q_2 doivent être positifs pour que (A.28) soit vérifiée.

Étude de l'inéquation (A.18)_{i=3} lorsque les paramètres Q_1 et Q_2 sont positifs

$$\text{On a : } R_1 < Q_2 < R_2, \quad \text{donc } Q_1 R_1 < Q_1 Q_2 < Q_1 R_2 \quad (\text{A.31})$$

On pose

$$x = L_{21} L_{12}, \quad (\text{A.32})$$

$$y = \frac{-Q_1^2}{4L_{12}}, \quad \text{et suivant (A.14) } 0 < \frac{1}{y} < 1 \quad (\text{A.33})$$

Pour $0 < x < 1$ (tel que $1 - x > 0$) en multipliant (A.15) par Q_1 , on obtient

$$Q_1 R_1 = 2y(1 + x) - 2y(1 - x)\sqrt{1 - \frac{1}{y}} \quad (\text{A.34})$$

D'après l'équation (A.29) $Q_1 Q_2 < Q_1 \beta$, et sachant que $Q_2 \in]R_1, R_2[$, on en déduit que $Q_1 R_1 < Q_1 \beta$. Par conséquent :

$$\begin{aligned} 2y(1 + x) - 2y(1 - x)\sqrt{1 - \frac{1}{y}} &< 4xy \\ 2y(1 - x) - 2y(1 - x)\sqrt{1 - \frac{1}{y}} &< 0 \\ 2y(1 - x)\left(1 - \sqrt{1 - \frac{1}{y}}\right) &< 0 \end{aligned} \quad (\text{A.35})$$

Pour $0 < x < 1$, (A.35) n'est donc jamais vérifiée. Pour le cas $x > 1$, on doit tester si l'inéquation $(A.18)_{i=3}$ est satisfaite. (A.21) montre que δ_{33} doit être positif :

$$(1+x)^2 - xQ_1Q_2 > 0 \quad \text{soit} \quad Q_1Q_2 < \frac{(1+x)^2}{x} \quad (A.36)$$

D'après (A.31),(A.34) et (A.36) , on doit avoir :

$$2y(1+x) - 2y(x-1)\sqrt{1-\frac{1}{y}} < \frac{(1+x)^2}{x} \quad (A.37)$$

Il s'agit de montrer que cette relation n'est pas vérifiée pour $x > 1$, on sait par ailleurs d'après (A.33) que $y > 1$, on doit donc montrer que :

$$2y(1+x) - 2y(x-1)\sqrt{1-\frac{1}{y}} - \frac{(1+x)^2}{x} > 0, \quad y > 1 \quad x > 1 \quad (A.38)$$

En posant

$$P = 2y(1+x) - 2y(x-1)\sqrt{1-\frac{1}{y}} - \frac{(1+x)^2}{x} = 0 \quad (A.39)$$

on obtient :

$$(2y - 2y\sqrt{1-\frac{1}{y}} - 1)x^2 + (2y + 2y\sqrt{1-\frac{1}{y}} - 2)x - 1 = 0 \quad (A.40)$$

On pose alors $z = \sqrt{1-\frac{1}{y}} \Rightarrow y = \frac{1}{1-z^2}$, $0 < z < 1$.

(A.40) met donc en jeu un polynôme en x de la forme $ax^2 + 2b'x + c$ avec :

$$\begin{cases} a = \frac{1-z}{1+z} \\ b' = \frac{z}{1-z} \\ c = -1 \end{cases} \quad (A.41)$$

Le discriminant réduit de ce polynôme est :

$$\delta' = \left(\frac{z}{1-z}\right)^2 + \frac{1-z}{1+z} > 0, \quad \forall 0 < z < 1 \quad (A.42)$$

Le polynôme admet deux points d'intersection, P_1 et P_2 , avec l'axe des abscisses, car $a > 0$, $\forall 0 < z < 1$. Il suffit donc de montrer que $P_2 < 1$. Nous avons :

$$P_2 = \frac{-b' + \sqrt{\delta'}}{a} = \frac{-\frac{z}{1-z} + \sqrt{\left(\frac{z}{1-z}\right)^2 + \frac{1-z}{1+z}}}{\frac{1-z}{1+z}} \quad (A.43)$$

Nous avons alors les relations suivantes :

$$P_2 < 1 \quad (A.44)$$

$$\sqrt{\left(\frac{z}{1-z}\right)^2 + \frac{1-z}{1+z}} < \frac{z}{1-z} + \frac{1-z}{1+z} \quad (A.45)$$

$$\left(\frac{z}{1-z}\right)^2 + \frac{1-z}{1+z} < \left(\frac{z}{1-z}\right)^2 + \left(\frac{1-z}{1+z}\right)^2 + \frac{2z}{1+z} \quad (A.46)$$

$$1 - z^2 < 3z^2 + 1 \quad (A.47)$$

soit finalement :

$$4z^2 > 0 \quad (\text{A.48})$$

qui est toujours vérifiée. Par conséquent, l'inéquation (A.18)_{i=3} n'est jamais vérifiée.

Nous pouvons également montrer que cette inégalité n'est pas vérifiée par simulations pour les valeurs des variables x et y appartenant à un certain intervalle (voir Figures A.1 et A.2 et le script MATLAB).

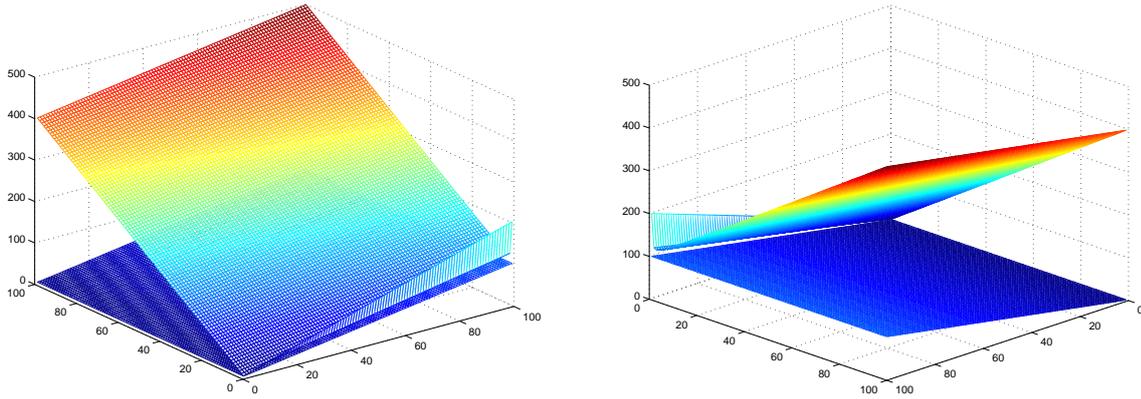


FIGURE A.1 – Tracés des fonctions $Z_1 = 2y(1+x) - 2y(x-1)\sqrt{1 - \frac{1}{y}}$ et $Z_2 = \frac{(1+x)^2}{x}$

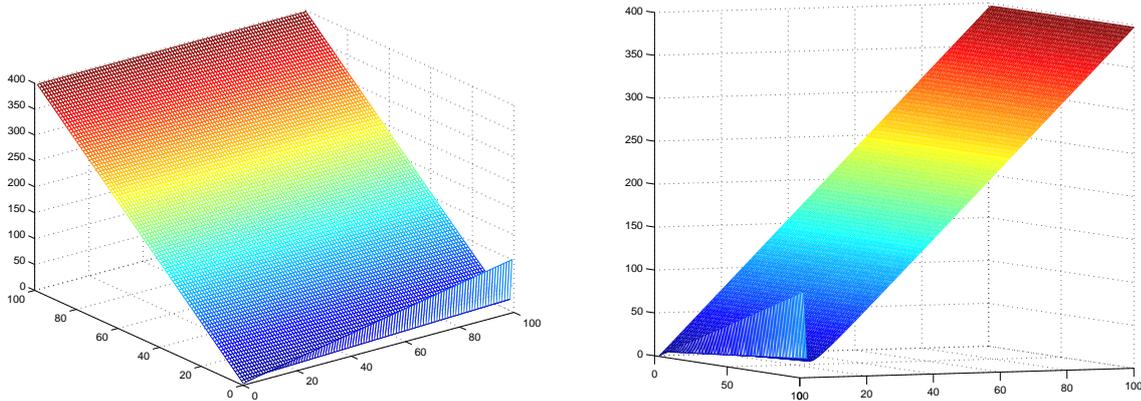


FIGURE A.2 – Tracés de la fonction $Z_3 = 2y(1+x) - 2y(x-1)\sqrt{1 - \frac{1}{y}} - \frac{(1+x)^2}{x}$

Script Matlab

```
[X,Y] = meshgrid(1 : 1 : 100, 1 : 1 : 100);
Z1 = 2 * Y .* ((1 + X) - (X - 1) .* sqrt(1 - 1./Y));
```

$$Z_2 = (1 + X)^2 ./ X;$$

$$Z_3 = 2 * Y .* ((1 + X) - (X - 1) .* sqrt(1 - 1./Y)) - (1 + X)^2 ./ X;$$

figure, mesh(Z₁); hold on; mesh(Z₂)

figure, mesh(Z₃);

A.2.2 Cas où les sources sont de signes contraires

En utilisant le même raisonnement que précédemment, nous aboutissons aux conclusions suivantes lorsque les sources sont de signes contraires :

- Pour $Q_1 > 0$, l'inéquation (A.18) _{$i=2$} n'est jamais vérifiée.
- Pour $Q_1, Q_2 < 0$ et $0 < x < 1$ l'inéquation (A.18) _{$i=2$} n'est jamais vérifiée. Pour $Q_1, Q_2 < 0$ et $x > 1$ l'inéquation (A.18) _{$i=3$} n'est pas vérifiée.

