


ECOLE DOCTORALE SCIENCES ET INGENIERIE
De l'Université de Cergy-Pontoise

THÈSE

Présentée pour obtenir le grade de docteur de l'université de Cergy-Pontoise
Spécialité : Traitement des Images et du Signal

Méthodes PARAFAC pour la séparation de signaux

par
Joséphine Castaing

Laboratoire des Equipes de Traitement des Images et du Signal - UMR 8051

20 Octobre 2006

Devant le jury composé de :

M. L. De Lathauwer,	Directeur de Thèse
Mme I. Fijalkow,	Directrice de Thèse
M. P. Comon,	Rapporteur
M. E. Moreau,	Rapporteur
Mme. A. Ferreol,	Examinatrice
M. J.-F. Cardoso,	Examinateur
M. G. Tantot,	Examinateur

Table des matières

Table des figures

Liste des tableaux

Acronymes

DS-CDMA Direct Sequence - Code Division Multiple Acces
UMTS Universal Mobile Telecommunication System
MMSE Minimum Mean Sqare Error
RSB Rapport Signal sur Bruit
TES Taux d'Erreur Symbole
PARAFAC Parallel Factor Analysis
QPSK Quadrature Phase Shift Keying

ICI Inter-Chip Interference
ISI Inter-Symbol Interference
MC Module Constant

Notations

$\mathbb{R}$ désigne l'ensemble des nombres réels. $\mathbb{C}$ désigne l'ensemble des nombres complexes.
On désignera par $[1:R]$ l'ensemble des entiers de 1 à R.
Les tenseurs seront notés par des lettres calligraphiques $\mathcal{T}$, les matrices par des majuscules grasses $\mathbf{M}$, les vecteurs par des majuscules en italique V et les scalaires par des minuscules en italique s. D'autre part, la $i$ème colonne d'une matrice $\mathbf{M}$ sera notée M_i, le $i$ème élément d'un vecteur V sera noté v_i et l'élément d'indice (i,j) d'une matrice $\mathbf{M}$ sera noté m_{ij}. Les majuscules italiques seront également utilisées pour noter le plus grand élément d'un ensemble d'entiers $i \in [1:I]$.
Par ailleurs, le produit de Kronecker sera noté $\otimes$. Le produit de Kronecker des matrices $\mathbf{A}$ et $\mathbf{B}$ de tailles respectives $I \times J$ et $K \times L$ est une matrice de taille $IK \times JL$ définie par :

$$\mathbf{A} \otimes \mathbf{B} = \begin{pmatrix} a_{11}\mathbf{B} & a_{12}\mathbf{B} & \dots & a_{1J}\mathbf{B} \\ a_{21}\mathbf{B} & a_{22}\mathbf{B} & & a_{2J}\mathbf{B} \\ \vdots & & \ddots & \vdots \\ a_{I1}\mathbf{B} & a_{I2}\mathbf{B} & \dots & a_{IJ}\mathbf{B} \end{pmatrix} \tag{1}$$

Le symbole $\odot$ représentera le produit de Khatri-Rao ou produit de Kronecker par colonnes. Le produit de Khatri-Rao des matrices $\mathbf{A}$ et $\mathbf{B}$ de tailles respectives $I \times J$ et $K \times J$ est une matrice de taille $IK \times J$ définie par :

$$\mathbf{A} \odot \mathbf{B} = (A_1 \otimes B_1, A_2 \otimes B_2, \dots, A_J \otimes B_J) \tag{2}$$

D'autre part nous définissons l'opérateur *vec* qui permet d'écrire une matrice sous forme de vecteur en rangeant ses colonnes les unes en dessous des autres. Soit $\mathbf{A}$ une matrice de taille $I \times J$. L'élément d'indice $i + (j-1)I$ du vecteur $vec(\mathbf{A})$ de taille IJ est l'élément d'indice (i,j) de $\mathbf{A}$. L'opérateur inverse de *vec* sera noté *unvec*.
De la même manière, nous définissons l'opérateur *mat*, qui transforme un tenseur en une matrice. Soit $\mathcal{T}$ un tenseur d'ordre trois de taille $I \times J \times K$. La $k$ème colonne de la matrice $mat(\mathcal{T})$ est obtenue en extrayant la matrice d'indice k dans la troisième dimension ($\mathcal{T}(:,:,k)$ en notations Matlab) et en lui appliquant l'opérateur *vec*. $(mat(\mathcal{T}))_{(i+(j-1)I,k)} = \mathcal{T}_{ijk}$. L'opérateur *vecdiag* extrait la diagonale d'une matrice carrée et la range dans un vecteur colonne. L'opérateur *diag* range un vecteur sur la diagonale d'une matrice diagonale.
Le conjugué d'un élément u est noté u^*. La transposée de la matrice $\mathbf{A}$ sera notée $\mathbf{A}^T$, sa transposée hermitienne $\mathbf{A}^H$ et sa pseudo-inverse $\mathbf{A}^\dagger$. La norme de Frobenius de $\mathbf{A}$, notée $\|\mathbf{A}\|_F$ est égale à la racine de la somme des modules au carrés des éléments de $\mathbf{A}$: $\|\mathbf{A}\|_F^2 = \sum_{i,j} |a_{ij}|^2$.

Chapitre 1

Introduction

1.1 Contexte de l'étude

La séparation aveugle de sources consiste à estimer des signaux inconnus à partir d'un mélange observé de ces signaux, sans information sur les signaux et/ou le mélange. Par exemple, dans une pièce où plusieurs personnes parlent en même temps, on place des microphones. On n'entend qu'un brouhaha sur chaque enregistrement, mais en utilisant tous les enregistrements, il est possible de retrouver le discours de chaque locuteur. De telles méthodes s'appliquent dans des situations où l'on n'a pas de moyen d'avoir des informations sur les sources et/ou le mélange. C'est le cas en particulier des signaux « naturels ». Il existe par exemple des applications dans le domaine biomédical : on peut appliquer des méthodes aveugles de séparation de sources pour séparer le battement de cœur d'un fœtus de celui de sa mère [22]. Il existe des applications en sismologie, ou encore en mécanique, pour détecter un défaut dans les machines tournantes [26,55]. Dans le cadre des communications numériques coopératives, l'émetteur envoie au récepteur des informations lui permettant d'estimer le canal et de retrouver le message qui lui est destiné. Il est cependant intéressant d'être capable de retrouver les signaux émis sans connaissance a priori. L'application militaire est manifeste, on peut utiliser ces méthodes pour intercepter discrètement des communications sur un théâtre d'opérations. Il existe également des applications civiles en communications numériques. En particulier, la séquence d'apprentissage envoyée par l'émetteur pour permettre au récepteur d'estimer le canal occupe une place importante dans le message, et ne constitue pas une information utile pour la personne qui le reçoit. En GSM, cette séquence d'apprentissage occupe 20 % d'une trame, en UMTS, elle occupe près de 40 % d'une trame. Une solution pour augmenter le débit d'information pourrait alors être de supprimer, complètement ou en partie, ces séquences et d'estimer le canal de manière aveugle ou semi-aveugle. Le mélange peut être linéaire ou non linéaire, instantané ou convolutif. Le cas le plus simple, et aussi celui qui a été le plus étudié, est le cas des mélanges linéaires instantanés.

Il existe différentes méthodes pour retrouver les sources à partir du mélange. On peut par exemple s'appuyer sur l'hypothèse que les sources appartiennent à un alphabet fini [28, 50, 52]. Il est également possible de s'appuyer sur l'hypothèse que les sources sont mutuellement indépendantes, on parle dans ce cas d'analyse en composantes indépendantes (ACI) [3, 7, 11].
Plus récemment, les méthodes d'algèbre multilinéaire ont retenu une attention particulière [19, 21]. Les données du problème (les observations) peuvent en effet dans certains cas être regar-

dées comme les éléments d'un tenseur d'ordre supérieur à trois. Il existe une décomposition des tenseurs appelée décomposition PARAFAC qui propose de décomposer un tenseur sous la forme d'une somme minimale de tenseurs de rang un. Résoudre le problème de séparation de sources revient alors à déterminer les paramètres de la décomposition. L'intérêt de cette décomposition réside dans son unicité sous certaines conditions. Une borne a été déterminée en dessous de laquelle la décomposition est toujours unique [32]. On peut cependant se demander s'il est possible de trouver des conditions moins restrictives dans certains cas particuliers. D'autre part, l'algorithme traditionellement utilisé pour déterminer les paramètres de la décomposition est un algorithme des moindres carrés alternés. On peut se demander s'il n'existe pas d'autres techniques, plus fiables ou plus rapides pour identifier les paramètres de la décomposition.

1.2 Contenu du document

Le document est organisé de la manière suivante. Dans le chapitre 2, nous présenterons la décomposition PARAFAC d'un tenseur. Nous expliciterons une condition d'unicité, puis nous détaillerons l'algorithme généralement employé pour déterminer les paramètres de la décomposition. Ensuite, dans le cas d'un tenseur d'ordre trois, nous proposerons une nouvelle technique pour identifier les paramètres de la décomposition si le rang du tenseur est inférieur à l'une des dimensions du tenseur et au produit des deux autres. Cette technique mène à un système de matrices à diagonaliser conjointement. Nous proposerons également dans ce cas une nouvelle borne sur le nombre maximum de paramètres dans la décomposition, moins restrictive que la borne précédente.

Le chapitre 3 est consacré à la séparation aveugle de signaux DS-CDMA. Nous verrons que la structure particulière de ces signaux, obtenue grâce à l'étalement de spectre nous permet d'utiliser la technique développée au chapitre précédent. Nous verrons qu'il est également possible d'exploiter le fait que nous possédons une information a priori sur les sources. En effet, en communications numériques, nous pouvons supposer que les sources appartiennent à une constellation circulaire. Cette hypothèse nous permet d'aboutir à un nouveau système de matrices à diagonaliser conjointement. Nous proposons alors de diagonaliser simultanément les deux systèmes à l'aide de différentes techniques.

Dans le chapitre 4, nous proposons une autre application de la technique présentée au chapitre 2. Nous verrons qu'elle nous permet de résoudre des problèmes d'analyse en composantes indépendantes dans le cas sous-déterminé, c'est-à-dire lorsqu'il y a plus de sources que de capteurs. Nous proposerons deux solutions, la première s'appuyant sur les statistiques d'ordre deux des données, la deuxième sur leurs statistiques d'ordre quatre.

Le chapitre 5 est consacré à quelques propositions pour évaluer le rang générique des tenseurs d'un ordre et d'une dimension donnée. Le rang générique est le rang obtenu en choisissant un tenseur aléatoirement selon une loi continue. Les résultats proposés sont valables pour les décompositions à valeurs dans l'ensemble des complexes. Nous proposerons d'autre part une borne sur le rang d'un tenseur en dessus de laquelle la décomposition PARAFAC ne peut plus être unique.

1.3 Contributions de l'auteur

Congrès

1. Joséphine Castaing, Lieven De Lathauwer, « An Algebraic Technique for the Blind Separation of DS-CDMA Signals », *Proc. 12th European Signal Processing Conference (EUSIPCO 2004), Vienne, Autriche, pp. 377–380, sept. 2004.*
2. Joséphine Castaing, Lieven De Lathauwer « Separation of DS-CDMA signals using PARAFAC based techniques », *Workshop on Tensor Decompositions and Applications, Luminy, Marseille, France, 29 août- 2 sept. 2005.*
3. Joséphine Castaing, Lieven De Lathauwer « Séparation aveugle de signaux DS-CDMA à l'aide de techniques algébriques », *20ème colloque GRETSI sur le traitement du signal et des images, pp. 965–968, Louvain-La-Neuve, Belgique, sept. 2005.*
4. Lieven De Lathauwer, Joséphine Castaing « Second-Order Blind Identification of Underdetermined Mixtures », *6th Int. Conference on Independent Component Analysis and Blind Signal Separation (ICA 2006), Charleston, SC, Etats-Unis, pp. 40–47, mar. 2006.*

Revues

1. Lieven De Lathauwer, Joséphine Castaing, « Tensor-based techniques for the blind separation of DS-CDMA signals », *Signal Procesing, Special Issue on Tensor Signal Processing, à paraître.*
2. Lieven De Lathauwer, Joséphine Castaing, Jean-François Cardoso, « Fourth-Order Cumulant Based Underdetermined Independent Component Analysis », *accepté pour soumission à IEEE Transactions on Signal Processing.*
3. Lieven De Lathauwer, Joséphine Castaing, « Blind Identification of Underdetermined Mixtures by Simultaneous Matrix Diagonalization », *soumis à IEEE Transactions on Signal Processing.*
4. Pierre Comon, Jos M.F. ten Berge, Joséphine Castaing, Lieven De Lathauwer, « Generic and Typical Ranks of Multi-Way Arrays », *en préparation.*

Chapitre 2

Parafac et lien avec la diagonalisation simultanée

2.1 Introduction

Dans un nombre croissant d'applications, il est nécessaire de manipuler des quantités indicées par plusieurs indices. Ces quantités sont appelées des tenseurs. Un tenseur d'ordre un est un vecteur, un tenseur d'ordre deux est une matrice. L'analyse des tenseurs d'ordre supérieur à trois est dite algèbre multilinéaire. On s'intéresse aux tenseurs d'ordre supérieur à trois car ils possèdent des propriétés qui ne sont pas valables pour les matrices ou les vecteurs.
Un tenseur peut se décomposer de différentes manières. Une première idée consiste à généraliser une décomposition à l'aide de matrices unitaires, on parle alors de décomposition en valeurs singulières d'ordre supérieur (HOSVD) [31]. Cette solution ne sera pas développée dans ce document. Nous nous intéressons ici à une décomposition des tenseurs connue sous le nom de Parallel Factor Analysis. La décomposition a été introduite de manière indépendante sous le nom de Canonical Decomposition (CANDECOMP) en psychométrie [9] et de Parallel Factor Analysis (PARAFAC) en phonétique [30] en 1970. Elle a ensuite été utilisée dans des domaines variés, comme la chimiométrie [5], ou les télécommunications [42, 43].
La décomposition PARAFAC propose de décomposer un tenseur sous la forme d'une somme minimale de tenseurs de rang un. Le rang d'un tenseur est une généralisation du rang des matrices, sa définition exacte sera donnée dans la suite.
Dans ce chapitre nous établissons le lien qui existe entre cette décomposition et la résolution d'un système de matrices à diagonaliser conjointement. La diagonalisation conjointe est devenue un outil important en traitement du signal depuis une dizaine d'années. De nombreux auteurs l'ont utilisée et ont proposé des solutions pour la résoudre, dans le cas où la matrice commune est orthogonale (unitaire) ou dans le cas général [3, 7, 29, 34, 37, 52, 54].

Le chapitre est organisé de la manière suivante : après avoir introduit dans le paragraphe 2.2 quelques définitions mathématiques indispensables, nous expliciterons dans le paragraphe 2.3 la décomposition, une condition d'unicité dite condition de Kruskal et les moyens de mettre en œuvre cette décomposition. Le paragraphe 2.4 est dédié au lien que nous avons établi entre la décomposition et la diagonalisation simultanée d'un système de matrices. Nous montrerons que l'algorithme permet d'estimer les paramètres de la décomposition même lorsque la condition de

Kruskal n'est pas vérifiée. D'autre part, nous proposerons une nouvelle borne sur le nombre maximum de paramètres pouvant être identifiés dans la décomposition. Nous présenterons quelques simulations numériques dans le paragraphe 2.5, enfin nous conclurons le chapitre par le paragraphe 2.6.

2.2 Définitions

Définition 1 *Produit externe*

Le produit externe sera noté $\circ$. Le produit externe de N vecteurs $U^{(1)}, U^{(2)}, \ldots, U^{(N)}$ de tailles respectives $I_1, I_2, \ldots, I_N$ est un tenseur d'ordre N de taille $I_1 \times I_2 \times \ldots \times I_N$ dont l'élément d'indice $(i_1, i_2, \ldots, i_N)$ est défini par :

$$\left(U^{(1)} \circ U^{(2)} \circ \ldots \circ U^{(N)}\right)_{i_1,i_2,\ldots,i_N} = U^{(1)}_{i_1} U^{(2)}_{i_2} \ldots U^{(N)}_{i_N}. \tag{2.1}$$

Définition 2 *Tenseur de rang un*

Un tenseur qui se décompose sous la forme d'un produit externe de vecteurs est dit de rang un.

Définition 3 *Tenseur de rang R*

Un tenseur qui se décompose sous la forme d'une somme minimale de R tenseurs de rang un est dit de rang R.

Après avoir pris connaissance de ces notions, nous pouvons maintenant introduire la décomposition PARAFAC.

2.3 Modèle PARAFAC

La décomposition PARAFAC d'un tenseur $\mathcal{Y}$ d'ordre N de taille $I_1 \times I_2 \times \ldots \times I_N$ est la décomposition de ce tenseur sous forme de la somme d'un nombre minimum de tenseurs de rang un :

$$\mathcal{Y} = \sum_{r=1}^{R} U^{(1)}_r \circ U^{(2)}_r \circ \ldots \circ U^{(N)}_r, \tag{2.2}$$

où $U^{(n)}_r \in \mathbb{C}^{I_n}$, $r \in [1:R]$, $n \in [1:N]$. On notera par ailleurs $\mathbf{U}^{(n)}$ la matrice de taille $I_n \times R$ obtenue en concaténant les vecteurs $U^{(n)}_1, U^{(n)}_2, \ldots, U^{(n)}_R$, $n \in [1:N]$.
Les tenseurs d'ordre 3 étant majoritairement utilisés dans la suite, nous introduisons tout de suite les notations qui seront employées dans le cas d'un tenseur de cet ordre.
La décomposition PARAFAC d'un tenseur $\mathcal{Y}$ d'ordre 3, de taille $I \times J \times K$ et de rang R s'écrit :

$$\mathcal{Y} = \sum_{r=1}^{R} A_r \circ H_r \circ S_r, \tag{2.3}$$

où $A_r \in \mathbb{C}^I$, $H_r \in \mathbb{C}^J$, $S_r \in \mathbb{C}^K$. Un schéma de la décomposition d'un tenseur d'ordre 3 est présenté dans la figure 2.1.

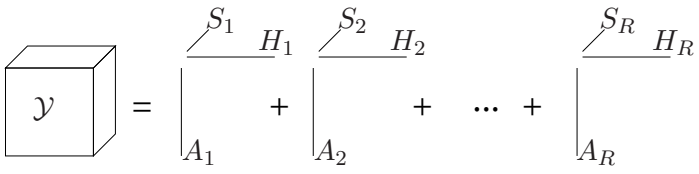

FIG. 2.1 – Schéma de la décomposition PARAFAC.

L'intérêt de cette décomposition réside dans son unicité sous certaines conditions. Si la décomposition (2.2) est unique et que l'on est capable de trouver une solution pour les paramètres $U_r^{(n)}$, $r \in [1:R], n \in [1:N]$, alors nous pouvons être sûrs que la solution trouvée est la bonne.
A ce niveau, nous pouvons faire la remarque suivante : la décomposition d'une matrice en une somme de matrices de rang un existe elle aussi, mais elle n'est pas unique, sauf si l'on impose certaines contraintes fortes sur les matrices de la décomposition, par exemple une condition d'orthogonalité.
D'autre part, la décomposition (2.2) ne peut être unique qu'à deux indéterminations près. En effet, les colonnes des matrices $\mathbf{U}^{(n)}$, $n \in [1:N]$ peuvent être permutées et multipliées par un scalaire. Si les vecteurs $U_r^{(n)}$ sont solutions, alors les vecteurs $\alpha_n U_r^{(n)}$, avec $\prod_{n=1}^{N} \alpha_n = 1$ le sont aussi. Par ailleurs, l'ordre des termes $U_r^{(1)} \circ U_r^{(2)} \circ \ldots \circ U_r^{(N)}$ est arbitraire.
Lorsque la décomposition est unique à ces deux indéterminations près, on dit qu'elle est *essentiellement unique*.

Le paragraphe 2.3.1 est consacré à une condition d'unicité de la décomposition PARAFAC. Dans le paragraphe 2.3.2 nous donnerons une solution pour évaluer les paramètres de la décomposition. Le paragraphe 2.3.3 est consacré à quelques améliorations qu'il est possible d'apporter afin de réduire le temps de calcul lors de l'estimation des paramètres.

2.3.1 Borne de Kruskal

Avant d'énoncer la condition d'unicité, nous avons besoin d'introduire le concept clé suivant :

Définition 4 *Rang de Kruskal*

Le rang de Kruskal de $\mathbf{A}$, noté $rank_k(\mathbf{A})$ est le nombre maximum m de colonnes de $\mathbf{A}$ tel que toute sous-matrice de $\mathbf{A}$ de m colonnes soit de rang plein.
Le k-rang d'une matrice est toujours inférieur ou égal à son rang.

Exemple : Soit la matrice $\mathbf{A}$ de taille 2×3 définie par :

$$\mathbf{A} = \begin{bmatrix} 3 & 2 & 1 \\ 5 & 4 & 2 \end{bmatrix} \tag{2.4}$$

$\mathbf{A}$ est de rang 2 mais de k-rang 1 car ses deux dernières colonnes ne sont pas indépendantes.

Théorème 1 *Théorème d'Unicité*

Soit $\mathcal{Y}$ un tenseur d'ordre N et de rang R, se décomposant sous la forme (2.2). Si

$$\sum_{n=1}^{N} rank_k(\mathbf{U}^{(n)}) \geq 2R + (N-1), \tag{2.5}$$

alors cette décomposition est essentiellement unique.

La preuve de ce théorème a d'abord été donnée à l'ordre 3 pour des tenseurs réels [32] puis plus tard à l'ordre 3 pour des tenseurs complexes [43] et à tout ordre pour des tenseurs complexes [41]. Nous ferons référence à la partie droite de l'inéquation sous le nom de *borne de Kruskal*.
Si le tenseur $\mathcal{Y}$ est d'ordre 3, en conservant les notations utilisées dans (2.3), la condition de Kruskal s'écrit :

$$rank_k(\mathbf{A}) + rank_k(\mathbf{H}) + rank_k(\mathbf{S}) \geq 2(R+1), \tag{2.6}$$

où $\mathbf{A}$, $\mathbf{H}$ et $\mathbf{S}$ désignent les matrices contenant respectivement les vecteurs $(A_r)_{r\in[1:R]}$, $(H_r)_{r\in[1:R]}$ et $(S_r)_{r\in[1:R]}$. Nous allons maintenant voir comment estimer les paramètres de la décomposition à l'aide d'un algorithme des moindres carrés alternés (ALS pour *Alternating Least Square*).

2.3.2 Algorithme ALS

Afin d'estimer les paramètres de la décomposition, nous devons minimiser la fonction de coût

$$f(\mathbf{A},\mathbf{H},\mathbf{S}) = \|\mathcal{Y} - \sum_{r=1}^{R} A_r \circ H_r \circ S_r\|^2. \tag{2.7}$$

Si une solution est trouvée pour $\mathbf{A}$, $\mathbf{H}$ et $\mathbf{S}$ et que R est en dessous de la borne de Kruskal, alors cette solution est la bonne. Le principe de l'algorithme ALS est de mettre à jour de manière alternée chacune des matrices en gardant les deux autres fixées. Si deux matrices parmi les trois sont fixées, le système à résoudre devient un problème simple de moindres carrés.
On définit trois modes de lecture du tenseur (un schéma de ces modes est donné dans la figure 2.2).

$$\tilde{\mathbf{X}} = [\mathbf{X}^{(1)T}, \mathbf{X}^{(2)T}, \dots \mathbf{X}^{(K)T}]^T \in \mathbb{C}^{JK\times I} \tag{2.8}$$
$$\tilde{\mathbf{Y}} = [\mathbf{Y}^{(1)T}, \mathbf{Y}^{(2)T}, \dots \mathbf{Y}^{(J)T}]^T \in \mathbb{C}^{IJ\times K} \tag{2.9}$$
$$\tilde{\mathbf{Z}} = [\mathbf{Z}^{(1)T}, \mathbf{Z}^{(2)T}, \dots \mathbf{Z}^{(I)T}]^T \in \mathbb{C}^{KI\times J} \tag{2.10}$$

Les matrices $\tilde{\mathbf{X}}$, $\tilde{\mathbf{Y}}$ et $\tilde{\mathbf{Z}}$ vérifient :

$$\tilde{\mathbf{X}} = (\mathbf{H}\odot\mathbf{S})\cdot\mathbf{A}^T \tag{2.11}$$
$$\tilde{\mathbf{Y}} = (\mathbf{A}\odot\mathbf{H})\cdot\mathbf{S}^T \tag{2.12}$$
$$\tilde{\mathbf{Z}} = (\mathbf{S}\odot\mathbf{A})\cdot\mathbf{H}^T \tag{2.13}$$

Chaque mise à jour des matrices est obtenue par une simple inversion matricielle. On décide que l'algorithme a convergé si la norme de Frobenius de la différence entre l'estimée de $\mathbf{A}$ à l'itération k et son estimée à l'itération $k+1$ est inférieure à une certaine tolérance ϵ. Finalement, nous obtenons ainsi une estimée de $\mathbf{F}$ et une estimée de $\mathbf{F}^T$. La $r$ème colonne de $\mathbf{F}$ peut alors être choisie égale au vecteur singulier dominant de la matrice de taille $R\times 2$ constituée de la $r$ème colonne de l'estimée de $\mathbf{F}$ et de la transposée de la $r$ème ligne de l'estimée de $\tilde{\mathbf{F}}$.
Un synopsis de l'algorithme ALS est donné dans la table 2.1.

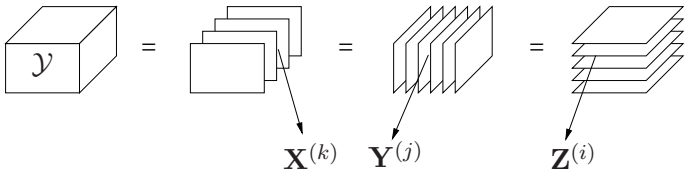

FIG. 2.2 – Modes de lecture du tenseur

1. Initialisation : on choisit $\mathbf{H}$ et $\mathbf{S}$ de manière aléatoire
2. Mise à jour de $\mathbf{A}$: $\mathbf{A} = ((\mathbf{H} \odot \mathbf{S})^{\dagger} \cdot \tilde{\mathbf{X}})^T$
3. Mise à jour de $\mathbf{H}$: $\mathbf{H} = ((\mathbf{A} \odot \mathbf{H})^{\dagger} \cdot \tilde{\mathbf{Y}})^T$
4. Mise à jour de $\mathbf{S}$: $\mathbf{S} = ((\mathbf{S} \odot \mathbf{A})^{\dagger} \cdot \tilde{\mathbf{Z}})^T$
5. Aller en (2) jusqu'à convergence

TAB. 2.1 – Résumé de l'algorithme ALS

2.3.3 Initialisation de l'ALS et Compression du tenseur des données

L'algorithme ALS converge généralement (ce n'est pas toujours le cas [24, 44]), cependant sa convergence peut être lente et il peut converger vers un minimum local. Il est possible d'initialiser l'algorithme si deux des dimensions sont plus grandes que le rang R du tenseur. Supposons par exemple que $R \leq I, J$. Notons $\mathbf{Y}_1$ et $\mathbf{Y}_2$ les deux premières tranches matricielles du tenseur $\mathcal{Y}$.

$$\mathbf{Y}_1 = \mathbf{A}\mathbf{D}_1\mathbf{H}^T \tag{2.14}$$
$$\mathbf{Y}_2 = \mathbf{A}\mathbf{D}_2\mathbf{H}^T, \tag{2.15}$$

où $\mathbf{D}_1$ et $\mathbf{D}_2$ sont des matrices diagonales de taille $R \times R$ dont les diagonales contiennent respectivement les éléments $s_{11}, s_{12}, \ldots, s_{1R}$ et les éléments $s_{21}, s_{22}, \ldots, s_{2R}$.
Le produit $\mathbf{Y}_1\mathbf{Y}_2^{\dagger}$ s'écrit alors :

$$\mathbf{Y}_1\mathbf{Y}_2^{\dagger} = \mathbf{A}\mathbf{D}_1\mathbf{D}_2^{-1}\mathbf{A}^{\dagger}. \tag{2.16}$$

La matrice $\mathbf{D_1}{\mathbf{D_2}}^{-1}$ est diagonale. La matrice $\mathbf{Y_1}{\mathbf{Y_2}}^{\dagger}$ de taille $I \times I$ est de rang R car $R \leq I, J$. On peut peut donc évaluer la matrice $\mathbf{A}$ à partir d'une décomposition en valeurs propres de la matrice $\mathbf{Y}_1\mathbf{Y}_2^{\dagger}$.
Par ailleurs, si l'une des dimensions est beaucoup plus grande que les autres (nous détaillerons dans la suite ce que signifie grand), le temps de calcul de l'algorithme peut devenir très important. Il est possible de compresser le tenseur des observations pour réduire le temps de calcul. Supposons par exemple $K >> I, J$. Le tenseur $\mathcal{Y}$ peut s'écrire sous la forme de la matrice $\mathbf{Y}$ de taille $IJ \times K$

$$\mathbf{Y} = (\mathbf{A} \odot \mathbf{H})\mathbf{S}^T. \tag{2.17}$$

Par ailleurs la décomposition en valeurs singulières de $\mathbf{Y}$ s'écrit : $\mathbf{Y} = \mathbf{U}\mathbf{D}\mathbf{V}^H$. Si $IJ \leq K$, alors la matrice $\mathbf{Y}_2 = \mathbf{U}\mathbf{D}$ est de taille $IJ \times IJ$. Soit $\mathcal{Y}_2$ sa représentation tensorielle, on peut chercher

la décomposition PARAFAC de ce tenseur réduit. Les paramètres de la décomposition de $\mathcal{Y}_2$ sont les mêmes que ceux du tenseur $\mathcal{Y}$ sauf dans la dimension dans laquelle il a été compressé [5] :

$$\mathcal{Y}_2 = \sum_{r=1}^{R} A_r \circ H_r \circ Z_r. \tag{2.18}$$

Les paramètres dans la dimension qui a été compressée peuvent être obtenus à partir des éléments dans cette dimension du tenseur $\mathcal{Y}_2$ et de la matrice des vecteurs singuliers de droite de $\mathbf{Y}$. En effet,

$$\begin{aligned} \mathbf{Y} &= \mathbf{Y}_2\mathbf{V}^H & (2.19)\\ &= (\mathbf{A} \odot \mathbf{H})(\mathbf{Z}^T\mathbf{V}^H). & (2.20) \end{aligned}$$

Nous pouvons donc réduire simplement le temps de calcul des paramètres de la décomposition si l'une des dimensions est supérieure au produit des deux autres.

Dans le paragraphe suivant, nous allons présenter une technique alternative permettant de trouver les paramètres de la décomposition sans appliquer d'algorithme ALS aux données brutes. Nous verrons que cette technique a de multiples avantages : il est possible de dépasser la borne de Kruskal, il est toujours possible de trouver une bonne initialisation.

2.4 Lien avec la diagonalisation simultanée

Nous avons vu qu'il était possible d'estimer les paramètres de la décomposition (2.3) à l'aide d'un algorithme ALS si la condition de Kruskal (2.6) est vérifiée. Cependant, si l'ALS converge généralement et de façon monotone, il peut converger vers un minimum local, en particulier si on ne possède pas une bonne initialisation. D'autre part, la condition sur R, rang maximum du tenseur, est assez contraignante.
Nous allons montrer qu'il est possible d'envisager le problème de manière différente. La solution proposée maintenant permet d'estimer les paramètres de la décomposition à l'aide d'une diagonalisation conjointe. D'autre part, nous allons montrer que cette solution mène à une nouvelle borne sur le rang du tenseur, moins contraignante que la borne de Kruskal sous certaines hypothèses. L'idée de cet algorithme a été inspirée par [6].

2.4.1 Reformulation du problème

Nous allons faire l'hypothèse forte suivante (la justification de la nécessité de cette hypothèse va être donnée dans la suite) :

$$R \leq min(IJ, K). \tag{2.21}$$

Dans un premier temps, nous allons ranger les éléments du tenseur $\mathcal{Y}$ dans une matrice $\mathbf{Y}$ à l'aide de l'opérateur mat : $\mathbf{Y} = mat(\mathcal{Y})$. Pour cela on extrait la tranche matricielle d'indice r dans la troisième dimension de $\mathcal{Y}$, on range ses colonnes les unes en dessous des autres puis on place le vecteur obtenu dans la colonne d'indice r dans la matrice $\mathbf{Y}$. L'élément d'indice $((j-1)I+i, k)$ de $\mathbf{Y}$ s'écrit alors :

$$y_{(j-1)I+i,k} = y_{ijk}.$$

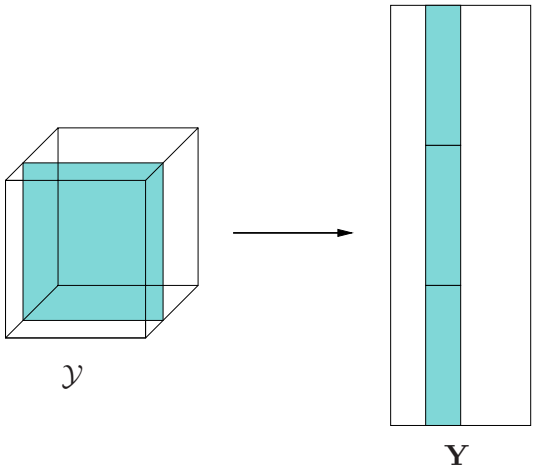

FIG. 2.3 – Construction de la matrice $\mathbf{Y}$ à partir du tenseur $\mathcal{Y}$.

Un schéma de cette transformation est donné dans la figure 2.3.
On peut également écrire $\mathbf{Y}$ à l'aide d'un produit de Khatri-Rao :

$$\mathbf{Y} = (\mathbf{A} \odot \mathbf{H})\ \mathbf{S}^T. \tag{2.22}$$

Si $R \leq min(IJ, K)$, la matrice $\mathbf{A} \odot \mathbf{H}$ est de rang plein avec une probabilité un si les éléments de $\mathbf{A}$ et de $\mathbf{H}$ suivent des lois continues [18]. Nous allons d'autre part supposer que $\mathbf{S}$ est aussi de rang plein. Si les éléments de $\mathbf{S}$ appartiennent à un alphabet fini, comme c'est le cas en communications numériques, il est en fait possible que $\mathbf{S}$ ne soit pas de rang plein, mais la probabilité qu'elle le soit augmente si K augmente.
Si $\mathbf{A} \odot \mathbf{H}$ et $\mathbf{S}$ sont de rang plein, alors le rang R du tenseur $\mathcal{Y}$ est égal au rang de la matrice $\mathbf{Y}$. R peut donc être estimé comme le nombre de valeurs singulières significatives de $\mathbf{Y}$. Nous pouvons constater que $\mathbf{Y}$ possède une structure très particulière. En effet, si on applique l'opérateur *unvec* aux colonnes de la matrice $\mathbf{A} \odot \mathbf{H}$, que l'on va noter $(\mathbf{A} \odot \mathbf{H})_r$, $r \in [1 : R]$, on obtient les matrices $A_r H_r^T$ qui sont des matrices de rang 1. C'est cette particularité de la structure de $\mathbf{Y}$ qui va nous servir dans la suite.

Revenons pour le moment à notre problème. Nous savons qu'il existe une autre décomposition de $\mathbf{Y}$, sous la forme du produit d'une matrice unitaire $\mathbf{U}$, d'une matrice diagonale positive $\mathbf{D}$ et de la transposée conjuguée d'une matrice unitaire $\mathbf{V}$ (décomposition en valeurs singulières ou SVD) :

$$\mathbf{Y} = \mathbf{U}\mathbf{D}\mathbf{V}^H. \tag{2.23}$$

Nous pouvons déduire des équations (2.22) et (2.23) qu'il existe une matrice $\mathbf{F}$ non singulière de taille $R \times R$ et a priori inconnue qui lie $\mathbf{A} \odot \mathbf{H}$ à $\mathbf{U}$ et à $\mathbf{D}$ et $\mathbf{S}$ à $\mathbf{V}$:

$$\begin{cases} \mathbf{A} \odot \mathbf{H} &= \mathbf{U}\mathbf{D}\mathbf{F} \\ \mathbf{S}^T &= \mathbf{F}^{-1}\mathbf{V}^H \end{cases}, \tag{2.24}$$

Supposons que nous connaissons $\mathbf{F}$, alors nous pouvons trouver facilement les matrices $\mathbf{S}$, $\mathbf{A}$ et $\mathbf{H}$. En effet, d'après la deuxième équation du système (2.24), on obtient $\mathbf{S}$ de manière directe : $\mathbf{S} = \mathbf{V}^*\mathbf{F}^{-T}$.
Notons d'autre part $\mathbf{N}_i$ la matrice de taille $R \times R$ obtenue en appliquant l'opérateur *unvec* à la colonne numéro i de $\mathbf{A} \odot \mathbf{H}$.

$$\mathbf{N}_i = unvec\,(A_i \otimes H_i) = H_i A_i^T.$$

Cette matrice est de rang un. A une multiplication par un scalaire près, H_i est le vecteur singulier dominant de gauche de la matrice $\mathbf{N}_i$, c'est à dire le vecteur singulier de gauche correspondant à sa plus grande valeur singulière, et A_i est le conjugué du vecteur singulier de droite dominant. Il s'agit maintenant d'évaluer la matrice $\mathbf{F}$. A cette fin, nous allons utiliser la première équation du système (2.24) et exploiter la structure particulière de la matrice $\mathbf{A} \odot \mathbf{H}$.