A.3 Conclusion

Dans cette annexe, nous avons raisonné en deux étapes. Dans un premier temps, nous avons déterminé dans quels cas il n'existe qu'un seul point d'équilibre du réseau (relation (A.1) vérifiée), puis dans un deuxième temps on étudie la stabilité locale du seul point d'équilibre.

Pour la première étape, il apparaît que le cas où il n'existe qu'un seul point d'équilibre est celui pour lequel $P(Q_2) < 0$ et $Q_2 \in]R_1, R_2[$, où R_1 et R_2 désignent les racines de (A.3). On précise que dans ce cas Q_1 et Q_2 sont de même signe.

Pour la deuxième étape, on étudie la stabilité locale au seul point d'équilibre, i.e. dans le cas cité précédemment. On étudie alors le système d'inéquations de stabilité locale défini par (A.18). Deux cas sont étudiés, d'abord le cas, seul détaillé, où les sources sont de même signe, puis celui pour lequel les sources sont de signes contraires.

Pour ce premier cas nous commençons par étudier les inéquations (A.18) _{$i=1,2$} . Il apparaît alors que Q_1 et Q_2 doivent être positifs, on passe alors à l'inéquation (A.18) _{$i=3$} avec Q_1 et Q_2 positifs. On démontre alors que cette inéquation n'est jamais vérifiée.

Le même raisonnement peut être tenu pour le cas de sources de signes contraires et nous aboutissons aux mêmes conclusions que précédemment.

La conclusion générale est que lorsqu'il n'existe qu'un seul point d'équilibre, celui-ci est toujours instable.

Annexe B

Détails des calculs pour les « cumulants sources »

Dans cette annexe nous complétons les résultats obtenus pour les cumulants associés aux sources d'ordre 4, en détaillant les calculs effectués afin de parvenir aux expressions données.

B.1 Cumulants associés aux sources d'ordre 4 non nuls

Nous devons montrer que :

$$\left\{ \begin{array}{l}
 C[(s_1^2)_4] = E[s_1^8] - 4E[s_1^6]E[s_1^2] - 6E[s_1^2]^4 \\
 \quad - 3E[s_1^4]^2 + 12E[s_1^4]E[s_1^2]^2, \\
 C[(s_2^2)_4] = E[s_2^8] - 4E[s_2^6]E[s_2^2] - 6E[s_2^2]^4 \\
 \quad - 3E[s_2^4]^2 + 12E[s_2^4]E[s_2^2]^2, \\
 C[(s_1 s_2)_4] = E[s_1^4]E[s_2^4] - 3E[s_1^2]^2E[s_2^2]^2, \\
 C[(s_1^2)_2, (s_1 s_2)_2] = E[s_1^6]E[s_2^2] - E[s_1^2]E[s_2^2](3E[s_1^4] - 2E[s_1^2]^2), \\
 C[(s_2^2)_2, (s_1 s_2)_2] = E[s_2^6]E[s_1^2] - E[s_2^2]E[s_1^2](3E[s_2^4] - 2E[s_2^2]^2), \\
 C[s_1^2, s_2^2, (s_1 s_2)_2] = E[s_1^4]E[s_2^4] - E[s_1^2]^2E[s_2^4] \\
 \quad - E[s_2^2]^2E[s_1^4] + E[s_1^2]^2E[s_2^2]^2.
 \end{array} \right. \quad (B.1)$$

Nous avons :

- Pour $C[(s_1^2)_4]$ les partitions suivantes :

- * $\{s_1^2, s_1^2, s_1^2, s_1^2\}$ une fois,
- * $\{\{s_1^2\}, \{s_1^2\}, \{s_1^2\}, \{s_1^2\}\}$ une fois,
- * $\{\{s_1^2, s_1^2\}, \{s_1^2, s_1^2\}\}$ trois fois,
- * $\{\{s_1^2\}, \{s_1^2, s_1^2, s_1^2\}\}$ quatre fois
- * et $\{\{s_1^2\}, \{s_1^2\}, \{s_1^2, s_1^2\}\}$ six fois.

Le nombre d'éléments de la partition sont respectivement 1, 4, 2, 2 et 3. Nous calculons alors le cumulants :

$$\begin{aligned}
 C[(s_1^2)_4] &= (-1)^0 0! E[s_1^2 s_1^2 s_1^2 s_1^2] + (-1)^{4-1} (4-1)! E[s_1^2] E[s_1^2] E[s_1^2] E[s_1^2] \\
 &\quad + 3(-1)^{2-1} (2-1)! E[s_1^2 s_1^2] E[s_1^2 s_1^2] + 4(-1)^{2-1} (2-1)! E[s_1^2] E[s_1^2 s_1^2 s_1^2] \\
 &\quad + 6(-1)^{3-1} (3-1)! E[s_1^2] E[s_1^2] E[s_1^2 s_1^2] \\
 &= E[s_1^8] - 6E^4[s_1^2] - 3E^2[s_1^4] - 4E[s_1^2] E[s_1^6] + 12E^2[s_1^2] E[s_1^4]
 \end{aligned}
 \tag{B.2}$$

- Pour $C[(s_2^2)_4]$ et $C[(s_1 s_2)_4]$ on procède de la même manière que précédemment.

- Pour $C[(s_1^2)_2, (s_1 s_2)_2]$ nous avons les partitions suivantes :

- * $\{s_1^2, s_1^2, s_1 s_2, s_1 s_2\}$ une fois,
- * $\{\{s_1^2\}, \{s_1^2\}, \{s_1 s_2\}, \{s_1 s_2\}\}$ une fois,
- * $\{\{s_1^2, s_1^2\}, \{s_1 s_2, s_1 s_2\}\}$ une fois,
- * $\{\{s_1^2, s_1 s_2\}, \{s_1^2, s_1 s_2\}\}$ deux fois,
- * $\{\{s_1^2\}, \{s_1^2, s_1 s_2, s_1 s_2\}\}$ deux fois,
- * $\{\{s_1 s_2\}, \{s_1^2, s_1^2, s_1 s_2\}\}$ deux fois,
- * $\{\{s_1^2\}, \{s_1^2\}, \{s_1 s_2, s_1 s_2\}\}$ une fois,
- * $\{\{s_1^2\}, \{s_1 s_2\}, \{s_1^2, s_1 s_2\}\}$ quatre fois,
- * et $\{\{s_1^2, s_1^2\}, \{s_1 s_2\}, \{s_1 s_2\}\}$ une fois.

Le nombre d'éléments pour chaque partition est 1, 4, 2, 2, 2, 2, 3, 3 et 3. Nous en déduisons :

$$\begin{aligned}
C[(s_1^2)_2, (s_1 s_2)_2] &= (-1)^0 0! E[s_1^2 s_1^2 s_1 s_2 s_1 s_2] \\
&+ (-1)^{4-1} (4-1)! E[s_1^2] E[s_1^2] E[s_1 s_2] E[s_1 s_2] \\
&+ (-1)^{2-1} (2-1)! E[s_1^2 s_1^2] E[s_1 s_2 s_1 s_2] \\
&+ 2(-1)^{2-1} (2-1)! E[s_1^2 s_1 s_2] E[s_1^2 s_1 s_2] \\
&+ 2(-1)^{2-1} (2-1)! E[s_1^2] E[s_1^2 s_1 s_2 s_1 s_2] \\
&+ 2(-1)^{2-1} (2-1)! E[s_1 s_2] E[s_1^2 s_1^2 s_1 s_2] \\
&+ (-1)^{3-1} (3-1)! E[s_1^2] E[s_1^2] E[s_1 s_2 s_1 s_2] \\
&+ 4(-1)^{3-1} (3-1)! E[s_1^2] E[s_1 s_2] E[s_1^2 s_1 s_2] \\
&+ (-1)^{3-1} (3-1)! E[s_1 s_2] E[s_1 s_2] E[s_1^2 s_1^2] \\
&= E[s_1^6] E[s_2^2] - E[s_1^2] E[s_2^2] (3E[s_1^4] - 2E[s_1^2]^2) \quad (\text{B.3})
\end{aligned}$$

- Pour $C[(s_2^2)_2, (s_1 s_2)_2]$ on procède de la même manière que précédemment en remplaçant s_1^2 par s_2^2 .