Notons $\mathbf{E}_r$ la matrice de taille $I \times J$ obtenue en extrayant la $r$ème colonne de la matrice $\tilde{\mathbf{U}} = \mathbf{UD}$ et en la réordonnant sous forme d'une matrice :

$$\mathbf{E}_r = unvec(\tilde{\mathbf{U}}_r) \tag{2.25}$$

D'après l'équation (2.24), $\mathbf{E}_r$ peut aussi s'écrire à l'aide de $\mathbf{A}$, $\mathbf{H}$ et $\mathbf{F}$.

$$\begin{aligned} \mathbf{E}_r &= unvec\left(\left((\mathbf{A} \odot \mathbf{H})\,\mathbf{F}^{-1}\right)_r\right) && (2.26)\\ &= \sum_{k=1}^{R} \left(H_k A_k^T\right)(\mathbf{F}^{-1})_{kr}. && (2.27) \end{aligned}$$

Chaque matrice $\mathbf{E}_r$, $r \in [1:R]$ s'écrit comme la somme des matrices de rang un $\left(H_k A_k^T\right)$ pondérées par les éléments du vecteur colonne $(\mathbf{F}^{-1})_r$. Pour trouver les éléments de la matrice $\mathbf{F}^{-1}$, nous allons donc chercher des combinaisons linéaires des $\mathbf{E}_r$ qui sont des matrices de rang un. Nous avons besoin d'un outil nous permettant de savoir si une matrice est de rang un. Un tel outil est proposé dans le théorème suivant [6, 18] :

Théorème 2 *Matrices de rang un*
Soit la fonction $\Phi : (\mathbf{X}, \mathbf{Y}) \in \mathbb{C}^{I\times J} \times \mathbb{C}^{I\times J} \longmapsto \Phi(\mathbf{X}, \mathbf{Y}) \in \mathbb{C}^{I\times J\times I\times J}$ *définie par :*
$(\Phi(\mathbf{X}, \mathbf{Y}))_{ijkl} = x_{ij}y_{kl} + y_{ij}x_{kl} - x_{il}y_{kj} - y_{il}x_{kj}$ *pour tout* $(i,k) \in [1:I], (j,l) \in [1:J]$.
Soit $\mathbf{X} \in \mathbb{C}^{I\times J}$, *alors* $\Phi(\mathbf{X}, \mathbf{X}) = 0$ *si et seulement si le rang de* $\mathbf{X}$ *est au plus un .*

Preuve :
Le cas $\mathbf{X} = 0$ est évident.
Soit $\mathbf{X}$ une matrice de rang un. Il existe deux vecteurs $\mathbf{u}$ et $\mathbf{v}$ tels que $x_{ij} = u_i v_j$. Alors $(\Phi(\mathbf{X}, \mathbf{X}))_{ijkl} = 2(u_i v_j u_k v_l - u_i v_l u_k v_j) = 0$.
Soit maintenant $\mathbf{X}$ une matrice vérifiant $\Phi(\mathbf{X}, \mathbf{X}) = 0$. La décomposition en valeurs singulières de $\mathbf{X}$ est donnée par $\mathbf{X} = \mathbf{U\Sigma V}^H$. Nous obtenons donc :

$$\begin{aligned} x_{ij}x_{kl} - x_{il}x_{kj} &= \sum_{r,s} \sigma_r\sigma_s u_{ir}u_{ks}(v_{jr}v_{ls} - v_{js}v_{lr})^* \\ &= \sum_{r\neq s} \sigma_r\sigma_s u_{ir}u_{ks}(v_{jr}v_{ls} - v_{js}v_{lr})^*. \end{aligned}$$

Les matrices $\mathbf{U}$ et $\mathbf{V}$ sont unitaires, donc les tenseurs dont les éléments sont les $u_{ir}u_{ks}(v_{jr}v_{ls})^*$ et $u_{ir}u_{ks}(v_{js}v_{lr})^*$, $r \neq s$ sont linéairement indépendants. Par conséquent, $\sigma_r\sigma_s = 0$ si $r \neq s$ et $\mathbf{X}$ et $\mathbf{\Sigma}$ sont de rang un. □

La fontion Φ permet de construire un ensemble de R^2 tenseurs Φ_{rs} définis par :

$$\begin{aligned}\Phi_{rs} &= \Phi\left(\mathbf{E}_r, \mathbf{E}_s\right) \\ &= \Phi\left(\sum_{p=1}^{R} H_p A_p^T (\mathbf{F}^{-1})_{pr},\ \sum_{q=1}^{R} H_q A_q^T (\mathbf{F}^{-1})_{qs}\right).\end{aligned}$$

En vertu de la bilinéarité de Φ, nous pouvons écrire :

$$\Phi_{rs} = \sum_{p,q=1}^{R} (\mathbf{F}^{-1})_{pr}(\mathbf{F}^{-1})_{qs}\Phi\left(H_p A_p^T,\ H_q A_q^T\right). \tag{2.28}$$

Supposons qu'il existe une matrice $\mathbf{B}$ symétrique de taille $R \times R$ (nous montrerons l'existence d'une telle matrice dans la suite) telle que :

$$\sum_{r,s=1}^{R} \Phi_{rs} b_{rs} = 0. \tag{2.29}$$

En remplaçant Φ_{rs} par son expression dans (2.28), nous obtenons :

$$\sum_{r,s=1}^{R}\sum_{p,q=1}^{R} (\mathbf{F}^{-1})_{pr}(\mathbf{F}^{-1})_{qs}\Phi\left(H_p A_p^T,\ H_q A_q^T\right) b_{rs} = 0.$$

D'après le théorème 2, $\Phi\left(H_p A_p^T,\ H_p A_p^T\right) = 0$ pour tout p dans $[1 : R]$, donc l'équation précédente s'écrit encore :

$$\sum_{r,s=1}^{R}\sum_{\substack{p,q=1\\ p\neq q}}^{R} (\mathbf{F}^{-1})_{pr}(\mathbf{F}^{-1})_{qs} b_{rs}\Phi\left(H_p A_p^T,\ H_q A_q^T\right) = 0.$$

D'autre part, Φ et $\mathbf{B}$ étant symétriques, nous pouvons écrire :

$$\sum_{r,s=1}^{R}\sum_{\substack{p,q=1\\ p<q}}^{R} (\mathbf{F}^{-1})_{pr}(\mathbf{F}^{-1})_{qs} b_{rs}\Phi\left(H_p A_p^T,\ H_q A_q^T\right) = 0. \tag{2.30}$$

Supposons maintenant que les tenseurs $\left(\Phi\left(H_p A_p^T,\ H_q A_q^T\right)\right)_{p<q}$ sont linéairement indépendants. Cette condition forte va imposer une contrainte sur le rang R du tenseur $\mathcal{Y}$ dont nous parlerons au paragraphe 2.4.2.
Si les tenseurs $\left(\Phi\left(H_p A_p^T,\ H_q A_q^T\right)\right)_{p<q}$ sont linéairement indépendants, alors nous pouvons déduire de l'équation (2.30) le résultat suivant :

$$\sum_{r,s=1}^{R} (\mathbf{F}^{-1})_{pr}(\mathbf{F}^{-1})_{qs} b_{rs} = 0, \quad \forall (p,q) \in [1:R], p < q. \tag{2.31}$$

Par symétrie, nous avons également :

$$\sum_{r,s=1}^{R} (\mathbf{F}^{-1})_{pr}(\mathbf{F}^{-1})_{qs} b_{rs} = 0, \quad \forall (p,q) \in [1:R], p > q. \tag{2.32}$$

Nous pouvons donc finalement écrire :

$$\sum_{r,s=1}^{R} (\mathbf{F}^{-1})_{pr}(\mathbf{F}^{-1})_{qs} b_{rs} = \lambda_{pq}\delta_{pq}, \tag{2.33}$$

où δ désigne le symbole de Kronecker ($\delta_{pq} = 1$ si $p = q$, $\delta_{pq} = 0$ sinon) et où les λ_{pq} sont des scalaires.
L'équation (2.33) peut être écrite sous forme matricielle :

$$\mathbf{B} = \mathbf{F}\mathbf{\Lambda}\mathbf{F}^T, \tag{2.34}$$

Dans cette expression, $\mathbf{\Lambda}$ est une matrice diagonale dont les éléments sont les λ_{pp}, $p \in [1:R]$.
Par ailleurs, toute matrice $\mathbf{B}$ s'écrivant sous la forme $\mathbf{F}\mathbf{\Lambda}\mathbf{F}^T$ avec $\mathbf{\Lambda}$ une matrice diagonale quelconque vérifie l'équation (2.29). L'hypothèse de l'existence d'une matrice satisfaisant (2.29) est donc vérifiée. Si l'on choisit R matrices diagonales indépendantes, nous obtenons R solutions indépendantes pour (2.29).
Le noyau de la matrice $[vec(\Phi_{11}), vec(\Phi_{12}), \ldots, vec(\Phi_{RR})]$ contient R matrices indépendantes pouvant se décomposer sous la forme (2.34).
Cependant, si l'on choisit R solutions indépendantes quelconques dans le noyau de cette matrice nous n'obtenons pas nécessairement des matrices solutions symétriques. En particulier, toute matrice antisymétrique vérifie l'équation (2.29). Nous allons voir comment obtenir des solutions symétriques.
En vertu de la symétrie de $\mathbf{B}$ et Φ, l'équation (2.29) peut s'écrire :

$$\sum_{\substack{r,s=1 \\ r<s}}^{R} \Phi_{rs} b_{rs} + \frac{1}{2}\sum_{r=1}^{R} \Phi_{rr} b_{rr} = 0. \tag{2.35}$$

Rangeons les tenseurs Φ_{rs}, $(r,s) \in [1:R], r \leq s$ dans les vecteurs $P_{rs} = vec(\Phi_{rs})$ de taille I^2J^2 et ces vecteurs P_{rs} dans la matrice $\mathbf{P} = [P_{11}, P_{12}, \ldots, P_{RR}]$ de taille $I^2J^2 \times R(R+1)/2$. L'équation précédente peut encore s'écrire :

$$[P_{11}, P_{12}, \ldots, P_{RR}] \begin{bmatrix} x_{11} \\ x_{12} \\ \vdots \\ x_{RR} \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix} \tag{2.36}$$

Avec $x_{rr} = \frac{1}{2} b_{rr}$ et $x_{rs} = b_{rs}$.

Les R vecteurs singuliers de droite de $\mathbf{P}$ correspondants aux R plus petites valeurs propres sont solutions. Après avoir rangés ces vecteurs solutions dans les matrices triangulaires supérieures $\mathbf{X}_t$, $t \in [1:R]$, nous pouvons trouver les matrices $\mathbf{B}_t$ simplement en calculant $\mathbf{B}_t = \mathbf{X}_t + \mathbf{X}_t^T$.
La matrice $\mathbf{F}$ peut alors être évaluée à l'aide d'une diagonalisation simultanée du système

$$\left\{ \begin{array}{rcl} \mathbf{B}_1 & = & \mathbf{F}\mathbf{\Lambda}_1\mathbf{F}^T \\ \mathbf{B}_2 & = & \mathbf{F}\mathbf{\Lambda}_2\mathbf{F}^T \\ & \vdots & \\ \mathbf{B}_r & = & \mathbf{F}\mathbf{\Lambda}_r\mathbf{F}^T \end{array} \right. , \tag{2.37}$$

dans lequel les matrices $\mathbf{\Lambda}_r,\ r \in [1:R]$ sont des matrices diagonales inconnues a priori.

Remarque. Les matrices $\mathbf{B}_r$ n'ont pas nécessairement toutes la même précision. Ces matrices sont issues du noyau de $\mathbf{P}$, donc plus la valeur propre $\sigma_{B,r}$ correspondant à la matrice $\mathbf{B}_r$ est petite, plus la confiance que l'on peut accorder à celle-ci est grande. Il est possible, pour améliorer les résultats, de pondérer chaque matrice $\mathbf{B}_r$ par exemple par l'inverse de la valeur $\sigma_{B,r}$ correspondante.

Il existe différentes techniques pour résoudre le système (2.37), nous détaillerons deux solutions possibles dans le paragraphe 2.4.3. Un synopsis de l'algorithme que nous noterons CD-SD pour Canonical Decomposition by Simultaneous Diagonalization est donné dans la table 2.2.

Nous allons maintenant nous intéresser à l'hypothèse que nous avons émise pour passer de l'équation (2.30) à l'équation (2.33), et chercher sous quelles conditions elle est vérifiée.

2.4.2 Borne

Nous avons montré que sous la condition (2.21), le problème peut être résolu si les tenseurs $\Phi(H_pA_p^T, H_qA_q^T)$ sont linéairement indépendants. On montre que si $\mathbf{A}$ et $\mathbf{H}$ suivent des distributions continues, la condition sur R pour que les tenseurs $\Phi(H_pA_p^T, H_qA_q^T)$ soient linéairement indépendants est la suivante :

$$R(R-1) \leq \frac{1}{2}I(I-1)J(J-1) \tag{2.38}$$

La démonstration de ce résultat est assez technique, elle a été présentée dans [18].
Néanmoins, on peut pressentir cette condition. Pour plus de lisibilité, nous notons dans ce paragraphe $\mathcal{T}^{(pq)} = (\Phi(H_pA_p^T, H_qA_q^T))$.
Les IJ éléments de la forme $\mathcal{T}^{(pq)}_{ijij}, i \in [1:I], j \in [1:J]$ sont nuls. De même les $IJ(J-1)$ éléments de la forme $\mathcal{T}^{(pq)}_{ijil}, i \in [1:I], (j,l) \in [1:J], j \neq l$ et les $IJ(I-1)$ éléments de la forme $\mathcal{T}^{(pq)}_{ijkj}, (i,k) \in [1:I], i \neq k, j \in [1:J]$ sont nuls.
Enfin $\mathcal{T}^{(pq)}_{ijkl} = \mathcal{T}^{(pq)}_{klij} = -\mathcal{T}^{(pq)}_{kjil} = -\mathcal{T}^{(pq)}_{ilkj}$.
Donc en raison de cette symétrie, parmi les $IJ(I-1)(J-1)$ éléments de la forme $\mathcal{T}^{(pq)}_{ijkl}, (i,k) \in [1:I], i \neq k, (j,l) \in [1:J], j \neq l$, seulement au plus $IJ(I-1)(J-1)/4$ ne sont ni égaux ni opposés.

1. Ranger $\mathcal{Y}$ dans la matrice $\mathbf{Y} = mat(\mathcal{Y})$.
2. Calculer la SVD de $\mathbf{Y} = \mathbf{U}\mathbf{D}\mathbf{V}^H$.
3. Ranger les colonnes $\mathbf{UD}$ dans les matrices $\mathbf{E}_r = unvec((\mathbf{UD})_r)$.
4. Evaluer les tenseurs $\Phi_{rs} = \Phi(\mathbf{E}_r, \mathbf{E}_s)$, $(r,s) \in [1:R], r \leq s$ et les ranger dans les vecteurs $P_{rs} = vec(\Phi_{rs})$.
5. Ranger les vecteurs P_{rs} dans la matrice $\mathbf{P} = [P_{11}, P_{12}, \ldots, P_{RR}]$ de taille $I^2J^2 \times R(R+1)/2$.
6. Calculer les R vecteurs singuliers de droite de $\mathbf{P}$ correspondants aux R plus petites valeurs singulières et les ranger dans les matrices triangulaires supérieures $\mathbf{X}_r, r \in [1:R]$.
7. Evaluer les matrices $\mathbf{B}_r = \mathbf{X}_r + \mathbf{X}_r^T$.
8. Résoudre le système suivant à l'aide de l'un des algorithmes présentés dans le paragraphe 2.4.3
$$\begin{cases} \mathbf{B}_1 &= \mathbf{F}\mathbf{\Lambda}_1\mathbf{F}^T \\ \mathbf{B}_2 &= \mathbf{F}\mathbf{\Lambda}_2\mathbf{F}^T \\ & \vdots \\ \mathbf{B}_R &= \mathbf{F}\mathbf{\Lambda}_R\mathbf{F}^T \end{cases}.$$
9. Evaluer $\mathbf{S} = \mathbf{V}^*\mathbf{F}^{-T}$.
10. Evaluer $\mathbf{H}$ et $\mathbf{A}$: la $r$ème colonne de $\mathbf{H}$ est le vecteur singulier dominant de gauche de la matrice $unvec((\mathbf{UDF})_r)$ et la $r$ème colonne de $\mathbf{A}$ est le conjugué du vecteur singulier dominant de droite de la matrice $unvec((\mathbf{UDF})_r)$.

TAB. 2.2 – Résumé de l'algorithme CD-SD

Rangeons les tenseurs $\mathcal{T}^{(pq)}$ sous forme de vecteurs dans la matrice $\tilde{\mathbf{P}}$ de taille $I^2J^2 \times R(R-1)/2$. La matrice $\tilde{\mathbf{P}}$ possède $IJ + IJ(J-1) + IJ(I-1) + 3IJ(I-1)(J-1)/4$ lignes nulles, égales ou opposées aux $IJ(I-1)(J-1)/4$ autres lignes. Donc $\tilde{\mathbf{P}}$ est au plus de rang $IJ(I-1)(J-1)/4$ si $R(R-1)/2 \leq IJ(I-1)(J-1)/4$. En réalité elle est dans ce cas exactement de rang $IJ(I-1)(J-1)/4$. Les colonnes de $\tilde{\mathbf{P}}$ représentant les tenseurs $\mathcal{T}^{(pq)}, (p,q) \in [1:R], p < q$ sont donc indépendantes seulement si $R(R-1) \leq I(I-1)J(J-1)/2$.

Nous pouvons voir que cette borne sur R est moins contraignante que la borne de Kruskal si I et J sont grands. En effet, dans (2.38) R est borné par une quantité qui dépend de manière quadratique de I et J alors que dans (2.6) il est borné par une quantité qui en dépend linéairement.

2.4.3 Résolution du système

Dans cette partie nous expliquons comment résoudre le problème (2.37). En théorie, en l'absence de bruit il est possible d'évaluer $\mathbf{F}$ à partir de deux équations de ce système, en effet $\mathbf{F}$ est la matrice propre de la matrice $\mathbf{B}_i\mathbf{B}_j^{-1}, i \neq j$:

$$\mathbf{B}_i\mathbf{B}_j^{-1} = \mathbf{F}\mathbf{\Lambda}_i\mathbf{\Lambda}_j^{-1}\mathbf{F}^{-1}$$

En pratique, il est plus judicieux de choisir de conserver l'ensemble des équations et de diagonaliser conjointement toutes les matrices $\mathbf{B}_r, r \in [1:R]$. Il existe différentes manières de procéder, nous proposons ici une solution de type ALS et une solution de type QZ étendu.

2.4.3.1 Résolution du système par un algorithme de type ALS

Nous allons montrer dans ce paragraphe qu'il est possible de résoudre le problème (2.37) à l'aide d'un algorithme des moindres carrés alternés. Nous ferons référence dans la suite à cet algorithme sous la notation *SD-ALS* (Simultaneous Diagonalization by ALS).
Le principe de l'algorithme est de minimiser la fonction de coût f définie par :

$$f(\mathbf{F}, \mathbf{\Lambda}_r, \tilde{\mathbf{F}}) = \sum_{r=1}^{R} \|\mathbf{B}_r - \mathbf{F}\mathbf{\Lambda}_r\tilde{\mathbf{F}}\|^2, \tag{2.39}$$

avec $\tilde{\mathbf{F}} = \mathbf{F}^T$, de manière alternée par rapport à $\mathbf{F}$, à $\tilde{\mathbf{F}}$ et à $\mathbf{\Lambda} = [vecdiag(\mathbf{\Lambda}_1), vecdiag(\mathbf{\Lambda}_2), \ldots, vecdiag(\mathbf{\Lambda}_R)]$ en gardant les deux autres matrices fixées. Si deux matrices sont fixées, alors les matrices $(\mathbf{B}_r)_{r\in[1:R]}$ dépendent de la troisième de manière linéaire. Estimer cette matrice revient donc simplement à résoudre un système linéaire :

$$\begin{cases} \mathbf{B}_1 &= \mathbf{F}\mathbf{\Lambda}_1\tilde{\mathbf{F}} \\ \mathbf{B}_2 &= \mathbf{F}\mathbf{\Lambda}_2\tilde{\mathbf{F}} \\ &\vdots \\ \mathbf{B}_R &= \mathbf{F}\mathbf{\Lambda}_R\tilde{\mathbf{F}} \end{cases} \tag{2.40}$$

Chaque itération se décompose en trois étapes :

1. Mise à jour de l'estimée de $\mathbf{\Lambda}_r$

   Afin d'écrire les diagonales des matrices $\mathbf{\Lambda}_r$ en fonctions des matrices $\mathbf{B}_r$, nous allons utiliser la propriété suivante de l'opérateur *vec* :

   Soient $\mathbf{X}$, $\mathbf{Y}$, $\mathbf{Z}$ trois matrices de tailles respectives $a\times b$, $b\times c$ et $c\times d$. Le vecteur $vec(\mathbf{XYZ})$ peut s'écrire :

   $$vec(\mathbf{XYZ}) = \left(\mathbf{Z}^T \otimes \mathbf{X}\right) vec(\mathbf{Y}). \tag{2.41}$$

   En appliquant l'opérateur *vec* aux matrices $\mathbf{B}_r$, nous obtenons alors d'après (2.41) :

   $$vec(\mathbf{B}_r) = \left(\tilde{\mathbf{F}}^T \otimes \mathbf{F}\right) vec(\mathbf{\Lambda}_r) \tag{2.42}$$

   Cette équation peut encore s'écrire :

   $$vec(\mathbf{B}_r) = \left(\tilde{\mathbf{F}}^T \odot \mathbf{F}\right) vecdiag(\mathbf{\Lambda}_r). \tag{2.43}$$

   D'où

   $$[vec(\mathbf{B}_1), vec(\mathbf{B}_2), \ldots, vec(\mathbf{B}_R)] = \left(\tilde{\mathbf{F}}^T \odot \mathbf{F}\right) \mathbf{\Lambda}. \tag{2.44}$$

   On tire $\mathbf{\Lambda}$ de ce système linéaire.

2. Mise à jour de l'estimée de $\mathbf{F}$
   Posons $\delta_1 = [\mathbf{\Lambda}_1\tilde{\mathbf{F}}, \mathbf{\Lambda}_2\tilde{\mathbf{F}}, \ldots, \mathbf{\Lambda}_R\tilde{\mathbf{F}}]$ et $\delta_2 = [\mathbf{B}_1, \mathbf{B}_2, \ldots, \mathbf{B}_R]$. D'après l'équation (2.40), $\delta_2 = \mathbf{F}\delta_1 = \mathbf{I}_R\mathbf{F}\delta_1$. En appliquant la propriété (2.41), nous obtenons donc :

$$vec(\delta_2) = \left(\delta_1^T \otimes \mathbf{I}_R\right) vec(\mathbf{F}). \tag{2.45}$$

   $\mathbf{F}$ se déduit facilement de cette équation par multiplication à gauche de $vec(\delta_2)$ par l'inverse de $\left(\delta_1^T \otimes \mathbf{I}_R\right)$.

3. Mise à jour de l'estimée de $\tilde{\mathbf{F}}$
   Posons $\delta_3 = \left[(\mathbf{F}\mathbf{\Lambda}_1)^T, (\mathbf{F}\mathbf{\Lambda}_2)^T, \ldots, (\mathbf{F}\mathbf{\Lambda}_R)^T\right]^T$ et $\delta_4 = \left[\mathbf{B}_1^T, \mathbf{B}_2^T, \ldots, \mathbf{B}_R^T\right]^T$. D'après l'équation (2.40), $\delta_4 = \delta_3\tilde{\mathbf{F}} = \delta_3\tilde{\mathbf{F}}\mathbf{I}_R$. En appliquant la propriété (2.41), nous obtenons une expression de $vec(\delta_4)$ en fonction de $vec(\tilde{\mathbf{F}})$:

$$vec(\delta_4) = (\mathbf{I}_R \otimes \delta_3)\, vec(\tilde{\mathbf{F}}), \tag{2.46}$$

   $\tilde{\mathbf{F}}$ se déduit facilement de cette équation par multiplication à gauche de $vec(\delta_4)$ par l'inverse de $(\mathbf{I}_R \otimes \delta_3)$.

Pour initialiser cet algorithme, nous pouvons choisir $\mathbf{F}_{init}$ égale à la matrice propre de $\mathbf{B}_1\mathbf{B}_2^{-1}$ et $\tilde{\mathbf{F}}_{init}$ égale à la transposée de $\mathbf{F}_{init}$.
On décide que l'algorithme a convergé lorsque la norme de Frobenius de la différence entre l'estimée de $\mathbf{F}$ à l'itération k et son estimée à l'itération $k+1$ est inférieure à une certaine tolérance ϵ. Un synopsis de SD-ALS est donné dans la table 2.3.

---

1. Initialisation : $\mathbf{F}$ est la matrice propre de $\mathbf{B}_1\mathbf{B}_2^{-1}$ et $\tilde{\mathbf{F}} = \mathbf{F}^T$
2. Mise à jour de $\mathbf{\Lambda}$: $\mathbf{\Lambda} = (\tilde{\mathbf{F}}^T \odot \mathbf{F})^\dagger [vec(\mathbf{B}_1), vec(\mathbf{B}_2), \ldots, vec(\mathbf{B}_R)]$
3. Mise à jour de $\mathbf{F}$: $\mathbf{F} = unvec((\delta_1^T \otimes \mathbf{I}_R)^\dagger vec(\delta_2))$
4. Mise à jour de $\tilde{\mathbf{F}}$: $\tilde{\mathbf{F}} = unvec((\mathbf{I}_R \otimes \delta_3)^\dagger vec(\delta_4))$
5. Aller en (2) jusqu'à convergence

TAB. 2.3 – Résumé de l'algorithme SD-ALS

2.4.3.2 Résolution du système par un algorithme de type QZ étendu

L'algorithme QZ étendu a été développé par Van der Veen et Paulraj [52]. L'idée de cet algorithme est de transformer le problème de diagonalisation conjointe en un problème de triangulation conjointe par des matrices unitaires. Nous noterons cette solution SD-QZ pour Simultaneous Diagonalization by extended QZ-iteration (diagonalisation simultanée par QZ étendu).
Voyons maintenant le principe de cet algorithme.
La décomposition QR de $\mathbf{F}$ et la décomposition RQ de $\mathbf{F}^T$ s'écrivent :

$$\mathbf{F} = \mathbf{Q}^H\mathbf{R}' \tag{2.47}$$
$$\mathbf{F}^T = \mathbf{R}''\mathbf{Z}^H \tag{2.48}$$

où $\mathbf{Q}$ et $\mathbf{Z}$ sont des matrices unitaires et $\mathbf{R}'$ et $\mathbf{R}''$ des matrices triangulaires supérieures.
En remplaçant dans (2.37) $\mathbf{F}$ et $\mathbf{F}^T$ par leur expression dans (2.47) et (2.48), nous obtenons le système suivant :

$$\begin{cases} \mathbf{Q}\mathbf{B}_1\mathbf{Z} &= \mathbf{R}_1 = \mathbf{R}'\mathbf{\Lambda}_1\mathbf{R}'' \\ \mathbf{Q}\mathbf{B}_2\mathbf{Z} &= \mathbf{R}_2 = \mathbf{R}'\mathbf{\Lambda}_2\mathbf{R}'' \\ & \vdots \\ \mathbf{Q}\mathbf{B}_R\mathbf{Z} &= \mathbf{R}_R = \mathbf{R}'\mathbf{\Lambda}_R\mathbf{R}'' \end{cases} \tag{2.49}$$

où les matrices $\mathbf{R}_1, \mathbf{R}_2, \ldots, \mathbf{R}_R$ sont triangulaires supérieures.
Il s'agit donc de trouver les matrices unitaires $\mathbf{Q}$ et $\mathbf{Z}$ telles que les matrices $\mathbf{Q}\mathbf{B}_i\mathbf{Z}$ pour $i \in [1:R]$ soient les plus triangulaires supérieures possibles. L'algorithme est itératif, on met à jour à chaque étape alternativement $\mathbf{Q}$ et $\mathbf{Z}$.

Considérons tout d'abord la mise à jour de $\mathbf{Q}$. Nous notons $\mathbf{R}_i^{(k)}$, $\mathbf{Q}^{(k)}$ et $\mathbf{Z}^{(k)}$ les estimées respectivement de $\mathbf{R}_i$, $\mathbf{Q}$ et $\mathbf{Z}$ après l'itération k. Nous avons

$$\mathbf{R}_i^{(k)} = \mathbf{Q}^{(k)}\mathbf{B}_i\mathbf{Z}^{(k)}, \forall i \in [1:R] \tag{2.50}$$

Nous cherchons une matrice unitaire $\tilde{\mathbf{Q}}$ telle que les matrices $\tilde{\mathbf{Q}}\mathbf{R}_i^{(k)}$ soient conjointement plus triangulaires supérieures que les matrices $\mathbf{R}_i^{(k)}$, $i \in [1:R]$. La matrice $\tilde{\mathbf{Q}}$ est construite comme le produit de matrices unitaires imposant une structure triangulaire supérieure sur la première, la deuxième, la $(R-1)$ème colonne de $\mathbf{R}_i^{(k)}$

$$\mathbf{Q}^{(k+1)} = \tilde{\mathbf{Q}}\mathbf{Q}^{(k)} = \left[\begin{array}{c|c} \mathbf{I}_{R-2} & \mathbf{0} \\ \hline \mathbf{0} & \mathbf{H}_{R-1} \end{array}\right] \cdots \left[\begin{array}{c|c} 1 & \mathbf{0} \\ \hline \mathbf{0} & \mathbf{H}_2 \end{array}\right] \mathbf{H}_1\mathbf{Q}^{(k)}, \tag{2.51}$$

où les matrices $\mathbf{H}_r$ de taille $(R-r+1) \times (R-r+1)$ sont unitaires.

Intéressons nous pour le moment au calcul de la matrice $\mathbf{H}_1$. On note $\mathbf{W}_1$ la matrice de taille $R \times R$ contenant la première colonne de chacune des matrices $\mathbf{R}_r^{(k)}$, $r \in [1:R]$. $\mathbf{H}_1$ est telle que le produit $\mathbf{H}_1\mathbf{W}_1$ est une matrice dont la première ligne est non nulle et dont les autres éléments sont les plus nuls possible (voir la figure 2.4).
La première ligne de $\mathbf{H}_1$ est le vecteur V^H qui maximise la fonction de coût

$$f(V) = V^H\mathbf{W}_1\mathbf{W}_1^HV. \tag{2.52}$$

Le vecteur singulier dominant de gauche de $\mathbf{W}_1$ est solution. Les autres lignes de $\mathbf{H}_1$ peuvent être choisies comme n'importe quelle base orthonormale orthogonale à la première ligne. La matrice $\mathbf{H}_1$ peut par conséquent être choisie comme la transposée hermitienne de la matrice des vecteurs singuliers de gauche de $\mathbf{W}_1$.

Nous appliquons maintenant le même principe aux deuxièmes colonnes des $\mathbf{R_i}$, $i \in [1:R]$. La matrice $\mathbf{H}_2$ de taille $(R-1) \times (R-1)$ est telle que le produit de $\mathbf{H}_2$ avec la matrice $\mathbf{W}_2$ de taille $R-1 \times R$, contenant les deuxièmes colonnes des $\mathbf{R}_i$, $i \in [1:R]$ de la deuxième ligne à la dernière, est une matrice dont la première ligne est non nulle et dont les autres éléments sont les plus nuls possible. $\mathbf{H}_2$ peut être choisie comme la transposée hermitienne de la matrice des vecteurs singuliers de gauche de $\mathbf{W}_2$.

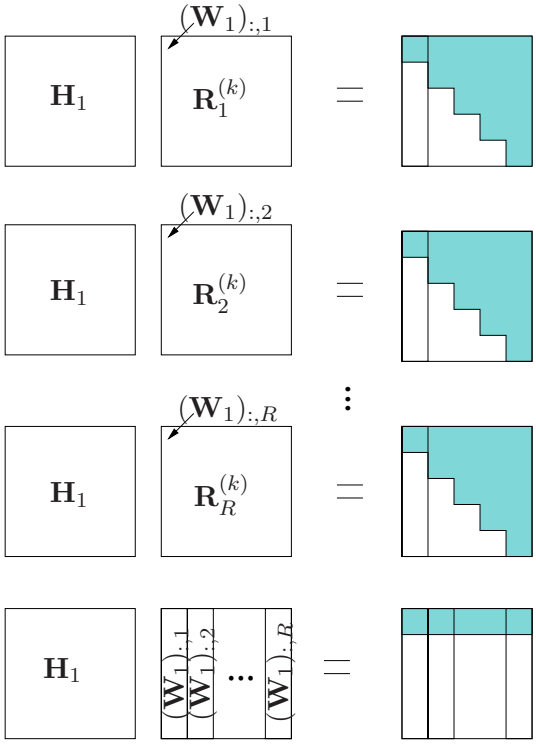

FIG. 2.4 – Construction de la matrice $\mathbf{W}_1$.

Nous pouvons ensuite appliquer le même principe pour identifier les matrices $\mathbf{H}_3, \ldots, \mathbf{H}_{R-1}$.

La mise à jour de $\mathbf{Z}$ suit le même principe. Nous posons

$$\mathbf{Z}^{(k+1)} = \mathbf{Z}^{(k)}\mathbf{G}_1 \left[\begin{array}{c|c} 1 & \mathbf{0} \\ \hline \mathbf{0} & \mathbf{G}_2 \end{array}\right] \ldots \left[\begin{array}{c|c} \mathbf{I}_{R-2} & \mathbf{0} \\ \hline \mathbf{0} & \mathbf{G}_{R-1} \end{array}\right], \tag{2.53}$$

où $\mathbf{G}_1$, $\mathbf{G}_2$, ..., $\mathbf{G}_{R-1}$ sont des matrices unitaires qui vont imposer une structure triangulaire respectivement sur la Rème, $R-1$ème, ..., 2ème ligne des $\tilde{Q}\mathbf{R}_r$, $r \in [1:R]$. A titre d'exemple, examinons la construction de $\mathbf{G}_1$. Notons Z la dernière colonne de $\mathbf{G}_1$. Nous cherchons le vecteur Z qui maximise la fonction de coût $\tilde{f}$ définie par :

$$\tilde{f}(Z) = Z^H \tilde{\mathbf{W}}_1^H \tilde{\mathbf{W}}_1 Z, \tag{2.54}$$

où $\tilde{\mathbf{W}}_1^H$ désigne la matrice contenant la dernière ligne de chacune des matrices $\tilde{\mathbf{Q}}\mathbf{R}_r^{(k)}$, $r \in [1:R]$. La matrice $\mathbf{G}_1$ est choisie égale à la matrice des vecteurs singuliers de droite de $\tilde{\mathbf{W}}_1$.
De la même manière, $\mathbf{G}_2$ est obtenue à partir des matrices $\tilde{\mathbf{Q}}\mathbf{R}_r^{(k)}\mathbf{G}_1$ auxquelles on a retiré la dernière ligne et la dernière colonne.

L'algorithme peut être initialisé à l'aide des matrices $\mathbf{B}_1$ et $\mathbf{B}_2$. En effet, le produit $\mathbf{B}_1^\dagger \mathbf{B}_2$ s'écrit :

$$\mathbf{B}_1^\dagger \mathbf{B}_2 = \mathbf{Z}\mathbf{R}_1^\dagger \mathbf{R}_2 \mathbf{Z}^H \tag{2.55}$$

Or, la décomposition de Schur de $\mathbf{B}_1^\dagger \mathbf{B}_2$ s'écrit :

$$\mathbf{B}_1^\dagger \mathbf{B}_2 = \mathbf{Z}_1 \mathbf{T}_1 \mathbf{Z}_1^H \tag{2.56}$$

où $\mathbf{Z}_1$ est une matrice unitaire et $\mathbf{T}_1$ est une matrice triangulaire supérieure.
Nous pouvons donc choisir $\mathbf{Z}_{init} = \mathbf{Z}_1$.
D'autre part, le produit $\mathbf{B}_2\mathbf{Z}$ s'écrit :

$$\mathbf{B}_2\mathbf{Z} = \mathbf{Q}^H \mathbf{R}_1 \tag{2.57}$$

Et la décompostion QR de ce produit $\mathbf{B}_2\mathbf{Z}$ s'écrit :

$$\mathbf{B}_2\mathbf{Z} = \mathbf{Q}_1^H \mathbf{T}_2 \tag{2.58}$$

où $\mathbf{Q}_1$ est une matrice unitaire et $\mathbf{T}_2$ est une matrice triangulaire supérieure.
Nous pouvons donc choisir $\mathbf{Q}_{init} = \mathbf{Q}_1^H$.

L'algorithme est arrêté lorsque la norme de Frobenius $\|\mathbf{Q}^{(k+1)} - \mathbf{Q}^{(k)}\|$ est inférieure à une certaine tolérance ϵ.

Une fois que l'on a estimé $\mathbf{Q}$, $\mathbf{Z}$ et $\mathbf{R}_r$, il s'agit d'estimer la matrice $\mathbf{F}$.
Soit $\mathbf{D}_1$ la matrice de taille $R \times R$ dont la $r$ème ligne, $r \in [1 : R]$ contient la diagonale de la matrice $\mathbf{R}_r$:

$$\mathbf{D}_1 = \begin{bmatrix} vecdiag(\mathbf{R}_1)^T \\ vecdiag(\mathbf{R}_2)^T \\ \vdots \\ vecdiag(\mathbf{R}_R)^T \end{bmatrix} \tag{2.59}$$

Il a été montré dans [52] que pour tout $k \in [1 : R]$, la $k$ème colonne de $\mathbf{F}$ vérifie :

$$\sum_{r=1}^{R} (\mathbf{D}^{-1})_{kr} \mathbf{B}_r = \alpha_k F_k F_k^T, \tag{2.60}$$

où α_k est un scalaire. En pratique, la matrice $\sum_{r=1}^{R} ({\mathbf{D}_1}^{-1})_{kr} \mathbf{B}_r$ dans (2.60) n'est pas exactement de rang un et l'on estime F_k comme le vecteur singulier dominant de gauche de cette matrice.
Un synopsis de l'algorithme SD-QZ est donné dans la table 2.4.

2.5 Simulations

Dans une première simulation, nous avons comparé les performances de l'algorithme ALS lorsqu'il est initialisé de manière aléatoire (nous le noterons DALS pour Direct ALS afin de le différencier de l'algorithme présenté dans le paragraphe 2.4.3) et lorsqu'il est initialisé à l'aide de la technique présentée dans le paragraphe 2.3.3. On notera cet algorithme DALS-WI pour « DALS Well Initialised ». Nous avons considéré $I = J = 4$, $K = 2000$, $R = 4$. La tolérance est fixée à $\epsilon_{\text{ALS}} = 10^{-7}$. Les éléments des matrices $\mathbf{A}$ et $\mathbf{H}$ suivent une loi gaussienne de moyenne nulle et de variance un et les éléments de la matrice $\mathbf{S}$ appartiennent à une constellation QPSK. L'erreur estimée est le nombre d'éléments de $\mathbf{S} - \hat{\mathbf{S}}$ différents de 0 divisé par le nombre total d'éléments de

1. Initialisation :
   On calcule la décomposition de Schur de $\mathbf{B}_1^\dagger\mathbf{B}_2$: $[\mathbf{Z}_1, \mathbf{T}] = schur(\mathbf{B}_1^\dagger\mathbf{B}_2)$
   $\mathbf{Z}_{init}$ est choisi tel que $\mathbf{Z}_{init} = \mathbf{Z}_1$
   On calcule la décomposition QR de $\mathbf{B}_2\mathbf{Z}_{init}$: $[\mathbf{Q}_1, \mathbf{T}_2] = qr(\mathbf{B}_2\mathbf{Z}_{init})$
   $\mathbf{Q}_{init}$ est choisi tel que $\mathbf{Q}_{init} = \mathbf{Q}_1^H$
2. Mise à jour de $\mathbf{Q}$:
   pour $r = 1 : R-1$
   – $\mathbf{W}_r = [\mathbf{R}_1(r:R,r), \mathbf{R}_2(r:R,r), \ldots, \mathbf{R}_R(r:R,r)] \in \mathbb{C}^{(R-r+1)\times R}$
   – $\mathbf{H}_r$ est la transposée hermitienne de la matrice propre de gauche de $\mathbf{W}_r$
   fin
   $\mathbf{Q} = \left[\begin{array}{c|c} \mathbf{I}_{R-2} & \mathbf{0} \\ \hline \mathbf{0} & \mathbf{H}_{R-1} \end{array}\right] \ldots \left[\begin{array}{c|c} 1 & \mathbf{0} \\ \hline \mathbf{0} & \mathbf{H}_2 \end{array}\right] \mathbf{H}_1\mathbf{Q}$
3. Mise à jour de $\mathbf{Z}$:
   pour $r = 1 : R-1$
   – $\tilde{\mathbf{W}}_r = [\mathbf{R}_1(R-r+1, 1:R-r+1)^T, \mathbf{R}_2(R-r+1, 1:R-r+1)^T, \ldots, \mathbf{R}_R(R-r+1, 1:R-r+1)^T]^T \in \mathbb{C}^{R\times(R-r+1)}$
   – $\mathbf{G}_r$ est la matrice propre de droite de $\tilde{\mathbf{W}}_r$
   fin
   $\mathbf{Z} = \mathbf{Z}\mathbf{G}_1 \left[\begin{array}{c|c} 1 & \mathbf{0} \\ \hline \mathbf{0} & \mathbf{G}_2 \end{array}\right] \ldots \left[\begin{array}{c|c} \mathbf{I}_{R-2} & \mathbf{0} \\ \hline \mathbf{0} & \mathbf{G}_{R-1} \end{array}\right]$
4. Aller en (2) jusqu'à convergence

TAB. 2.4 – Résumé de l'algorithme SD-QZ

$\mathbf{S}$ (c'est-à-dire KR), où $\hat{\mathbf{S}}$ désigne l'estimée de la matrice $\mathbf{S}$ après décision. L'algorithme DALS est initialisé à l'aide de trois valeurs initiales distinctes. Seule la meilleure performance est retenue.

La figure 2.5 représente l'erreur moyenne sur 100 simulations. Pour un Rapport Signal sur Bruit grand, l'algorithme DALS-WI donne de meilleurs résultats. Si le RSB est faible, on peut supposer que l'initialisation que nous avons calculée n'est pas meilleure qu'une initialisation aléatoire. La figure 2.6 montre le temps de calcul nécessaire dans les deux cas. Lorsque l'ALS est bien initialisé, la fonction de coût est proche de son minimum et l'ALS converge plus vite. La figure 2.7 présente le pourcentage de cas où l'ALS initialisé aléatoirement converge vers un minimum local : l'algorithme ne trouve pas son minimum global dans environ 10% des cas.

Il est également possible d'accélérer la convergence de l'algorithme ALS en utilisant une recherche linéaire optimisée (enhanced linesearch, ELS) [39].

Dans la deuxième simulation, nous avons comparé les performances des algorithmes présentés dans le paragraphe 2.4.3. Nous avons considéré $R = 7$, $I = 4$, $J = 4$, $K = 200$. Les éléments des matrices $\mathbf{A}$ et $\mathbf{H}$ suivent une loi gaussienne de moyenne nulle et de variance un et les éléments de la matrice $\mathbf{S}$ appartiennent à une constellation QPSK.

Les performances ont été moyennées sur 100 essais.

Nous avons comparé les performances de (1) SD-QZ avec une tolérance $\epsilon_{\text{SD-QZ}} = 10^{-1}$, (2) SD-ALS avec une tolérance $\epsilon_{\text{SD-ALS}} = 10^{-1}$, (3) DALS avec une tolérance $\epsilon_{\text{DALS}} = 10^{-7}$, en partant

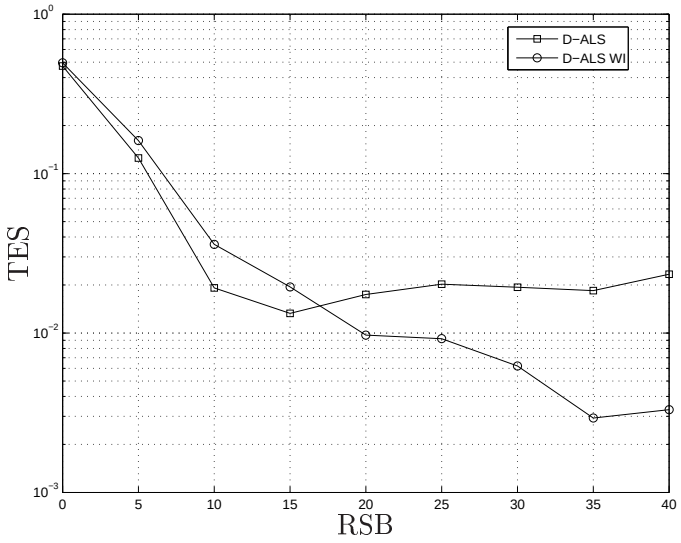

FIG. 2.5 – Erreur moyenne en fonction du RSB dans la première simulation ($I = J = 4$, $K = 2000$, $R = 4$).

de trois valeurs initiales aléatoires, et (4) du filtre Minimum Mean Square Error (MMSE), en considérant $\mathbf{A}$ et $\mathbf{H}$ et la puissance du bruit connues. D'autre part, afin de réduire le temps de calcul, l'algorithme DALS été appliqué au tenseur réduit des observations, de taille $(I \times J \times IJ)$ comme cela été présenté dans le pararaphe 2.3.3.

Les résultats obtenus sont présentés dans les figures 2.8–2.10.

Sur la figure 2.8 nous avons tracé le taux d'erreur symbole *médian* (TES) en fonction du rapport signal sur bruit (RSB). Les trois algorithmes montrent des performances équivalentes, proches des performances du filtre non-aveugle MMSE. La tolérance de l'algorithme DALS a été choisie égale à 10^{-7}, si l'on augmente la tolérance on augmente en conséquence le taux d'erreur médian. Sur la figure 2.9, nous avons tracé le taux d'erreur *moyen*. Nous pouvons constater que l'algorithme DALS semble avoir atteint un palier vers 10 dB. L'allure de la courbe s'explique par le fait que l'algorithme n'a pas trouvé l'optimum global dans environ 10% des essais. Ce pourcentage augmente losque ϵ_{DALS} augmente.

Dans la figure 2.10 nous avons tracé le temps de calul en fonction du Rapport Signal sur Bruit. Pour l'ALS, nous avons choisi le temps nécessaire pour l'initialisation donnant la meilleure performance. Par conséquent, le temps de calcul total, pour 3 valeurs initiales, est au moins 3 fois plus grand. Par ailleurs, l'algorithme DALS semble rencontrer des difficultés dès que le rang R du tenseur est supérieur à I et J. Lorsque le rang du tenseur est grand, les algorithmes SD-ALS et SD-QZ présentent des performances supérieures et sont beaucoup moins coûteux que l'algorithme DALS. L'algorithme SD-QZ est par ailleurs légèrement plus rapide que l'algorithme SD-ALS et présente des performances similaires.

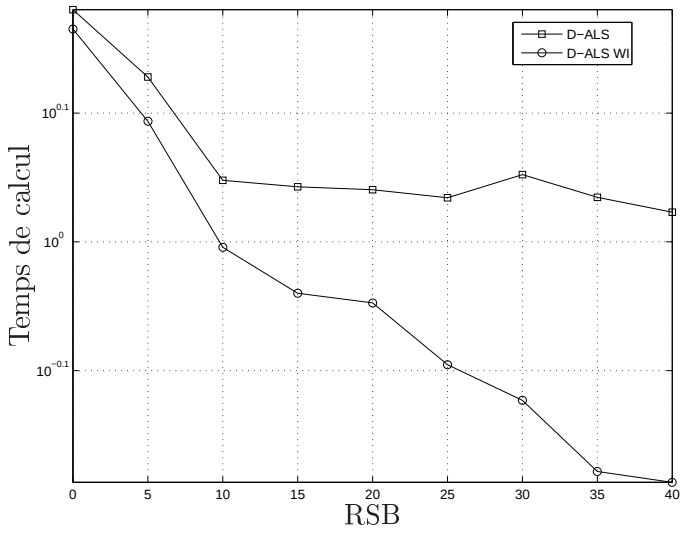

FIG. 2.6 – Temps de calcul moyen en fonction du RSB dans la première simulation ($I = J = 4$, $K = 2000$, $R = 4$).

2.6 Conclusion

Dans ce chapitre, nous avons présenté la décomposition PARAFAC d'un tenseur, une condition d'unicité, dite condition de Kruskal, et les moyens d'estimer ses paramètres. Nous avons montré qu'il existait un lien entre cette décomposition et la résolution d'un système de matrices à diagonaliser conjointement.
Nous avons montré qu'il était possible par là d'estimer les paramètres de la décomposition PARAFAC même lorsque la condition de Kruskal n'est pas vérifiée et nous avons proposé une nouvelle borne sur le rang du tenseur sous laquelle il est possible de les estimer. Cette borne dépend de manière quadratique de deux des dimensions du tenseur tandis que la borne de Kruskal en dépend de manière linéaire. Nous avons également montré qu'il était possible d'estimer le rang du tenseur simplement à l'aide d'une SVD.
Nous avons enfin proposé un algorithme constructif pour déterminer les paramètres de la décomposition. Cet algorithme mène à un système de matrices à diagonaliser conjointement, et nous avons proposé deux algorithmes pour résoudre ce système.

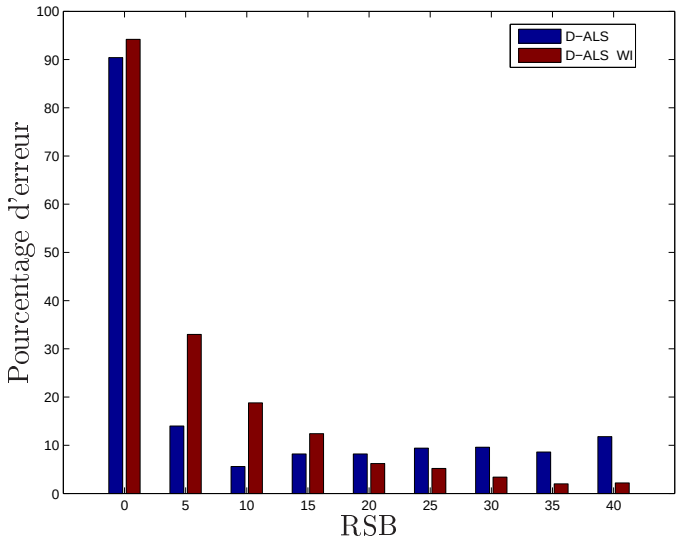

FIG. 2.7 – Pourcentage de convergence vers un minimum local en fonction du RSB dans la première simulation ($I = J = 4$, $K = 2000$, $R = 4$).

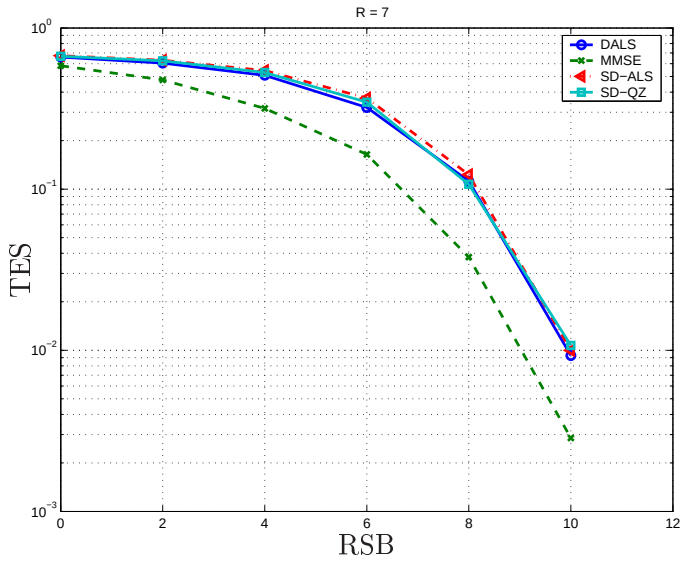

FIG. 2.8 – TES médian en fonction du RSB dans la deuxième simulation ($I = J = 4$, $K = 200$, $R = 7$).

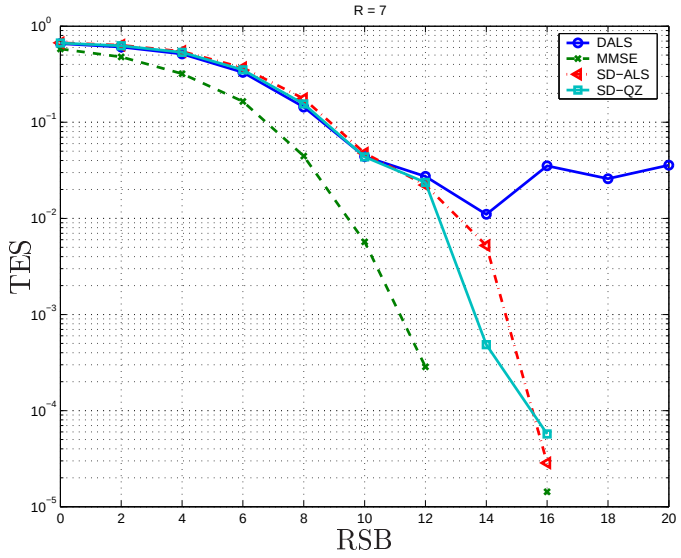

FIG. 2.9 – TES moyen en fonction du RSB dans la deuxième simulation ($I = J = 4$, $K = 200$, $R = 7$).

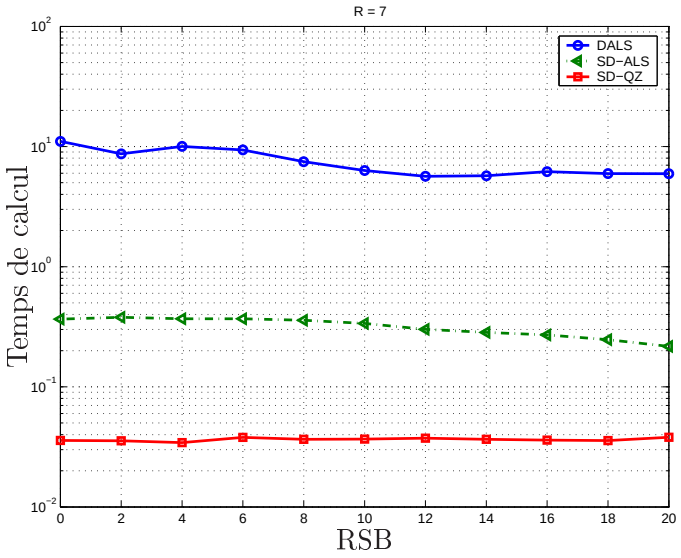

FIG. 2.10 – Temps de calcul moyen en fonction du RSB dans la deuxième simulation ($I = J = 4$, $K = 200$, $R = 7$).

Chapitre 3

Application aux systèmes CDMA

3.1 Introduction

Les premiers systèmes de radiocommunication mobiles, des postes radio militaires tactiques, sont nés aux Etats-Unis pendant la deuxième Guerre Mondiale. La fin de la guerre sonne le démarrage des radiocommunications civiles au travers de ce que l'on appelle les PMR (Private Mobile Radiocommunications). D'abord cantonnées aux usages des services de sécurité, les PMR se démocratisent et gagnent toute sorte d'activités professionnelles. Essentiellement analogiques, les premiers standards PMR sont assez limités : ils ne bénéficient que de couvertures géographiques limitées, et ne sont pas interopérables. Parallèlement à ces standards, et souvent en concurrence avec la PMR, se développent des standards dits « cellulaires », à partir des années 70, avec le déploiement d'un des premiers réseaux mobiles en Suède. En France, c'est le système analogique utilisant un multiplexage fréquentiel (FDMA) Radiocom 2000 qui se développe à partir de 1986. La deuxième génération de systèmes cellulaires utilise un multiplexage temporel et fréquentiel (F/TDMA), comme la norme GSM déployée à partir de 1991, ou déjà un multiplexage par code (CDMA) comme la norme IS-95 déployée aux Etats-Unis à partir de 1993. Le besoin de proposer aux utilisateurs des débits plus importants conduit les constructeurs et équipementiers à s'associer au sein du groupe 3GPP (Third Generation Partnership Project) pour élaborer la troisième génération de système cellulaire. C'est de ce groupe de normalisation que naît l'UMTS (Universal Mobile Telecommunication System [40]), destiné à remplacer le GSM et ses évolutions GPRS et EDGE en Europe. D'autres standards ou normes 3G apparaissent, comme le CDMA 2000 (combinaison d'une technologie OFDM et d'un accès par codes), ou le FOMA japonais, première norme 3G déployée. Tous ces systèmes utilisent aujourd'hui une technologie CDMA, ou WCDMA.

L'étalement de spectre par séquence directe ou DS-CDMA, a tout d'abord été utilisé dans des applications militaires : radios tactiques américaines, puis implémentation pour le système global de positionnement GPS. Cette technologie a ensuite été déployée dans des systèmes civils de radiocommunications mobiles, tout d'abord 2G (IS-95 / CDMA-One), puis 3G (UMTS). L'avantage des techniques CDMA est de présenter une grande résistance au brouillage bande étroite, et une certaine robustesse aux interférences inter-utilisateurs. Elles présentent de plus un intérêt en terme de discrétion en raison des largeurs de bande employées.

Dans le cas d'une liaison coopérative, l'émetteur envoie à son destinataire des séquences d'appren-

tissage lui permettant d'estimer le canal. D'autre part, le récepteur connaît le code d'étalement de l'émetteur, ce qui va lui permettre d'annuler la contribution des autres émetteurs et d'estimer le signal qui lui est destiné. Le cas aveugle est différent. Nous supposons que le récepteur n'a pas connaissance de la séquence d'étalement et ne possède pas d'information sur le canal. Dans ce chapitre nous allons montrer que nous pouvons estimer les signaux émis de manière aveugle en s'appuyant sur la structure PARAFAC des données CDMA.

Nous commencerons, dans le paragraphe 3.2, par présenter le modèle des signaux, le paragraphe 3.3 est consacré à la méthode qui est employée pour estimer les signaux sources dans le cas coopératif. Dans le paragraphe 3.4, nous montrerons que nous pouvons estimer les sources de manière aveugle, en s'appuyant sur la structure PARAFAC des signaux CDMA. Dans le paragraphe 3.5, nous exposerons une autre technique pour extraire les sources qui s'appuie sur l'hypothèse que les sources sont de module constant et dans le paragraphe 3.6, nous montrerons comment combiner les résultats apportés par la structure PARAFAC des signaux et par la contrainte du module constant afin d'améliorer les performances du système. Nous présenterons quelques résultats de simulations dans le paragraphe 3.7 et enfin nous conclurons ce chapite par la paragraphe 3.8.

3.2 Modèle

Le modèle qui va suivre est applicable à la liaison montante (du mobile à la station de base) ou à la liaison descendante (de la station de base au mobile). Il existe évidemment en pratique des différences, en particulier car les signaux sont synchrones en liaison descendante et ne le sont pas en liaison montante.
Un schéma de la transmission est donné dans la figure 3.1 [38, 53]. Nous supposons pour le moment que le canal est sans bruit et sans mémoire.
Nous supposons que R utilisateurs émettent des séquences d'information de longueur K. Ces séquences sont étalées à l'aide d'un code d'étalement de longueur J. On notera s_{kr} le $k$ème symbole d'information à transmettre par le $r$ème utilisateur et c_{jr} le $j$ème symbole de sa séquence d'étalement. Le signal x_{kjr} émis par l'utilisateur r à l'instant $kJ+j$ s'écrit comme le produit de son $k$ème symbole d'information s_{kr} et du $j$ème chip de sa séquence d'étalement c_{jr} :

$$x_{kjr} = s_{kr}c_{jr}, \tag{3.1}$$

Les signaux sont reçus sur un réseau de I antennes. Nous supposons que le mélange produit par le canal est linéaire et instantané. Le signal y_{ijk} reçu à l'instant $kJ+j$ sur l'antenne i s'écrit donc comme une combinaison linéaire des signaux émis par les R utilisateurs :

$$\begin{aligned} y_{ijk} &= \sum_{r=1}^{R} a_{ir}x_{kjr} \\ &= \sum_{r=1}^{R} a_{ir}s_{kr}c_{jr}, \quad i \in [1:I], j \in [1:J], k \in [1:K]. \end{aligned} \tag{3.2}$$

Finalement, le signal reçu sur l'antenne i à l'instant $kJ+j$ s'écrit comme la somme sur les utilisateurs du produit du coefficient de l'antenne i, du $k$ème symbole d'information et du $j$ème chip de la séquence d'étalement.

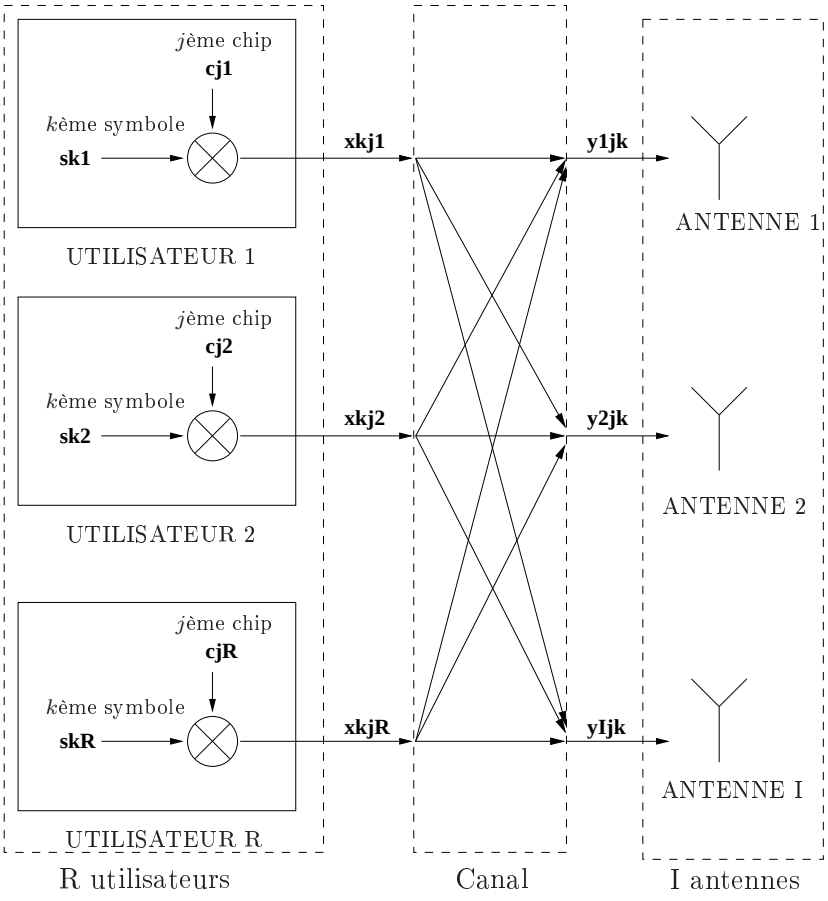

FIG. 3.1 – Schéma de la transmission des signaux CDMA

Ce modèle reste valable dans le cas où il y a de l'interférence entre chips (ICI) mais pas d'interférence entre symboles (ISI) [43]. Il suffit alors de remplacer dans (3.2) c_{jr} par h_{jr}, où h_{jr} désigne le produit de convolution entre la séquence d'étalement du $r$ème utilisateur et la réponse impulsionnelle du canal correspondant. Le modèle s'écrit maintenant :

$$y_{ijk} = \sum_{r=1}^{R} a_{ir} h_{jr} s_{kr}, \quad i \in [1:I], j \in [1:J], k \in [1:K]. \tag{3.3}$$

3.3 Cas coopératif

Ce paragraphe est consacré à l'estimation des signaux dans le cas coopératif. Le principe consiste à s'appuyer sur l'orthogonalité des codes d'étalement des différents utilisateurs.

Nous supposons donc les codes orthogonaux (ou pseudo-orthogonaux), au sens du produit scalaire

$< a, b >= (1/J)ab^T$. Si deux codes $\mathbf{c}^{(i)}$ et $\mathbf{c}^{(j)}$ sont orthogonaux, alors

$$< \mathbf{c}^{(i)}, \mathbf{c}^{(j)} > = 0 \text{ si } i \neq j \tag{3.4}$$

$$= 1 \text{ sinon.} \tag{3.5}$$

Notons $\mathbf{c}^{(r)} = [c_{1r}, c_{2r}, \ldots, c_{Jr}]$ la séquence d'étalement de l'émetteur r et $\mathbf{x}^{(r,k)} = [x_{k1r}, x_{k2r}, \ldots, x_{kJr}]$ la séquence qu'il émet entre les instants $kJ+1$ et $kJ+J$. On note d'autre part $\mathbf{y}^{(i,k)} = [y_{i1k}, y_{i2k}, \ldots, y_{iJk}]$ la séquence de longueur J reçue par l'utilisateur i entre les instants $kJ+1$ et $kJ+J$. Le récepteur effectue le produit scalaire entre le signal $\mathbf{y}^{(i,k)}$ qu'il reçoit et son code d'étalement.

$$< \mathbf{y}^{(i,k)}, \mathbf{c}^{(i)} > = \frac{1}{J}\sum_{j=1}^{J} y_{ijk}c_{ji} = \frac{1}{J}\sum_{j=1}^{J}\sum_{r=1}^{R} a_{ir}x_{kjr}c_{ji} \tag{3.6}$$

$$= \frac{1}{J}\sum_{j=1}^{J}\sum_{r=1}^{R} a_{ir}s_{kr}c_{jr}c_{ji} = \sum_{r=1}^{R} a_{ir}s_{kr}\frac{1}{J}\sum_{j=1}^{J} c_{jr}c_{ji} \tag{3.7}$$

$$= \sum_{r=1}^{R} a_{ir}s_{kr} < \mathbf{c}^{(r)}, \mathbf{c}^{(i)} > \tag{3.8}$$

$$= a_{ii}s_{ki} \tag{3.9}$$

Le récepteur retrouve donc le symbole s_{ki} à une amplitude près qui ne dépend que du canal entre l'émetteur et le récepteur.
Les hypothèses choisies sont très fortes : le récepteur doit connaître les séquences d'apprentissage, et celles-ci doivent être orthogonales (ou pseudo-orthogonales). Dans le cas aveugle que nous allons aborder maintenant, on suppose évidemment inconnues les séquences d'étalement, mais nous allons aussi supposer qu'elles ne sont pas nécessairement orthogonales.

3.4 Cas aveugle : application de la décomposition PARAFAC

Reprenons le modèle présenté en (3.3) :

$$y_{ijk} = \sum_{r=1}^{R} a_{ir}h_{jr}s_{kr}, \quad i \in [1:I], j \in [1:J], k \in [1:K]. \tag{3.10}$$

Ecrivons cette équation sous forme tensorielle. L'élément y_{ijk} peut être regardé comme l'élément d'indice ijk d'un tenseur $\mathcal{Y}$ de taille $I \times J \times K$:

$$\mathcal{Y} = \sum_{r=1}^{R} A_r \circ H_r \circ S_r, \tag{3.11}$$

où A_r, H_r, S_r représentent respectivement le vecteur des coefficients d'antenne, la séquence d'étalement (ou la séquence convoluée au canal) et le vecteur d'information de l'utilisateur r. Cette équation est une décomposition en facteurs parallèles du tenseur $\mathcal{Y}$. La structure PARAFAC des données DS-CDMA a été relevée pour la première fois dans [43].
Nous avons vu dans le chapitre 2 qu'il était également possible d'écrire cette équation sous forme matricielle :

$$\mathbf{Y} = (\mathbf{A} \odot \mathbf{H})\ \mathbf{S}^T, \tag{3.12}$$

où les colonnes de $\mathbf{A}$ sont les vecteurs A_r, les colonnes de $\mathbf{H}$ sont les vecteurs H_r et les colonnes de $\mathbf{S}$ les vecteurs S_r. Nous avons vu d'autre part qu'il existe une matrice $\mathbf{F}$ non singulière de taille $R \times R$ et a priori inconnue qui lie $\mathbf{A} \odot \mathbf{H}$ à $\mathbf{U}$ et $\mathbf{D}$ et $\mathbf{S}$ à $\mathbf{V}$:

$$\left\{ \begin{array}{rcl} \mathbf{A} \odot \mathbf{H} & = & \mathbf{UDF} \\ \mathbf{S}^T & = & \mathbf{F}^{-1}\mathbf{V}^H \end{array} \right. , \tag{3.13}$$

En exploitant la première équation de ce système, nous avons vu qu'il était possible d'estimer $\mathbf{F}$ à partir d'un système de matrices à diagonaliser conjointement :

$$\left\{ \begin{array}{rcl} \mathbf{B}_1 & = & \mathbf{F}\mathbf{\Lambda}_1\mathbf{F}^T \\ \mathbf{B}_2 & = & \mathbf{F}\mathbf{\Lambda}_2\mathbf{F}^T \\ & \vdots & \\ \mathbf{B}_R & = & \mathbf{F}\mathbf{\Lambda}_R\mathbf{F}^T \end{array} \right. \tag{3.14}$$

Les matrices recherchées $\mathbf{A}$, $\mathbf{H}$ et $\mathbf{S}$ peuvent ensuite se déduire de $\mathbf{F}$ à partir de l'équation (3.13).

Il est donc possible d'estimer les matrices $\mathbf{A}$, $\mathbf{H}$ et $\mathbf{S}$ en s'appuyant sur la structure des données. Nous allons maintenant voir qu'il est possible d'exploiter le fait que nous possédons d'autres hypothèses sur les sources.

3.5 Contrainte du Module Constant

Dans le paragraphe précédent, nous avons utilisé la première équation du système (3.13) et exploité la structure de la matrice $\mathbf{A} \odot \mathbf{H}$ pour estimer la matrice $\mathbf{F}$. Mais il est également possible d'utiliser la deuxième équation de ce système si nous avons certaines informations sur les sources. D'après (3.13) la matrice $\mathbf{V}$ des vecteurs singuliers de gauche de $\mathbf{Y}$ vérifie :

$$\mathbf{V}^H = \mathbf{F}\mathbf{S}^T \tag{3.15}$$

Cette expression est celle d'un mélange instantané par une matrice de taille $R \times R$. Il existe différentes méthodes pour résoudre ce problème, et nous avons retenu l'algorithme ACMA (Analytical Constant Modulus Algorithm) [51,52], qui s'applique à des sources de module constant. L'hypothèse de constance des sources est envisageable en communications numériques, où les symboles appartiennent souvent à une constellation circulaire. Nous avons choisi cet algorithme car, comme nous allons le voir, il mène à un système de matrices à diagonaliser conjointement très similaire au système que nous avons obtenu dans le paragraphe précédent. Nous allons maintenant expliquer brièvement le fonctionnement de cet algorithme.
Soit $\mathbf{w}^H$ une ligne quelconque de la matrice $\mathbf{F}^{-1}$. Notons d'autre part X_k la $k$ème colonne de la matrice $\mathbf{V}^H$. Nous supposons les sources de module constant égal à un, donc $\|s_{kr}\|^2 = 1$, $\forall k \in [1:K], \forall r \in [1:R]$. Ce qui peut encore s'écrire :

$$\mathbf{w}^H X_k X_k^H \mathbf{w} = 1, \forall k \in [1:K] \tag{3.16}$$

Cette équation peut encore s'écrire à l'aide du produit de Kronecker :

$$(X_k^* \otimes X_k)^H (\mathbf{w}^* \otimes \mathbf{w}) = 1 \tag{3.17}$$

Soit $\mathbf{P}$ la matrice de taille $K \times R^2$ définie par :

$$\mathbf{P} = \begin{bmatrix} (X_1^* \otimes X_1)^H \\ (X_2^* \otimes X_2)^H \\ \vdots \\ (X_K^* \otimes X_K)^H \end{bmatrix} \tag{3.18}$$

Nous cherchons le vecteur $\mathbf{y}$ tel que

$$\mathbf{P}\mathbf{y} = \mathbf{1} \text{ et } \mathbf{y} = (\mathbf{w}^* \otimes \mathbf{w}) \tag{3.19}$$

Comme il y a R sources, l'équation (3.19) possède au moins R solutions indépendantes $\mathbf{w}_1$, $\mathbf{w}_2, \ldots, \mathbf{w}_R$ qui sont les lignes de la matrice $\mathbf{F}^{-1}$.
Tout vecteur de la forme suivante est solution :

$$\lambda_1(\mathbf{w}_1^* \otimes \mathbf{w}_1) + \lambda_2(\mathbf{w}_1^* \otimes \mathbf{w}_1) + \ldots + \lambda_R(\mathbf{w}_R^* \otimes \mathbf{w}_R) \tag{3.20}$$

avec $\sum_{r=1}^{R} \lambda_r = 1$.
Sélectionnons R solutions indépendantes $\mathbf{y}_1, \mathbf{y}_2, \ldots, \mathbf{y}_R$ de l'équation (3.19) définies par

$$\mathbf{y}_r = \lambda_{r1}(\mathbf{w}_1^* \otimes \mathbf{w}_1) + \lambda_{r2}(\mathbf{w}_1^* \otimes \mathbf{w}_1) + \ldots + \lambda_{rR}(\mathbf{w}_R^* \otimes \mathbf{w}_R), \forall r \in [1:R] \tag{3.21}$$

et appliquons leur l'opérateur *unvec*. Nous obtenons le système suivant :

$$\left\{ \begin{array}{rcl} \mathbf{C}_1 & = & \lambda_{11}\mathbf{w}_1\mathbf{w}_1^H + \lambda_{12}\mathbf{w}_2\mathbf{w}_2^H + \ldots, \lambda_{1R}\mathbf{w}_R^H\mathbf{w}_R \\ \mathbf{C}_2 & = & \lambda_{21}\mathbf{w}_1\mathbf{w}_1^H + \lambda_{22}\mathbf{w}_2\mathbf{w}_2^H + \ldots, \lambda_{2R}\mathbf{w}_R^H\mathbf{w}_R \\ & \vdots & \\ \mathbf{C}_R & = & \lambda_{R1}\mathbf{w}_1\mathbf{w}_1^H + \lambda_{R2}\mathbf{w}_2\mathbf{w}_2^H + \ldots, \lambda_{RR}\mathbf{w}_R\mathbf{w}_R^H \end{array} \right. \tag{3.22}$$

avec $\mathbf{C}_r = \mathit{unvec}(\mathbf{y}_r)$ pour tout $r \in [1:R]$.
Soit $\mathbf{\Omega}_r$ la matrice diagonale dont les éléments diagonaux sont les $\lambda_{r1}, \lambda_{r2}, \ldots, \lambda_{rR}$, $r \in [1:R]$.
L'équation précédente peut encore s'écrire :

$$\left\{ \begin{array}{rcl} \mathbf{C}_1 & = & \mathbf{F}^{-H}\mathbf{\Omega}_1\mathbf{F}^{-1} \\ \mathbf{C}_2 & = & \mathbf{F}^{-H}\mathbf{\Omega}_2\mathbf{F}^{-1} \\ & \vdots & \\ \mathbf{C}_R & = & \mathbf{F}^{-H}\mathbf{\Omega}_R\mathbf{F}^{-1} \end{array} \right. . \tag{3.23}$$

L'hypothèse de sources de module constant nous permet donc d'aboutir à un système (3.23) très similaire au sytème (3.14) obtenu en exploitant la structure PARAFAC des données DS-CDMA dans le paragraphe 3.4. Dans le paragraphe suivant, nous allons proposer trois algorithmes permettant de résoudre conjointement ces deux systèmes.