- Pour $C[s_1^2, s_2^2, (s_1 s_2)_2]$, nous avons les partitions

- * $\{s_1^2, s_2^2, s_1 s_2, s_1 s_2\}$ une fois,
- * $\{\{s_1^2\}, \{s_2^2\}, \{s_1 s_2\}, \{s_1 s_2\}\}$ une fois,
- * $\{\{s_1^2, s_2^2\}, \{s_1 s_2, s_1 s_2\}\}$ une fois,
- * $\{\{s_1^2, s_1 s_2\}, \{s_1 s_2, s_2^2\}\}$ deux fois,
- * $\{\{s_1^2\}, \{s_2^2, s_1 s_2, s_1 s_2\}\}$ une fois,
- * $\{\{s_2^2\}, \{s_1^2, s_1 s_2, s_1 s_2\}\}$ une fois,
- * $\{\{s_1 s_2\}, \{s_1^2, s_1 s_2, s_2^2\}\}$ deux fois,
- * $\{\{s_1^2\}, \{s_2^2\}, \{s_1 s_2, s_1 s_2\}\}$ une fois,
- * $\{\{s_1^2\}, \{s_1 s_2\}, \{s_2^2, s_1 s_2\}\}$ deux fois,
- * $\{\{s_2^2\}, \{s_1 s_2\}, \{s_1^2, s_1 s_2\}\}$ deux fois,
- * et $\{\{s_1 s_2\}, \{s_1 s_2\}, \{s_1^2, s_2^2\}\}$ une fois,

avec les nombres d'éléments respectifs 1, 4, 2, 2, 2, 2, 2, 3, 3, 3 et 3. On en déduit que :

$$\begin{aligned}
C[s_1^2, s_2^2, (s_1 s_2)_2] &= (-1)^0 0! E[s_1^2 s_2^2 s_1 s_2 s_1 s_2] \\
&+ (-1)^{4-1} (4-1)! E[s_1^2] E[s_2^2] E[s_1 s_2] E[s_1 s_2] \\
&+ (-1)^{2-1} (2-1)! E[s_1^2 s_2^2] E[s_1 s_2 s_1 s_2] \\
&+ 2(-1)^{2-1} (2-1)! E[s_1^2 s_1 s_2] E[s_2^2 s_1 s_2] \\
&+ (-1)^{2-1} (2-1)! E[s_1^2] E[s_2^2 s_1 s_2 s_1 s_2] \\
&+ (-1)^{2-1} (2-1)! E[s_2^2] E[s_1^2 s_1 s_2 s_1 s_2] \\
&+ 2(-1)^{2-1} (2-1)! E[s_1 s_2] E[s_1^2 s_2^2 s_1 s_2] \\
&+ (-1)^{3-1} (3-1)! E[s_1^2] E[s_2^2] E[s_1 s_2 s_1 s_2] \\
&+ 2(-1)^{3-1} (3-1)! E[s_1^2] E[s_1 s_2] E[s_2^2 s_1 s_2] \\
&+ 2(-1)^{3-1} (3-1)! E[s_2^2] E[s_1 s_2] E[s_1^2 s_1 s_2] \\
&+ (-1)^{3-1} (3-1)! E[s_1 s_2] E[s_1 s_2] E[s_1^2 s_2^2] \\
&= E[s_1^4] E[s_2^4] - E[s_1^2]^2 E[s_2^4] \\
&\quad - E[s_2^2]^2 E[s_1^4] + E[s_1^2]^2 E[s_2^2]^2. \tag{B.4}
\end{aligned}$$

B.2 Cumulants associés aux sources d'ordre 4 nuls

Il s'agit de démontrer les relations suivantes :

$$\begin{cases}
C[(s_1^2)_3, s_2^2] = C[(s_2^2)_3, s_1^2] = C[(s_1 s_2)_3, s_1^2] = 0. \\
C[(s_1 s_2)_3, s_2^2] = C[(s_1^2)_3, s_1 s_2] = C[(s_2^2)_3, s_1 s_2] = 0. \\
C[s_1^2, s_1 s_2, (s_2^2)_2] = C[s_2^2, s_1 s_2, (s_1^2)_2] = C[(s_1^2)_2, (s_2^2)_2] = 0.
\end{cases} \tag{B.5}$$

Nous avons :

- Pour $C[(s_1^2)_3, s_2^2]$ les partitions suivantes :

- * $\{s_1^2, s_1^2, s_1^2, s_2^2\}$ une fois,
- * $\{\{s_1^2\}, \{s_1^2\}, \{s_1^2\}, \{s_2^2\}\}$ une fois,
- * $\{\{s_1^2, s_1^2\}, \{s_1^2, s_2^2\}\}$ trois fois,
- * $\{\{s_1^2\}, \{s_1^2, s_1^2, s_2^2\}\}$ trois fois,

- * $\{\{s_2^2\}, \{s_1^2, s_1^2, s_1^2\}\}$ une fois,
- * $\{\{s_1^2\}, \{s_1^2\}, \{s_1^2, s_2^2\}\}$ trois fois,
- * et $\{\{s_1^2\}, \{s_2^2\}, \{s_1^2, s_1^2\}\}$ trois fois.

Le nombre d'éléments correspondant est 1, 4, 2, 2, 2, 3, et 3. Nous avons donc le résultat suivant :

$$\begin{aligned}
 C[(s_1^2)_3, s_2^2] &= (-1)^0 0! E[s_1^2 s_1^2 s_1^2 s_2^2] + (-1)^{4-1} (4-1)! E[s_1^2] E[s_1^2] E[s_1^2] E[s_2^2] \\
 &\quad + 3(-1)^{2-1} (2-1)! E[s_1^2 s_1^2] E[s_1^2 s_2^2] \\
 &\quad + 3(-1)^{2-1} (2-1)! E[s_1^2] E[s_1^2 s_1^2 s_2^2] \\
 &\quad + (-1)^{2-1} (2-1)! E[s_2^2] E[s_1^2 s_1^2 s_1^2] \\
 &\quad + 3(-1)^{3-1} (3-1)! E[s_1^2] E[s_1^2] E[s_2^2 s_1^2] \\
 &\quad + 3(-1)^{3-1} (3-1)! E[s_1^2] E[s_2^2] E[s_1^2 s_1^2] \\
 &= 0.
 \end{aligned} \tag{B.6}$$

- Pour $C[(s_2^2)_3, s_1^2]$ nous procédons de la même manière que ci-dessus en intervertissant s_1 et s_2 .

- Pour $C[(s_1 s_2)_3, s_1^2]$ nous avons les partitions :

- * $\{s_1 s_2, s_1 s_2, s_1 s_2, s_1^2\}$ une fois,
- * $\{\{s_1 s_2\}, \{s_1 s_2\}, \{s_1 s_2\}, \{s_1^2\}\}$ une fois,
- * $\{\{s_1 s_2, s_1 s_2\}, \{s_1 s_2, s_1^2\}\}$ trois fois,
- * $\{\{s_1 s_2\}, \{s_1 s_2, s_1 s_2, s_1^2\}\}$ trois fois,
- * $\{\{s_1^2\}, \{s_1 s_2, s_1 s_2, s_1 s_2\}\}$ une fois,
- * $\{\{s_1 s_2\}, \{s_1 s_2\}, \{s_1 s_2, s_1^2\}\}$ trois fois,
- * et $\{\{s_1^2\}, \{s_1 s_2\}, \{s_1 s_2, s_1 s_2\}\}$ trois fois.

Le nombre d'éléments est respectivement 1, 4, 2, 2, 2, 3 et 3. Chaque membre correspondant est nul car les sources sont statistiquement indépendantes, centrées et de densité symétrique, et nous obtenons :

$$C[(s_1 s_2)_3, s_1^2] = 0. \tag{B.7}$$

- Nous procédons de même pour $C[(s_1 s_2)_3, s_2^2]$ en remplaçant s_1 par s_2 .
- Pour $C[(s_1^2)_3, s_1 s_2]$ nous avons les partitions suivantes :
 - * $\{s_1^2, s_1^2, s_1^2, s_1 s_2\}$ une fois,
 - * $\{\{s_1^2\}, \{s_1^2\}, \{s_1^2\}, \{s_1 s_2\}\}$ une fois,
 - * $\{\{s_1^2, s_1^2\}, \{s_1^2, s_1 s_2\}\}$ trois fois,
 - * $\{\{s_1^2\}, \{s_1^2, s_1^2, s_1 s_2\}\}$ trois fois,
 - * $\{\{s_1 s_2\}, \{s_1^2, s_1^2, s_1^2\}\}$ une fois,
 - * $\{\{s_1^2\}, \{s_1^2\}, \{s_1^2, s_1 s_2\}\}$ trois fois,
 - * $\{\{s_1^2\}, \{s_1^2, s_1^2\}, \{s_1 s_2\}\}$ trois fois.

Le nombre d'éléments est 1, 4, 2, 2, 2, 3, et 3. Chaque membre correspondant est nul car les sources sont statistiquement indépendantes et centrées, d'où le résultat :

$$C[(s_1^2)_3, s_1 s_2] = 0. \quad (\text{B.8})$$

- Nous procédons de même pour $C[(s_2^2)_3, s_1 s_2]$ en remplaçant s_1 par s_2 .
- Pour $C[(s_1^2), s_1 s_2, (s_2^2)_2]$ nous avons les partitions
 - * $\{s_1^2, s_1 s_2, s_2^2, s_2^2\}$ une fois,
 - * $\{\{s_1^2\}, \{s_1 s_2\}, \{s_2^2\}, \{s_2^2\}\}$ une fois,
 - * $\{\{s_1^2, s_1 s_2\}, \{s_2^2, s_2^2\}\}$ une fois,
 - * $\{\{s_2^2, s_1 s_2\}, \{s_1^2, s_2^2\}\}$ deux fois,
 - * $\{\{s_1^2\}, \{s_1 s_2, s_2^2, s_2^2\}\}$ une fois,
 - * $\{\{s_2^2\}, \{s_1 s_2, s_1^2, s_2^2\}\}$ deux fois,
 - * $\{\{s_1 s_2\}, \{s_2^2, s_1^2, s_2^2\}\}$ une fois,
 - * $\{\{s_2^2\}, \{s_2^2\}, \{s_1 s_2, s_1^2\}\}$ une fois,
 - * $\{\{s_1^2\}, \{s_1 s_2\}, \{s_2^2, s_2^2\}\}$ une fois,
 - * $\{\{s_1^2\}, \{s_2^2\}, \{s_1 s_2, s_2^2\}\}$ deux fois,
 - * et $\{\{s_2^2\}, \{s_1 s_2\}, \{s_1^2, s_2^2\}\}$ deux fois.

Chaque membre correspondant est nul car les sources sont statistiquement indépendantes, centrées et de densité symétrique, et nous obtenons :

$$C[(s_1^2), s_1 s_2, (s_2^2)_2] = 0. \quad (\text{B.9})$$

- Nous raisonnons comme précédemment pour $C[s_2^2, s_1 s_2, (s_1^2)_2]$ en intervertissant s_1^2 et s_2^2 .

- Enfin pour le cumulante $C[(s_1^2)_2, (s_2^2)_2]$ nous avons comme partitions :

- * $\{s_1^2, s_1^2, s_2^2, s_2^2\}$ une fois,
- * $\{\{s_1^2\}, \{s_1^2\}, \{s_2^2\}, \{s_2^2\}\}$ une fois,
- * $\{\{s_1^2, s_1^2\}, \{s_2^2, s_2^2\}\}$ une fois,
- * $\{\{s_1^2, s_2^2\}, \{s_1^2, s_2^2\}\}$ deux fois,
- * $\{\{s_1^2\}, \{s_1^2, s_2^2, s_2^2\}\}$ deux fois,
- * $\{\{s_2^2\}, \{s_1^2, s_1^2, s_2^2\}\}$ deux fois,
- * $\{\{s_1^2\}, \{s_1^2\}, \{s_2^2, s_2^2\}\}$ une fois,
- * $\{\{s_1^2\}, \{s_2^2\}, \{s_1^2, s_2^2\}\}$ quatre fois,
- * et $\{\{s_2^2\}, \{s_2^2\}, \{s_1^2, s_1^2\}\}$ une fois

avec le nombre d'éléments respectifs 1, 4, 2, 2, 2, 2, 3, 3 et 3. Nous obtenons alors :

$$\begin{aligned}
 C[(s_1^2)_2, (s_2^2)_2] &= (-1)^0 0! E[s_1^2 s_1^2 s_2^2 s_2^2] + (-1)^{4-1} (4-1)! E[s_1^2] E[s_1^2] E[s_2^2] E[s_2^2] \\
 &\quad + (-1)^{2-1} (2-1)! E[s_1^2 s_1^2] E[s_2^2 s_2^2] \\
 &\quad + 2(-1)^{2-1} (2-1)! E[s_1^2 s_2^2] E[s_1^2 s_2^2] \\
 &\quad + 2(-1)^{2-1} (2-1)! E[s_1^2] E[s_2^2 s_1^2 s_2^2] \\
 &\quad + 2(-1)^{2-1} (2-1)! E[s_2^2] E[s_1^2 s_1^2 s_2^2] \\
 &\quad + (-1)^{3-1} (3-1)! E[s_1^2] E[s_1^2] E[s_2^2 s_2^2] \\
 &\quad + 4(-1)^{3-1} (3-1)! E[s_1^2] E[s_2^2] E[s_1^2 s_2^2] \\
 &\quad + (-1)^{3-1} (3-1)! E[s_2^2] E[s_2^2] E[s_1^2 s_1^2] \\
 &= 0. \quad (\text{B.10})
 \end{aligned}$$

Annexe C

Méthodes de Gauss-Newton et Levenberg-Marquardt

Les méthodes itératives figurent parmi les méthodes numériques puissantes permettant la résolution approchée d'équations de toute nature.

Cette annexe a été réalisée à partir de [7, 62].

C.1 Position du problème

Soit R une application de $\mathbb{R}^m \rightarrow \mathbb{R}^n$ de classe \mathcal{C}^2 . Nous nous intéressons à la résolution du système d'équations :

$$R(\mathbf{v}) = 0 \Leftrightarrow \begin{cases} r_1(v_1, \dots, v_m) = 0 \\ r_2(v_1, \dots, v_m) = 0 \\ \vdots \\ r_n(v_1, \dots, v_m) = 0 \end{cases} \quad (\text{C.1})$$

où les $r_i(\mathbf{v})_{i=1..n}$ de $\mathbb{R}^m \rightarrow \mathbb{R}$, sont des fonctions non-linéaires, appelées ici résidus et R sera le vecteur résiduel.

Optimisation au sens des moindres carrés (MC)

Dans le problème des moindres carrés, une fonction $f(\mathbf{v})$ doit être minimisée, qui est une somme de carrés de fonctions de la forme :

$$\min_{x \in \mathbb{R}^m} f(\mathbf{v}) = \frac{1}{2} \sum_i r_i(\mathbf{v})^2 = \frac{1}{2} \|R(\mathbf{v})\|_2^2 \quad (\text{C.2})$$

Certaines caractéristiques du problème peuvent souvent être exploitées afin d'améliorer l'efficacité itérative de la procédure de résolution. Le gradient et la matrice hessienne de (C.2) ont une structure particulière. En effet, les dérivées de $f(\mathbf{v})$ peuvent être exprimées en termes du jacobien $J(\mathbf{v})$ qui est la matrice $m \times n$ des dérivées partielles d'ordre 1 des résidus, définie par :

$$J(\mathbf{v}) = \left(\frac{\partial r_j}{\partial v_i} \right)_{i=1, \dots, m; j=1, \dots, n} = \begin{bmatrix} \nabla r_1(\mathbf{v})^T \\ \vdots \\ \nabla r_n(\mathbf{v})^T \end{bmatrix} \quad (\text{C.3})$$

où $\nabla r_j(\mathbf{v})$, $j = 1, \dots, m$ est le gradient de r_j . Le gradient et le hessien de f peuvent alors être exprimés comme suit :

$$\nabla f(\mathbf{v}) = \sum_{j=1}^m r_j(\mathbf{v}) \nabla r_j(\mathbf{v}) = J(\mathbf{v})^T R(\mathbf{v}) \quad (\text{C.4})$$

$$\begin{aligned} \nabla^2 f(\mathbf{v}) &= \sum_{j=1}^m \nabla r_j(\mathbf{v}) \nabla r_j(\mathbf{v})^T + \sum_{j=1}^m r_j(\mathbf{v}) \nabla^2 r_j(\mathbf{v}) \\ &= J(\mathbf{v})^T J(\mathbf{v}) + \sum_{j=1}^m r_j(\mathbf{v}) \nabla^2 r_j(\mathbf{v}) \end{aligned} \quad (\text{C.5})$$

Dans la plupart des applications, les premières dérivées partielles des résidus, et donc la matrice jacobienne $J(\mathbf{v})$, sont relativement faciles à calculer. Nous pouvons alors obtenir le gradient $\nabla f(\mathbf{v})$ (formule C.4). En utilisant $J(\mathbf{v})$, nous pouvons calculer le premier terme $J(\mathbf{v})^T J(\mathbf{v})$ dans le hessien $\nabla^2 f(\mathbf{v})$, sans évaluer les dérivées secondes des r_j . En effet le deuxième terme dans (C.5) est généralement très coûteux à calculer en pratique, vu qu'il implique les dérivées secondes des r_j . Il est alors judicieux de l'ignorer et d'approcher la matrice hessienne par le premier terme uniquement, i.e. $J(\mathbf{v})^T J(\mathbf{v})$.