3.6 Combinaison MC-PARAFAC

Dans ce paragraphe, nous proposons trois algorithmes permettant de résoudre le système combinant les systèmes (3.14) et (3.23). Les deux premiers algorithmes proposés décrits dans les paragraphes 3.6.1 et 3.6.2 sont des généralisations des algorithmes ALS et QZ étendu dont il a été question au chapitre 2. Le troisième algorithme, décrit dans le paragraphe 3.6.4 est un algorithme du type Jacobi.

3.6.1 Généralisation de l'ALS

Posons $\tilde{\mathbf{F}} = \mathbf{F}^T$. D'après les équations (3.14) et (3.23), nous devons résoudre le système suivant :

$$\left\{\begin{array}{rcl} \mathbf{B}_1 & = & \mathbf{F}\mathbf{\Lambda}_1\tilde{\mathbf{F}} \\ \mathbf{B}_2 & = & \mathbf{F}\mathbf{\Lambda}_2\tilde{\mathbf{F}} \\ \vdots & & \\ \mathbf{B}_R & = & \mathbf{F}\mathbf{\Lambda}_R\tilde{\mathbf{F}} \\ \mathbf{C}_1 & = & \tilde{\mathbf{F}}^{-*}\mathbf{\Omega}_1\mathbf{F}^{-1} \\ \mathbf{C}_2 & = & \tilde{\mathbf{F}}^{-*}\mathbf{\Omega}_2\mathbf{F}^{-1} \\ \vdots & & \\ \mathbf{C}_R & = & \tilde{\mathbf{F}}^{-*}\mathbf{\Omega}_R\mathbf{F}^{-1} \end{array}\right. \tag{3.24}$$

La fonction de coût à minimiser est :

$$f\left(\mathbf{F}, \tilde{\mathbf{F}}, \{\mathbf{\Lambda}_r\}, \{\mathbf{\Omega}_r\}\right) = \sum_{r=1}^{R}\left(\left\|\mathbf{B}_r - \mathbf{F}\mathbf{\Lambda}_r\tilde{\mathbf{F}}\right\|^2 + \left\|\tilde{\mathbf{F}}^*\mathbf{C}_r\mathbf{F} - \mathbf{\Omega}_r\right\|^2\right). \tag{3.25}$$

L'algorithme ALS généralisé consiste à mettre à jour de manière alternée les matrices $\mathbf{\Omega}_r$ et $\mathbf{\Lambda}_r$, la matrice $\mathbf{F}$ et la matrice $\tilde{\mathbf{F}}$.

1. Mise à jour des matrices $\mathbf{\Omega}_r$ et $\mathbf{\Lambda}_r$, $r \in [1:R]$
   En appliquant l'opérateur *vec* aux matrices $\mathbf{B}_r = \mathbf{F}\mathbf{\Lambda}_r\tilde{\mathbf{F}}$ et $\mathbf{C}_r = \tilde{\mathbf{F}}^{-*}\mathbf{\Omega}_r\mathbf{F}^{-1}$, $r \in [1:R]$, nous obtenons :

   $$vec(\mathbf{B}_r) = \left(\tilde{\mathbf{F}}^T \odot \mathbf{F}\right) vecdiag(\mathbf{\Lambda}_r) \tag{3.26}$$

   et

   $$\mathbf{\Omega}_r = diag\left(vecdiag(\tilde{\mathbf{F}}^*\mathbf{C}_r\mathbf{F})\right). \tag{3.27}$$

   Les matrices $\mathbf{\Lambda}_r$ et $\mathbf{\Omega}_r$, $r \in [1:R]$, peuvent ête évaluées à partir de ces équations linéaires.

2. Mise à jour de la matrice $\mathbf{F}$
   Notons $\Delta_1 = [\mathbf{\Lambda}_1\tilde{\mathbf{F}}, \mathbf{\Lambda}_2\tilde{\mathbf{F}}, \ldots, \mathbf{\Lambda}_R\tilde{\mathbf{F}}]$, $\Delta_2 = [\mathbf{B}_1, \mathbf{B}_2, \ldots, \mathbf{B}_R]$, $\Gamma_1 = [(\tilde{\mathbf{F}}^*\mathbf{C}_1)^T, (\tilde{\mathbf{F}}^*\mathbf{C}_2)^T, \ldots, (\tilde{\mathbf{F}}^*\mathbf{C}_R)^T]^T$ et $\Gamma_2 = [\mathbf{\Omega}_1^T, \mathbf{\Omega}_2^T, \ldots, \mathbf{\Omega}_R^T]^T$.
   Le système (3.24) peut être écrit à l'aide des matrices Δ_1, Δ_2, Γ_1 et Γ_2 de la manière suivante :

   $$\left\{\begin{array}{rcl} \Delta_2 & = & \mathbf{F}\Delta_1 = \mathbf{I}_R\mathbf{F}\Delta_1 \\ \Gamma_2 & = & \Gamma_1\mathbf{F} = \Gamma_1\mathbf{F}\mathbf{I}_R \end{array}\right. \tag{3.28}$$

   D'après la propriété (2.41), nous pouvons écrire :

   $$\left[\begin{array}{c} vec(\Delta_2) \\ vec(\Gamma_2) \end{array}\right] = \left[\begin{array}{c} \Delta_1^T \otimes \mathbf{I}_R \\ \mathbf{I}_R \otimes \Gamma_1 \end{array}\right] vec(\mathbf{F}). \tag{3.29}$$

   La matrice $\mathbf{F}$ peut ensuite être estimée à partir de ce système.

3. Mise à jour de la matrice $\tilde{\mathbf{F}}$

   Posons $\Delta_3 = \left[(\mathbf{F}\mathbf{\Lambda}_1)^T, (\mathbf{F}\mathbf{\Lambda}_2)^T, \ldots, (\mathbf{F}\mathbf{\Lambda}_R)^T\right]^T$, $\Delta_4 = \left[\mathbf{B}_1^T, \mathbf{B}_2^T, \ldots, \mathbf{B}_R^T\right]^T$, $\Gamma_3 = [\mathbf{C}_1\mathbf{F}, \mathbf{C}_2\mathbf{F}, \ldots, \mathbf{C}_R\mathbf{F}]$ et $\Gamma_4 = [\mathbf{\Omega}_1, \mathbf{\Omega}_2, \ldots, \mathbf{\Omega}_R]$. Le système (3.24) peut être écrit à l'aide des matrices Δ_3, Δ_4, Γ_3 et Γ_4 sous la forme :

$$\left\{ \begin{array}{l} \Delta_4 = \Delta_3\tilde{\mathbf{F}} = \Delta_3\tilde{\mathbf{F}}\mathbf{I}_R \\ \Gamma_4^* = \tilde{\mathbf{F}}\Gamma_3^* = \mathbf{I}_R\tilde{\mathbf{F}}\Gamma_3^* \end{array} \right. \tag{3.30}$$

   D'après la propriété (2.41), nous pouvons écrire :

$$\left[\begin{array}{c} vec(\Delta_4) \\ vec(\Gamma_4^*) \end{array} \right] = \left[\begin{array}{c} \mathbf{I}_R \otimes \Delta_3 \\ \Gamma_3^H \otimes \mathbf{I}_R \end{array} \right] vec(\tilde{\mathbf{F}}). \tag{3.31}$$

   La matrice $\tilde{\mathbf{F}}$ peut être estimée à partir de cette équation.

Cet algorithme sera noté dans la suite CSD-ALS pour Coupled Simultaneous Diagonalization by Alternating Least Squares (diagonalisation simultanée couplée par moindres carrés alternés). Comme dans le paragraphe 2.4.3, nous pouvons initialiser $\mathbf{F}$ avec la matrice propre de $\mathbf{B}_1\mathbf{B}_2^{-1}$ et $\tilde{\mathbf{F}}$ avec sa transposée. Par ailleurs, nous décidons que l'algorithme a convergé lorsque la norme de Frobenius de la différence entre l'estimée de $\mathbf{F}$ à l'itération k et son estimée à l'itération $k+1$ est inférieure à une certaine tolérance $\epsilon_{CSD-ALS}$.

3.6.2 Itération QZ généralisée

Dans cette partie, nous présentons une autre solution pour résoudre le problème de la diagonalisation simultanée des systèmes de matrices (3.14) et (3.23). Cet algorithme est une généralisation de l'algorithme SD-QZ, que nous noterons CSD-QZ pour Coupled Simultaneous Diagonalization by generalized QZ iteration (diagonalisation simultanée couplée par itération de QZ généralisée). En conservant les notations choisies dans les équations (2.47) et (2.48), nous pouvons écrire :

$$\mathbf{F}^{-H} = \mathbf{Q}^H(\mathbf{R}')^{-H} \tag{3.32}$$

$$\mathbf{F}^{-1} = (\mathbf{R}'')^{-T}\mathbf{Z}^T \tag{3.33}$$

Posons $\mathbf{L}_r = (\mathbf{R}')^{-H}\mathbf{\Omega}_r(\mathbf{R}'')^{-T}$, $r \in [1:R]$. Le système (3.24) s'écrit maintenant :

$$\left\{ \begin{array}{rcl} \mathbf{Q}\mathbf{B}_1\mathbf{Z} &=& \mathbf{R}_1 \\ \mathbf{Q}\mathbf{B}_2\mathbf{Z} &=& \mathbf{R}_2 \\ & \vdots & \\ \mathbf{Q}\mathbf{B}_R\mathbf{Z} &=& \mathbf{R}_R \\ \mathbf{Q}\mathbf{C}_1\mathbf{Z}^* &=& \mathbf{L}_1 \\ \mathbf{Q}\mathbf{C}_2\mathbf{Z}^* &=& \mathbf{L}_2 \\ & \vdots & \\ \mathbf{Q}\mathbf{C}_R\mathbf{Z}^* &=& \mathbf{L}_R \end{array} \right. \tag{3.34}$$

où les matrices $\mathbf{Q}$ et $\mathbf{Z}$ de taille $R \times R$ sont unitaires, où les matrices $\mathbf{R}_r$ de taille $R \times R$ sont triangulaires supérieures et où les matrices $\mathbf{L}_r$ de taille $R \times R$ sont triangulaires inférieures. Un schéma de la décomposition simultanée est donné dans la figure 3.2.

De manière similaire à ce qui été présenté dans le paragraphe 2.4.3, les matrices $\mathbf{Q}$ et $\mathbf{Z}$ sont mises à jour de manière alternées de façon à rendre les matrices $\mathbf{R}_r$, $r \in [1:R]$, les plus triangulaires supérieures possible et les matrices $\mathbf{L}_r$, $r \in [1:R]$, les plus triangulaires inférieures possible. Les matrices $\mathbf{Q}$ et $\mathbf{Z}$ possèdent les structures (2.51) et (2.53).
Nous notons $\mathbf{R}_i^{(k)}$, $\mathbf{L}_i^{(k)}$, $\mathbf{Q}^{(k)}$ et $\mathbf{Z}^{(k)}$ les estimées respectivement de $\mathbf{R}_i$, $\mathbf{L}_i$, $\mathbf{Q}$ et $\mathbf{Z}$ après l'itération k. Nous avons :

$$\mathbf{R}_i^{(k)} = \mathbf{Q}^{(k)}\mathbf{B}_i\mathbf{Z}^{(k)}, \forall i \in [1:R] \tag{3.35}$$
$$\mathbf{L}_i^{(k)} = \mathbf{Q}^{(k)}\mathbf{C}_i\mathbf{Z}^{(k)*}, \forall i \in [1:R]. \tag{3.36}$$

Pour $p \in [1:R-1]$, la matrice $\mathbf{H}_p$ va imposer une structure triangulaire supérieure à la $p$ème colonne de chaque matrice $\mathbf{R}_r^{(k)}$ et une structure triangulaire inférieure à la $(p+1)$ème colonne de $\mathbf{L}_r^{(k)}$. De manière analogue, la matrice $\mathbf{G}_p$ va imposer une structure triangulaire supérieure à la $R-p+1$ème ligne de $\mathbf{Q}\mathbf{R}_r^{(k)}$ et une structure triangulaire inférieure à la $(R-p)$ème ligne de $\mathbf{Q}\mathbf{L}_r^{(k)}$.

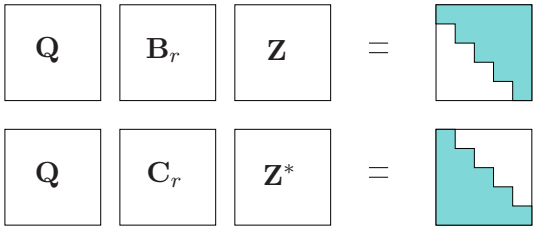

FIG. 3.2 – Visualisation du système (3.34).

Commençons par la mise à jour de $\mathbf{Q}$ et intéressons nous pour le moment au calcul de la matrice $\mathbf{H}_1$. $\mathbf{H}_1$ doit être telle que le produit de $\mathbf{H}_1$ avec la première colonne de chaque matrice $\mathbf{R}_r^{(k)}$ soit un vecteur dont le premier élément est non nul et dont les autres éléments sont les plus petits possible tandis que le produit de $\mathbf{H}_1$ avec la seconde colonne de chaque matrice $\mathbf{L}_r^{(k)}$ soit un vecteur dont le premier élément est le plus petit possible et les autres sont non nuls. Notons V^H la première ligne de la matrice $\mathbf{H}_1$, $\mathbf{W}_1$ la matrice de taille $R \times R$ contenant la première colonne de chaque matrice $\mathbf{R}_r^{(k)}$ et $\mathbf{W'}_1$ la matrice de taille $R \times R$ contenant la deuxième colonne de chaque matrice $\mathbf{L}_r^{(k)}$. Si nous imposons uniquement la contrainte PARAFAC, c'est-à-dire si nous n'exploitons que le système de matrices $\mathbf{B}_r$, alors nous pouvons estimer V en *maximisant* la fonction f_1 définie par :

$$f_1(V) = V^H\mathbf{W}_1\mathbf{W}_1^HV. \tag{3.37}$$

Si nous imposons uniquement la contrainte MC, c'est-à-dire si nous n'exploitons que le sytème de matrices $\mathbf{C}_r$, alors nous pouvons estimer V en *minimisant* la fonction f_2 définie par :

$$f_2(V) = V^H\mathbf{W'}_1\mathbf{W'}_1^HV. \tag{3.38}$$

Si l'on souhaite imposer les deux contraintes, alors nous pouvons estimer V en maximisant la fonction f définie par :

$$f(V) = V^H \mathbf{W}_1 \mathbf{W}_1^H V - V^H \mathbf{W'}_1 \mathbf{W'}_1^H V + \left\|\mathbf{W'}_1\right\|^2. \tag{3.39}$$

Dans cette équation, le terme de régularisation $\|\mathbf{W'}_1\|^2$ permet de garder f toujours positive. Le vecteur V optimal est le vecteur dominant de gauche de $\mathbf{W}_1\mathbf{W}_1^H - \mathbf{W'}_1\mathbf{W'}_1^H + \|\mathbf{W'}_1\|^2 \mathbf{I}_R$. La matrice $\mathbf{H}_1$ peut être choisie égale à la transposée hermitienne de la matrice propre de $\mathbf{W}_1\mathbf{W}_1^H - \mathbf{W'}_1\mathbf{W'}_1^H + \|\mathbf{W'}_1\|^2 \mathbf{I}_R$, avec les valeurs propres rangées par ordre décroissant.
La matrice $\mathbf{H}_2$, peut être obtenue en utilisant la même technique, à partir des matrices $\mathbf{H}_1\mathbf{R}_r^{(k)}$ et $\mathbf{H}_1\mathbf{L}_r^{(k)}$, auxquelles on aura ôté la première ligne et la première colonne. Nous pouvons appliquer le même principe pour identifier les matrices $\mathbf{H}_3, \ldots, \mathbf{H}_{R-1}$.

Analysons maintenant la mise à jour de $\mathbf{Z}$.
De manière similaire, la matrice $\mathbf{G}_1$ doit être telle que le produit de la dernière ligne de chaque matrice $\mathbf{Q}^{(k+1)}\mathbf{R}_r^{(k)}$ avec $\mathbf{G}_1$ soit un vecteur ligne dont le dernier élément est non nul et dont les autres éléments sont les plus petits possible tandis que le produit de l'avant-dernière ligne de chaque matrice $\mathbf{Q}^{(k+1)^*}\mathbf{L}_r^{(k)^*}$ avec $\mathbf{G}_1$ soit un vecteur ligne dont le dernier élément est est le plus petit possible. Notons $\tilde{\mathbf{W}}_1$ la matrice contenant la dernière ligne de chaque matrice $\mathbf{Q}^{(k+1)}\mathbf{R}_r^{(k)}$, ces lignes étant rangées les unes sous les autres et $\tilde{\mathbf{W}}'_1$ la matrice contenant l'avant-dernière ligne de chaque matrice $\mathbf{Q}^{(k+1)}\mathbf{L}_r^{(k)}$, ces lignes étant rangées les unes sous les autres. Soit Z la dernière colonne de $\mathbf{G}_1$. Z doit maximiser la fonction de coût $\tilde{f}$ définie par :

$$\tilde{f}(Z) = Z^H \tilde{\mathbf{W}}_1^H \tilde{\mathbf{W}}_1 Z - Z^H \left(\tilde{\mathbf{W}'}_1^H \tilde{\mathbf{W}'}_1\right)^* Z + \left\|\tilde{\mathbf{W}'}_1\right\|^2. \tag{3.40}$$

Le vecteur Z optimal est le vecteur propre de $\tilde{\mathbf{W}}_1^H\tilde{\mathbf{W}}_1 - \left(\tilde{\mathbf{W}'}_1^H\tilde{\mathbf{W}'}_1\right)^* + \left\|\tilde{\mathbf{W}'}_1\right\|^2 \mathbf{I}_R$, correspondant à sa plus grande valeur propre. Nous pouvons donc choisir la matrice $\mathbf{G}_1$ égale à la matrice propre unitaire de $\tilde{\mathbf{W}}_1^H\tilde{\mathbf{W}}_1 - \left(\tilde{\mathbf{W}'}_1^H\tilde{\mathbf{W}'}_1\right)^* + \left\|\tilde{\mathbf{W}'}_1\right\|^2 \mathbf{I}_R$, dont les valeurs propres sont ordonnées par ordre croissant. La matrice $\mathbf{G}_2$ est obtenue de la même manière à l'aide des matrices $\mathbf{Q}^{(k+1)}\mathbf{R}_r^{(k)}\mathbf{G}_1$ et $\mathbf{Q}^{(k+1)}\mathbf{L}_r^{(k)}\mathbf{G}_1$, auxquelles on aura ôté la dernière ligne et la dernière colonne. Nous pouvons ensuite appliquer le même principe pour identifier les matrices $\mathbf{G}_3, \ldots, \mathbf{G}_{R-1}$.
L'algorithme est arrêté lorsque la norme de Frobenius $\|\mathbf{Q}^{(k+1)} - \mathbf{Q}^{(k)}\|$ est inférieure à une certaine tolérance ϵ.
Une fois que l'on a estimé $\mathbf{Q}$ et $\mathbf{Z}$, il est possible d'obtenir deux estimées de $\mathbf{F}$ à partir de (3.34). En effet, on peut obtenir $\mathbf{F}$ à partir des matrices $\mathbf{B}_r$, comme dans le paragraphe 2.4.3.2. On peut égalemant obtenir $\mathbf{F}^{-H}$, et par conséquent $\mathbf{F}$ à partir des matrices $\mathbf{C}_r$. Dans ce cas, par analogie avec (2.59), les diagonales des matrices $\mathbf{L}_r$ sont rangées dans une matrice $\mathbf{D}_2$. Les colonnes de $\mathbf{F}^{-H}$ sont obtenues par analogie avec (2.60). En vertu de la structure de (3.34), les colonnes des deux estimées de $\mathbf{F}$ sont rangées dans le même ordre. L'estimée de $\mathbf{F}$ peut finalement être choisie comme la moyenne des deux solutions obtenues, éventuellement pondérées en fonction de la confiance accordée à chacun des systèmes (3.14) et (3.23), comme cela va être présenté dans le prochain paragraphe.

3.6.3 Pondération

Lorsque nous résolvons les systèmes (3.14) et (3.23), il peut être intéressant de prendre en compte la confiance que nous accordons à chacun d'eux. En réalité, il est difficile d'évaluer la précison de ces deux systèmes, qui dépend des caractéristiques du canal, du bruit, du nombre d'échantillons, etc. Nous adoptons ici une approche heuristique. La précision des deux systèmes est évaluée a posteriori.
La résolution de (3.14), nous donne une estimée $\hat{\mathbf{F}}_B$ de $\mathbf{F}$ et des estimées de $\mathbf{\Lambda}_r$, $r \in [1:R]$ que nous noterons $\hat{\mathbf{\Lambda}}_r$. Les matrices $\mathbf{B}_r$, $r \in [1:R]$ s'écrivent :

$$\mathbf{B}_r = \hat{\mathbf{F}}_B \hat{\mathbf{\Lambda}}_r \hat{\mathbf{F}}_B^T. \tag{3.41}$$

Dans cette équation, les matrices $\hat{\mathbf{\Lambda}}_r$ ne sont pas nécessairement diagonales. Notons $\hat{\mathbf{B}}_r$ les matrices obtenues en imposant la diagonalité des $\hat{\mathbf{\Lambda}}_r$:

$$\hat{\mathbf{B}}_r = \hat{\mathbf{F}}_B \; diag(vecdiag(\hat{\Lambda}_r))\hat{\mathbf{F}}_B^T, \tag{3.42}$$

où $diag(vecdiag(\hat{\mathbf{\Lambda}}_r))$ désigne la matrice dont les éléments diagonaux sont les éléments diagonaux de $\hat{\mathbf{\Lambda}}_r$ et dont les autres éléments sont nuls.
L'erreur relative totale associée à (3.14) est alors donnée par :

$$e_B^2 = \frac{\sum_{r=1}^{R} \|\mathbf{B}_r - \hat{\mathbf{B}}_r\|^2}{\sum_{r=1}^{R} \|\mathbf{B}_r\|^2}. \tag{3.43}$$

De la même manière, l'erreur relative totale associée à (3.23) est donnée par :

$$e_C^2 = \frac{\sum_{r=1}^{R} \|\mathbf{C}_r - \hat{\mathbf{C}}_r\|^2}{\sum_{r=1}^{R} \|\mathbf{C}_r\|^2}, \tag{3.44}$$

avec

$$\hat{\mathbf{C}}_r = \hat{\mathbf{F}}_C^{-H} \; diag(vecdiag(\hat{\mathbf{\Omega}}_r))\hat{\mathbf{F}}_C^{-1}. \tag{3.45}$$

Si l'un des système est beaucoup plus précis, alors on retient simplement l'estimée de $\mathbf{F}$ correspondante. Si e_B et e_C sont du même ordre, alors on divise les matrices $\mathbf{B}_r$ et $\mathbf{C}_r$ respectivement par e_B et e_C et on leur applique l'algorithme ALS généralisé développé dans le paragraphe 3.6.1 ou encore l'itération de QZ généralisée développée dans la paragraphe 3.6.2.

3.6.4 Utilisation d'un algorithme de type Jacobi

L'algorithme que nous allons détailler maintenant permet d'estimer $\mathbf{F}$ à partir de matrices qui sont des produits des matrices issues de chacun des sytèmes. Nous allons chercher $\mathbf{F}$ à partir du système suivant :

$$\mathbf{B}_r \mathbf{C}_s^* = \mathbf{F}\mathbf{\Lambda}_r \mathbf{\Omega}_s^* \mathbf{F}^{-*}, \; (r,s) \in [1:R] \tag{3.46}$$

La décomposition QR de $\mathbf{F}$ s'écrit :

$$\mathbf{F} = \mathbf{Q}^H \mathbf{R}_\mathbf{F} \tag{3.47}$$

Par conséquent, l'équation (3.46) peut s'écrire sous la forme :

$$\mathbf{B}_r\mathbf{C}_s^* = \mathbf{Q}^H\mathbf{R_F}\mathbf{\Lambda}_r\mathbf{\Omega}_s^*\mathbf{R_F}^{-*}\mathbf{Q}^{-T} \tag{3.48}$$

Posons $\mathbf{R}_{rs} = \mathbf{R_F}\mathbf{\Lambda}_r\mathbf{\Omega}_s^*\mathbf{R_F}^{-*}$. Cette matrice est triangulaire supérieure. D'après l'équation précédente, $\mathbf{R}_{rs}$ peut s'écrire :

$$\mathbf{R}_{rs} = \mathbf{Q}\mathbf{B}_r\mathbf{C}_s^*\mathbf{Q}^T \tag{3.49}$$

Nous cherchons donc une matrice $\mathbf{Q}$ unitaire telle que les produits $\mathbf{Q}\mathbf{B}_r\mathbf{C}_s^*\mathbf{Q}^T, r \in [1:R], s \in [1:R]$ soient des matrices les plus triangulaires supérieures possible.
Le système (3.49) contient R^2 équations. R matrices diagonales indépendantes génèrent l'ensemble des matrices de la forme 3.46. Nous pouvons donc choisir de ne retenir que R équations, par exemple, les équations correspondant à $r = s$. Notons $\mathbf{\Gamma}_r = \mathbf{B}_r\mathbf{C}_r^*$, le système s'écrit maintenant :

$$\left\{\begin{array}{lcl} \mathbf{R}_{11} & = & \mathbf{Q}\mathbf{\Gamma}_1\mathbf{Q}^T \\ \mathbf{R}_{22} & = & \mathbf{Q}\mathbf{\Gamma}_2\mathbf{Q}^T \\ & \vdots & \\ \mathbf{R}_{RR} & = & \mathbf{Q}\mathbf{\Gamma}_R\mathbf{Q}^T \end{array}\right. \tag{3.50}$$

La matrice $\mathbf{Q}$ peut se décomposer comme le produit de rotations élémentaires de Jacobi Θ_{qp} :

$$\mathbf{Q} = \prod_{q=1}^{N}\prod_{p=1}^{q-1}\Theta_{pq} \tag{3.51}$$

Une rotation élémentaire Θ_{qp} est une matrice unitaire dont tous les éléments diagonaux sont égaux à 1, sauf l'élément placé sur la $p$ème ligne et la $p$ème colonne et l'élément placé sur la $q$ème ligne et la $q$ème colonne, tous deux égaux à $c = \cos(\theta)$, et dont tous les éléments hors-diagonaux sont nuls, sauf l'élément d'indice (p,q) égal à $s = \sin(\theta)exp(j\phi)$ et l'élément d'indice (q,p), égal à $-s^*$:

$$\Theta_{qp} = \begin{bmatrix} 1 & \dots & 0 & \dots & 0 & \dots & 0 \\ \vdots & \ddots & \vdots & & \vdots & & \vdots \\ 0 & \dots & c & \dots & s & \dots & 0 \\ \vdots & & \vdots & \ddots & \vdots & & \vdots \\ 0 & \dots & -s^* & \dots & c & \dots & 0 \\ \vdots & & \vdots & & \vdots & \ddots & \vdots \\ 0 & \dots & 0 & \dots & 0 & \dots & 1 \end{bmatrix} \tag{3.52}$$

Il s'agit de trouver les angles θ et ϕ tels que les matrices $\mathbf{\Gamma}_r^{'}$, $r \in [1:R]$ définies par :

$$\mathbf{\Gamma}_r^{'} = \Theta_{qp}^T\mathbf{\Gamma}_r\Theta_{qp} \tag{3.53}$$

soient plus triangulaires supérieures que les matrices $\mathbf{\Gamma}_r$.
On note $v_{ij}^{(r)}$ l'élément d'indice (i,j) de $\mathbf{\Gamma}_r$ et $v_{ij}^{(r)'}$ l'élément d'indice (i,j) de $\mathbf{\Gamma}_r^{'}$. La transformation (3.53) n'a pas d'influence sur les éléments dont les indices sont tous deux différents de p et de q. Les autres éléments sont donnés par :

$$
\begin{aligned}
v_{pk}^{(r)'} &= cv_{pk}^{(r)} - s^* v_{qk}^{(r)}, \quad \text{si } 1 \leq k < p && (3.54)\\
v_{qk}^{(r)'} &= sv_{pk}^{(r)} + cv_{qk}^{(r)}, \quad \text{si } 1 \leq k < q, k \neq p && (3.55)\\
v_{kp}^{(r)'} &= cv_{kp}^{(r)} - s^* v_{kq}^{(r)}, \quad \text{si } p < k \leq N && (3.56)\\
v_{kq}^{(r)'} &= sv_{kp}^{(r)} + cv_{kq}^{(r)}, \quad \text{si } q < k \leq N, k \neq q && (3.57)\\
v_{qp}^{(r)'} &= c(sv_{pp}^{(r)} + cv_{qp}^{(r)}) - s^*(sv_{pq}^{(r)} + cv_{qp}^{(r)}) && (3.58)
\end{aligned}
$$

Soit $f(\theta, \phi)$ la fonction de coût égale à la somme des modules au carré des éléments sous diagonaux de toutes les matrices Γ_r', $r \in [1:R]$:

$$f(\theta, \phi) = \sum_{r=1}^{R} \sum_{i<j} |v_{ij}^{(r)'}|^2. \tag{3.59}$$

Nous avons vu que si $i \neq p,\ i \neq q,\ j \neq p$ et $j \neq q$, alors $v_{ij}^{(r)'} = v_{ij}^{(r)}$, donc la fonction de coût peut encore s'écrire :

$$
\begin{aligned}
f(\theta, \phi) = &\sum_{r=1}^{R} \sum_{k=1}^{p-1} \left(|v_{pk}^{(r)'}|^2 + |v_{qk}^{(r)'}|^2 \right) + \sum_{k=q+1}^{N} \left(|v_{kp}^{(r)'}|^2 + |v_{kq}^{(r)'}|^2 \right) + \\
&\sum_{k=p+1}^{q-1} \left(|v_{kp}^{(r)'}|^2 + |v_{qk}^{(r)'}|^2 \right) + |v_{qp}^{(r)}|^2 + cste \qquad (3.60)
\end{aligned}
$$

D'après (3.54) et (3.55), nous pouvons écrire :

$$|v_{pk}^{(r)'}|^2 + |v_{qk}^{(r)'}|^2 = |v_{pk}^{(r)}|^2 + |v_{qk}^{(r)}|^2, \text{ si } 1 \leq k < p \tag{3.61}$$

D'après (3.56) et (3.57), nous avons :

$$|v_{kp}^{(r)'}|^2 + |v_{kq}^{(r)'}|^2 = |v_{kp}^{(r)}|^2 + |v_{kq}^{(r)}|^2, \text{ si } q < k \leq N \tag{3.62}$$

Posons

$$V = \begin{bmatrix} \cos(2\theta) \\ \sin(2\theta)\cos(\phi) \\ \sin(2\theta)\sin(\phi) \end{bmatrix} \tag{3.63}$$

D'après (3.55) et (3.56), nous pouvons écrire :

$$|v_{qk}^{(r)'}|^2 + |v_{kp}^{(r)'}|^2 = a_{qk,kp}^{(r)} + \mathbf{h}_{qk,kp}^{(r)T} V, \text{ si } p < k < q \tag{3.64}$$

avec

$$a_{qk,kp}^{(r)} = (1/2) \left(|v_{qk}^{(r)}|^2 + |v_{kp}^{(r)}|^2 + |v_{pk}^{(r)}|^2 + |v_{kq}^{(r)}|^2 \right)$$

et

$$\mathbf{h}_{qk,kp}^{(r)} = \begin{bmatrix} (1/2) \left(|v_{qk}^{(r)}|^2 + |v_{kp}^{(r)}|^2 - |v_{pk}^{(r)}|^2 - |v_{kq}^{(r)}|^2 \right) \\ Re \left(v_{qk}^{(r)} v_{pk}^{(r)*} - v_{kq}^{(r)} v_{kp}^{(r)*} \right) \\ Im \left(v_{qk}^{(r)} v_{pk}^{(r)*} - v_{kq}^{(r)} v_{kp}^{(r)*} \right) \end{bmatrix}.$$

Enfin, d'après (3.55), en gardant la définition de V donnée en (3.63), nous avons :

$$|v_{qp}^{(r)'}|^2 = a_{qp}^{(r)} + \mathbf{h}_{qp}^{(r)T} V + V^T \mathbf{G}_{qp}^{(r)} V, \tag{3.65}$$

avec

$$a_{qp}^{(r)} = |v_{pp}^{(r)}|^2 + |v_{qp}^{(r)}|^2 + |v_{pq}^{(r)}|^2 + |v_{qq}^{(r)}|^2,$$

$$\mathbf{h}_{qp}^{(r)} = (1/2) \begin{bmatrix} |v_{qp}^{(r)}|^2 - |v_{pq}^{(r)}|^2 \\ Re\left(v_{pp}^{(r)} v_{qp}^{(r)*} + v_{pq}^{(r)} v_{qq}^{(r)*} - v_{pp}^{(r)} v_{pq}^{(r)*} - v_{qp}^{(r)} v_{qq}^{(r)*}\right) \\ Im\left(-v_{pp}^{(r)} v_{qp}^{(r)*} - v_{pq}^{(r)} v_{qq}^{(r)*} + v_{pp}^{(r)} v_{pq}^{(r)*} + v_{qp}^{(r)} v_{qq}^{(r)*}\right) \end{bmatrix}$$

et

$$\begin{array}{lcll} \mathbf{G}_{qp}^{(r)}(1,1) & = & (1/4)\left(|v_{qp}^{(r)}|^2 + |v_{pq}^{(r)}|^2 - |v_{pp}^{(r)}|^2 - |v_{qq}^{(r)}|^2\right) & \\ \mathbf{G}_{qp}^{(r)}(1,2) & = & (1/2)Re\left(v_{pp}^{(r)} v_{qp}^{(r)*} - v_{pq}^{(r)} v_{qq}^{(r)*} - v_{qp}^{(r)} v_{qq}^{(r)*} + v_{pp}^{(r)} v_{pq}^{(r)*}\right) & = & \mathbf{G}_{qp}^{(r)}(2,1) \\ \mathbf{G}_{qp}^{(r)}(1,3) & = & (1/2)Im\left(-v_{pp}^{(r)} v_{qp}^{(r)*} + v_{pq}^{(r)} v_{qq}^{(r)*} + v_{qp}^{(r)} v_{qq}^{(r)*} - v_{pp}^{(r)} v_{pq}^{(r)*}\right) & = & \mathbf{G}_{qp}^{(r)}(3,1) \\ \mathbf{G}_{qp}^{(r)}(2,2) & = & -(1/2)Re\left(v_{qp}^{(r)} v_{pq}^{(r)*} + v_{pp}^{(r)} v_{qq}^{(r)*}\right) & \\ \mathbf{G}_{qp}^{(r)}(2,3) & = & Im\left(v_{pp}^{(r)} v_{qq}^{(r)*}\right) & = & \mathbf{G}_{qp}^{(r)}(3,2) \\ \mathbf{G}_{qp}^{(r)}(3,3) & = & (1/2)Re\left(v_{pp}^{(r)} v_{qq}^{(r)*} - v_{qp}^{(r)} v_{pq}^{(r)*}\right) & \end{array}$$

Posons $\mathbf{h} = \sum_{r=1}^{R} h_{qp}^{(r)} + h_{qk,kp}^{(r)}$ et $\mathbf{G} = \sum_{r=1}^{R} G_{qp}^{(r)}$. La fonction de coût s'écrit en utilisant le vecteur $V = [\cos(2\theta), \sin(2\theta)\cos(\phi), \sin(2\theta)\sin(\phi)]^T$ défini précédemment :

$$f(V) = V^T \mathbf{G} V + \mathbf{h}^T V + c \tag{3.66}$$

où c désigne une constante.
La minimisation de f sous contrainte $\|V\|^2 = 1$ à l'aide du multiplieur de Lagrange nous mène à la résolution de l'équation suivante :

$$2(\mathbf{G} + \lambda \mathbf{I})V + \mathbf{h} = 0, \tag{3.67}$$

et donc :

$$V = -\frac{1}{2}(\mathbf{G} + \lambda \mathbf{I})^{-1}\mathbf{h}, \tag{3.68}$$

où λ désigne un scalaire choisi tel que $\|V\|^2 = 1$, soit encore :

$$\frac{1}{4}\sum_{i=1}^{3} \frac{\left(E_i^T \mathbf{h}\right)^2}{\left(\lambda_i + \lambda\right)^2} = 1 \tag{3.69}$$

où E_i et λ_i, $i \in [1:3]$ désignent respectivement les vecteurs propres et les valeurs propres de $\mathbf{G}$. Cette équation peut encore s'écrire :

$$c_6\lambda^6 + c_5\lambda^5 + c_4\lambda^4 + c_3\lambda^3 + c_2\lambda^2 + c_1\lambda^1 + c_0\lambda^0 = 0 \tag{3.70}$$

avec

$$c_6 = 4 \tag{3.71}$$

$$c_5 = 8\sum_i \lambda_i \tag{3.72}$$

$$c_4 = 4\left(\sum_i \lambda_i\right)^2 + 8\sum_i \prod_{j>i} \lambda_i\lambda_j - \sum_i a_i^2 \tag{3.73}$$

$$c_3 = 8\prod_i \lambda_i + 8\left(\sum_i \lambda_i\right)\left(\sum_i \prod_{j>i} \lambda_i\lambda_j\right) - \left(\sum_i a_i^2\left(2\sum j \neq i\lambda_j\right)\right) \tag{3.74}$$

$$c_2 = 4\left(\sum_i \prod j > i\lambda_i\lambda_j\right)^2 + 8\left(\sum_i \lambda_i\right)\left(\prod_i \lambda_i\right) - \sum_i a_i^2\left(\left(\sum_{j\neq i} \lambda_j\right)^2 + 2\prod_{j\neq i} \lambda_j\right) \tag{3.75}$$

$$c_1 = 8\left(\sum_i \prod_{j>i} \lambda_i\lambda_j\right)\left(\prod_i \lambda_i\right) - \sum_i \left(a_i^2 2\sum_{j\neq i} \lambda_j\right)\left(\prod_{j\neq i} \lambda_j\right) \tag{3.76}$$

$$c_0 = 4\left(\prod_i \lambda_i\right)^2 - \sum_i (a_i^2)\left(\prod j \neq i\lambda_j\right)^2 \tag{3.77}$$

et $a_i = E_i^T\mathbf{h}$.
Le problème de minimisation revient à évaluer les racines d'un polynôme de degré 6. Parmi ces racines, nous sélectionnons la racine réelle dont le vecteur V correspondant minimise la fonction f.