C.2 Méthode de Gauss-Newton (GN)

Il s'agit d'une méthode pour minimiser la fonction coût (C.2), qui exploite la structure du gradient $\nabla f(\mathbf{v})$ (C.4) et du hessien $\nabla^2 f(\mathbf{v})$ (C.5). Cette méthode peut être vue comme une version modifiée de la méthode de Newton avec ligne de recherche : au lieu de résoudre les équations standard de Newton (voir (C.4)) :

$$\nabla^2 f(\mathbf{x}_k) d_k = -\nabla f(\mathbf{x}_k) \quad (\text{C.6})$$

nous résolvons le système suivant afin d'obtenir la direction de recherche \mathbf{d}_k^{GN} :

$$J_k^T J_k \mathbf{d}_k^{\text{GN}} = -J_k \mathbf{R}_k \quad (\text{C.7})$$

Cette simple modification donne un certain nombre d'avantages par rapport à la méthode de Newton :

- Notre utilisation de l'approximation du hessien

$$\nabla^2 f(\mathbf{x}_k) \approx J_k^T J_k \quad (\text{C.8})$$

nous permet de ne pas avoir à calculer les résidus individuels du hessien $\nabla^2 r_j$, $j = 1, \dots, m$ nécessaires dans le second terme de (C.5).

En effet, si nous calculons le jacobien J_k afin d'évaluer $\nabla f(\mathbf{x}_k) = J_k^T \mathbf{R}_k$, l'approximation (C.8) ne nécessitera pas une évaluation supplémentaire des dérivées.

- Il existe plusieurs situations où le premier terme $J_k^T J_k$ de (C.5) domine le second terme (au moins au voisinage de la solution \mathbf{x}^*), donc $J_k^T J_k$ est une bonne approximation de $\nabla^2 f(\mathbf{x}_k)$.

Le premier terme de (C.5) sera dominant lorsque la norme de chaque terme de second ordre ($|r_j(\mathbf{v})| |\nabla^2 r_j(\mathbf{v})|$) est significativement inférieure aux valeurs propres de $J^T J$.

- L'autre avantage de la méthode de GN est que si J_k est de rang plein et le gradient $\nabla f(\mathbf{x}_k) \neq 0$, la direction \mathbf{d}_k^{GN} est une direction de descente pour f , et donc une direction souhaitable comme ligne de recherche. En effet nous avons à partir de (C.4) et (C.7) :

$$(\mathbf{d}_k^{\text{GN}})^T \nabla f(\mathbf{x}_k) = (\mathbf{d}_k^{\text{GN}})^T J_k^T \mathbf{R}_k = -(\mathbf{d}_k^{\text{GN}})^T J_k^T J_k (\mathbf{d}_k^{\text{GN}}) = -\|J_k (\mathbf{d}_k^{\text{GN}})\|^2. \quad (\text{C.9})$$

- le terme \mathbf{d}_k^{GN} est la solution du problème des MC linéaires :

$$\min_{\mathbf{d}} \frac{1}{2} \|J_k \mathbf{d} + \mathbf{R}_k\|^2 \quad (\text{C.10})$$

C.3 Méthode de Levenberg-Marquardt (LM)

Nous rappelons que la méthode de GN est équivalente à la méthode de Newton avec ligne de recherche, excepté que nous utilisons l'approximation plus appropriée (C.8) pour le hessien.

La méthode de Levenberg-Marquardt peut être obtenue à partir de la même approximation du hessien, mais en remplaçant la ligne de recherche par une stratégie de *région de confiance*. L'utilisation d'une région de confiance permet d'éviter l'une des faiblesses de l'algorithme de GN spécifiquement son comportement lorsque le jacobien $J(v)$ est déficient au niveau du rang. Comme les approximations du hessien sont utilisées, les propriétés de convergence locale des deux algorithmes sont les mêmes.

Pour une région de confiance sphérique, le sous-problème à résoudre à chaque itération est le suivant :

$$\min_{\mathbf{d}} \frac{1}{2} \|J_k \mathbf{d} + \mathbf{R}_k\|^2, \text{ sous la contrainte } \|\mathbf{d}\| < \delta_k \quad (\text{C.11})$$

où $\delta_k > 0$ est le rayon de la région de confiance.

En effet, si nous choisissons un modèle de fonction $m_k(\cdot)$

$$m_k(\mathbf{d}) = \frac{1}{2} \|\mathbf{R}_k\|^2 + \mathbf{d}^T J_k^T \mathbf{R}_k + \frac{1}{2} \mathbf{d}^T J_k^T J_k \mathbf{d} \quad (\text{C.12})$$

Lorsque la solution $(\mathbf{d}_k^{\text{GN}})^T$ de l'algorithme de GN appartient à la région de confiance (i.e. $\|\mathbf{d}_k^{\text{GN}}\| < \delta_k$), alors cette étape résout également le problème (C.11). C'est-à-dire qu'il existe un $\lambda > 0$ tel que la solution $\mathbf{d} = \mathbf{d}_k^{\text{LM}}$ de (C.11) satisfait $\|\mathbf{d}\| = \delta_k$ et

$$(J_k^T J_k + \lambda I) \mathbf{d} = -J_k^T \mathbf{R}_k \quad (\text{C.13})$$

Nous avons alors le lemme suivant :

Lemme :

Le vecteur \mathbf{d}_k^{LM} est une solution du sous-problème de région de confiance :

$$\min_{\mathbf{d}} \|J\mathbf{d} + R\|^2, \quad \|\mathbf{d}\| \leq \delta \quad (\text{C.14})$$

ssi \mathbf{d}_k^{LM} est réalisable et qu'il existe un scalaire $\lambda > 0$ tel que

$$(J^T J + \lambda I) \mathbf{d}_k^{\text{LM}} = -J^T R \quad (\text{C.15})$$

$$\lambda(\delta - \|\mathbf{d}_k^{\text{LM}}\|) = 0 \quad (\text{C.16})$$

C.4 Rappel sur la méthode de Newton locale

L'algorithme peut être résumé comme suit :

Objectif : Trouver (l'approximation de) la solution du système :

$$\nabla f = 0 \tag{C.17}$$

Entrées :

- gradient de la fonction $\nabla f : \mathbb{R}^n \rightarrow \mathbb{R}^n$
- Hessian de la fonction $\nabla^2 f : \mathbb{R}^n \rightarrow \mathbb{R}^{n \times n}$
- première approximation de la solution $x_0 \in \mathbb{R}^n$
- la précision demandée $\epsilon \in \mathbb{R}, \epsilon > 0$

Sortie : Approximation de la solution $x^* \in \mathbb{R}^n$.

Initialisation : $k = 0$

Itération :

1. calculer d_k solution de $\nabla^2 f(x_k)d_k = -\nabla f(x_k)$
2. $x_{k+1} = x_k + d_k$
3. $k = k + 1$

Condition d'arrêt : Si $\|\nabla f(x_k)\| \leq \epsilon$, alors $x^* = x_k$

Annexe D

Fonctions MATLAB

Dans cette annexe nous donnons la description rapide des différentes fonctions Matlab utilisées dans le Chapitre 3. Plus de détails peuvent être trouvés dans [20, 76].

Cette annexe a été construite à partir de [20, 76].

D.1 Estimation des cumulants

Les fonctions Matlab utilisées pour estimer les cumulants des observations sont issues de la toolbox « HOSA », et sont : *cumest*, pour estimer les auto-cumulants ainsi que *cum2x*, *cum3x* et *cum4x* pour estimer les cumulants croisés des observations. Leur descriptif est le suivant :

- *cumest* :

(a) **But** *cumest* calcule les estimées de cumulants en utilisant la méthode du recouvrement de segments.

(b) **Syntaxe**

$$\begin{cases} cvec = cumest(y) \\ cvec = cumest(y, norder, maxlag, samp_{seg}, overlap, flag, k1, k2) \end{cases} \quad (D.1)$$

(c) **Description** : *cumest* calcule les estimées 1-D des cumulants du processus y .

(d) **Arguments d'entrée** :

- y est une matrice de données ou un vecteur.

- *norder* spécifie l'ordre du cumulants, et peut être 2, 3, ou 4 ; la valeur par défaut est 2.
- *maxlag* spécifie le décalage maximum du cumulants à calculer ; sa valeur par défaut est 0.
- *samp_{seg}* donne le nombre d'échantillons par segment ; la valeur par défaut est la longueur de la série temporelle.
- *overlap* donne le pourcentage de recouvrement entre segment ; la valeur maximum autorisée est 99 ; la valeur par défaut est 0.
- Si *flag* est biaisé, alors des estimées biaisées sont calculées (par défaut) ; si la première lettre n'est pas le 'b', des estimées non biaisées sont calculées. Si *y* est une matrice, les colonnes sont supposées correspondre à des réalisations indépendantes ; dans ce cas *overlap* est mis à zéro, et *samp_{seg}* est mis à la dimension de la ligne. Les cumulants sont estimés pour chaque réalisation, puis moyennés.
- Les paramètres *k1* et *k2* contrôlent quelle partie 1-D de la fonction cumulants est calculée ; leurs valeurs par défaut sont nulles.