Dans le cas où $\mathbf{h}$ est orthogonal à l'un des vecteurs propres de $\mathbf{G}$, alors la racine λ_i correspondante est une racine double de (3.69). Dans ce cas, la solution de (3.67) n'est plus donnée par (3.68), mais par :

$$V = -\frac{1}{2}(\mathbf{G} - \lambda_i\mathbf{I})^\dagger\mathbf{h} + c_i E_i, \tag{3.78}$$

où c_i est un scalaire défini par :

$$c_i = \pm\sqrt{1 - \frac{1}{4}\|(\mathbf{G} - \lambda_i\mathbf{I})^\dagger\mathbf{h}\|^2} \tag{3.79}$$

et où le signe de c_i est choisi de telle manière que le vecteur V minimise la fonction f.

Remarque 1 Nous avons choisi de retenir ici les matrices $\mathbf{B}_r\mathbf{C}_s^*$, telles que $r = s$. On peut procéder de la manière suivante pour retenir les R équations les plus intéressantes. On range les matrices $\mathbf{B}_r\mathbf{C}_s^*$, $(r,s) \in [1:R]$ dans R^2 vecteurs de taille R^2. On calcule la décomposition en valeurs singulières de la matrice contenant ces vecteurs rangés en colonnes. On ne retient ensuite que les R vecteurs singuliers dominants, qui sont réordonnés sous forme de matrices à l'aide de l'opérateur *unvec*. On peut pondérer leur importance par exemple en multipliant chacun de ces vecteurs singuliers par la valeur singulière correspondante.

3.7 Simulations

Une première simulation illustre les performances des algorithmes présentés dans les paragraphes 3.6.1 à 3.6.3. On considère que $R = 6$ utilisateurs émettent seulement $K = 50$ symboles QPSK, qui sont étalés avec un facteur d'étalement de $J = 4$. Ces signaux sont reçus sur un réseau de $I = 4$ antennes. Les éléments de $\mathbf{A}$ et de $\mathbf{H}$ suivent une loi gaussienne de moyenne nulle et de variance 1. Les résultats ont été moyennés sur 100 simulations.
Nous avons comparé les performances de (1) SD-QZ avec une tolérance $\epsilon_{\text{SD-QZ}} = 10^{-1}$, en exploitant seulement la structure PARAFAC, (2) ACMA, en exploitant seulement la contrainte CM, (3) CSD-ALS avec une tolérance $\epsilon_{\text{CSD-ALS}} = 10^{-1}$, en exploitant à la fois la structure PARAFAC et la contrainte CM et (4) CSD-QZ, avec une tolérance $\epsilon_{\text{CSD-QZ}} = 10^{-1}$, en exploitant également les deux contraintes.
Les matrices $\mathbf{B}_r$ et $\mathbf{C}_r$ ont été pondérées respectivement par e_B^{-1} et e_C^{-1}. Les résultats sont présentés dans les figures 3.3 et 3.4. Dans la figure 3.3, nous avons tracé le TES moyen en fonction du RSB. Dans la figure 3.4 nous avons tracé le temps de calcul moyen en fonction du RSB. En combinant les deux contraintes, nous avons pu diminuer le TES à un coût de calcul supplémentaire faible.

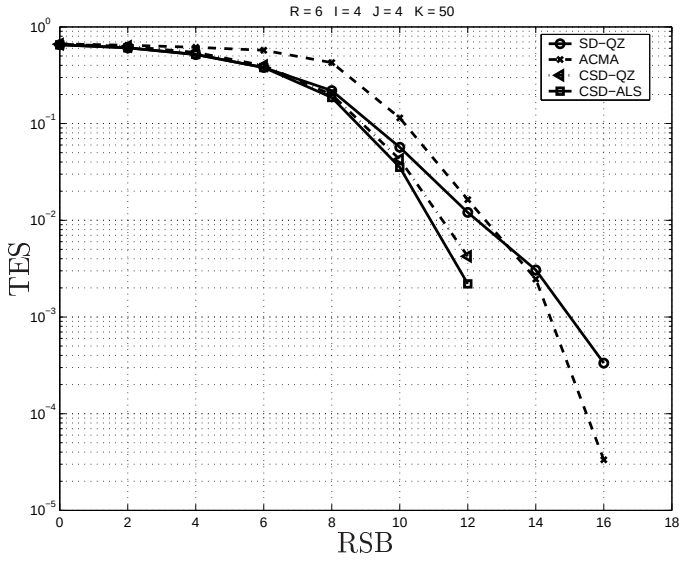

FIG. 3.3 – TES moyen en fonction du RSB dans la première simulation ($I = J = 4$, $K = 50$, $R = 6$).

La deuxième simulation présente les performances de l'algorithme de type Jacobi présenté dans le paragraphe 3.6.4. Nous avons considéré la présence de $R = 3$ utilisateurs émettant $K = 50$ symboles QPSK, qui sont étalés avec un facteur d'étalement de $J = 3$. Ces signaux sont reçus sur un réseau de $I = 3$ antennes. Les éléments de $\mathbf{A}$ et de $\mathbf{H}$ suivent encore une loi gaussienne de moyenne nulle et de variance 1. Les matrices $\mathbf{B}_r\mathbf{C}_s^*$ ont été choisies de la manière proposée dans la remarque 1. Les résultats ont été moyennés sur 100 simulations.
Nous avons comparé les performance de (1) l'algorithme de type Jacobi et (2) l'algorithme CSD-QZ. Dans la figure 3.5, nous avons tracé la norme de Frobenius de la différence entre la matrice

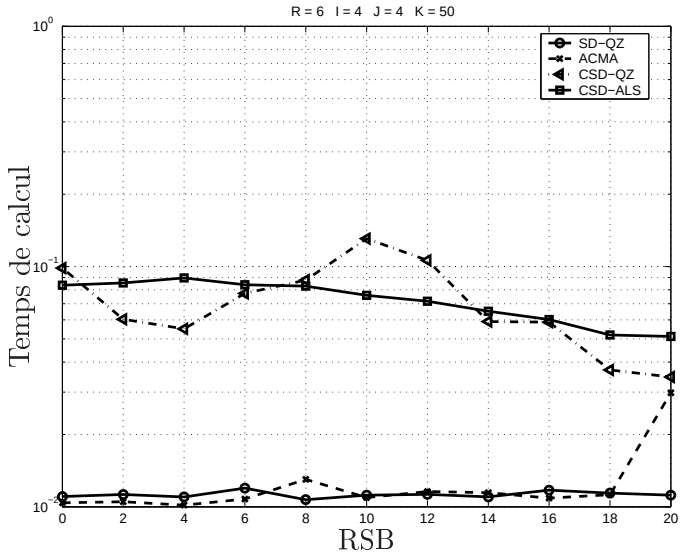

FIG. 3.4 – Temps de calcul moyen en fonction du RSB dans la première simulation ($I = J = 4$, $K = 50$, $R = 6$).

F et la matrice **F** estimée en fonction du RSB. Dans la figure 3.6, nous avons tracé le temps de calcul moyen en fonction du RSB. Nous pouvons voir que l'algorithme de type Jacobi est moins robuste au bruit que CSD-QZ, et il est plus lent. On préfèrera donc utiliser l'algorithme CSD-QZ.

3.8 Conclusion

Dans ce chapitre, nous avons présenté une méthode de séparation des signaux DS-CDMA utilisant les algorithmes présentés au chapitre précédent. Nous avons par ailleurs montré qu'il était possible de combiner les résultats de cette technique et ceux de l'algorithme ACMA. Nous avons proposé pour cela différentes solutions. L'exploitation conjointe de la contrainte PARAFAC et de la contrainte CM nous permet d'améliorer la précision des résultats.

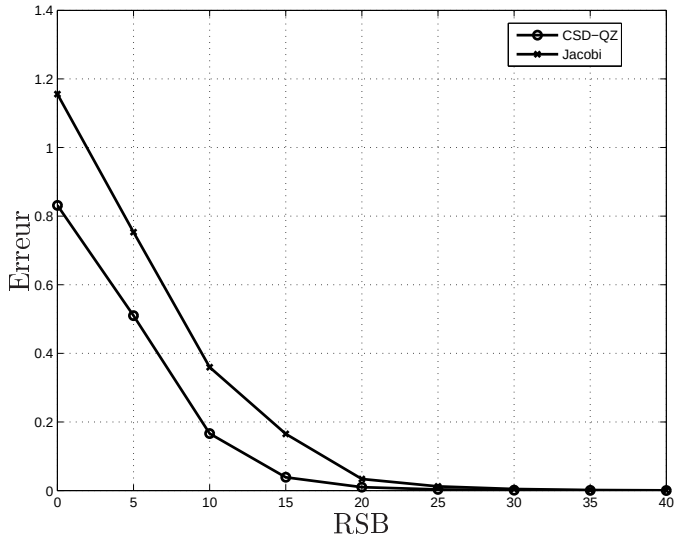

FIG. 3.5 – Erreur moyenne sur $\mathbf{F}$ en fonction du RSB dans la deuxième simulation ($I = J = 3$, $K = 50$, $R = 3$).

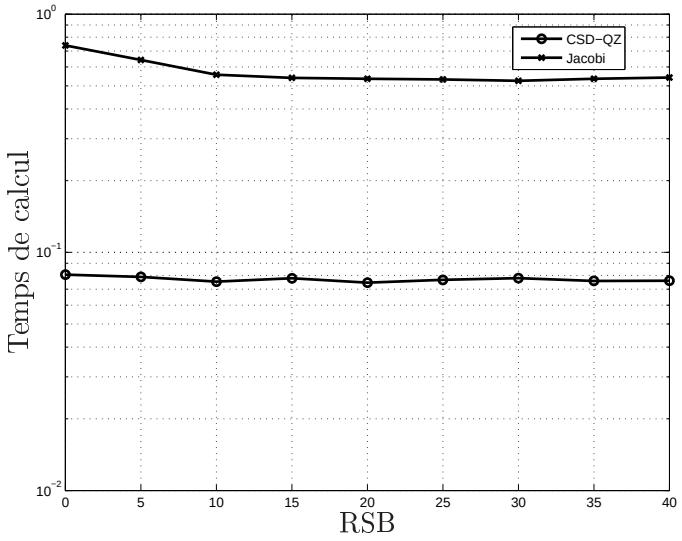

FIG. 3.6 – Temps de calcul moyen en fonction du RSB dans la deuxième simulation ($I = J = 3$, $K = 50$, $R = 3$).

Chapitre 4

Application à l'Analyse en Composantes Indépendantes

4.1 Introduction

Dans ce chapitre nous considérons le problème de l'Analyse en Composantes Indépendantes (ACI). Nous supposons que nous observons des signaux qui sont des mélanges linéaires de signaux sources inconnus. L'objectif de l'ACI est d'estimer la matrice de mélange et/ou les sources à partir des observations, en supposant l'indépendance statistique des sources. Le modèle est donc le suivant. Le vecteur observation $\mathbf{x}_t$ reçu sur le réseau de J capteurs à un instant t s'écrit comme la somme du vecteur bruit $\mathbf{n}_t$ à l'instant t et du produit de la matrice de mélange $\mathbf{A}$ et du vecteur source $\mathbf{s}_t$ à l'instant t :

$$\mathbf{x}_t = \mathbf{A}\mathbf{s}_t + \mathbf{n}_t \tag{4.1}$$

Nous supposons que le vecteur observations $\mathbf{x}_t \in \mathbb{C}^J$, le vecteur bruit $\mathbf{b}_t \in \mathbb{C}^J$, et le vecteur sources $\mathbf{x}_t \in \mathbb{C}^R$ sont de moyenne nulle. La matrice de mélange $\mathbf{A}$ appartient à $\mathbb{C}^{J\times R}$.
D'autre part, nous nous intéressons en particulier au cas dit *sous-déterminé*, c'est à dire le cas où le nombre de sources R est supérieur au nombre d'observations J.
Afin d'estimer les sources, dans le cas sur-déterminé, on multiplie généralement la matrice des observations par la pseudo-inverse de la matrice de mélange estimée. Il n'est plus possible de procéder de cette manière dans le cas sous-déterminé, puisque la matrice de mélange n'est plus inversible à gauche. On peut choisir d'estimer simultanément la matrice de mélange et les sources, ou encore d'estimer la matrice, puis les sources en s'appuyant sur une connaissance a priori sur les sources. Si les sources appartiennent à un alphabet fini, il est possible de faire une recherche exhaustive sur toutes les combinaisons possibles. Si au plus $J-1$ sources sont actives au même instant, alors on peut déterminer les vecteurs de mélanges actifs à cet instant et inverser le mélange correspondant. Dans ce chapitre, nous traitons uniquement le problème de l'estimation de la matrice de mélange.
Une large classe d'algorithmes pour l'ACI sous-déterminée exploite le fait que les sources sont parcimonieuses, ou encore que les données obtenues par une transformation linéaire des sources le sont [4,27,33,49]. Nous nous appuyons ici sur des techniques algébriques. De telles techniques ont été utilisée pour résoudre le problème dans le cas particulier de deux sources et trois observations [13,23]. Dans [2] les auteurs ont exploité la structure du tenseur des cumulants d'ordre 6. Lorsque

les sources sont individuellement corrélées, on peut également exploiter la structure d'un tenseur contenant les matrices de quadricovariance des observations pour différents retards [25].

Dans ce chapitre nous proposons deux méthodes algébriques pour résoudre le problème de l'ACI sous-déterminée. Dans le paragraphe 4.2, nous présenterons une technique s'appuyant sur les statistiques d'ordre deux des signaux lorsqu'ils sont individuellement corrélés, dans le paragraphe 4.3 nous présenterons une autre manière de procéder, s'appuyant uniquement sur leurs statistiques d'ordre quatre. Nous conclurons ce chapitre par le paragraphe 4.4.

4.2 Statistiques d'ordre deux

Dans [3], les auteurs ont proposé une méthode, appelée SOBI (Second Order Blind Identification), s'appuyant sur les statistiques d'ordre deux et exploitant la diversité apportée par les matrices de covariance pour différents retards. SOBI repose sur un blanchiment des données, et une analyse en composantes principales. Il n'est pas applicable dans le cas sous-déterminé. L'algorithme qui va être présenté part des mêmes hypothèses que SOBI, mais il fonctionne dans le cas sous-déterminé. Il sera noté SOBIUM pour "Second Order Blind Identification of Underdetermined Mixtures".
Nous supposons que les sources sont mutuellement indépendantes et individuellement corrélées en temps. La matrice de covariance des observations pour un retard τ_1 s'écrit :

$$\begin{aligned} \mathbf{C}_1 &= E\{\mathbf{x}_t \mathbf{x}_{t-\tau_1}^H\} && (4.2)\\ &= \mathbf{A} E\{\mathbf{s}_t \mathbf{s}_{t-\tau_1}^H\} \mathbf{A}^H && (4.3)\\ &= \mathbf{A}\mathbf{D}_1\mathbf{A}^H && (4.4) \end{aligned}$$

Les sources sont mutuellement décorrélées et individuellement corrélées en temps, donc la matrice $\mathbf{D}_1$ est une matrice diagonale.
Les matrices de covariances pour différents retards $\tau_1, \tau_2, \ldots, \tau_K$ s'écrivent alors :

$$\left\{ \begin{array}{rcl} \mathbf{C}_1 &=& \mathbf{A}\mathbf{D}_1\mathbf{A}^H \\ \mathbf{C}_2 &=& \mathbf{A}\mathbf{D}_2\mathbf{A}^H \\ & \vdots & \\ \mathbf{C}_K &=& \mathbf{A}\mathbf{D}_K\mathbf{A}^H \end{array} \right. , \qquad (4.5)$$

où les matrices $\mathbf{D}_1, \mathbf{D}_2, \ldots, \mathbf{D}_K$ sont diagonales. Soit $\mathbf{D}$ la matrice de taille $K \times R$ contenant les diagonales des matrices $\mathbf{D}_k$, $k \in [1:K]$ rangées en colonnes, $(\mathbf{D})_{kr} = (\mathbf{D}_k)_{rr}$ pour tout k dans $[1:K]$ et tout r dans $[1:R]$.

Nous pouvons définir un tenseur $\mathcal{C} \in \mathbb{C}^{J \times J \times K}$ dont l'élément d'indice (i,j,k) représente la covariance entre l'observation reçue sur la $i$ème antenne et l'observation reçue sur la $j$ème antenne après un retard τ_k et s'écrit :

$$c_{ijk} = \sum_{r=1}^{R} a_{ir} a_{jr}^* d_{kr} \qquad (4.6)$$

$\mathcal{C}$ peut s'écrire :

$$\mathcal{C} = \sum_{r=1}^{R} A_r \circ A_r^* \circ D_r \tag{4.7}$$

L'équation (4.7) est une décomposition PARAFAC du tenseur $\mathcal{C}$ dans laquelle chaque terme de rang un représente la contribution d'une source. Nous avons vu au chapitre 2 qu'il est possible d'estimer les vecteurs A_r et D_r sous certaines conditions. En particulier, nous pouvons distinguer deux cas : le cas où le nombre de matrices de covariance disponibles K est strictement inférieur au nombre de sources R et le cas où le nombre de matrices de covariance disponibles K est supérieur ou égal au nombre de sources R.

Cas 1 : $K < R$
Nous avons vu au paragraphe 2.3.1 que la décomposition (4.7) est unique si

$$2\ rank_k(\mathbf{A}) + rank_k(\mathbf{D}) \geq 2(R+1). \tag{4.8}$$

Une matrice est génériquement de rang plein et de k-rang plein. Nous pouvons alors supposer $rank_k(\mathbf{A}) = min(J,R) = R$ et $rank_k(\mathbf{D}) = min(K,R) = K$ si $K < R$. Finalement, dans ce cas, la décomposition est unique si :

$$R \leq J - 1 + K/2. \tag{4.9}$$

Pour estimer les paramètres, nous pouvons appliquer un algorithme DALS qui minimise la fonction de coût $f(\mathbf{U},\mathbf{V},\mathbf{W}) = \|\mathcal{C} - \sum_{r=1}^{R} U_r \circ V_r \circ W_r\|^2$. L'algorithme ne suppose pas de lien entre les vecteurs U_r et V_r qui convergent respectivement vers A_r et A_r^* . Pour tout $r \in [1:R]$, le vecteur A_r peut être estimé comme le vecteur propre dominant de la matrice $[U_r, V_r^*]$ de taille $J \times 2$.
Les dimensions du tenseur ne permettent pas de trouver une bonne initialisation à l'aide de la technique présentée dans le paragraphe 2.3.3 mais il est possible d'accélérer la convergence de l'algorithme en utilisant une recherche linéaire optimisée (enhanced linesearch, ELS) [39]. Un synopsis de cette méthode est donné dans la table 4.2.

Cas 2 : $R \geq K$
Dans ce cas, il est toujours possible d'utiliser la technique mise en œuvre dans le cas 1. Cependant, si $R \geq K$ et $R < J^2$, nous pouvons également choisir d'utiliser l'algorithme SD-ALS ou l'algorithme SD-QZ présentés dans le paragraphe 2.4. Le tenseur des observations ayant une structure légèrement différente du cas général décrit dans le chapitre 2 en raison de la symétrie des matrices de covariance, la borne sur le nombre de sources doit alors être telle que :

$$2R(R-1) \leq J^2(J-1)^2/2 \tag{4.10}$$

dans le cas complexe et $R < R_{max}$ défini par le tableau 4.1 dans le cas réel [45].

4.2.1 Simulations

Dans une première simulation, nous avons examiné le cas où le nombre de matrices de covariance K est inférieur au nombre de sources (cas 1). Nous avons considéré la présence de $R = 5$ signaux sources de longueur $T = 10000$ reçus sur un réseau de $J = 4$ capteurs. Pour générer des sources individuellemnt corrélées, nous avons procédé de la manière suivante : nous avons tout d'abord

J	2	3	4	5	6	7	8
R_{max}	2	4	6	10	15	20	26

TAB. 4.1 – Nombre maximal de sources pour SOBIUM dans le cas réel si $K \leq R$.

généré R signaux dont les éléments suivaient une loi gausienne complexe standardisée (la partie réelle et la partie imaginaire de chaque élément suivait une loi gaussienne de moyenne nulle et de variance 1/2). Puis nous avons filtré chacun de ces signaux par un filtre dont les coefficients sont les éléments d'une ligne d'une matrice de Hadamard de taille 16×16 (nous avons considéré les lignes 1, 2, 4, 7 et 8 de la matrice de Hadamard générée par défaut par Matlab.).
Nous avons considéré que les signaux sont reçus sur une antenne circulaire, les coefficients de la matrice de mélange $\mathbf{A}$ s'écrivent donc :

$$a_{kr} = \exp\left(\frac{2\pi j}{\lambda}(x_k \cos(\theta_r)\cos(\phi_r) + y_k \cos(\theta_r)\sin(\phi_r))\right) \tag{4.12}$$

avec

$$x_k = R_a \cos(2\pi(k-1)/J) \tag{4.13}$$
$$y_k = R_a \sin(2\pi(k-1)/J) \tag{4.14}$$

avec $R_a/\lambda = .55$.
Les directions d'arrivée des signaux sont données par $\theta_1 = 3\pi/10$, $\theta_2 = 3\pi/10$, $\theta_3 = 2\pi/5$, $\theta_4 = 0$, $\theta_5 = \pi/10$, $\phi_1 = 7\pi/10$, $\phi_2 = 9\pi/10$, $\phi_3 = 3\pi/5$, $\phi_4 = 4\pi/5$, $\phi_5 = 3\pi/5$.
Les matrices de covariance sont calculées pour un retard $\tau = 0, 1 \ldots, 3$. Les éléments de la matrice bruit suivent une loi complexe standardisée.
La matrice de mélange est évaluée à partir de l'algorithme DALS (cas 1) initialisé à l'aide de 5 valeurs initiales. Seule la meilleure performance parmi les cinq est retenue.
L'erreur relative est égale à la norme de la différence entre la matrice de mélange $\mathbf{A}$ et la matrice de mélange estimée $\hat{\mathbf{A}}$ (dont les colonnes ont été ordonnées et multipliées par un scalaire de manière optimale) divisée par la norme de la matrice de mélange, moyennée sur 100 essais : $err = E\{\|\mathbf{A} - \hat{\mathbf{A}}\|/\|\mathbf{A}\|\}$.
Dans la figure 4.1, nous avons tracé l'erreur relative moyenne en fonction du RSB.

Dans la seconde simulation, nous avons considéré la présence de $R = 5$ ou 6 sources de longueur $T = 5000$ échantillons. Nous avons conservé les mêmes paramètres pour la matrice de mélange et les sources dans le cas de $R = 5$ sources. Dans le cas de 6 sources, les directions d'arrivée des signaux sont données par $\theta_1 = 3\pi/10$, $\theta_2 = 3\pi/10$, $\theta_3 = 2\pi/5$, $\theta_4 = 0$, $\theta_5 = \pi/10$, $\theta_6 = 3\pi/5$, $\phi_1 = 7\pi/10$, $\phi_2 = 9\pi/10$, $\phi_3 = 3\pi/5$, $\phi_4 = 4\pi/5$, $\phi_5 = 3\pi/5$, $\phi_6 = \pi/5$. Les signaux ont été filtrés à l'aide des lignes 1, 2, 4, 7, 8 et 13 de la matrice de Hadamard de taille 16×16 générée par défaut par Matlab.
Nous avons pris en compte 12 matrices de covariance ($\tau = 0, 1, \ldots, 11$). Le problème est dans ce cas mieux conditionné, et il est possible d'appliquer l'algorithme décrit dans le cas 2. La diagonalisation simultanée a été réalisée à l'aide de l'algorithme QZ. Dans la figure 4.2, nous

1. Ranger les matrices de covariance des observations dans la matrice $\mathbf{C}$ de taille $J^2 \times K$.
2. Calculer la svd de $\mathbf{C}$, $\mathbf{C} = \mathbf{U\Sigma V}^H$.
3. Pour tout r dans $[1 : R]$, ranger la $r$ème colonne de $\mathbf{U\Sigma}$ dans la matrice $\mathbf{E}_r = unvec((\mathbf{U\Sigma})_r)$.
4. Pour tout (r, s) dans $[1 : R]$, $r \leq s$, évaluer le tenseur $\Phi_{rs} = \Phi(\mathbf{E}_r, \mathbf{E}_s)$ et le ranger dans le vecteur $P_{rs} = vec(\Phi_{rs})$ de taille J^4.
5. Construire la matrice $\mathbf{P} = [P_{11}, P_{12}, ..., P_{RR}]$ de taille $J^4 \times R(R+1)/2$.
6. Calculer les R vecteurs singuliers associés aux R plus petites valeurs singulières de $\mathbf{P}$ et les ranger dans les matrices triangulaires supérieures $\mathbf{X}_r$, $r \in [1 : R]$.
7. Calculer $\mathbf{B}_r = \mathbf{X}_r + \mathbf{X}_r^T$.
8. Evaluer la matrice $\mathbf{F}$, solution du système
$$\begin{cases} \mathbf{B}_1 &= \mathbf{F\Lambda_1 F}^T \\ \mathbf{B}_2 &= \mathbf{F\Lambda_2 F}^T \\ &\vdots \\ \mathbf{B}_R &= \mathbf{F\Lambda_R F}^T \end{cases} . \tag{4.11}$$
9. Construire les matrices $\mathbf{N}_r = unvec((\mathbf{U\Sigma F})_r)$, $r \in [1 : R]$.
10. Evaluer $\mathbf{A}$: la $r$ème colonne de $\mathbf{A}$ est le vecteur propre dominant de $\mathbf{N}_r$.

TAB. 4.2 – Résumé de SOBIUM

avons tracé l'erreur relative moyenne. Pour $R = 5$ et $RSB = 30$, l'erreur relative est de -35dB, cela signifie que l'erreur sur chaque élément de la matrice de mélange est d'environ 1.8 %, pour $R = 6$ et $RSB = 30$, l'erreur sur chaque élément de la matrice de mélange est d'environ 5.6 %.

4.2.2 Application à des signaux de parole

Dans ce paragraphe, nous allons illustrer l'algorithme en l'appliquant à des signaux de parole. Nous supposons que quatre personnes parlent en même temps et que l'on dispose de trois capteurs. Le mélange est centré gaussien. Nous avons supposé l'absence de bruit.
Les signaux de parole sont des signaux corrélés sur des intervalles de temps petits, et le signal émis par une personne est indépendant de celui qui est émis par une autre. Les signaux de parole possèdent d'autre part la propriété d'être parcimonieux, il y a plus de blancs que de sons dans un signal de parole. Nous pouvons exploiter cette particularité pour estimer les sources une fois la matrice de mélange évaluée. Il faut alors pour cela évaluer quelles personnes parlent et quelles personnes se taisent à chaque instant. Cela est possible si le nombre d'observations est supérieur au nombre de sources moins un. Pour plus de simplicité, nous avons pris l'hypothèse que le premier et le deuxième parlent ensemble, pendant que le troisième et le quatrième se taisent durant 20000 échantillons, puis c'est au tour du deuxième et du troisième de parler durant 5000 échantillons, puis au troisième et au quatrième jusqu'à la fin (c'est-dire environ 20000 échantillons).

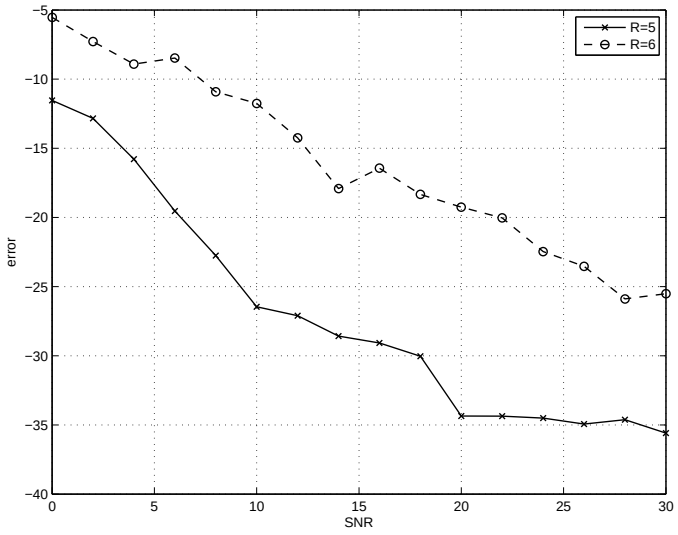

FIG. 4.1 – Erreur relative moyenne en fonction du RSB ($K = 4$).

Nous avons estimé la matrice de mélange à l'aide de SOBIUM sur la longueur totale des signaux mélangés. Puis, pour retrouver les signaux du premier interlocuteur et du deuxième durant les 20000 premiers échantillons, nous avons inversé la sous-matrice de la matrice de mélange constitué de ses deux premières colonnes. Nous avons procédé de la même manière avec les colonnes 2 et 3 pour retrouver les signaux du deuxième et du troisième interlocuteurs durant les 5000 échantillons suivant, et avec les colonnes 3 et 4 pour retrouver les signaux du troisième et du quatrième interlocuteurs ensuite. Les signaux sources sont représentés sur la figure 4.2.2, les signaux mélangés sur la figure 4.2.2, et les signaux estimés sur la figure 4.2.2.

4.3 Statistiques d'ordre quatre

Dans ce paragraphe nous proposons de nous appuyer cette fois uniquement sur les statistiques d'ordre quatre. Nous allons supposer que les kurtosis des sources sont non nuls. Dans le paragraphe 4.3.1, nous présenterons un algorithme qui conduit à une diagonalisation simultanée par une matrice orthogonale. Il est noté FOOBI pour *Fourth Order Only Blind Identification* et suit la stratégie employée dans [6]. Le paragraphe 4.3.2 est consacré à une variante de FOOBI qui conduit quant à elle à une zéro-diagonalisation, on la notera FOOBI-2. Nous présenterons quelques simulations numériques dans le paragraphe 4.3.3.

4.3.1 FOOBI

Le tenseur de la quadricovariance des sources $\mathcal{C}^{\mathbf{x}} \in \mathbb{C}^{J\times J\times J\times J}$ est défini par

$$\mathcal{C}^{\mathbf{x}} = Cum\{\mathbf{x}, \mathbf{x}^*, \mathbf{x}^*, \mathbf{x}\} \tag{4.15}$$

et au niveau des éléments :

$$c^{\mathbf{x}}_{ijkl} = Cum\{x_i, x^*_j, x^*_k, x_l\} \tag{4.16}$$

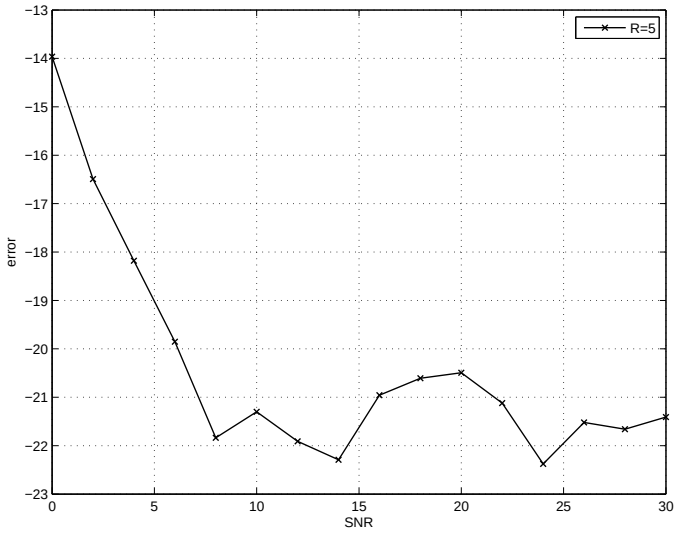

FIG. 4.2 – Erreur relative moyenne en fonction du RSB ($K = 12$).

En raison de la propriété de multilinéarité des cumulants, nous pouvons écrire :

$$c^{\mathbf{x}}_{ijkl} = \sum_{r=1}^{R} \kappa_r a_{ir} a^*_{jr} a^*_{kr} a_{lr} \tag{4.17}$$

où κ_r désigne le kurtosis de la $r$ème source. Cette expression est la décomposition d'un tenseur d'ordre 4 et de rang R en une somme de tenseurs d'ordre 4 et de rang 1.
Nous pouvons représenter le tenseur $\mathcal{C}^{\mathbf{x}}$ sous forme matricielle. Soit $\mathbf{C}^{\mathbf{x}}$ la matrice de taille $J^2 \times J^2$ définie par :

$$\mathbf{C}^{\mathbf{x}}((i-1)J + j, (k-1)J + l) = c^{\mathbf{x}}_{ijkl}$$

pour toutes les valeurs des indices. Posons d'autre par ailleurs $\tilde{\mathbf{C}}^{\mathbf{s}} = diag(\kappa_1, \kappa_2, \ldots, \kappa_R) \in \mathbb{C}^{R \times R}$.
Nous pouvons écrire l'équation (4.17) sous forme matricielle :

$$\mathbf{C}^{\mathbf{x}} = (\mathbf{A} \odot \mathbf{A}^*)\tilde{\mathbf{C}}^{\mathbf{s}}(\mathbf{A} \odot \mathbf{A}^*)^H \tag{4.18}$$

Nous avons une première décomposition de $\mathbf{C}^{\mathbf{x}}$. Dans le chapitre 2, nous avions également représenté le tenseur des données sous forme matricielle, puis nous avions proposé une seconde décomposition, en valeurs singulières. Ici, la structure de la matrice $\mathbf{C}^{\mathbf{x}}$ est plus complexe et nous pouvons exploiter le fait qu'elle possède de nombreuses symétries. Nous allons donc utiliser une décomposition légèrement différente, qui est introduite dans le théorème suivant.

Théorème 3 *Décomposition d'un tenseur supersymétrique*
Soit $\mathcal{T}$ un tenseur vérifiant $t_{klij} = t^*_{ijkl}$ et $t_{jikl} = t^*_{ijkl}$, alors $\mathcal{T}$ peut se décomposer sous la forme :

$$t_{ijkl} = \sum_{r=1}^{R} \lambda_r (\mathbf{E}_r)_{ij} (\mathbf{E}_r)^*_{kl}, \tag{4.19}$$

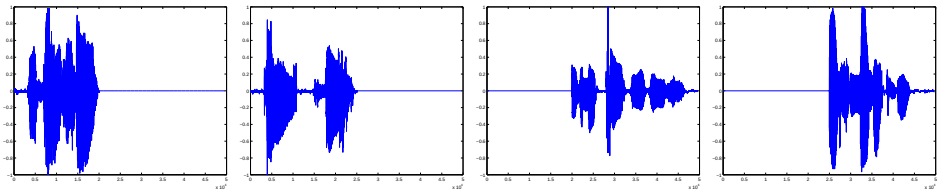

FIG. 4.3 – Signaux sources

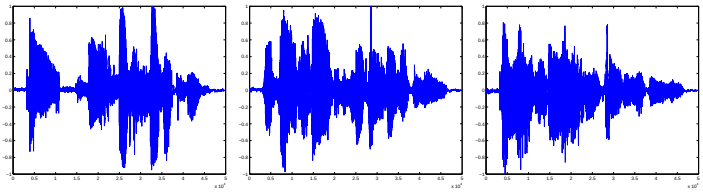

FIG. 4.4 – Signaux mélangés

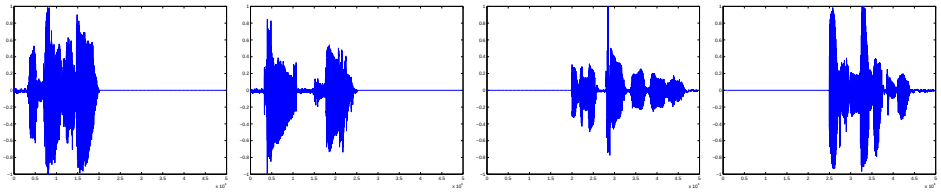

FIG. 4.5 – Signaux estimés

où les matrices $\mathbf{E}_r, r \in [1:R]$ *sont hermitiennes et mutuellement orthogonales au sens du produit vectoriel euclidien et où les* $\lambda_r, r \in [1:R]$ *sont des scalaires réels.*R *désigne le rang de la matrice* $\mathbf{T} = mat(\mathcal{T})$ *de taille* $J^2 \times J^2$.
Par ailleurs, posons $\mathbf{e}_r = vec(\mathbf{E}_r)$, $\mathbf{E} = [\mathbf{e}_1, \mathbf{e}_2, \ldots, \mathbf{e}_r] \in \mathbb{C}^{J^2 \times J^2}$ *et* $\mathbf{\Lambda} = diag([\lambda_1, \lambda_2, \ldots, \lambda_R]) \in \mathbb{R}^{J^2 \times J^2}$. *Alors nous pouvons écrire l'équation (4.19) sous forme matricielle :*

$$\mathbf{T} = \mathbf{E}\mathbf{\Lambda}\mathbf{E}^H \tag{4.20}$$

où les colonnes de $\mathbf{E}$ *sont orthogonale et de norme 1 et où* $e_{(i-1)J+j,r} = e^*_{(j-1)J+i,r}$, $(i,j) \in [1:J]$.