(e) **Argument de sortie** : *cvec* contiendra $C2(m)$, $C3(m, k1)$ ou $C4(m, k1, k2)$, $m = -maxlag \dots maxlag$, dépendant de l'ordre du cumulants spécifié. Il faut noter que *cumest* estime une tranche 1-D du cumulants.

- cum2x :

(a) **But** cum2x calcule le cumulants croisé de deux signaux.

(b) **Syntaxe**

$$cvec = cum2x(x, y, maxlag, samp_{seg}, overlap, flag) \quad (D.2)$$

(c) **Description** : Calcule le cumulants croisé d'ordre deux des deux signaux, *x* et *y*.

(d) **Arguments d'entrée** :

- *x*, *y* doivent avoir même dimension.
- *maxlag* spécifie le décalage maximum du cumulants à calculer ; sa valeur par défaut est 0.

- $samp_{seg}$ spécifie le nombre d'échantillons par segment, sa valeur par défaut est la dimension de la ligne y ; si y est un vecteur ligne, la dimension de la colonne est utilisée comme valeur par défaut.
- $overlap$ donne le pourcentage de recouvrement entre segments; les valeurs autorisées sont comprises entre $[0,99]$; la valeur par défaut est 0.
- Si $flag$ est biaisé, alors des estimées biaisées sont calculées (par défaut); si la première lettre n'est pas le 'b', des estimées non biaisées sont calculées.

(e) **Argument de sortie** : $cvec$ contiendra les estimées de

$$E[(x^*(n) - \mu_x^*)(y(n+m) - \mu_y)], \quad m = -maxlag \dots maxlag \quad (D.3)$$

où $E[\cdot]$ désigne l'opérateur espérance, μ_x est la moyenne du processus x , et $*$ dénote le complexe conjugué.

Si x , y sont des matrices, les colonnes sont supposées correspondre à différentes réalisations; dans ce cas, $overlap$ est mis à zéro, et $samp_{seg}$ à la dimension ligne de la matrice; le cumulants croisés est estimé à partir de chaque réalisation, puis moyenné.

- **cum3x** :

(a) **But** cum3x calcule le cumulants croisés d'ordre trois de trois signaux.

(b) **Syntaxe**

$$cvec = cum3x(x, y, z, maxlag, samp_{seg}, overlap, flag, k1) \quad (D.4)$$

(c) **Description** calcule le cumulants croisés d'ordre trois des trois signaux. x , y , et z , qui doivent avoir mêmes dimensions.

(d) **Arguments d'entrée** :

- $maxlag$, $samp_{seg}$, $overlap$, $flag$: voir ci-dessus.
- le paramètre k_1 contrôle quel segment 1-D est calculé. Sa valeur par défaut est 0. En faisant varier k_1 , nous pouvons obtenir tous les cumulants croisés d'ordre trois.

- (e) **Argument de sortie** : *cvec* contiendra les estimées de

$$E[(x^*(n) - \mu_x^*)(y(n+m) - \mu_y)(z(n+k1) - \mu_z)], \quad (D.5)$$

où, $m = -maxlag, \dots, maxlag$. Les μ_r désignent les moyennes des processus r . Si x, y, z sont des matrices, les colonnes sont supposées correspondre à différentes réalisations ; dans ce cas, *overlap* est mis à zéro, et *samp_{seg}* à la dimension ligne de la matrice ; le cumulante croisé est estimé à partir de chaque réalisation, puis moyenné.

- cum4x :

- (a) **But** cum4x calcule le cumulante croisé d'ordre quatre de quatre signaux.

- (b) **Syntaxe**

$$cvec = cum4x(w, x, y, z, maxlag, samp_{seg}, overlap, flag, k1, k2) \quad (D.6)$$

- (c) **Description** calcule le cumulante croisé d'ordre quatre des quatre signaux. w, x, y , et z , qui doivent avoir mêmes dimensions.

- (d) **Arguments d'entrée** :

- *maxlag, samp_{seg}, overlap, flag* : voir ci-dessus.
- Les paramètres $k1$ et $k2$ contrôlent quel segment 1-D du cumulante croisé est calculé, la valeur par défaut est pour les deux paramètres 0. En faisant varier $k1$ et $k2$ nous pouvons obtenir tous les cumulants croisés d'ordre quatre.

- (e) **Argument de sortie** : *cvec* contiendra les estimées de

$$cum(w^*(n), x(n+m), y(n+k1), z^*(n+k2)), \quad (D.7)$$

avec $m = -maxlag \dots maxlag$, * le complexe conjugué. Si w, x, y, z sont des matrices, les colonnes sont supposées correspondre à différentes réalisations ; dans ce cas, *overlap* est mis à zéro, et *samp_{seg}* à la dimension ligne de la matrice ; le cumulante croisé est estimé à partir de chaque réalisation, puis moyenné.

D.2 Résolution du système non-linéaire

La fonction Matlab utilisée pour résoudre le système d'équations non linéaires est la fonction FSOLVE de la toolbox « Optimization », dont le bref descriptif est le suivant :

1. **But** : résout des systèmes d'équations non-linéaires de plusieurs variables.

La fonction FSOLVE tente de résoudre des équations de la forme :

$$F(x) = 0 \quad (\text{D.8})$$

en x , où x est un vecteur et $F(x)$ est une fonction qui retourne un vecteur de valeurs.

2. **Syntaxe** : les différents appels de la fonction FSOLVE peuvent être

$$\left\{ \begin{array}{l} x = fsolve(fun, x_0) \\ x = fsolve(fun, x_0, options) \\ x = fsolve(fun, x_0, options, P1, P2, \dots) \\ [x, fval] = fsolve(\dots) \\ [x, fval, exitflag] = fsolve(\dots) \\ [x, fval, exitflag, output] = fsolve(\dots) \\ [x, fval, exitflag, output, jacobian] = fsolve(\dots) \end{array} \right. \quad (\text{D.9})$$

3. **Description** : FSOLVE trouve les racines (zéros) du système d'équations non-linéaires.

4. **Arguments d'entrée** : Nous donnons les détails des arguments pour fun , x_0 , et $options$.

- L'entrée fun est la fonction qui doit être minimisée.
- L'entrée x_0 est le point d'initialisation de la fonction FSOLVE.
- L'entrée $options$ contient les paramètres d'optimisation par défaut et remplacés par les valeurs contenues dans la structure OPTIONS, argument pouvant être créé ou modifié avec la fonction OPTIMSET (voir l'aide OPTIMSET pour les détails). Les options utilisées sont Display, TolX,

TolFun, DerivativeCheck, Diagnostics, FunValCheck, Jacobian, JacobMult, JacobPattern, LineSearchType, NonlEqnAlgorithm, MaxFunEvals, MaxIter, PlotFcns, OutputFcn, DiffMinChange and DiffMaxChange, LargeScale, MaxPCGIter, PrecondBandWidth, TolPCG, and TypicalX.

5. **Arguments de sortie** : les arguments de sortie peuvent être x , $fval$, $exitflag$, et $output$.

- La sortie x est la solution du système.
- La sortie $fval$ est la valeur de la fonction F au point x .
- La sortie $output$ est une structure avec entre autres : le nombre d'itérations contenu dans OUTPUT.iterations, le nombre d'évaluations de la fonction contenu dans OUTPUT.funcCount, l'algorithme utilisé dans OUTPUT.algorithm, et le message de sortie dans OUTPUT.message.
- La sortie $exitflag$ décrit les conditions de sortie de FSOLVE. Les valeurs possibles d'EXITFLAG et les conditions de sortie correspondantes sont :
 - * 1 : FSOLVE a convergé vers une solution X .
 - * 2 : Changement en X plus petit que la tolérance spécifiée.
 - * 3 : Changement dans le résidu plus petit que la tolérance spécifiée.
 - * 4 : Longueur de la direction de recherche plus petite que la tolérance spécifiée.
 - * 0 : Nombre maximum d'évaluations de la fonction atteint.
 - * -1 : L'algorithme se termine par la fonction de sortie.
 - * -2 : L'algorithme semble converger vers un point qui n'est pas une racine.
 - * -3 : Le rayon de la région de confiance est trop petit.
 - * -4 : La ligne de recherche ne peut décroître suffisamment le résidu le long de la direction de recherche courante.

Bibliographie

- [1] K. Abed-Meraim, A. Belouchrani, and Y. Hua: *Blind identification of a linear-quadratic mixture of independent components based on joint diagonalization procedure*. IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP-96), 1996.
- [2] S. Achard and C. Jutten: *Identifiability of post-nonlinear mixtures*. IEEE Signal Processing Letters, vol. 12(5):pp 423–426, 2005.
- [3] L. B. Almeida: *Linear and nonlinear ICA based on mutual information – the MISEP method*. Signal Processing, vol. 84:pp. 231–245, 2004.
- [4] L. B. Almeida: *Separating a real-life nonlinear image mixture*. Journal of Machine Learning Research, vol. 6, 2005.
- [5] S. I. Amari, A. Cichocki, and H. H. Yang: *A new learning algorithm for blind source separation*. Proceedings of advances in Neural Information Processing Systems 8, NIPS, pages 757–763, 1995.
- [6] A. Belouchrani and J. F. Cardoso: *Maximum likelihood source separation for discrete sources*. Proceedings of the European Signal Processing Conference (EUSIPCO94), Vol.2:pp. 768–771, 1994.
- [7] M. Berliaire: *Introduction à l'optimisation différentiable*. Presses Polytechniques et Universitaires Romandes, 2006.
- [8] I. V. Blagouchine and E. Moreau: *Unbiased adaptive estimations of the fourth-order cumulant for real random zero-mean signals*. IEEE Trans. on Signal Processing,, vol. 57 (9):pp 3330–3346, 2009.
- [9] G. Burel: *Blind separation of sources: a nonlinear neural algorithm*. Neural Networks, vol. 5(6):pp. 937–947, 1992.

- [10] A. Taleb et C. Jutten: *Séparation de sources dans les mélanges post-non-linéaires*. Seizième colloque du Grets, Sept. 1997.
- [11] P. Comon et C. Jutten: *Séparation de sources 1 concepts de base et analyse en composantes indépendantes*. Hermès, 2007.
- [12] P. Comon et C. Jutten: *Séparation de sources 2 au-delà de l'aveugle et applications*. Hermès, 2007.
- [13] J. F. Cardoso: *A blind source separation technique using second order statistics*. IEEE Trans. on Signal Processing, 45(2):434–444, 1997.
- [14] J. F. Cardoso: *Infomax and maximum likelihood for blind source separation*. IEEE Signal Processing Letters, 1997.
- [15] J. F. Cardoso: *Blind signal separation: statistical principles*. Proceedings of the IEEE, vol. 86(10):pp. 2009–2025, Oct. 1998.
- [16] J. F. Cardoso and A. Souloumiac: *Blind beamforming for non gaussian signals*. IEE Proceedings-F, 140(6):362–370, 1993.
- [17] M. Castella: *Inversion of polynomials systems and separation of nonlinear mixtures of finite-alphabet sources*. IEEE Trans. on Signal Processing, vol. 56, Aug. 2008.
- [18] A. Cichocki, W. Kasprzak, and S. I. Amari: *Multi-layer neural networks with a local adaptive learning rule for blind separation of source signals*. Proceedings of NOLTA'95, pages 61–65, 1995.
- [19] A. Cichocki, R. Zdunek, A. H. Phan, and S. I. Amari: *Nonnegative Matrix and Tensor Factorizations: Applications to Exploratory Multi-Way Data Analysis and Blind Source Separation*. John Wiley and Sons, 2009.
- [20] T. Coleman, M. A. Branch, and A. Grace: *Optimization toolbox user's guide— for use with matlab*.
- [21] P. Comon: *Independent component analysis*. Proceedings of International Signal Processing Workshop on Higher-Order Statistics, pages 111–120, 1991.

-
- [22] P. Comon: *Independent component analysis, a new concept?* Signal Processing, vol. 36:pp. 287–314, 1994.
- [23] P. Comon and C. Jutten: *Handbook of Blind Source separation - Independent Component Analysis and applications*. Academic Press, 2010.
- [24] Y. Deville: *Méthode et système de caractérisation et de séparation de signaux mélangés*. Note technique pour brevet,, TE 5250.
- [25] Y. Deville: *System for estimating signals received in the form of mixed signals*. United States Patent US005909646A.
- [26] Y. Deville: *Méthode de séparation de sources pour mélanges linéaires-quadratiques (A source separation method for linear-quadratic mixtures), in french*. Private Communication, Sept. 2000.
- [27] Y. Deville and S. Hosseini: *Blind operation of a recurrent neural network for linear-quadratic source separation: fixed points, stabilization and adaptation scheme*. LVA/ICA 2010, LNCS 6365:237–244, 2010.
- [28] Y. Deville and S. Hosseini: *Recurrent networks for separating extractable-target nonlinear mixtures. Part I: Non-blind configurations*. Signal Processing, vol. 89, no. 4:pp. 378–393, April 2009.
- [29] Y. Deville and S. Hosseini: *Blind identification and separation methods for linear-quadratic mixtures and/or linearly independent non-stationary signals*. Proceedings of the 9th International Symposium on Signal Processing and its Applications (ISSPA 2007), Feb. 2007.
- [30] Y. Deville, S. Hosseini, and A. Deville: *Effect of indirect dependencies on maximum likelihood and information theoretic blind source separation for nonlinear mixtures*. Signal Processing, Vol.91(4):pp. 793–800, 2011.
- [31] L. T. Duarte and C. Jutten: *Blind source separation of a class of nonlinear mixtures*. Proc.2007 ICA, London, UK, September 9-12, pages 41–48, 2007.
- [32] L. T. Duarte and C. Jutten: *A mutual information minimization approach for a class of nonlinear recurrent separating systems*. IEEE Workshop on Machine Learning for Signal Processing, pages 122 – 127, 2007.

- [33] J. Eriksson and V. Koivunen: *Blind separation of a class of nonlinear ICA models*. IN Proc. IEEE Int. Symp. Circuits Systems (ISCAS), vol. 6:pp.5890–5893, May 2005.
- [34] J. Eriksson and V. Koivunen: *Blind identifiability of class of nonlinear instantaneous ica models*. Proc. of XI European Signal Processing Conference (EUSIPCO 2002), pages 7–10, Sept. 2002.
- [35] F.Merrikh-Bayat, M. Babaie-Zadeh, and C. Jutten: *A nonlinear blind source separation solution for removing the show-through effect in the scanned documents*. European Signal Processing Conference, EUSIPCO 2008, 2008.
- [36] J. C Fort: *Stabilité de l’algorithme de séparation de sources Jutten et Héroult*. Traitement du signal, vol. 8:pp. 35–42, 1991.
- [37] M. Gaeta and J. L. Lacoume: *Estimateurs du maximum de vraisemblance étendus à la séparation de sources non gaussiennes*. Traitement du signal, vol. 7(5):pp. 419–434, 1990.
- [38] M. Gaeta and J. L. Lacoume: *Source separation without prior knowledge: The maximum likelihood solution*. Proceedings of the European Signal Processing Conference (EUSIPCO’90), pages 621–624, 1990.
- [39] R. Guidara: *Méthodes markoviennes pour la séparation aveugle de signaux et images*. PhD thesis, Université Toulouse III- Paul Sabatier, 2009.
- [40] S. Hosseini and Y. Deville: *Blind separation of linear-quadratic mixtures of real sources using a recurrent structure*. Proc. of the 7th International Work-Conference on Artificial and Natural Neural Networks (IWANN 2003), vol. 2:pp. 241–248, 2003.
- [41] S. Hosseini and Y. Deville: *Blind maximum likelihood separation of a linear-quadratic mixture*. Proc. of ICA 2004., vol. LNCS 3195:pp. 694–701, Erratum:<http://arxiv.org/abs/1001.0863>, 2004.
- [42] A. Hyvarinen, J. Karhunen, and E. Oja: *Independent Component Analysis*. Wiley-Interscience Publication, 2001.

-
- [43] A. Hyvarinen and P. Pajunen: *Nonlinear independent component analysis: Existence and uniqueness results*. Neural Networks, vol. 12(3):pp. 429–439, 1999.
- [44] C. Jutten et J. Héroult: *Une solution neuromimétique au problème de séparation de sources*. Traitement du Signal – Machines Numériques, vol. 5(6):pp. 389–403, 1988.
- [45] C. Jutten et J. Héroult: *Blind separation of sources, Part I: An adaptive algorithm based on neuromimetic architecture*. Signal Processing, vol. 24:pp. 1–10, 1991.
- [46] C. Jutten, P. Comon, and J. Héroult: *Blind separation of sources, Part II: Problem statement*. Signal Processing, vol. 24:pp. 11–20, 1991.
- [47] C. Jutten and J. Karhunen: *Advances in nonlinear blind source separation*. 4th International Symposium on Independent Component Analysis and Blind Source Separation (ICA 2003), 2003.
- [48] C. Jutten and A. Taleb: *Source separation: From dusk till dawn*. ICA 2000 (invited paper), pages 15–26, 2000.
- [49] M. Krob and M. Benidir: *Blind identification of a linear-quadratic model using higher-order statistics*. IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP-93), vol. 4:pp. 440 – 443, 1993.
- [50] T. W. Lee: *Nonlinear approaches to independent component analysis*. Proc. of the American Institute of Physics, Oct. 1999.
- [51] T. W. Lee, B. Koehler, and R. Orglmeister: *Blind source separation of nonlinear mixing models*. in IEEE NNSP, pages 406–415, 1997.
- [52] O. Macchi and E. Moreau: *Self-adaptive source separation, Part I: Convergence analysis of a direct linear network controlled by the Héroult-Jutten algorithm*. IEEE Transactions on Signal Processing, vol. 45(4):pp. 918–926, April 1997.
- [53] M. Babaie Zadeh Malmiri: *On blind source separation in convolutive and non-linear mixtures*. PhD thesis, INPG and SHARIF University, 2002.