Preuve :
En vertu de la symétrie $t_{klij} = t^*_{ijkl}$, la matrice $\mathbf{T}$ est hermitienne. Par conséquent, sa décomposition en valeurs propres s'écrit sous la forme $\mathbf{T} = \mathbf{E}\mathbf{\Lambda}\mathbf{E}^H$, avec $\mathbf{\Lambda}$ une matrice diagonale réelle. Le tenseur $\mathcal{T}$ peut alors s'écrire sous la forme (4.19) avec les $\mathbf{E}_r$ mutuellement orthogonales. Soit $\mathcal{S}$ un tenseur d'ordre 4 défini au niveau élémentaire par $s_{ijkl} = t_{jilk}$. On note $\mathbf{S} = mat(\mathcal{S})$ sa

représentation matricielle. D'après (4.19), les éléments de $\mathcal{S}$ s'écrivent :

$$s_{ijkl} = \sum_{r=1}^{R} \lambda_r (\mathbf{E}_r)_{ji} (\mathbf{E}_r)^*_{lk}, \tag{4.21}$$

et la matrice $\mathbf{S}$ s'écrit :

$$\mathbf{S} = \tilde{\mathbf{E}} \mathbf{\Lambda} \tilde{\mathbf{E}}^H \tag{4.22}$$

où la matrice $\tilde{\mathbf{E}} = [\mathbf{e}_1, \mathbf{e}_2, \ldots, \mathbf{e}_R]$ est définie par $\tilde{e}_{(j-1)J+i,r} = e_{(i-1)J+j,r}$, $(i,j) \in [1:J], r \in [1:R]$. D'autre part, la décomposition en valeurs propres de $\mathbf{T}^*$ s'écrit :

$$\mathbf{T}^* = \mathbf{E}^* \mathbf{\Lambda} \mathbf{E}^T \tag{4.23}$$

et

$$t^*_{ijkl} = \sum_{r=1}^{R} \lambda_r (\mathbf{E}_r)^*_{ij} (\mathbf{E}_r)_{kl}. \tag{4.24}$$

En raison de la deuxième symétrie de $\mathcal{T}$, $t_{jilk} = t^*_{ijkl}$, les tenseurs $\mathcal{S}$ et $\mathcal{T}^*$ sont égaux et les matrices $\mathbf{S}$ et $\mathbf{T}^*$ sont égales. Si toutes les valeurs propres sont différentes, alors les projecteurs correspondants aux mêmes valeurs propres sont égaux :

$$\mathbf{e}_r * \mathbf{e}_r^T = \tilde{\mathbf{e}}_r \tilde{\mathbf{e}}_r^H, \forall r \in [1:R], \tag{4.25}$$

d'où

$$(\mathbf{E}_r)_{ji} (\mathbf{E}_r)^*_{lk} = (\mathbf{E}_r)_{ij} (\mathbf{E}_r)_{kl} \tag{4.26}$$

Si certaines valeurs propres sont d'ordre multiple, on peut toujours choisir les projecteurs $\mathbf{e}_r * \mathbf{e}_r^T$ et $\tilde{\mathbf{e}}_r \tilde{\mathbf{e}}_r^H$ égaux. Nous voyons donc que tous les projecteurs possèdent les mêmes symétrie que $\mathcal{T}$. Nous allons maintenant montrer que les matrices $\mathbf{E}_r$ peuvent toujours être choisies hermitiennes. Soit α_r un scalaire de module 1. Si on multiplie $\mathbf{E}_r$ par ce saclaire, le projecteur associé reste inchangé : $\alpha_r \mathbf{e}_r \alpha_r^* \mathbf{e}_r^H = \mathbf{e}_r \mathbf{e}_r^H$. Posons $\mathbf{E}_r^{'} = \alpha_r \mathbf{E}_r$.
Si un élément diagonal de $\mathbf{E}_r$ (par exemple $(\mathbf{E}_r)_{pp}$) est non nul, on peut choisir α_r de telle sorte que $(\mathbf{E}_r^{'})_{pp}$ soit réel, par exemple $\alpha_r = .(\mathbf{E}_r)^*_{pp} / |(\mathbf{E}_r)^*_{pp}|$. Alors, d'après (4.26), $(\mathbf{E}_r)_{ij} (\mathbf{E}_r^{'})_{pp} = (\mathbf{E}_r^{'})_{ji} (\mathbf{E}_r^{'})^*_{pp} = (\mathbf{E}_r^{'})_{ji} (\mathbf{E}_r^{'})_{pp}$ pour tout $(i,j) \in [1:J]$, donc la matrice $\mathbf{E}_r^{'}$ est hermitienne.
Si tous les éléments diagonaux de $\mathbf{E}_r$ sont nuls, nous procédons de manière différente. D'après l'équation (4.26), nous avons $|(\mathbf{E}_r)_{ji}| = |(\mathbf{E}_r)_{ij}|$. Sélectionnons un élément non nul de $\mathbf{E}_r$, par exemple $(\mathbf{E}_r)_{pq}$. Nous pouvons choisir α_r de telle sorte que $(\mathbf{E}_r^{'})_{pq}$ soit égal à $(\mathbf{E}_r^{'})^*_{qp}$, par exemple $\alpha_r = exp((-j/2) Arg((\mathbf{E}_r)_{pq}) + Arg((\mathbf{E}_r)_{qp}))$. Nous avons alors $(\mathbf{E}_r^{'})^*_{ij} (\mathbf{E}_r^{'})_{pq} = (\mathbf{E}_r^{'})_{ji} (\mathbf{E}_r^{'})^*_{qp} = (\mathbf{E}_r^{'})_{ji} (\mathbf{E}_r^{'})_{pq}$ pour pour tout $(i,j) \in [1:J]$, donc la matrice $\mathbf{E}_r^{'}$ est hermitienne. □

Nous insistons ici sur le fait que les matrices $\mathbf{E}_r$ ne sont pas hermitiennes par défaut, en effet, la matrice $\mathbf{E}^{''} = j\mathbf{E}$ est aussi une matrice propre de $\mathbf{T}$. Mais alors $(\mathbf{E}_r^{''})_{ij} = -(\mathbf{E}_r^{''})^*_{ji}$.
D'autre part, les matrices $\mathbf{T}$ sont de taille $J^2 \times J^2$, le rang R maximum est J^2 (taille de l'espace vectoriel des matrices hermitiennes de taille $J \times J$ dans $\mathbb{C}$). Dans le cas où les données sont réelles, le théorème reste valable si l'on supprime les conjugaisons. Le rang R maximum est alors $J(J+1)/2$, taille de l'espace vectoriel des matrices symétriques de taille $J \times J$ dans $\mathbb{R}$.

Le tenseur $\mathcal{C}^{\mathbf{x}}$ vérifie les hypothèses du théorème 3, par conséquent, $\mathcal{C}^{\mathbf{x}}$ et $\mathbf{C}^{\mathbf{x}}$ se décomposent sous la forme :

$$(\mathcal{C}^{\mathbf{x}})_{ijkl} = \sum_{r=1}^{R} \lambda_r (\mathbf{E}_r)_{ij} (\mathbf{E}_r)^*_{kl} \tag{4.27}$$

et

$$\mathbf{C}^{\mathbf{x}} = \mathbf{E}\mathbf{\Lambda}\mathbf{E}^H \tag{4.28}$$

où la matrice $\mathbf{E}$ est unitaire et où les matrices $\mathbf{E}_r = unvec(E_r)$ sont hermitiennes.
On peut obtenir la matrice $\mathbf{E}$ en calculant la décomposition en valeurs propres de $\mathbf{C}^{\mathbf{x}}$, les vecteurs propres obtenus doivent ensuite être normalisés de manière à rendre les matrices $\mathbf{E}_r$ hermitiennes, comme cela a été expliqué dans la preuve du théorème 3.

Nous allons chercher à lier les équations (4.18) et (4.28).
Supposons que les kurtosis de toutes les sources sont positifs. La matrice $\mathbf{C}^{\mathbf{x}}$ est alors définie positive. Si les kurtosis sont tous négatifs, comme c'est généralement le cas en communications numériques, on prendra la matrice $-\mathbf{C}^{\mathbf{x}}$ au lieu de la matrice $\mathbf{C}^{\mathbf{x}}$. Si les kurtosis des sources ne sont pas tous de même signe, le raisonnement qui va suivre reste le même, mais la matrice $\mathbf{Q}$ dont il sera question à partir de l'équation (4.30) n'est plus orthogonale mais non singulière. Le système (4.47) ne sera plus un système de matrices à diagonaliser conjointement par une matrice orthogonale mais par une matrice non singulière. Les méthodes que l'on peut utiliser pour résoudre un tel système ont été évoquées au chapitre 2.

$\mathbf{C}^{\mathbf{x}}$ s'écrit d'après (4.18) et (4.28) :

$$\mathbf{C}^{\mathbf{x}} = \left((\mathbf{A} \odot \mathbf{A}^*)(\tilde{\mathbf{C}}^{\mathbf{s}})^{1/2}\right)\left((\mathbf{A} \odot \mathbf{A}^*)(\tilde{\mathbf{C}}^{\mathbf{s}})^{1/2}\right)^H = \left(\mathbf{E}\mathbf{\Lambda}^{1/2}\right)\left(\mathbf{E}\mathbf{\Lambda}^{1/2}\right)^H \tag{4.29}$$

Les matrices $(\mathbf{A}\odot\mathbf{A}^H)(\tilde{\mathbf{C}}^{\mathbf{s}})^{1/2}$ et $\mathbf{E}\mathbf{\Lambda}^{1/2}$ sont des racines carrées de $\mathbf{C}^{\mathbf{x}}$. Il existe donc une matrice $\mathbf{Q}$ unitaire les reliant :

$$(\mathbf{A} \odot \mathbf{A}^*)(\tilde{\mathbf{C}}^{\mathbf{s}})^{1/2} = \mathbf{E}\mathbf{\Lambda}^{1/2}\mathbf{Q} \tag{4.30}$$

Nous allons montrer que $\mathbf{Q}$ est par ailleurs réelle. Pour cela nous allons exploiter le fait que les colonnes de $\mathbf{A} \odot \mathbf{A}^H$ et celles de $\mathbf{E}$ peuvent s'écrire sous la forme de matrices hermitiennes

En effet, soit $\mathbf{\Pi}$ une matrice de permutation de taille $J^2 \times J^2$ telle que

$$\begin{cases} \mathbf{\Pi}_{(i-1)J+j,(j-1)J+i} &= 1 \\ \mathbf{\Pi}_{ij} &= 0 \text{ ailleurs} \end{cases} \tag{4.31}$$

Alors

$$\mathbf{\Pi}(\mathbf{A} \odot \mathbf{A}^*) = (\mathbf{A} \odot \mathbf{A}^*)^* \tag{4.32}$$

$$\mathbf{\Pi}\mathbf{E} = \mathbf{E}^* \tag{4.33}$$

Nous avons par conséquent :

$$\mathbf{\Pi}(\mathbf{A} \odot \mathbf{A}^*)(\tilde{\mathbf{C}}^{\mathbf{s}})^{1/2} = \mathbf{\Pi}\mathbf{E}\mathbf{\Lambda}^{1/2}\mathbf{Q} \tag{4.34}$$

$$= (\mathbf{A} \odot \mathbf{A}^*)^*(\tilde{\mathbf{C}}^{\mathbf{s}})^{1/2} = \mathbf{E}^*\mathbf{\Lambda}^{1/2}\mathbf{Q} \tag{4.35}$$

$$= \mathbf{E}^*\mathbf{\Lambda}^{1/2}\mathbf{Q}^* \tag{4.36}$$

La matrice $\mathbf{Q}$ est donc réelle. □

Les matrices $\mathbf{E}$ et $\mathbf{\Lambda}$ sont connues, donc si la matrice $\mathbf{Q}$ était connue on pourrait évaluer $(\mathbf{A} \odot \mathbf{A}^*)\mathbf{C}^{\mathbf{s}1/2}$. En effet, la colonne r de cette matrice s'écrit comme le produit du scalaire $\kappa_r^{1/2} = (\mathbf{C}^{\mathbf{s}1/2})_{rr}$ avec $\mathbf{A}_r \otimes \mathbf{A}_r^*$. Soit $\mathbf{N}_r$, $r \in [1 : R]$ la matrice de taille $J \times J$ définie par $\mathbf{N}_r = unvec(((\mathbf{A} \odot \mathbf{A}^*)\mathbf{C}^{\mathbf{s}1/2})_r)$. $\mathbf{N}_r$ peut s'écrire :

$$\mathbf{N}_r = \kappa_r^{1/2} A_r A_r^H \tag{4.37}$$

$\mathbf{N}_r$ est une matrice de rang un. A un facteur d'échelle près, le vecteur A_r peut être estimé comme le vecteur propre dominant de $\mathbf{N}_r$.
Il s'agit maintenant d'estimer la matrice $\mathbf{Q}$. Nous allons pour cela exploiter la structure particulière de la matrice $\mathbf{A} \odot \mathbf{A}^*$.
Posons $\mathbf{H} = \mathbf{E}^*\mathbf{\Lambda}^{1/2}$, et notons $\mathbf{H}_r$ la matrice hermitienne obtenue en extrayant la $r$ème colonne de $\mathbf{H}$ et en la réordonnant sous la forme d'une matrice : $\mathbf{H}_r = unvec(H_r)$. D'après l'équation (4.30), la matrice $\mathbf{H}_r$ peut s'écrire en fonction de $\mathbf{A}$, $\mathbf{C}^{\mathbf{s}1/2}$ et $\mathbf{Q}$:

$$\mathbf{H}_r = \sum_{k=1}^{R} A_k A_k^H \kappa_k^{1/2} q_{rk}. \tag{4.38}$$

Les matrices $\mathbf{H}_r$ s'écrivent comme des combinaisons linéaires de matrices de rang un. Nous allons donc chercher des combinaisons linéaires des matrices $\mathbf{H}_r$ qui sont des matrices de rang un. Nous avons pour cela besoin d'un outil nous permettant de savoir si une matrice hermitienne est de rang un. Le théorème suivant nous propose un tel outil. Ce théorème est très similaire au théorème 2 présenté au chapitre 2 au détail près qu'il s'applique à des matrices hermitiennes.

Théorème 4 *Matrices hermitiennes de rang un*
Soit la fonction $\Phi : (\mathbf{X}, \mathbf{Y}) \in \mathbb{H}^{J\times J} \times \mathbb{H}^{J\times J} \longmapsto \Phi(\mathbf{X}, \mathbf{Y}) \in \mathbb{C}^{J\times J\times J\times J}$, *où* $\mathbb{H}^{J\times J}$ *désigne l'ensemble des matrices hermitiennes de taille* $J \times J$, *définie par :*

$$(\Phi(\mathbf{X}, \mathbf{Y}))_{ijkl} = x_{ij} y_{kl}^* + y_{ij} x_{kl}^* - x_{ik} y_{jl}^* - y_{ik} x_{jl}^* \tag{4.39}$$

Soit $\mathbf{X} \in \mathbb{H}^{J\times J}$, *alors* $\Phi(\mathbf{X}, \mathbf{X}) = 0$ *si et seulement si le rang de* $\mathbf{X}$ *est au plus un.*

La preuve de ce théorème est analogue à la preuve qui a été donnée pour le théorème 2 au chapitre 2.

La technique qui va être employée pour déterminer $\mathbf{Q}$ est proche de la technique que nous avons employé au chapitre 2 pour déterminer la matrice non singulière $\mathbf{F}$.
Notons Φ_{rs} le tenseur d'ordre 4 défini par $\Phi_{rs} = \Phi(\mathbf{H}_r, \mathbf{H}_s)$. Soit $\mathbf{W}$ une matrice symétrique réelle telle que :

$$\sum_{r,s=1}^{R} \Phi_{rs} w_{rs} = 0 \tag{4.40}$$

Cette expression peut encore s'écrire :

$$\sum_{r,s=1}^{R} \sum_{t,u=1}^{R} \Phi(A_t A_t^H, A_u A_u^H) q_{rt} q_{su} \kappa_t^{1/2} \kappa_u^{1/2} w_{rs} = 0 \tag{4.41}$$

Ou encore, en vertu de la symétrie de $\mathbf{W}$ et de Φ et d'après le théorème 4 :

$$\sum_{r,s=1}^{R} \sum_{\substack{t,u=1 \\ t<u}}^{R} \Phi(A_t A_t^H, A_u A_u^H) q_{rt} q_{su} \kappa_t^{1/2} \kappa_u^{1/2} w_{rs} = 0 \tag{4.42}$$

Supposons que les tenseurs $(\Phi(A_t A_t^H, A_u A_u^H))_{t<u}$ soient linéairement indépendants. Alors l'équation (4.42) nous donne :

$$\sum_{r,s=1}^{R} q_{rt} q_{su} w_{rs} = \lambda_{tu} \delta_{tu}, \quad \forall t, u, \tag{4.43}$$

où δ désigne le symbole de kronecker et où λ_{tu} est un scalaire. Cette équation s'écrit sous forme matricielle :

$$\mathbf{W} = \mathbf{Q}\mathbf{D}\mathbf{Q}^T, \tag{4.44}$$

où la matrice $\mathbf{D}$ est une matrice diagonale dont les éléments diagonaux sont les δ_{tt}, $t \in [1:R]$. Toute matrice symétrique réelle s'écrivant sous la forme $\mathbf{W} = \mathbf{Q}\mathbf{D}\mathbf{Q}^T$ où $\mathbf{D}$ est une matrice diagonale vérifie (4.40). L'hypothèse d'existence d'une matrice satisfaisant (4.40) est donc vérifiée. Par ailleurs, un ensemble de R matrices diagonales réelles indépendantes $\mathbf{D}_r$ génère R matrices symétriques réelles indépendantes $\mathbf{W}_r$ vérifiant (4.40). Donc il existe au plus R matrices symétriques réelles indépendantes $\mathbf{W}_r$ satisfaisant (4.40).
Les vecteurs $vec(W_r)$, $r \in [1:R]$ appartiennent au noyau de la matrice $\tilde{\mathbf{P}}$ de taille $J^4 \times R^2$ définie par $\tilde{\mathbf{P}} = [vec(\Phi_{11}), vec(\Phi_{12}), \ldots, vec(\Phi_{RR})]$. Toute matrice antisymétrique $\mathbf{M}$ vérifie elle aussi l'équation (4.40) et le vecteur la représentant $vec(\mathbf{M})$ appartient aussi au noyau de $\tilde{\mathbf{P}}$. Nous devons donc chercher les solutions symétriques dans le noyau de $\tilde{\mathbf{P}}$.
Soit $\mathbf{X}$ une matrice triangulaire supérieure telle que $x_{rs} = w_{rs}$ si $r < s$, $x_{rs} = \frac{1}{2} w_{rs}$ si $r = s$ et $x_{rs} = 0$ sinon. En vertu de la symétrie de Φ et de $\mathbf{W}$, l'équation (4.40) peut s'écrire :

$$\sum_{\substack{r,s=1 \\ r \leq s}}^{R} \Phi_{rs} x_{rs} + \sum_{r=1}^{R} \Phi_{rr} x_{rr} = 0 \tag{4.45}$$

Rangeons les tenseurs Φ_{rs}, $(r,s) \in [1:R], r \leq s$ dans les vecteurs $P_{rs} = vec(\Phi_{rs})$ de taille J^4 et ces vecteurs P_{rs} dans la matrice $\mathbf{P} = [P_{11}, P_{12}, \ldots, P_{RR}]$ de taille $J^4 \times R(R+1)/2$. L'équation précédente peut encore s'écrire :

$$[P_{11}, P_{12}, \ldots, P_{RR}] \begin{bmatrix} x_{11} \\ x_{12} \\ \vdots \\ x_{RR} \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix} \tag{4.46}$$

Les R vecteurs singuliers de droite de $\mathbf{P}$ correspondants aux R plus petites valeurs singulières sont solutions. Après avoir rangés ces vecteurs solutions dans les matrices triangulaires supérieures $\mathbf{X}_t$, $t \in [1:R]$, nous pouvons trouver les matrices $\mathbf{W}_t$ simplement en calculant $\mathbf{W}_t = \mathbf{X}_t + \mathbf{X}_t^T$. La matrice $\mathbf{Q}$ peut alors être évaluée à l'aide d'une diagonalisation simultanée par une matrice orthogonale du système suivant :

$$\left\{\begin{array}{rcl} \mathbf{W}_1 & = & \mathbf{Q}\mathbf{D_1}\mathbf{Q}^T \\ \mathbf{W}_2 & = & \mathbf{Q}\mathbf{D_1}\mathbf{Q}^T \\ & \vdots & \\ \mathbf{W}_R & = & \mathbf{Q}\mathbf{D_R}\mathbf{Q}^T \end{array}\right. \tag{4.47}$$

Ce système peut par exemple être résolu à l'aide de l'algorithme de Jacobi proposé dans [7,8].

Pour obtenir ce résultat, nous avons supposé que les tenseurs $\Phi(A_tA_t^T, A_uA_u^T)$, $t < u$ étaient linéairement indépendants. Les limites de validité de cette hypothèse sont données par le théorème suivant :

Théorème 5 *Indépendance des tenseurs* $\Phi(A_tA_t^T, A_uA_u^T)$
Les tenseurs $\Phi(A_tA_t^T, A_uA_u^T)$, $t < u$ *sont génériquement indépendants si*

$$2R(R-1) \leq J^2(J-1)^2/2 \tag{4.48}$$

dans le cas complexe et $R < R_{max}$ *défini par le tableau 4.1 dans le cas réel [45].*

Remarque Il a été établi dans [10] que pour les applications de traitement d'antennes, les caractéristiques de l'antenne (réseau linéaire, circulaire) peuvent introduire une structure particulière dans le tenseur des observations, qui limite le nombre de sources. Pour ces applications, le nombre de sources pouvant être présentes n'est pas donc borné par le théorème 5 mais par le minimum entre le nombre maximal de sources donné par le théorème 5 et le nombre maximal de capteurs virtuels donné par [10].
Un synopsis de l'algorithme FOOBI est donné dans la table 4.3.

1. Estimer le tenseur des cumulants des observations $\mathcal{C}^{\mathbf{x}} = cum(\mathbf{x}, \mathbf{x}^*, \mathbf{x}^*, \mathbf{x})$.
2. Ranger $\mathcal{C}^{\mathbf{x}}$ dans la matrice $\mathbf{C}^{\mathbf{x}}$.
3. Calculer la décomposition en valeurs propres de $\mathbf{C}^{\mathbf{x}} = \mathbf{E}\mathbf{\Lambda}\mathbf{E}^H$, poser $\mathbf{H} = \mathbf{E}\mathbf{\Lambda}^{1/2}$ et normaliser les colonnes de $\mathbf{H}$ de façon à ce que les matrices $unvec((\mathbf{H})_r)$, $r \in [1:R]$ soient hermitiennes. Le nombre de sources R est égal au rang de la matrice $\mathbf{C}^{\mathbf{x}}$.
4. Calculer les tenseurs $\Phi_{rs} = \Phi(\mathbf{H}_r, \mathbf{H}_s)$ pour $(r,s) \in [1:R], r \leq s$ et les ranger dans les vecteurs $P_{rs} = vec(\mathcal{P}_{rs})$.
5. Ranger les vecteurs P_{rs} dans la matrice $\mathbf{P} = [P_{11}, P_{12}, \ldots, P_{RR}]$ de taille $J^4 \times R(R+1)/2$.
6. Calculer les R vecteurs singuliers de droite de $\mathbf{P}$ correspondants aux R plus petites valeurs singulières et les ranger dans les matrices triangulaires supérieures $\mathbf{X}_r, r \in [1:R]$.
7. Evaluer les matrices $\mathbf{B}_r = \mathbf{X}_r + \mathbf{X}_r^T$.
8. Résoudre le système $\mathbf{B}_r = \mathbf{Q}\mathbf{\Lambda}_r\mathbf{Q}^T$, $r \in [1:R]$, avec $\mathbf{Q}$ orthogonale réelle.
9. Estimer les colonnes de $\mathbf{A}$: pour tout $r \in [1:R]$, la $r$ème colonne de $\mathbf{A}$ est le vecteur propre dominant de la matrice $\mathbf{N}_r = unvec((\mathbf{HQ})_r)$.

TAB. 4.3 – Résumé de FOOBI

4.3.2 FOOBI-2

Nous allons proposer une variante de l'algorithme FOOBI, qui sera notée FOOBI-2. Comme nous allons le voir, cet algorithme fonctionne pour un nombre de sources R plus grand que FOOBI. Nous supposons ici que les kurtosis des sources sont tous positifs et nous partons de l'équation (4.38). Nous n'allons pas utiliser l'outil proposé dans le théorème 4 pour déterminer si une matrice hermitienne est de rang un, mais l'outil proposé dans le théorème suivant :

Théorème 6 *Matrices hermitiennes de rang un*
Soit la fonction Ψ *:* $(\mathbf{X},\mathbf{Y}) \in \mathbb{H}^{J\times J} \times \mathbb{H}^{J\times J} \longmapsto \Psi(\mathbf{X},\mathbf{Y}) \in \mathbb{C}^{J\times J\times J\times J}$*, définie par :*

$$\Psi(\mathbf{X},\mathbf{Y})) = \mathbf{XY} + \mathbf{YX} - trace(\mathbf{X})\mathbf{Y} - trace(\mathbf{Y})\mathbf{X}. \tag{4.49}$$

Soit $\mathbf{X} \in \mathbb{H}^{J\times J}$*, alors* $\Psi(\mathbf{X},\mathbf{X}) = 0$ *si et seulement si le rang de* $\mathbf{X}$ *est au plus un.*

Preuve : Le cas =0 est évident.
Soit $\mathbf{X}$ une matrice de rang un. Il existe un vecteur U tel que $\mathbf{X} = UU^H$. Alors $\Psi(\mathbf{X},\mathbf{X}) = 2UU^HUU^H - 2U^HUUU^H = 0$.
Soit maintenant $\mathbf{X}$ une matrice hermitienne telle que $\Psi(\mathbf{X},\mathbf{X}) = 0$. La décomposition en valeurs propres de $\mathbf{X}$ s'écrit $\mathbf{X} = \mathbf{U\Lambda U}^H$. $\Psi(\mathbf{X},\mathbf{X}) = 0$ si et seulement si

$$\mathbf{XX} = trace(\mathbf{X})\mathbf{X} \tag{4.50}$$

Soit encore

$$\begin{aligned} \mathbf{U\Lambda}^2\mathbf{U}^H &= trace(\mathbf{\Lambda})\mathbf{U\Lambda U}^H & (4.51)\\ \mathbf{\Lambda}^2 &= trace(\mathbf{\Lambda})\mathbf{\Lambda} & (4.52)\\ \mathbf{\Lambda} &= trace(\mathbf{\Lambda})\mathbf{I}. & (4.53) \end{aligned}$$

Par conséquent au plus une valeur propre est différente de zéro. □

Posons $\mathbf{H} = \mathbf{E}^*\mathbf{\Lambda}^{1/2}$, et notons $\mathbf{H}_r$ la matrice hermitienne obtenue en extrayant la $r$ème colonne de $\mathbf{H}$ et en la réordonnant sous la forme d'une matrice : $\mathbf{H}_r = unvec(H_r)$. Notons Ψ_{st} le tenseur d'ordre 4 défini par $\Psi_{st} = \Psi(\mathbf{H}_s,\mathbf{H}_t)$. Nous avons pour tout r dans $[1:R]$:

$$\begin{aligned} \sum_{s,t=1}^{R} q_{sr}q_{tr}\Psi_{st} &= \sum_{s,t=1}^{R} q_{sr}q_{tr} \sum_{u,v=1}^{R} \kappa_u^{1/2}\kappa_v^{1/2} q_{su}q_{tv}\Psi(A_uA_u^H, A_vA_v^H) & (4.54)\\ &= \sum_{u,v=1}^{R} \Psi(A_uA_u^H, A_vA_v^H)\kappa_u^{1/2}\kappa_v^{1/2} \sum_{s,t=1}^{R} q_{sr}q_{tr}q_{su}q_{tv} & (4.55)\\ &= \sum_{u,v=1}^{R} \Psi(A_uA_u^H, A_vA_v^H)\kappa_u^{1/2}\kappa_v^{1/2}(\mathbf{Q}^T\mathbf{Q})_{ru}(\mathbf{Q}^T\mathbf{Q})_{rv}. & (4.56) \end{aligned}$$

Si $r \neq u$ et $r \neq v$, alors $(\mathbf{Q}^T\mathbf{Q})_{ru}(\mathbf{Q}^T\mathbf{Q})_{rv} = 0$ en vertu de l'orthogonalité de $\mathbf{Q}$. Si $r = u = v$, alors $\Psi(A_uA_u^H, A_vA_v^H) = 0$ en vertu du théorème 6. Finalement :

$$\sum_{s,t=1}^{R} q_{sr}q_{tr}\Psi_{st} = 0. \tag{4.57}$$

Soit $\mathbf{B}_{ij}$ les matrices symétriques de taille $R \times R$ définie par $(\mathbf{B}_{ij})_{st} = (\Psi_{st})_{ij}$. L'équation précédente peut encore s'écrire à l'aide des matrices $\mathbf{B}_{ij}$:

$$diag(\mathbf{Q}^T \mathbf{B}_{ij} \mathbf{Q}) = 0 \quad 1 \leq i \leq j \leq J. \tag{4.58}$$

Par conséquent,

$$\begin{aligned} diag(\mathbf{Q}^T Re(\mathbf{B}_{ij})\mathbf{Q}) &= 0 \quad 1 \leq i \leq j \leq J & (4.59)\\ diag(\mathbf{Q}^T Im(\mathbf{B}_{ij})\mathbf{Q}) &= 0 \quad 1 \leq i < j \leq J & (4.60) \end{aligned}$$

La matrice $\mathbf{Q}$ peut donc être estimée à l'aide d'une zéro-diagonalisation conjointe de $J(J+1)/2$ (dans le cas réel) ou de J^2 (dans le cas complexe) matrices symétriques réelles. Cette zéro-diagonalisation conjointe peut être conduite à l'aide d'une variante de l'algorithme de Jacobi proposé dans [7,8]. Il faut choisir à chaque étape la rotation de Jacobi qui *minimise* (au lieu de *maximise*) la somme des carrés des éléments diagonaux.
Si la décomposition (4.17) est unique, il est possible d'appliquer FOOBI-2 tant que $R \leq J^2$ dans le cas complexe ou $R \leq J(J+1)/2$ dans le cas réel. Il a été établi dans [14] qu'une décomposition en une somme de termes de rang un est génériquement unique si le nombre de paramètres dans la décomposition est strictement inférieur au nombre d'éléments indépendants dans le tenseur. Si ces deux nombres sont égaux, alors il existe un nombre fini de décompositions possibles.
Dans le cas complexe, le nombre total de parties réelles et de parties imaginaires distinctes des éléments d'un tenseur générique $\mathcal{T}$ vérifiant les symétries $t_{klij} = t_{jilk} = t^*_{ijkl}$ et $t_{ljki} = t_{ijkl}$ est égal à (voir chapitre 5) :

$$D(J) = 6\binom{J}{4} + 4\binom{J}{1}\binom{J-1}{2} + 3\binom{J}{2} + 2\binom{J}{1}\binom{J-1}{1} + \binom{J}{1}, \tag{4.61}$$

où $\binom{J}{n} = 0$ si $n > J$. D'autre part, le nombre d'éléments réels dans la décomposition $t_{ijkl} = \sum_{r=1}^{R} \kappa_r a_{ir} a^*_{jr} a^*_{kr} a_{lr}$, avec κ_r réel est égal à $2JR$. La décomposition est donc unique si $R \leq R_{uc}$ avec R_{uc} donné dans la table 4.4.

Dans le cas réel, le nombre d'éléments indépendants dans le tenseur générique supersymétrique $\mathcal{T}$ vérifiant les symétries $t_{ijkl} = t_{\sigma(ijkl)}$, où σ désigne une permutation quelconque des indices, est égal à

$$D(J) = J(J+1)(J+2)(J+3)/4!, \tag{4.62}$$

et le nombre d'éléments réels dans la décomposition $t_{ijkl} = \sum_{r=1}^{R} \kappa_r a_{ir} a_{jr} a_{kr} a_{lr}$, avec κ_r réel est égal à JR. La décomposition est donc génériquement unique si $R \leq R_{ur}$ avec R_{ur} donné dans la table 4.4 (la remarque faite sur les applications en traitement d'antennes dans le paragraphe 4.3.1 est toujours valable).
Un synopsis de l'algorithme FOOBI-2 est donné dans la table 4.5.

4.3.3 Simulations

Dans une première simulation, nous avons supposé la présence de R sources reçues sur une antenne circulaire de rayon R_a constituée de $J = 4$ capteurs. On suppose une propagation en champ libre. La matrice de mélange s'écrit donc :

J	2	3	4	5	6	7	8	9
R_{uc}	2	5	12	22	36	55	80	112
R_{ur}	2	4	8	13	20	29	41	54

TAB. 4.4 – Rang maximal du tenseur $\mathcal{T}$ de taille $J \times J \times J \times J$ dans le cas complexe (R_{uc}) et dans le cas réel (R_{ur})

1. Estimer le tenseur des cumulants des observations $\mathcal{C}^{\mathbf{x}} = cum(\mathbf{x}, \mathbf{x}^*, \mathbf{x}, \mathbf{x}^*)$.
2. Ranger $\mathcal{C}^{\mathbf{x}}$ dans la matrice $\mathbf{C}^{\mathbf{x}}$.
3. Calculer la décomposition en valeurs propres de $\mathcal{C}^{\mathbf{x}} = \mathbf{E}\mathbf{\Lambda}\mathbf{E}^H$, poser $\mathbf{H} = \mathbf{E}\mathbf{\Lambda}^{1/2}$ et normaliser les colonnes de $\mathbf{H}$ de façon à ce que les matrices $unvec((\mathbf{H})_r)$, $r \in [1:R]$ soient hermitiennes. Le nombre de sources R est égal au rang de la matrice $\mathbf{C}^{\mathbf{x}}$.
4. Calculer les tenseurs Ψ_{st}, $(s,t) \in [1:R]$, $s \leq t$ et les ranger dans les matrices symétriques réelles $\mathbf{B}_{ij}$ définies par $(\mathbf{B}_{ij})_{st} = (\Psi_{st})_{ij}$, $(i,j) \in [1:J], i \leq j$.
5. Evaluer la matrice orthogonale réelle $\mathbf{Q}$, qui zéro-diagonalise les matrices $Re(\mathbf{B}_{ij})$ et $Im(\mathbf{B}_{ij})$.
6. Estimer les colonnes de $\mathbf{A}$: pour tout $r \in [1:R]$ la $r$ème colonne de $\mathbf{A}$ est le vecteur propre dominant de la matrice $\mathbf{N}_r = unvec((\mathbf{HQ})_r)$.

TAB. 4.5 – Résumé de FOOBI-2

$$a_{jr} = \exp\left(\frac{2\pi i}{\lambda}(x_j \cos(\theta_r)\cos(\phi_r) + y_j \cos(\theta_r)\sin(\phi_r))\right) \tag{4.63}$$

avec

$$x_j = R_a \cos(2\pi(j-1)/J) \tag{4.64}$$
$$y_j = R_a \sin(2\pi(j-1)/J) \tag{4.65}$$

avec i la racine carrée de -1 et où $R_a/\lambda = .55$.
On considérera deux cas, $R = 5$ et $R = 6$. Si $R = 6$, les directions d'arrivée des signaux sont données par $\theta_1 = 3\pi/10$, $\theta_2 = 3\pi/10$, $\theta_3 = 2\pi/5$, $\theta_4 = 0$, $\theta_5 = \pi/10$, $\theta_6 = 3\pi/5$ $\phi_1 = 7\pi/10$, $\phi_2 = 9\pi/10$, $\phi_3 = 3\pi/5$, $\phi_4 = 4\pi/5$, $\phi_5 = 3\pi/5$, $\phi_6 = \pi/5$. Si $R = 5$, on considère les cinq premières directions d'arrivée du cas $R = 6$.
Les éléments de la matrice sources appartiennent à une constellation QPSK. Le bruit est supposé suivre une loi gaussienne complexe de moyenne nulle.
La matrice de mélange est estimée à parir de (1) l'algorithme FOOBI, (2) l'algorithme FOOBI-2, et (3) l'algorithme BIRTH [1,2], qui s'appuie sur les statistiques d'ordre 6 des observations.
L'erreur relative est égale à la norme de la différence entre la matrice de mélange $\mathbf{A}$ et la matrice de mélange estimée $\hat{\mathbf{A}}$ (dont les colonnes ont été ordonnées et multipliées par un scalaire de

manière optimale) divisée par la norme de la matrice de mélange, moyennée sur 100 essais : $err = E\{\|\mathbf{A} - \hat{\mathbf{A}}\|/\|\mathbf{A}\|\}$.
Sur la figure 4.6, nous avons représenté l'erreur en fonction du RSB pour des signaux de longueur $T = 5000$ échantillons et $R = 5$ ou $R = 6$ sources. Les courbes de FOOBI et de FOOBI-2 coïncident pratiquement. BIRTH est légèrement moins précis. Nous avons également comparé les résultats à ceux qui sont obtenus à l'aide de l'algorithme AC-DC [54] appliqué aux matrices propres hermitiennes $\mathbf{H}_r$ du cumulant d'ordre 4 définies par (4.38), mais l'algorithme semble ne pas fonctionner dans le cas sous-déterminé.