- [54] A. Mansour: *Contribution à la séparation aveugle de sources*. PhD thesis, INPG, 1997.
- [55] A. Mansour, A. Kardec Barros, and N. Ohnishi: *Comparison among three estimators for higher order statistics*. Fifth International Conference on Neural Information Processing, pages 899–902, Oct. 1998.
- [56] A. Mansour and C. Jutten: *A direct solution for blind separation of sources*. IEEE Transactions on Signal Processing, vol. 44(3):pp 746–748, 1996.
- [57] A. Mertins and I. Russel: *An extended acdc algorithm for the blind estimation of convolutive mixing systems*. Proceedings of the Seventh International Symposium on Signal Processing and its Applications (ISSPA 2003), vol. 2:pp 527–530, 2003.
- [58] E. Moreau: *Une généralisation de l’algorithme jade*. Dix-septième colloque du Grets, 1999.
- [59] E. Moreau and O. Macchi: *Self-adaptive source separation, Part II: Comparison of the direct, feedback, and mixed linear network*. IEEE Transactions on Signal Processing, vol. 46(1):pp. 39–50, Jan. 1998.
- [60] Y. Naudet: *Application de la Séparation de Sources au traitement de la réponse d’un capteur à courants de Foucault, étude et compensation de perturbations liées à la température*. PhD thesis, Université de Reims Champagne-Ardenne, 2001.
- [61] C. L. Nikias and J. M. Mendel: *Signal processing with higher-order spectra*. IEEE Signal Processing Magazine, July 1993.
- [62] J. Nocedal and S. J. Wright: *Numerical optimization*. Second edition, Springer series in Operations Research, 2006.
- [63] J. L. Lacoume et P.-O. Amblard et P. Comon: *Statistiques d’ordre supérieur pour le traitement de signal*. Masson, Paris, 1997.
- [64] P. Pajunen, A. Hyvarinen, and J. Karhunen: *Nonlinear blind source separation by Self-Organizing Maps*. Proceedings of International Conference on Neural Information, vol. 2:pp. 1207–1212, 1996.

-
- [65] P. Pajunen and J. Karhunen: *A maximum likelihood approach to nonlinear blind source separation*. In proceedings of the Int. Conf. on Artificial Neural Networks (ICANN'97), 1997.
- [66] L. Parra, G. Deco, and S. Miesbach: *Redundancy reduction with information preserving nonlinear maps*. in Networks, vol. 6:pp. 61–72, 1995.
- [67] L. Parra, C. Spence, and B. De Vries: *Convolutive source separation and signal modeling with ML*. Proceedings of the International Symposium on Intelligent Systems (ISIS 1997), 1997.
- [68] B. Pearlmutter and L. Parra: *A context sensitive generalization of independent component analysis*. Proceedings of the International Conference on Neural Information Processing (ICONIP1996), 1996.
- [69] D. T. Pham: *Blind separation of instantaneous mixture of sources via an independent component analysis*. IEEE Trans. on Signal Processing, Vol.44(11):pp. 2768–2779, 1996.
- [70] D. T. Pham: *Fast algorithms for estimating mutual information entropies and score functions*. ICA2003, pages 17–22, 2003.
- [71] D. T. Pham: *Bayesian blind separation of generalized hyperbolic processes in noisy and underdeterminate mixtures*. IEEE Trans. on Signal Processing, Vol.54(9):pp. 3257–3269, 2006.
- [72] D. T. Pham: *Fast algorithms for mutual information based independent component analysis*. IEEE Trans. on Signal Processing, Vol.52(10):pp. 2690–2700, October 2004.
- [73] D. T. Pham and P. Garat: *Blind separation of mixture of independent sources through a quasi-maximum likelihood approach*. IEEE Transactions on Signal Processing, vol. 45 (7):pp. 1712–1725, July 1997.
- [74] M. Puigt: *Méthodes de séparation aveugle de sources fondées sur des transformées temps-fréquence. Application à des signaux de parole*. PhD thesis, Université Toulouse III- Paul Sabatier, 2007.

- [75] E. Soroucyari: *Blind separation of sources, Part III: Stability analysis*. Signal Processing, vol. 24:pp. 21–29, 1991.
- [76] A. Swami, J. M. Mendel, and C. L. Nikias: *Higher-order spectral analysis toolbox user's guide— for use with matlab*.
- [77] A. Taleb and C. Jutten: *Nonlinear source separation: The post-nonlinear mixtures*. ESANN'1997, pages 279–284, 1997.
- [78] A. Taleb and C. Jutten: *Source separation in post-nonlinear mixtures*. IEEE Transactions on Signal Processing, pages 2807–2820, 1999.
- [79] F. J. Theis and W. Nakamura: *Quadratic independent component analysis*. IE-ICE Trans. Fundamentals, Special Section on Nonlinear Theory and its Applications, vol. E87-A (9), 2004.
- [80] J. Thomas: *Algorithmes temporels rapides à point fixe pour la séparation aveugle de mélanges convolutifs et/ou sous-déterminés*. PhD thesis, Université Toulouse III- Paul Sabatier, 2007.
- [81] J.M.T. Thompson and H.B. Stewart: *Nonlinear dynamics and chaos*. Wiley, Chichester, England, 2002.
- [82] L. Tong, V.C. Soon, Y. F. Huang, and R. Liu: *Amuse: a new blind identification algorithm*. Proc. IEEE ISCAS, pages 1784–1787, 1990.
- [83] H. Valpola, X. Giannakopoulos, A. Honkela, and J. Karhunen: *Nonlinear independent component analysis using ensemble learning: Experiments and discussion*. In Proceeding of International Workshop on Independent Component Analysis and Blind Signal Separation (ICA 2000), pages 351–356, 2000.
- [84] H. H. Yang, S. I. Amari, and A. Cichocki: *Information back-propagation for blind separation of sources in non-linear mixture*. International Conference on Neural Networks, vol. 4:pp. 2141–2146, 1997.
- [85] A. Yeredor: *Non-orthogonal joint diagonalization in the least-squares sense with application in blind source separation*. IEEE Trans. on Signal Processing, vol. 50 (7):pp 1545–1551, 2002.

-
- [86] L. Zhang, S.I. Amari, and A. Cichocki: *Equi-convergence algorithm for blind separation of sources with arbitrary distributions*. IWANN'01, pages 826–833, 2001.

Publications scientifiques de l'auteur

Actes de conférences internationales à comité de lecture

C. Chaouchi, Y. Deville and S. Hosseini, *Nonlinear Source Separation : a Quadratic Recurrent Inversion Structure*, Proceedings of the 9th International Workshop on Electronics, Control, Modelling, Measurement and Signals (ECMS 2009), Mondragon, Spain, July 8-10, 2009.

C. Chaouchi, Y. Deville and S. Hosseini, *Nonlinear Source Separation : A Maximum Likelihood Approach for Quadratic Mixtures*, 30th International Workshop on Bayesian Inference and Maximum Entropy Methods in Science and Engineering (MaxEnt 2010), Chamonix, France, July 4-9, 2010.

C. Chaouchi, Y. Deville and S. Hosseini, *Cumulant-Based Estimation of Quadratic Mixture Parameters for blind source separation*, European Signal Processing Conference (EUSIPCO 2010), Aalborg, Denmark, August 23-27 2010.

Actes de conférences nationales à comité de lecture

C. Chaouchi, Y. Deville et S. Hosseini, *Une structure récurrente pour la séparation de mélanges quadratiques*, (GRETSI'09), Dijon, France, sept. 2009.

Résumé

Méthodes de séparation aveugle de sources non-linéaires , étude du modèle quadratique

Résumé

Cette thèse présente des méthodes de séparation aveugle de sources pour un modèle de mélange non-linéaire particulier, le cas quadratique avec auto-termes et termes croisés.

Dans la première partie, nous présentons la structure de séparation étudiée qui se décline sous deux formes : étendue ou basique. Les propriétés de ce réseau récurrent sont ensuite analysées (points d'équilibre, stabilité locale).

Nous proposons alors deux méthodes de séparation aveugle de sources. La première exploite les cumulants des observations en un bloc placé en amont de la structure récurrente. La deuxième méthode est basée sur une approche par maximum de vraisemblance. Le tout est validé par des simulations numériques.

Mots-clés : Séparation Aveugle de Sources (SAS), Analyse en Composantes Indépendantes (ACI), mélange non-linéaire, mélange quadratique, cumulants, réseau récurrent, stabilité locale, maximum de vraisemblance.