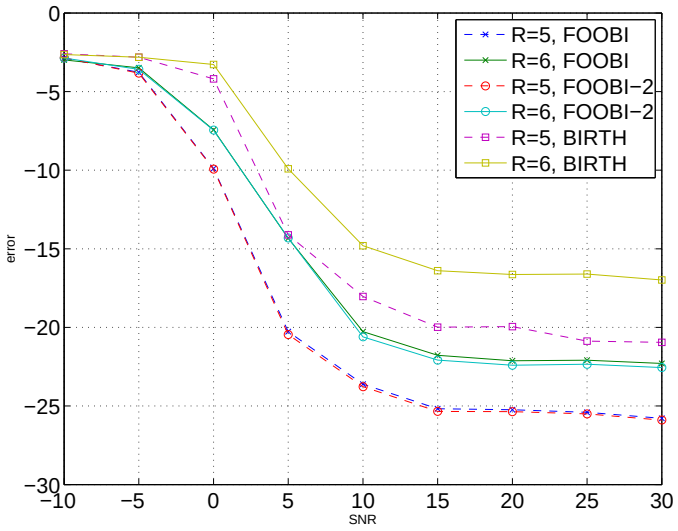

FIG. 4.6 – Erreur en fonction du RSB ($J = 4$,$R = 5, 6$, $T = 5000$).

Sur la figure 4.7, nous avons représenté l'erreur en fonction du nombre d'échantillons T dans le cas de $R = 5$ sources. Le RSB a été choisi égal à 16dB. Encore une fois, FOOBI et FOOBI-2 donnent des résultats comparables, tandis que BIRTH est un peu moins précis.
Sur la figure 4.8, nous avons représenté le coût de calcul en en fonction du nombre d'échantillons T. FOOBI et FOOBI-2 sont à peu près aussi coûteux tandis que BIRTH est environ 40 fois plus lent. Cela peut s'expliquer par le fait que BIRTH nécessite l'estimation des $O(J^6)$ éléments du cumulant d'ordre six. Le coût de l'estimation du cumulant d'ordre six représente plus de 90% du coût de calcul total. Le coût de l'estimation du cumulant d'ordre quatre représente 10% du coût total de FOOBI et FOOBI-2 pour 200 échantillons et 70% pour 5000 échantillons. Le coût de calcul varie peu en fonction du RSB.
Sur la figure 4.9, nous avons représenté l'erreur en fonction du conditionnement du problème. Le RSB est égal à 16dB et le nombre d'échantillons à $T = 5000$. θ_1 étant égal à θ_2, et ϕ_2 étant égal à $9\pi/10$, nous faisons varier ϕ_1 de $7\pi/10$ à $8.9\pi/10$. Les courbes de FOOBI et FOOBI-2 coïncident encore. Tant que le problème est bien conditionné, FOOBI et FOOBI-2 sont plus précis que BIRTH. Lorsque les deux premières colonnes de la matrice de mélange deviennent très proches (c'est-à-dire que ϕ_1 est très proche de ϕ_2), les performances de BIRTH deviennent meilleures que celles de FOOBI et FOOBI-2. Ce phénomène peut être expliqué par le fait que les vecteurs $A_1 \otimes A_1^*$ et $A_2 \otimes A_2^*$ sont plus proches l'un de l'autre que les vecteurs $A_1 \otimes A_1 \otimes A_1^*$ et $A_2 \otimes A_2 \otimes A_2^*$ [10].

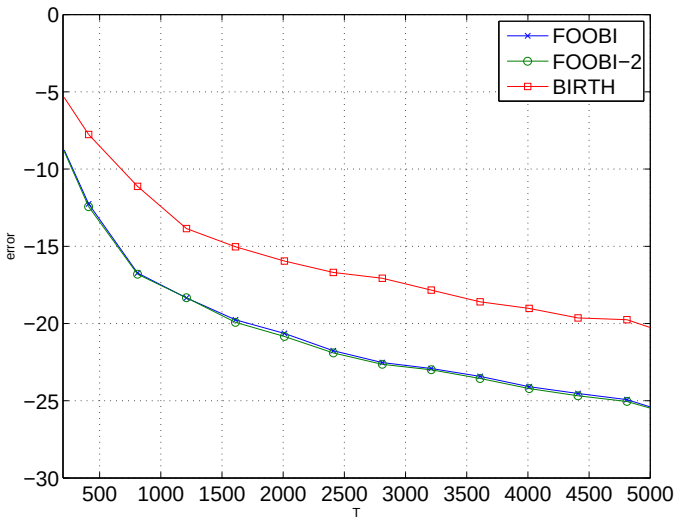

FIG. 4.7 – Erreur en fonction du nombre d'échantillons ($J = 4$, $R = 5$, $RSB = 16dB$).

Dans une seconde simulation, nous avons considéré le problème de $R = 5$ sources reçues sur un réseau circulaire de $J = 3$ antennes. Le nombre de sources étant supérieur au nombre de sources maximal autorisé pour FOOBI d'après le théorème 6, nous avons estimé la matrice de mélange uniquement à l'aide de (1) FOOBI-2 et de (2) BIRTH. Les paramètres de cette simulation sont les mêmes que ceux de la première simulation.
Sur la figure la figure 4.10, nous avons représenté l'erreur en fonction du RSB pour des signaux de longueur $T = 5000$ échantillons. FOOBI-2 est plus précis que BIRTH. Par ailleurs, les résultats obtenus en terme de performances et de coût de calcul en fonction du nombre d'échantillons pour un RSB fixé égal à 16 dB mènent aux mêmes conclusions que dans la simulation 1.

4.4 Conclusion

Dans ce chapitre, nous avons montré comment les résultats du chapitre 2 pouvaient être appliqués à l'Analyse en Composantes Indépendantes. Nous avons proposé tout d'abord un algorithme permettant d'estimer la matrice de mélange à partir de la décomposition d'un tenseur d'ordre trois contenant les matrices de covariance des observations pour différents retards. Puis nous avons proposé deux algorithmes utilisant la décomposition d'un tenseur d'ordre quatre contenant les cumulants d'ordre quatre des observations. Nous avons indiqué à chaque fois le nombre de sources maximal admissible en fonction du nombre de capteurs.

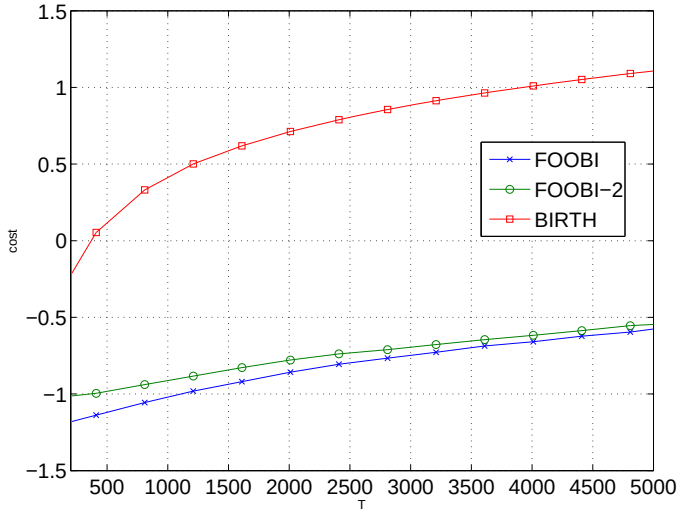

FIG. 4.8 – Coût de calcul en fonction du nombre d'échantillons ($J = 4$, $R = 5$, $RSB = 16dB$).

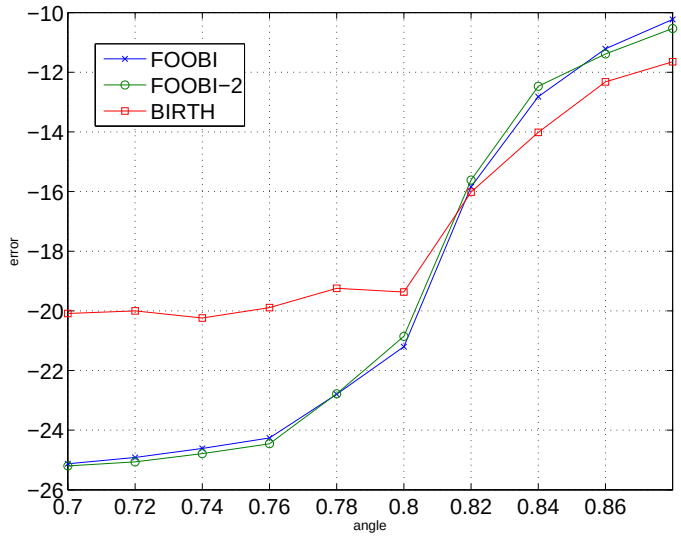

FIG. 4.9 – Erreur en fonction de l'angle d'élévation du premier vecteur de mélange ($J = 4$, $R = 5$, $RSB = 16dB$).

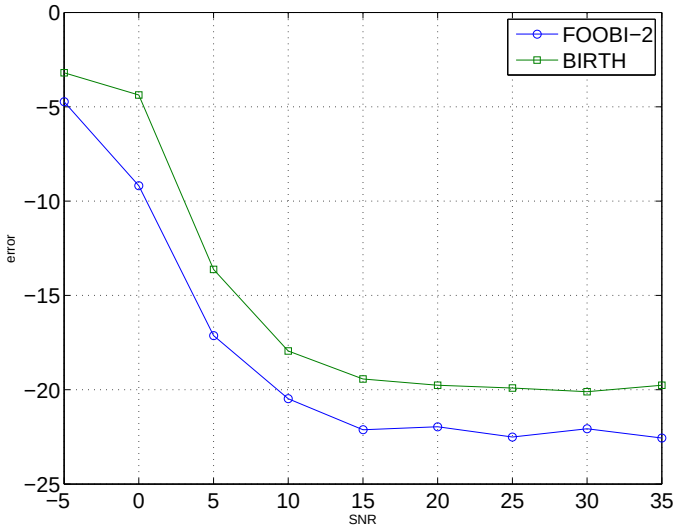

FIG. 4.10 – Erreur en fonction du RSB ($J = 3$).

Chapitre 5

Contribution au calcul du rang générique des tenseurs

5.1 Introduction

Nous avons introduit au chapitre 2 la notion de rang. Un certains nombre de questions peuvent se poser. Soit $\mathcal{T}$ un tenseur dont l'ordre et la taille sont donnés, quel rang maximal ce tenseur peut-il avoir ? Dans le cas des matrices, le rang est limité par la plus petite dimension de la matrice. Le rang maximal d'une matrice et donc toujours égal à la plus petite de ses dimensions. Le cas des tenseurs est différent. Le rang peut tout à fait être plus grand que la plus petite dimension du tenseur.

Exemple : Soit $\mathcal{T}$ le tenseur d'ordre 3 et de taille $3 \times 3 \times 3$ défini par [12] :

$$\begin{array}{rcl} \mathcal{T}(1,3,3) & = & \mathcal{T}(3,1,3) = \mathcal{T}(3,3,1) = 1 \\ \mathcal{T}(2,2,3) & = & \mathcal{T}(2,3,2) = \mathcal{T}(3,2,2) = 1 \\ \mathcal{T}(i,j,k) & = & 0 \text{ ailleurs} \end{array} \tag{5.1}$$

$\mathcal{T}$ est de rang 5.

Par ailleurs, un tenseur réel peut se décomposer sous la forme d'une somme de tenseurs *réels* de rang un ou d'une somme de tenseurs *complexes* de rang un. Le rang complexe, noté $R_{\mathbb{C}}$ est inférieur ou égal au rang réel noté $R_{\mathbb{R}}$. Après l'introduction de ce chapitre, nous ne nous intéresserons plus qu'au rang complexe.

Supposons maintenant que l'on choisisse aléatoirement selon une loi continue un tenseur dans l'ensemble des tenseurs d'ordre d et de taille $N_1 \times N_2 \times \ldots \times N_d$. Quel est le rang de ce tenseur ? Regardons tout d'abord le cas matriciel, le rang d'une matrice choisie aléatoirement est toujours égal au rang maximum, lui même égal comme nous l'avons vu à la plus petite dimension de la matrice. Dans le cas des tenseurs, le rang obtenu avec une probabilité non nulle est appelé *rang typique*. Il peut y avoir plusieurs rangs typiques, s'il n'y en a qu'un seul (c'est en particulier le cas dans $\mathbb{C}$ [14]), on parle de *rang générique*.

Exemple : Les rangs réels typiques pour un tenseur (réel) de taille $4 \times 4 \times 2$ sont 4 et 5 [48]. Un certain nombre de résultats ont été trouvés pour les rangs typiques de tenseurs $I \times J \times 2$ et $I \times J \times 3$, le rang générique des tenseurs symétriques de taille $2 \times 2 \times \ldots \times 2$, le rang générique des tenseurs symétriques d'ordre 3, le rang générique des tenseurs d'ordre 3 de taille $N \times N \times N$ [14, 16, 46–48].

Dans le paragraphe 5.2, nous nous intéresserons au cas particulier des tenseurs symétriques. Pour cette classe de tenseurs, il a été proposé dans [16] un algorithme pour évaluer le rang générique. Nous proposons d'appliquer un algorithme s'en inspirant pour évaluer le rang générique des tenseurs d'ordre 4 à symétrie complexe dans le paragraphe 5.2 et le rang générique des tenseurs quelconques dans le paragraphe 5.4. Nous conclurons par le paragraphe 5.5.

5.2 Tenseurs symétriques

On s'intéresse dans un premier temps au cas des tenseurs symétriques. Un tenseur $\mathcal{T}$ d'ordre d est dit symétrique si ses dimensions sont égales (on parlera dans la suite de la taille du tenseur par abus de langage) et que $\mathcal{T}_{(i_1,i_2,\ldots,i_d)} = \mathcal{T}_{\sigma(i_1,i_2,\ldots,i_d)}$, où $i_k \in [1:N]$ pour tout $k \in [1:d]$ et où σ est une permutation quelconque des indices.
Les résultats qui vont être présentés dans ce paragraphe ne sont pas nouveaux. Ils ont été étudiés en particulier dans [14–17]. Les auteurs ont regardé les tenseurs symétriques comme des formes polynomiales. Nous garderons ici la forme tensorielle pour conserver une cohérence avec la suite.
Le nombre total d'éléments dans un tenseur symétrique d'ordre d et de taille N, est N^d et le nombre d'éléments indépendants est $D(d,N) = \binom{N+d-1}{d}$ (voir [16], et A).
Dans la suite, on utilisera la notation suivante :

$$V^{\circ d} = \underbrace{V \circ V \circ \ldots \circ V}_{d \text{ fois}} \tag{5.2}$$

Nous allons voir qu'il est toujours possible d'écrire un tenseur $\mathcal{T}$ symétrique d'ordre d et de taille N sous la forme :

$$\mathcal{T} = \sum_{r=1}^{R} V_r^{\circ d} \tag{5.3}$$

avec $V_r \in \mathbb{C}, r \in [1:R]$.
Preuve :
Notons $\mathcal{S}_1(d,N)$ l'ensemble des tenseurs d'ordre d et de taille N s'écrivant sous la forme (5.3) et $\mathcal{S}(d,N)$ l'ensemble des tenseurs symétriques d'ordre d et de taille N. $\mathcal{S}(d,N)$ est un espace vectoriel et $\mathcal{S}_1(d,N)$ est un sous-espace vectoriel de $\mathcal{S}(d,N)$. Choisissons aléatoirement selon une loi continue des vecteurs V_r de taille N, et construisons avec les tenseurs symétriques $\mathcal{T}_r = V_r^{\circ d}$ et les vecteurs $U_r = vec(\mathcal{T}_r)$. Chacun de ces vecteurs, de taille N^d, représente un élément de $\mathcal{S}_1(d,N)$. Ils sont indépendants tant que la famille $(\mathcal{T}_r)$ est libre. Par calcul numérique, le rang de la matrice $[U_1, U_2, \ldots U_k]$ est égal à $D(d,N)$ si $k \geq D(d,N)$, donc la dimension de $\mathcal{S}_1(d,N)$ est supérieure ou égale à $D(d,N) = dim(\mathcal{S}(d,N))$. Comme $\mathcal{S}_1(d,N)$ est inclus dans $\mathcal{S}(d,N))$, on en déduit que ces deux ensembles sont égaux. □
Une démonstration plus rigoureuse de ce résultat est donnée dans [15].

L'entier R minimum qui permet d'écrire un tenseur $\mathcal{T}$ sous la forme (5.3) est son *rang symétrique*. On le note $Rs(\mathcal{T})$. Le rang symétrique maximum est noté Rs_{max} et le rang symétrique générique Rs_{gen}.

On a les inégalités suivantes :

$$\begin{aligned} R(\mathcal{T}) &\leq Rs(\mathcal{T}) && (5.4)\\ Rs_{gen} &\leq R_{gen} && (5.5)\\ Rs_{max} &\leq R_{max} && (5.6) \end{aligned}$$

D'après [17], $Rs(\mathcal{T})$ serait en fait génériquement égal à $R(\mathcal{T})$, et cela a été démontré dans certains cas spécifiques.

5.2.1 Calcul du rang symétrique générique

Le rang symétrique générique est borné à droite et à gauche [16] :

$$\frac{1}{N}\binom{N+d-1}{d} \leq Rs_{gen} \leq \binom{N+d-2}{d-1} \tag{5.7}$$

Nous allons voir comment évaluer numériquement la valeur exacte du rang symétrique générique.

On définit la fonction Φ qui à un ensemble de R vecteurs $V_1, V_2, \ldots, V_R$ de taille N associe le tenseur symétrique $\sum_{r=1}^{R} V_r^{\circ d}$.

$$\begin{array}{rccc} \Phi : & \mathbb{C}^{(NR)} & \rightarrow & \mathcal{S}(d,N) \\ & \mathbf{V} & \rightarrow & \sum_{r=1}^{R} V_r^{\circ d} \end{array} . \tag{5.8}$$

$\mathbb{C}^{(NR)}$ est un $\mathbb{C}$-espace vectoriel de dimension NR et $\mathcal{S}(d,N)$ est un $\mathbb{C}$-espace vectoriel de dimension $D(d,N)$. Φ est $\mathbb{C}$-différentiable.
Si $R = Rs_{max}$, alors l'image de $\mathbb{C}^{(NR)}$ par Φ est $\mathcal{S}(d,N)$. Si $R = Rs_{gen}$, alors l'adhérence de l'image de $\mathbb{C}^{(NR)}$ par Φ est $\mathcal{S}(d,N)$. Si le rang du Jacobien de Φ est égal à la dimension de $\mathcal{S}(d,N)$, c'est-à-dire $D(d,N)$, alors l'image de $\mathbb{C}^{(NR)}$ par Φ est dense dans $\mathcal{S}(d,N)$ et $R = Rs_{gen}$.

Le Jacobien de $\Phi : V_1, V_2, \ldots, V_R \rightarrow \mathcal{C} = \sum_{r=1}^{R} V_r^{\circ d}$ peut être évalué de la manière suivante.
Notons I_k le vecteur de longueur N nul partout sauf à la position k, où il vaut 1 : $I_k = (\mathbf{I_N})_k$.
Le tenseur des dérivées partielles des éléments de $\mathcal{C}$ par rapport à v_{nr} s'écrit :

$$\frac{\partial \mathcal{C}}{\partial v_{nr}} = I_n \circ V_r \circ V_r \circ \ldots \circ V_r + V_r \circ I_n \circ V_r \circ \ldots \circ V_r + \ldots + V_r \circ V_r \circ \ldots \circ V_r \circ I_n \tag{5.9}$$

Ce tenseur ne dépend que de V_r et pas des vecteurs $V_k, k \neq r$. Notons $\gamma(V_r, n)$ le vecteur ligne contenant les dérivées de tous les éléments de $\mathcal{C}$ par rapport à v_{nr}.

$$\gamma(V_r, n) = \left(vec\left(I_n \circ V_r \circ V_r \circ \ldots \circ V_r + V_r \circ I_n \circ V_r \circ \ldots \circ V_r + \ldots + V_r \circ V_r \circ \ldots \circ V_r \circ I_n\right)\right)^T \tag{5.10}$$

Le Jacobien de Φ est la matrice $\mathbf{J}$ de taille $RN \times N^d$, définie par :

$$\mathbf{J} = \left[\gamma(V_1,1)^T, \gamma(V_1,2)^T, \ldots, \gamma(V_1,N)^T, \gamma(V_2,1)^T, \ldots, \gamma(V_R,N)^T\right]^T . \tag{5.11}$$

Le théorème suivant nous donne alors le moyen de trouver le rang générique.

Théorème 7 *Lasker-Wakeford*
Un tenseur symétrique générique d'ordre d et de taille N se décompose en une somme minimale de R tenseurs s'écrivant sous la forme (5.2) si et seulement s'il existe R vecteurs $V_1, V_2, \ldots, V_R$, *tels que le rang de* $M = \left[\gamma(V_1,1)^T, \gamma(V_1,2)^T, \ldots, \gamma(V_1,N)^T, \ldots, \gamma(V_2,1)^T, \ldots, \gamma(V_R,N)^T\right]^T$ *est égal à* $D(d,N)$.

Un algorithme itératif pour trouver le rang générique a été proposé dans [16] :

1. $r = 1$, $M = [\,]$.
2. On choisit un vecteur aléatoire V_r.
3. On construit $\gamma(V_r, n)$ pour $n \in [1 : N]$ et on calcule le rang de la matrice $\mathbf{M} := [\mathbf{M}^T, \gamma(V_r,1)^T, \gamma(V_r,2)^T, \ldots, \gamma(V_r,N)^T]^T$.
4. Si le rang de $\mathbf{M}$ est $D(d,N)$, on s'arrête et le rang générique est $Rs_{gen} = r$, sinon, $r = r+1$ et on retourne en 2.

N	2	3	4	5	6	7	8
$D(4,N)$	5	15	35	70	128	210	330
Rs_{gen}	3	6	10	15	21	30	42

TAB. 5.1 – Rang symétrique générique dans le cas d'un tenseur d'ordre 4

Récemment, il a été montré dans [15] que le rang générique est en fait toujours égal à sa borne inférieure $\lceil D(d,N)/N \rceil$, sauf dans quelques cas particuliers qui y sont énumérés.

5.2.2 Unicité de la décomposition

Il a été établi dans [14] que la décomposition est unique si le nombre de paramètres dans la décomposition, c'est-à-dire NR est strictement inférieur au nombre d'éléments indépendants dans le tenseur, c'est-à-dire $D(d,N)$. Si ces deux nombres sont égaux, alors il existe un nombre fini de décompositions possibles. Le rang maximal pour qu'il y ait unicité générique de la décomposition est donc donné par $\lceil D(d,N)/N - 1 \rceil$, où $\lceil\ \rceil$ désigne la partie entière supérieure.
Le tableau 5.2 donne le rang maximal pour qu'il y ait unicité de la décomposition dans le cas d'un tenseur symétrique d'ordre 4.

N	2	3	4	5	6	7	8	9
R_{uni}	2	4	8	13	20	29	41	54

TAB. 5.2 – Rang maximal permettant d'avoir unicite de la décomposition (cas des tenseurs symétriques)

5.3 Tenseurs d'ordre quatre à symétrie complexe

5.3.1 Définition

Nous nous intéressons dans ce paragraphe aux tenseurs $\mathcal{T}$ d'ordre 4 et de taille N possédant les symétries suivantes :

$$\mathcal{T}_{ijkl} = \mathcal{T}_{kjil} = \mathcal{T}_{ilkj} = \mathcal{T}^*_{jilk} \tag{5.12}$$

En particulier, le tenseur des cumulants défini par

$$\mathcal{C}_{ijkl} = Cum(\mathbf{x}_i, \mathbf{x}^*_j, \mathbf{x}_k, \mathbf{x}^*_l) \tag{5.13}$$

possède cette symétrie complexe.
Nous noterons $\mathcal{S}_c(N)$ l'ensemble des tenseurs d'ordre 4 et de taille N à symétrie complexe.
Le nombre d'éléments réels indépendants d'un tenseur $\mathcal{T} \in \mathcal{S}_c(N)$ est (annexe B) :

$$D_c(N) = 6\binom{N}{4} + 4\binom{N}{1}\binom{N-1}{2} + 3\binom{N}{2} + 2\binom{N}{1}\binom{N-1}{1} + \binom{N}{1} \tag{5.14}$$

$\mathcal{S}_c(N)$ est l'ensemble des tenseurs $\mathcal{T}$ s'écrivant sous la forme :

$$\mathcal{T} = \sum_{r=1}^{R} \epsilon_r V_r \circ V_r^* \circ V_r \circ V_r^*, \tag{5.15}$$

où $\epsilon_r = \pm 1$. Pour montrer ce résultat, on peut procéder comme dans le cas des tenseurs symétriques. L'ensemble $\mathcal{S}_{c1}(N)$ des tenseurs s'écrivant sous la forme (5.15) est un sous-espace vectoriel de $\mathcal{S}_c(N)$. Choisissons aléatoirement selon une loi continue des vecteurs V_r de taille N, et construisons avec les vecteurs $U_r = vec(\mathcal{T}_r) = vec(V_r \circ V_r^* \circ V_r \circ V_r^*)$. Posons ensuite $W_r = [Re(U_r)^T, Im(U_r)^T]^T$.
Chacun des vecteurs W_r, de taille $2N^4$, représente un élément de $\mathcal{S}_{c1}(N)$. Les vecteurs W_r sont indépendants tant que la famille $(\mathcal{T}_r)$ est libre. Par calcul numérique, le rang de la matrice $[W_1, W_2, \dots W_k]$ est égal à $D_c(N)$ si $k \geq D_c(N)$, donc la dimension de $\mathcal{S}_{c1}(N)$ est supérieure ou égale à $D_c(N) = dim_{\mathbb{R}}(\mathcal{S}_c(N))$. Comme $\mathcal{S}_{c1}(N)$ est inclus dans $\mathcal{S}_c(N))$, on en déduit que ces deux ensembles sont égaux.

5.3.2 Calcul du rang générique

On définit la fonction Φ qui à un ensemble de R vecteurs $V_1, V_2, \dots, V_R$ de taille N associe le tenseur à symétrie complexe $\sum_{r=1}^{R} V_r \circ V_r^* \circ V_r \circ V_r^*$.

$$\begin{array}{rccc} \Phi : & \mathbb{C}^{(NR)} & \rightarrow & \mathcal{S}_c(N) \\ & \mathbf{V} & \rightarrow & \sum_{r=1}^{R} \epsilon_r V_r \circ V_r^* \circ V_r \circ V_r^* \end{array} . \tag{5.16}$$

$\mathbb{C}^{(NR)}$ est un $\mathbb{R}$-espace vectoriel de dimension $2NR$ et $\mathcal{S}(N)$ est un $\mathbb{R}$-espace vectoriel de dimension $D_c(N)$. Φ est $\mathbb{R}$-différentiable. Le signe ϵ_r est nécessaire pour que $\mathcal{S}_c(N)$ soit un espace vectoriel, mais il n'intervient pas dans le calcul du rang générique.
Si le rang du Jacobien de Φ est égal à la dimension de $\mathcal{S}_c(N)$, c'est-à-dire $D_c(N)$, alors l'image de $\mathbb{C}^{(NR)}$ par Φ est dense dans $\mathcal{S}_c(N)$ et $R = Rs_{gen}$.

Soit $\mathcal{C}$ un élément de $\Phi(\mathbb{C}^{(NR)})$.

$$\mathcal{C} = \sum_{r=1}^{R} V_r \circ V_r^* \circ V_r \circ V_r^* \tag{5.17}$$

On pose $V_r = A_r + iB_r$, avec $A_r, B_r \in \mathbb{R}^N$ pour tout r dans $[1 : R]$.
Le tenseur des dérivées partielles de $\mathcal{C}$ par rapport à la partie réelle a_{nr} de v_{nr} s'écrit :

$$\begin{aligned} \frac{\partial \mathcal{C}}{\partial a_{nr}} &= I_n \circ V_r^* \circ V_r \circ V_r^* + V_r \circ I_n \circ V_r \circ V_r^* \\ &+ V_r \circ V_r^* \circ I_n \circ V_r^* + V_r \circ V_r^* \circ V_r \circ I_n, \end{aligned} \tag{5.18}$$

et le tenseur des dérivées partielles de $\mathcal{C}$ par rapport à la partie imaginaire b_{nr} de v_{nr} s'écrit :

$$\begin{aligned} \frac{\partial \mathcal{C}}{\partial b_{nr}} &= iI_n \circ V_r^* \circ V_r \circ V_r^* + V_r \circ -iI_n \circ V_r \circ V_r^* \\ &+ V_r \circ V_r^* \circ iI_n \circ V_r^* + V_r \circ V_r^* \circ V_r \circ -iI_n \end{aligned} \tag{5.19}$$

Ces deux tenseurs ne dépendent que de V_r et pas des vecteurs $V_k, k \neq r$.
On pose :

$$\begin{aligned} \gamma_r(V_r, n) &= I_n \circ V_r^* \circ V_r \circ V_r^* + V_r \circ I_n \circ V_r \circ V_r^* \\ &+ V_r \circ V_r^* \circ I_n \circ V_r^* + V_r \circ V_r^* \circ V_r \circ I_n \end{aligned} \tag{5.20}$$

$$\begin{aligned} \gamma_i(V_r, n) &= iI_n \circ V_r^* \circ V_r \circ V_r^* + V_r \circ -iI_n \circ V_r \circ V_r^* \\ &+ V_r \circ V_r^* \circ iI_n \circ V_r^* + V_r \circ V_r^* \circ V_r \circ -iI_n \end{aligned} \tag{5.21}$$

Soit $\mathbf{J}_1(V_r, n)$ la matrice de taille $2 \times 2N^4$ définie par

$$\mathbf{J}_1(V_r, n) = \begin{bmatrix} Re(\gamma_r(V_r, n)) & Im(\gamma_r(V_r, n)) \\ Re(\gamma_i(V_i, n)) & Im(\gamma_i(V_r, n)) \end{bmatrix} \tag{5.22}$$

Le Jacobien de Φ est la matrice $\mathbf{J}$ de taille $2NR \times 2N^4$ définie par

$$\mathbf{J} = [\mathbf{J}_1(V_1, 1)^T, \mathbf{J}_1(V_1, 2)^T, \ldots, \mathbf{J}_1(V_1, N)^T, \mathbf{J}_1(V_2, 1)^T, \ldots, \mathbf{J}_1(V_2, N)^T, \ldots, \mathbf{J}_1(V_R, N)^T]^T \tag{5.23}$$

On peut alors établir un équivalent du théorème de Lasker-Wakeford pour les tenseurs à symétrie complexe.

Théorème 8 *Rang générique symétrique dans le cas d'un tenseur d'ordre 4 à symétrie complexe*
Un tenseur d'ordre 4 et de taille N à symétrie complexe se décompose en une somme minimale de R tenseurs s'écrivant sous la forme $(V \circ V^* \circ V \circ V^*)$ si et seulement s'il existe R vecteurs $V_1, V_2, \ldots, V_R$, tel que le rang de $\mathbf{M} = [\mathbf{J}_1(V_1, 1)^T, \mathbf{J}_1(V_1, 2)^T, \ldots, \mathbf{J}_1(V_1, N)^T, \mathbf{J}_1(V_2, 1)^T, \ldots, \mathbf{J}_1(V_2, N)^T, \ldots, \mathbf{J}_1(V_R$ est égal à $D_c(N)$.

Un algorithme peut alors être mis en œuvre pour déterminer le rang symétrique générique :

1. $r = 1$, $\mathbf{M} = [\,]$.

2. On choisit un vecteur aléatoire V_r.
3. On construit $\mathbf{J}_1(V_r, n)$ défini par l'équation (5.22) pour $n \in [1 : N]$.
4. On calcule le rang de la matrice $\mathbf{M} := [\mathbf{M}^T, \mathbf{J}_1(V_1, 1)^T, \mathbf{J}_1(V_1, 2)^T, \ldots, \mathbf{J}_1(V_1, N)^T, \mathbf{J}_1(V_2, 1)^T, \ldots, \mathbf{J}_1(V_2, N)^T, \ldots, \mathbf{J}_1(V_r, N)^T]^T$.
5. Si le rang de $\mathbf{M}$ est $D_c(d, N)$, on s'arrête et le rang générique est $Rc_{gen} = r$, sinon, $r = r+1$ et on retourne en 2.

Le rang symétrique générique pour les tenseurs d'ordre 4 à symétrie complexe de taille $N = 2$ à $N = 8$ est reporté dans le tableau 5.3.

N	2	3	4	5	6	7	8
$D_c(N)$	9	36	100	225	441	784	1296
Rc_{gen}	4	9	16	25	41	61	87

TAB. 5.3 – Rang symétrique générique dans le cas d'un tenseur d'ordre 4 à symétrie complexe

5.3.3 Unicité de la décomposition

Nous avons vu plus haut que la décomposition est unique si le nombre de paramètres dans la décomposition, c'est-à-dire $2NR$ est strictement inférieur au nombre d'éléments indépendants dans le tenseur, c'est-à-dire $D_c(d, N)$. Le rang maximal est donc donné par $\lceil D_c(d, N)/(2N) - 1 \rceil$. Le tableau 5.4 donne le rang maximal pour qu'il y ait unicité de la décomposition dans le cas d'un tenseur d'ordre 4 à symétrie complexe.

N	2	3	4	5	6	7	8	9
R_{uni}	2	5	12	22	36	55	80	112

TAB. 5.4 – Rang maximal permettant d'avoir unicite de la décomposition (cas des tenseurs à symétrie complexe)

5.4 Cas général

5.4.1 Calcul du rang générique

Pour déterminer le rang générique dans le cas général, nous allons procéder de la même manière que dans les cas des tenseurs symétriques et des tenseurs à symétrie complexe.
On note $\mathcal{A}(N_1, N_2, \ldots, N_d)$ l'ensemble des tenseurs d'ordre d et de taille $(N_1 \times N_2 \times \ldots \times N_d)$.
Un tenseur $\mathcal{T}$ appartenant à l'ensemble $\mathcal{A}(N_1, N_2, \ldots, N_d)$ peut s'écrire sous la forme :

$$\mathcal{T} = \sum_{r=1}^{R} V_r^{(1)} \circ V_r^{(2)} \circ \ldots \circ V_r^{(d)} \tag{5.24}$$

avec $V_r^{(p)} \in \mathbb{C}_p^N, p \in [1:d]$.
Soit *cat* la fonction de $\mathbb{C}^k$ dans $\mathbb{C}^{k+1}$ définie par $cat(V) = [V^T, 1]^T$, pour tout $V \in \mathbb{C}^k$.
$\mathcal{T}$ peut également s'écrire sous la forme :

$$\sum_{r=1}^{R} V_r^{(1)} \circ cat(V_r^{(2)}) \circ \ldots \circ cat(V_r^{(d)}) \tag{5.25}$$

avec $V_r^{(1)} \in \mathbb{C}^{N_1}$ et $V_r^{(p)} \in \mathbb{C}^{N_p - 1}, p \in [2:d]$.

Exemple : la matrice $\mathbf{A}$ de taille 2×2 définie par

$$\mathbf{A} = \begin{bmatrix} ac & ad \\ bc & bd \end{bmatrix} = \begin{bmatrix} a \\ b \end{bmatrix} . \begin{bmatrix} c & d \end{bmatrix} \tag{5.26}$$

peut aussi s'écrire :

$$\mathbf{A} = \begin{bmatrix} ad \\ bd \end{bmatrix} . \begin{bmatrix} c/d & 1 \end{bmatrix} \tag{5.27}$$

On définit la fonction Φ qui à un ensemble de Rd vecteurs $(V_r^{(p)})$ avec $(V_r^{(1)}) \in \mathbb{C}^{N_1}$, et $(V_r^{(p)}) \in \mathbb{C}^{N_p - 1}$ $p \in [2:d]$, $r \in [1:R]$, associe le tenseur $\sum_{r=1}^{R} V_r^{(1)} \circ cat(V_r^{(2)}) \circ \ldots \circ cat(V_r^{(d)}) \in \mathcal{A}(N_1, N_2, \ldots, N_d)$.

$$\begin{array}{rccc} \Phi : & \mathbb{C}^{(RN_1)} \times \mathbb{C}^{(R(N_2-1))} \times \ldots, \mathbb{C}^{(R(N_d-1))} & \rightarrow & \mathcal{A}(N_1, N_2, \ldots, N_d) \\ & \left(\mathbf{V}^{(1)}, \mathbf{V}^{(2)}, \ldots, \mathbf{V}^{(d)}\right) & \rightarrow & \sum_{r=1}^{R} V_r^{(1)} \circ cat(V_r^{(2)}) \circ \ldots \circ cat(V_r^{(d)}) \end{array} . \tag{5.28}$$

$\mathbb{C}^{(RN_1)} \times \mathbb{C}^{(R(N_2-1))} \times \ldots \times \mathbb{C}^{(R(N_d-1))}$ est un $\mathbb{C}$-espace vectoriel de dimension $R(1 + \sum_{k=1}^{d}(N_k - 1))$ et $\mathcal{A}(N_1, N_2, \ldots, N_d)$ est un $\mathbb{C}$-espace vectoriel de dimension $\prod_{k=1}^{d} N_k$. Φ est $\mathbb{C}$-différentiable.
Pour plus de commodité, nous allons noter dans la suite $U_r^{(1)} = V_r^{(1)}$ et $U_r^{(p)} = cat(V_r^{(p)})$ si $p \neq 1$.

Soit $\mathcal{T}$ un tenseur de taille $N_1 \times N_2 \times \ldots \times N_d$.

$$\mathcal{T} = \sum_{r=1}^{R} U_r^{(1)} \circ U_r^{(2)} \circ \ldots \circ U_r^{(d)} \tag{5.29}$$

Le tenseur des dérivées partielles des éléments de $\mathcal{T}$ par rapport à $u_{nr}^{(p)}$ s'écrit :

$$\frac{\partial \mathcal{T}}{\partial u_{nr}^{(p)}} = U_r^{(1)} \circ U_r^{(2)} \circ \ldots \circ U_r^{(p-1)} \circ I_n \circ U_r^{(p+1)} \circ U_r^{(d)} \tag{5.30}$$

Ce tenseur ne dépend que des vecteurs $U_r^{(q)}, q \neq p$, et pas des vecteurs $U_k^{(s)}, k \neq r, s \in [1:d]$.
Notons $\gamma(U_r^{(p)}, n)$ le vecteur ligne de longueur $\prod_{k=1}^{d} N_k$ contenant les dérivées de tous les éléments de $\mathcal{C}$ par rapport à $u_{nr}^{(p)}$, $p \in [1:d], n \in [1:N_p]$:

$$\gamma(U_r^{(p)}, n) = \left(vec\left(U_r^{(1)} \circ U_r^{(2)} \circ \ldots \circ U_r^{(p-1)} \circ I_n \circ U_r^{(p+1)} \circ U_r^{(d)} \right) \right)^T . \tag{5.31}$$

Notons $\mathbf{J}_1(\mathbf{U}^{(p)})$ la matrice de taille $(N_p \times \prod_{k=1}^{d} N_k)$ définie par

$$\mathbf{J}_1(U_r^{(p)}) = [\gamma(U_r^{(p)}, 1)^T, \gamma(U_r^{(p)}, 2)^T, \ldots, \gamma(U_r^{(p)}, N_p)^T]^T \tag{5.32}$$

et notons $\mathbf{J}_2(U_r^{(1)}, U_r^{(2)}, \ldots, U_r^{(d)})$ la matrice de taille $\left(\sum_{k=1}^{d} N_k \times \prod_{k=1}^{d} N_k\right)$ définie par :

$$\mathbf{J}_2(U_r^{(1)}, U_r^{(2)}, \ldots, U_r^{(d)}) = \left[\mathbf{J}_1(U_r^{(1)})^T, \mathbf{J}_1(U_r^{(2)})^T, \ldots, \mathbf{J}_1(U_r^{(d)})^T\right]^T \tag{5.33}$$

Le jacobien de Φ est la matrice $\mathbf{J}$ de taille $\left(R\sum_{k=1}^{d} N_k \times \prod_{k=1}^{d} N_k\right)$, définie par :

$$\mathbf{J} = \begin{bmatrix} \mathbf{J}_2(U_1^{(1)}, U_1^{(2)}, \ldots, U_1^{(d)}) \\ \mathbf{J}_2(U_2^{(1)}, U_2^{(2)}, \ldots, U_2^{(d)}) \\ \vdots \\ \mathbf{J}_2(U_R^{(1)}, U_R^{(2)}, \ldots, U_R^{(d)}) \end{bmatrix} \tag{5.34}$$

On peut alors établir un équivalent du théorème de Lasker-Wakeford pour les tenseurs quelconques.

Théorème 9 *Rang générique*
Un tenseur d'ordre d et de taille $N_1 \times N_2 \times \ldots \times N_d$ se décompose en une somme minimale de R tenseurs s'écrivant sous la forme $(V^{(1)} \circ cat(V^{(2)}) \circ \ldots \circ cat(V^{(d)}))$ si et seulement s'il existe R p-uplets de vecteurs $(V_r^{(1)}, V_r^{(2)}, \ldots, V_r^{(d)}) \in (\mathbb{C}^{N_1} \times \mathbb{C}^{N_2-1} \times \ldots \times \mathbb{C}^{N_d-1})$, $r \in [1:R]$ tel que le rang de $\mathbf{M} = [(\mathbf{J}_2(U_1^{(1)}, U_1^{(2)}, \ldots, U_1^{(d)}))^T, (\mathbf{J}_2(U_2^{(1)}, U_2^{(2)}, \ldots, U_2^{(d)}))^T, \ldots, (\mathbf{J}_2(U_R^{(1)}, U_R^{(2)}, \ldots, U_R^{(d)}))^T]^T$ est égal à $\prod_{k=1}^{d} N_k$, avec $\mathbf{J}_2(U_r^{(1)}, U_r^{(2)}, \ldots, U_r^{(d)})$ défini par (5.33), $U_r^{(1)} = V_r^{(1)}$ pour tout $r \in [1:R]$ et $U_r^{(p)} = cat(V_r^{(p)})$ pour tout $r \in [1:R]$ et tout $p \in [2:d]$.

Un algorithme peut alors être mis en œuvre pour déterminer le rang générique :

1. $r = 1$, $\mathbf{M} = [\,]$.
2. On choisit d vecteurs aléatoires $V_r^{(1)}, V_r^{(2)}, \ldots, V_r^{(d)}$.
3. On construit $\mathbf{J}_2(U_r^{(1)}, U_r^{(2)}, \ldots, U_r^{(d)})$ défini par l'équation (5.33).
4. On calcule le rang de la matrice $\mathbf{M} := [\mathbf{M}^T, (\mathbf{J}_2(U_1^{(1)}, U_1^{(2)}, \ldots, U_1^{(d)}))^T, \ldots, (\mathbf{J}_2(U_r^{(1)}, U_r^{(2)}, \ldots, U_r^{(d)}))^T]^T$.
5. Si le rang de $\mathbf{M}$ est $\prod_{k=1}^{d} N_k$, on s'arrête et le rang générique est $R_{gen} = r$, sinon, $r = r+1$ et on retourne en 2.

Le rang générique pour les tenseurs d'ordre 4 de taille $N \times N \times N \times N$ pour $N = 2$ à 6 est reporté dans le tableau 5.5, le rang générique pour les tenseurs d'ordre 3 de taille $N \times 5 \times 3$, pour $N = 2$ à 12 est reporté dans le tableau 5.6.

N	2	3	4	5	6	7
R_{gen}	4	9	20	37	62	97

TAB. 5.5 – Rang générique dans le cas d'un tenseur quelconque de taille $N \times N \times N \times N$

N	2	3	4	5	6	7	8	9	10	11	12
R_{gen}	5	5	6	8	8	9	9	9	10	11	12

TAB. 5.6 – Rang générique dans le cas d'un tenseur quelconque de taille $N \times 5 \times 3$

5.4.2 Unicité de la décomposition

Nous avons vu que la décomposition est unique si le nombre de paramètres dans la décomposition est strictement inférieur au nombre d'éléments indépendants dans le tenseur, c'est-à-dire si $R(1 + \sum_{k=1}^{d}(N_k - 1))$ est strictement inférieur à $\prod_{k=1}^{d} N_k$. Le rang maximal est donc donné par $\lceil (\prod_{k=1}^{d} N_k)/(1 + \sum_{k=1}^{d}(N_k - 1)) - 1 \rceil$.
Le tableau 5.7 donne le rang maximal pour qu'il y ait unicité de la décomposition dans le cas d'un tenseur d'ordre 4 de taille $N \times N \times N \times N$.

N	2	3	4	5	6	7	8	9
R_{uni}	3	8	19	36	61	96	141	198

TAB. 5.7 – Rang maximal permettant d'avoir unicite de la décomposition dans le cas d'un tenseur d'ordre 4 de taille $N \times N \times N \times N$

5.5 Conclusion

Nous avons vu comment il était possible dans le cas général et dans deux cas particuliers d'évaluer le rang générique des tenseurs d'un ordre et d'une taille donnés à l'aide d'un algorithme itératif. Nous avons également vu qu'il était possible de calculer le rang maximum pour laquelle la décomposition d'un tenseur est essentiellement unique. Ce rang maximum pour avoir unicité de la décomposion dépend du nombre de paramètres indépendants dans cette décomposition. Les résultats présentés sont valables dans $\mathbb{C}$, il serait intéressant à l'avenir de pouvoir évaluer les rangs typiques des tenseurs dans $\mathbb{R}$.

Chapitre 6

Conclusion

6.1 Contributions

Les travaux présentés dans ce document s'inscrivent dans le cadre des méthodes algébriques pour la séparation aveugle de sources. Ces méthodes s'appuient sur la structure des données, qui peuvent s'écrire sous forme de tenseurs possédant une forme algébrique particulière. Nous nous sommes en particulier intéressé à une décomposition des tenseurs connue sous le nom de décomposition PARAFAC. Cette décomposition propose de décomposer un tenseur de rang R en une somme de R tenseurs de rang un. Il existe une borne appelée borne de Kruskal en dessous de laquelle la décomposition est toujours unique et un algorithme des moindres carrés alternés (ALS) permettant d'évaluer les paramètres de la décomposition. Cependant, nous savons que la décomposition peut être unique pour une valeur du rang plus grande que la borne de Kruskal, et l'algorithme ALS pose différents problèmes.

La première partie de cette thèse est consacrée à une nouvelle technique permettant d'obtenir les paramètres de la décomposition PARAFAC d'un tenseur d'ordre trois dans le cas où le rang du tenseur est inférieur à l'une des dimensions du tenseur et au produit des deux autres. Cette méthode utilise une diagonalisation simultanée de matrices, que l'on peut résoudre à l'aide de différents procédés qui sont détaillés. D'autre part, nous avons établi dans ce cas une borne sur le rang du tenseur en dessous de laquelle la décomposition est unique. Cette borne est généralement moins restrictive que la borne de Kruskal. En effet, elle dépend du produit des deux plus petites dimensions du tenseur tandis que la borne de Kruskal dépend de leur somme. D'autre part, nous avons montré que le rang du tenseur pouvait être évalué à l'aide d'une simple décomposition en valeurs singulières.

Nous avons proposé différentes applications de cette technique en séparation de sources. La première application proposée concerne la séparation aveugle de signaux DS-CDMA. En effet, les données reçues à chaque instant peuvent être vues comme des éléments d'un tenseur d'ordre trois. Ce tenseur s'écrit comme la somme de tenseurs de rang un, correspondant chacun à la contribution d'un seul utilisateur, et pouvant s'écrire comme le produit externe de trois vecteurs, le premier contenant les symboles d'information de cet utilisateur, le second contenant son code d'étalement, et le troisième contenant les coefficients du canal entre cet utilisateur et les antennes de réception.

Nous avons d'autre part proposé de combiner l'information obtenue par la structure des signaux et l'information apportée par le fait que les sources sont de module constant. La structure PARAFAC des signaux nous permet d'aboutir à un système de matrices à diagonaliser conjointement, la contrainte du module constant nous mène à un autre système, qui est assez similaire au premier. Nous avons proposé trois algorithmes permettant de diagonaliser ensemble ces deux systèmes de matrices.

La seconde application que nous avons proposée concerne l'analyse en composantes indépendantes dans le cas sous-déterminé. Nous avons proposé de résoudre ce problème à l'aide de deux techniques. La première solution développée s'appuie sur les statistiques d'ordre deux des données. Nous partons de l'hypothèse que les sources sont mutuellement indépendantes et individuellement corrélées. Nous pouvons chercher à décomposer le tenseur d'ordre trois contenant les matrices de covariance des observations pour différents retards. Ce tenseur se décompose en une somme de tenseur de rang un représentant chacun la contribution d'une seule source. La deuxième solution présentée ici s'appuie sur les statistiques d'ordre quatre des données, nous décomposons alors le tenseur d'ordre quatre contenant les cumulants d'ordre quatre des observations. Nous avons décrit deux algorithmes, le premier menant à une diagonalisation conjointe de matrices, le second à une zéro-diagonalisation de matrices. Dans les deux cas, nous avons établi le nombre maximum de sources en fonction du nombre de capteurs.

Le dernier chapitre de ce document est consacré au calcul du rang générique des tenseurs. En nous inspirant des travaux réalisés dans le cas des tenseurs symétriques, nous avons proposé deux algorithmes itératifs pour déterminer le rang générique dans $\mathbb{C}$ des tenseurs d'ordre quatre à symétrie complexe (dont font partie les tenseurs des cumulants d'ordre quatre) et des tenseurs d'ordre et de dimension quelconques ne possédant pas de symétrie particulière. Nous avons également calculé dans ces différents cas le rang maximum d'un tenseur au dessus duquel sa décomposition n'est plus unique.

6.2 Perspectives

Nous avons étudié le cas où les signaux peuvent se modéliser sous la forme d'un tenseur s'écrivant comme une somme de tenseurs de rang un contribuant chacun pour une source (ou un utilisateur). Dans certains cas, cette décomposition n'a plus de sens physique. En particulier, dans le cas des signaux CDMA, nous avons considéré le cas de codes « courts », c'est-à-dire qu'un code d'étalement était de même longueur qu'un symbole. Ce n'est pas le cas des codes « longs », qui peuvent être plus longs qu'un symbole. Le problème peut dans ce cas être modélisé à l'aide d'une décomposition canonique généralisées : on ne décompose plus le tenseur des données sous la forme d'une somme de tenseurs de rang un, mais de tenseurs de rang 2, 3 etc. selon le problème. Une telle décomposition peut également être considérée par exemple lorsque les signaux se propagent selon différents multitrajets arrivant avec des retards tels que des interférences entre symboles surviennent (nous nous étions limité dans notre travail à l'interférences entre chips). Les travaux de [20,35,36] présentent de telles approches, ainsi que des applications en communications numériques.

Dans le cadre de l'analyse en composantes indépendantes, nous avons proposé une solution

s'appuyant sur les statistiques d'ordre deux des données et une solution s'appuyant sur leurs statistiques d'ordre quatre. Nous pouvons également imaginer de s'intéresser aux statistiques d'ordre six ou encore de combiner plusieurs ordres.

Les résultats concernant le rang générique que nous avons développés au chapitre 5 sont valables dans $\mathbb{C}$. Il serait intéressant de trouver quels sont les rangs typiques pour un tenseur dans $\mathbb{R}$. D'autre part, excepté dans quelques cas particulier, on ne connaît pas le rang maximal pouvant être atteint pour un ordre et une dimension donnée. Enfin, nous avons montré (au chapitre 2) qu'il était possible dans certains cas de déterminer le rang d'un tenseur donné à l'aide d'un décomposition en valeurs singulières, mais il n'existe pas de méthode dans le cas général. De larges perspectives s'ouvrent donc dans ce domaine.

Annexe A

Nombre d'éléments indépendants dans un tenseur d'ordre 4 symétrique

Un tenseur $\mathcal{C}$ symétrique vérifie :

$$\mathcal{C}_{ijkl} = \mathcal{C}_{\sigma(ijkl)} \tag{A.1}$$

quelque soit la permutation σ.

Dans le tableau suivant, on récapitule le nombre d'éléments distincts pour des indices donnés.

indices	nombre d'éléments
$ijkl, i \neq j \neq k \neq l$	$\binom{N}{4}$
$iikl, i \neq k \neq l$	$\binom{N}{1}\binom{N-1}{2}$
$iikk, i \neq k$	$\binom{N}{2}$
$iiil, i \neq l$	$\binom{N}{1}\binom{N-1}{1}$
$iiii$	$\binom{N}{1}$

Le nombre d'éléments indépendants est donc :

$$D_s(N,4) = \binom{N}{4} + \binom{N}{1}\binom{N-1}{2} + \binom{N}{2} + \binom{N}{1}\binom{N-1}{1} + \binom{N}{1} \tag{A.2}$$

$$= \binom{N+3}{4} \tag{A.3}$$

Annexe B

Nombre d'éléments indépendants dans un tenseur d'ordre 4 à symétrie complexe

Il y a 24 permutations possibles des indices d'un tenseur d'ordre 4. Un tenseur $\mathcal{C}$ à symétrie complexe vérifie :

$$\mathcal{C}_{ijkl} = \mathcal{C}_{kjil} = \mathcal{C}_{ilkj} = \mathcal{C}_{klij} = \mathcal{C}^*_{jilk} = \mathcal{C}^*_{jkli} = \mathcal{C}^*_{lijk} = \mathcal{C}^*_{lkji} \tag{B.1}$$

$$\mathcal{C}_{jikl} = \mathcal{C}_{jlki} = \mathcal{C}_{kijl} = \mathcal{C}_{klji} = \mathcal{C}^*_{ijlk} = \mathcal{C}^*_{ljik} = \mathcal{C}^*_{iklj} = \mathcal{C}^*_{lkij} \tag{B.2}$$

$$\mathcal{C}_{ikjl} = \mathcal{C}_{jkil} = \mathcal{C}_{iljk} = \mathcal{C}_{jlik} = \mathcal{C}^*_{kilj} = \mathcal{C}^*_{kjli} = \mathcal{C}^*_{likj} = \mathcal{C}^*_{ljki} \tag{B.3}$$

$$\tag{B.4}$$

Dans le tableau figurant sur la page suivante, on récapitule le nombre d'éléments distincts pour des indices donnés.

Le nombre d'éléments réels indépendants est donc :

$$D_c(N) = 6\binom{N}{4} + 4\binom{N}{1}\binom{N-1}{2} + 3\binom{N}{2} + 2\binom{N}{1}\binom{N-1}{1} + \binom{N}{1} \tag{B.5}$$

indices	nombre d'éléments	réel/complexe
$ijkl, i \neq j \neq k \neq l$	$\binom{N}{4}$	complexe
$jikl, i \neq j \neq k \neq l$	$\binom{N}{4}$	complexe
$ikjl, i \neq j \neq k \neq l$	$\binom{N}{4}$	complexe
$iikl, i \neq k \neq l$	$\binom{N}{1}\binom{N-1}{2}$	complexe
$ijil, i \neq j \neq l$	$\binom{N}{1}\binom{N-1}{2}$	complexe
$iikk, i \neq k$	$\binom{N}{2}$	réel
$ijij, i \neq j$	$\binom{N}{2}$	réel
$iiil, i \neq l$	$\binom{N}{1}\binom{N-1}{1}$	complexe
$iiii$	$\binom{N}{1}$	réel

Annexe C

Calcul des dérivées partielles des éléments d'un tenseur à symétrie complexe

On considère le tenseur $\mathcal{C}$ de taille N à symétrie complexe :

$$\mathcal{C} = V \circ V^* \circ V \circ V^*. \tag{C.1}$$

avec $V = A + iB$, $(A, B) \in \mathbb{R}^N$.
Nous récapitulons les valeurs des dérivées partielles des éléments de $\mathcal{C}$ par rapport à a_k et b_k.
Cas 0 : $(p \neq k, q \neq k, r \neq k, s \neq k)$

$$\frac{\partial C_{pqrs}}{\partial a_k} = 0 \tag{C.2}$$

$$\frac{\partial C_{pqrs}}{\partial b_k} = 0 \tag{C.3}$$

Cas 1 $(p = k, q \neq k, r \neq k, s \neq k)$

$$\frac{\partial C_{pqrs}}{\partial a_k} = (A(q) - i\ B(q))\ (A(r) + i\ B(r))\ (A(s) - i\ B(s)); \tag{C.4}$$

$$\frac{\partial C_{pqrs}}{\partial b_k} = i\ (A(q) - i\ B(q))\ (A(r) + i\ B(r))\ (A(s) - i\ B(s)); \tag{C.5}$$

Cas 2 $(r = k, q \neq k, p \neq k, s \neq k)$

$$\frac{\partial C_{pqrs}}{\partial a_k} = (A(q) - i\ B(q))\ (A(p) + i\ B(p))\ (A(s) - i\ B(s)); \tag{C.6}$$

$$\frac{\partial C_{pqrs}}{\partial b_k} = i\ (A(q) - i\ B(q))\ (A(p) + i\ B(p))\ (A(s) - i\ B(s)); \tag{C.7}$$

Cas 3 $(q = k, p \neq k, r \neq k, s \neq k)$

$$\frac{\partial C_{pqrs}}{\partial a_k} = (A(p) + i\ B(p))\ (A(r) + i\ B(r))\ (A(s) - i\ B(s)); \tag{C.8}$$

$$\frac{\partial C_{pqrs}}{\partial b_k} = -i\ (A(p) + i\ B(p))\ (A(r) + i\ B(r))\ (A(s) - i\ B(s)); \tag{C.9}$$

Cas 4 ($s = k, p \neq k, r \neq k, q \neq k$)

$$\frac{\partial C_{pqrs}}{\partial a_k} = (A(p) + i\ B(p))\ (A(r) + i\ B(r))\ (A(q) - i\ B(q)); \tag{C.10}$$

$$\frac{\partial C_{pqrs}}{\partial b_k} = -i\ (A(p) + i\ B(p))\ (A(r) + i\ B(r))\ (A(q) - i\ B(q)); \tag{C.11}$$

Cas 5 ($p = k, q = k, r \neq k, s \neq k$)

$$\frac{\partial C_{pqrs}}{\partial a_k} = 2\ A(k)\ (A(r) + i\ B(r))\ (A(s) - i\ B(s)); \tag{C.12}$$

$$\frac{\partial C_{pqrs}}{\partial b_k} = 2\ B(k)\ (A(r) + i\ B(r))\ (A(s) - i\ B(s)); \tag{C.13}$$

Cas 6 ($r = k, s = k, p \neq k, q \neq k$)

$$\frac{\partial C_{pqrs}}{\partial a_k} = 2\ A(k)\ (A(p) + i\ B(p))\ (A(q) - i\ B(q)); \tag{C.14}$$

$$\frac{\partial C_{pqrs}}{\partial b_k} = 2\ B(k)\ (A(p) + i\ B(p))\ (A(q) - i\ B(q)); \tag{C.15}$$

Cas 7 ($q = k, r = k, p \neq k, s \neq k$)

$$\frac{\partial C_{pqrs}}{\partial a_k} = 2\ A(k)\ (A(p) + i\ B(p))\ (A(s) - i\ B(s)); \tag{C.16}$$

$$\frac{\partial C_{pqrs}}{\partial b_k} = 2\ B(k)\ (A(p) + i\ B(p))\ (A(s) - i\ B(s)); \tag{C.17}$$

Cas 8 ($p = k, s = k, q \neq k, r \neq k$)

$$\frac{\partial C_{pqrs}}{\partial a_k} = 2\ A(k)\ (A(r) + i\ B(r))\ (A(q) - i\ B(q)); \tag{C.18}$$

$$\frac{\partial C_{pqrs}}{\partial b_k} = 2\ B(k)\ (A(r) + i\ B(r))\ (A(q) - i\ B(q)); \tag{C.19}$$

Cas 9 ($p = k, r = k, q \neq k, s \neq k$)

$$\frac{\partial C_{pqrs}}{\partial a_k} = 2\ (A(k) + i\ B(k))\ (A(q) - i\ B(q))\ (A(s) - i\ B(s)); \tag{C.20}$$

$$\frac{\partial C_{pqrs}}{\partial b_k} = 2\ i\ (A(k) + i\ B(k))\ (A(q) - i\ B(q))\ (A(s) - i\ B(s)); \tag{C.21}$$

Cas 10 ($q = k, s = k, p \neq k, r \neq k$)

$$\frac{\partial C_{pqrs}}{\partial a_k} = 2\ (A(k) - i\ B(k))\ (A(p) + i\ B(p))\ (A(r) + i\ B(r)); \tag{C.22}$$

$$\frac{\partial C_{pqrs}}{\partial b_k} = -2\ i\ (A(k) - i\ B(k))\ (A(p) + i\ B(p))\ (A(r) + i\ B(r)); \tag{C.23}$$

Cas 11 ($p = k, q = k, r = k, s \neq k$)

$$\frac{\partial C_{pqrs}}{\partial a_k} = (2\ A(k)\ (A(k) + i\ B(k)) + A(k)\ A(k) \tag{C.24}$$

$$+B(k)\ B(k))\ (A(s) - i\ B(s)); \tag{C.25}$$

$$\frac{\partial C_{pqrs}}{\partial b_k} = (2\ i\ (A(k)\ A(k) + B(k)\ B(k)) \tag{C.26}$$

$$-i\ (A(k) + i\ B(k))^2)\ (A(s) - i\ B(s)); \tag{C.27}$$

Cas 12 ($q \neq k, p = k, r = k, s = k$)

$$\frac{\partial C_{pqrs}}{\partial a_k} = (2\ A(k)\ (A(k) + i\ B(k)) + A(k)\ A(k) \tag{C.28}$$

$$+B(k)\ B(k))\ (A(q) - i\ B(q)); \tag{C.29}$$

$$\frac{\partial C_{pqrs}}{\partial b_k} = (2\ i\ (A(k)\ A(k) + B(k)\ B(k)) \tag{C.30}$$

$$-i\ (A(k) + i\ B(k))^2)\ (A(q) - i\ B(q)); \tag{C.31}$$

Cas 13 ($p \neq k, q = k, r = k, s = k$)

$$\frac{\partial C_{pqrs}}{\partial a_k} = (2\ A(k)\ (A(k) - i\ B(k)) + A(k)\ A(k) \tag{C.32}$$

$$+B(k)\ B(k))\ (A(p) + i\ B(p)); \tag{C.33}$$

$$\frac{\partial C_{pqrs}}{\partial b_k} = (2\ B(k)\ (A(k) - i\ B(k)) \tag{C.34}$$

$$-i\ (A(k)\ A(k) + B(k)\ B(k)))\ (A(p) + i\ B(p)); \tag{C.35}$$

Cas 14 ($r \neq k, p = k, q = k, s = k$)

$$\frac{\partial C_{pqrs}}{\partial a_k} = (2\ A(k)\ (A(k) - i\ B(k)) + A(k)\ A(k) \tag{C.36}$$

$$+B(k)\ B(k))\ (A(r) + i\ B(r)); \tag{C.37}$$

$$\frac{\partial C_{pqrs}}{\partial b_k} = (2\ B(k)\ (A(k) - i\ B(k)) \tag{C.38}$$

$$-i\ (A(k)\ A(k) + B(k)\ B(k)))\ (A(r) + i\ B(r)); \tag{C.39}$$

Cas 15 ($p = k, q = k, r = k, s = k$)

$$\partial C_{pqrs}/\partial a_k = 4\ A(k)\ (A(k)\ A(k) + B(k)\ B(k)); \tag{C.40}$$

$$\partial C_{pqrs}/\partial b_k = 4\ B(k)\ (A(k)\ A(k) + B(k)\ B(k)); \tag{C.41}$$

Bibliographie

[1] L. Albera, A. Ferréol, P. Comon, P. Chevalier, "Sixth order blind identification of underdetermined mixtures (BIRTH) of sources", *ICA 03, Fourth International Symposium on Independent Component Analysis and Blind Signal Separation*, Nara, Japan, avr. 2003.

[2] L. Albera, A. Ferréol, P. Comon, P. Chevalier, "Blind identification of overcomplete mixture of sources", *Lin. Alg. Appl.*, vol. 391, pp. 1–30, nov. 2004.

[3] A. Belouchrani, K. Abed-Meraim, J.-F. Cardoso, E. Moulines, "A blind source separation technique using second order statistics", *IEEE Trans. on Signal Processing*, vol. 45 (2), pp. 434–444, fev. 1997.

[4] P. Bofill, M. Zibulevsky, "Underdetermined blind source separation using sparse representations", *Signal Processing*, vol. 81, pp. 2353–2362, 2001.

[5] R. Bro, "PARAFAC. tutorial and applications", *Chemometrics and Intelligent Laboratory Systems*, vol. 38 (2), pp. 149–171, 1997.

[6] J.-F. Cardoso, "Super-symmetric decomposition of the fourth-order cumulant tensor. Blind identification of more sources than sensors", *Proc. IEEE ICASSP*, Toronto, Canada, 1991, vol. 5, pp. 3109–3112.

[7] J.-F. Cardoso, A. Souloumiac, "Blind beamforming for non-gaussian signals", *IEE Proceedings-F*, vol. 140 (6), pp. 362–370, 1993.

[8] J.-F. Cardoso, A. Souloumiac, "Jacobi angles for simultaneous diagonalization", *SIAM J. Matrix Anal. Appl.*, vol. 17 (1), pp. 161–164, jan. 1996.

[9] J. Carroll, J. Chang, "Analysis of individual differences in multidimensional scaling via an n-way generalization of "Eckart-Young" decomposition", *Psychometrika*, vol. 35, pp. 283–319, 1970.

[10] P. Chevalier, L. Albera, A. Ferréol, P. Comon, "On the virtual array concept for the higher order array processing", *IEEE Trans. on Signal Processing*, vol. 53 (4), pp. 1254–1271, avr. 2005.

[11] P. Comon, "Independent Component Analysis, a new concept ?", *Signal Processing, Elsevier*, vol. 36 (3), pp. 287–314, avr. 1994.

[12] P. Comon, "Tensor decompositions," *Mathematics in Signal Processing V*, J. G. McWhirter, I. K. Proudler, Eds., pp. 1–24. Clarendon Press, Oxford, UK, 2002.

[13] P. Comon, "Blind identification and source separation in 2×3 underdetermined mixtures", *IEEE Trans. on Signal Processing*, vol. 52 (1), pp. 11–22, jan. 2004.

[14] P. Comon, "Canonical tensor decompositions," Tech. rep. rr-2004-17, Lab. I3S, Sophia-Antipolis, France, juin 2004.

[15] P. Comon, G. Golub, L.-H. Lim, B. Mourrain, "Symmetric tensors and symmetric tensor rank", *SIAM J. Matrix Anal. Appl.*, 2006, soumis.

[16] P. Comon, B. Mourrain, "Decomposition of quantics in sums of powers of linear forms", *Signal Processing*, vol. 53 (2), pp. 96–107, sep. 1996.

[17] P. Comon, B. Mourrain, L.-H. Lim, G. Golub, "Genericity and rank deficiency of high order symmetric tensors", *ICASSP'06*, Toulouse, mai 2006.

[18] L. De Lathauwer, "A link between the canonical decomposition in multilinear algebra and simultaneous matrix diagonalization", *SIAM J. Matrix Anal. Appl.*, accepté.

[19] L. De Lathauwer, *Signal processing based on multilinear algebra*, Thèse de doctorat, KU Leuven, Leuven, Belgique, 1997.

[20] L. De Lathauwer, "Decomposition of higher-order tensor in block terms", *SIAM J. Matrix Anal. Appl.*, 2006, soumis.

[21] L. De Lathauwer, B. De Moor, "From matrix to tensor : multilinear algebra and signal processing," *Mathematics in Signal Processing IV*, pp. 1–15. J. MacWhirter and I. Proudler Eds., 1998.

[22] L. De Lathauwer, B. De Moor, Vandewalle J., "Fetal electrocardiogram extraction by blind source subspace separation", *IEEE Transactions on Biomedical Engineering, Special Topic Section on Advances in Statistical Signal Processing for Biomedicine*, vol. 47 (5), pp. 567–572, mai 2000.

[23] L. De Lathauwer, B. De Moor, J. Vandewalle, "An algebraic ICA algorithm for 3 sources and 2 sensors", *Proc. EUSIPCO 2000*, Tampere, Finland, sep. 2000.

[24] V. de Silva, L.-H. Lim, "Tensor rank and the ill-posedness of the best low-rank approximation problem," SCCM tech. rep. 05-03, Standford University, 2005, preprint.

[25] A. Ferréol, L. Albera, P. Chevalier, "Fourth order blind identification of underdetermined mixtures of sources (fobium)", *IEEE Trans. on Signal Processing*, vol. 53 (5), pp. 1640–1653, mai 2005.

[26] G. Gelle, M. Colas, C. Servière, "Blind source separation : A new preprocessing tool for rotating machines monitoring ? theoretical background and practical considerations", *IEEE Trans. on Instrumentation and Measurement*, vol. 52 (3), pp. 790–796, mar. 2003.

[27] P. Georgiev, F. Theis, A. Cichocki, "Sparse component analysis and blind source separation of underdetermined mixtures", *IEEE Trans. on Neural Networks*, vol. 16 (4), pp. 992–996, juil. 2005.

[28] D.N. Godard, "Self-recovering equalization and carrier tracking in two-dimensional data communication systems", *IEEE Trans. on Commun.*, vol. 28, pp. 1867–1875, nov. 1980.

[29] M. Haardt, J.A. Nossek, "Simultaneous Schur decomposition of several nonsymmetric matrices to achieve automatic pairing in multidimensional harmonic retrieval problems", *IEEE Trans. on Signal Processing*, vol. 46 (1), pp. 161–169, 1998.

[30] R. Harshman, "Foundations of the PARAFAC procedure : model and conditions for an "explanatory" multi-mode factor analysis", *UCLA Working Papers in Phonetics*, vol. 16, pp. 1–84, 1970.

[31] R.T. Kolda, "Orthogonal tensor decompositions", *SIAM Journal on Matrix Analysis and Applications*, vol. 23 (1), pp. 243–255, 2001.

[32] J.B. Kruskal, "Three-way arrays : rank and uniqueness of trilinear decompositions, with applications to arithmetic complexity and statistics", *Lin. Alg. Appl.*, vol. 18, pp. 95–138, 1977.

[33] M.S. Lewicki, T.J. Sejnowski, "Learning overcomplete representations", *Neural Computation*, vol. 12 (2), pp. 337–365, 2000.

[34] E. Moreau, "A generalization of joint-diagonalization criteria for source separation", *IEEE Trans. on Signal Processing*, vol. 49 (3), pp. 530–541, 2001.

[35] D. Nion, L. De Lathauwer, "A block factor analysis based receiver for blind multi-user access in wireless communications", *Proc. of ICASSP'06, IEEE Int. Conf. on Acoustics, Speech and Signal Processing*, Toulouse, France, mai 2006, pp. 825–828.

[36] D. Nion, L. De Lathauwer, "Line search computation of the block factor model for blind multi-user access in wireless communications", *Proc. of SPAWC 06, IEEE Workshop on Signal Processing Advances in Wireless Communications*, Cannes, France, juil. 2006.

[37] D.T. Pham, J.-F. Cardoso, "Blind separation of instantaneous mixtures of non stationary sources", *IEEE Trans. on Signal Processing*, vol. 49 (9), pp. 1837–1848, 2001.

[38] J.G. Proakis, *Digital Communications*, McGraw-Hill international editions, 1995.

[39] M. Rajih, P. Comon, "Enhanced line search : A novel method to accelerate parafac", *Proc. EUSIPCO 2005*, Antalya, Turquie, sep. 2005.

[40] J. Sanchez, M. Thioune, *UMTS*, Hermes Science Publications, 2004.

[41] N. Sidiropoulos, R. Bro, "On the uniqueness of multilinear decomposition of N-way arrays", *Journal of Chemometrics*, vol. 14, pp. 229–239, 2000.

[42] N. Sidiropoulos, R. Bro, G. Giannakis, "Parallel factor analysis in sensor array processing", *IEEE Trans. on Signal Processing*, vol. 48 (2), pp. 2377–2388, août 2000.

[43] N. Sidiropoulos, G. Giannakis, R. Bro, "Blind PARAFAC receivers for DS-CDMA systems", *IEEE Trans. on Signal Processing*, vol. 48 (3), pp. 810–823, mar. 2000.

[44] A. Stegeman, "Degeneracy in Candecomp/Parafac and Indscal explained for several three-sliced arrays with a two-valued typical rank", *Psychometrika*, 2005, soumis.

[45] A. Stegeman, J.M.F. ten Berge, L. De Lathauwer, "Sufficient conditions for uniqueness in Candecomp/Parafac and Indscal with random component matrices", *Psychometrika*, vol. 71, pp. 219–229, 2006.

[46] J.M.F. ten Berge, "Kruskal's polynomial for $2 \times 2 \times 2$ arrays and a generalization to $2 \times n \times n$ arrays", *Psychometrika*, vol. 56 (4), pp. 631–636, 1991.

[47] J.M.F. ten Berge, "The typical rank of tall three-way arrays", *Psychometrika*, vol. 65 (5), pp. 525–532, 2000.

[48] J.M.F. ten Berge, "Simplicity and typical rank of three-way arrays, with applications to TUCKER-3 analysis with simple cores", *Journal of Chemometrics*, vol. 18, pp. 17–21, 2004.

[49] F. Theis, E. Lang, C. Puntonet, "A geometric algorithm for overcomplete linear ICA", *Neurocomputing*, vol. 56, pp. 381–398, 2004.

[50] J.R. Treichler, M.G. Larimore, "New processing techniques based on constant modulus adaptative algorithm", *IEEE Transaction on Acoustics, Speech, Signal Processing*, vol. 33, pp. 420–431, avr. 1985.

[51] A.J. Van der Veen, *Signal Processing Advances in Wireless and Mobile Communications, "Algebraic Constant Modulus Algorithms"*, vol. 2, chapitre 3, G. Giannakis Ed., Prentice Hall, 2000.

[52] A.J. Van der Veen, A. Paulraj, "An analytical constant modulus algorithm", *IEEE Trans. on Signal Processing*, vol. 44 (5), pp. 1136–1155, sep. 1996.

[53] J.G. Verdú, *Multiuser Detection*, Cambridge Univ. Press, 1998.

[54] 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. 1545–1553, 2002.

[55] A. Ypma, P. Pajunen, "Rotating machine vibration analysis with second-order independent component analysis", *Proceedings of the First International Workshop on Independent Component Analysis and Signal Separation, ICA'99*, jan. 1999, pp. 37–42.