Décomposition canonique de tableaux hermitiens semi-définis positifs d'ordre pair par rotation procustéenne: application à l'ICA.

Laurent Albera, Ahmad Karfoul, Lieven de Lathauwer

▶ **To cite this version:**

Laurent Albera, Ahmad Karfoul, Lieven de Lathauwer. Décomposition canonique de tableaux hermitiens semi-définis positifs d'ordre pair par rotation procustéenne: application à l'ICA.. XXII colloque GRETSI, Sep 2009, Dijon, France. pp.488. hal-00913700

HAL Id: hal-00913700
https://hal.science/hal-00913700

Submitted on 14 Mar 2014

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

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

Décomposition canonique de tableaux hermitiens semi-définis positifs d'ordre pair par rotation procustéenne : application à l'ICA

Laurent Albera[1], Ahmad Karfoul[1], Lieven De Lathauwer[2]

[1]INSERM, U 642, Rennes, F-35000 France
Université de Rennes 1, LTSI, F-35000 France

[2]Laboratoire Traitement des Images
Katholieke Univ. Leuven, Subfaculty Science and Technology, B-8500 Kortrijk, Belgium
laurent.albera@univ-rennes1.fr,
ahmad.karfoul@univ-rennes1.fr,Lieven.DeLathauwer@kuleuven-kortrijk.be

Résumé – Nous présentons ici une nouvelle famille de méthodes itératives, nommée CanDeP (multi-way array Canonical Decomposition based on Procrustes rotation), afin de calculer la décomposition canonique d'un tableau hermitien semi-défini positif d'ordre pair. Cette approche alterne jusqu'à convergence la résolution du problème de Procrustes et la décomposition de tableaux de rang 1. Appliquée aux cumulants d'ordre $2q$, la méthode $2q$-CanDeP permet de réaliser une analyse en composantes indépendantes performante, offrant i) un gain important en résolution, ii) la capacité d'identifier un mélange sous-déterminé de sources et iii) une robustesse à un bruit Gaussien de covariance spatiale inconnue.

Abstract – A new family of iterative methods, named CandDep (multi-way array Canonical Decomposition based on Procrustes rotation) is proposed in this paper in order to canonically decompose an even higher order positive semi-definite array with hermitian symmetry. This approach switches, until the convergence, between the best rank-1 approximation of a given HO array and the resolution of the well-known Procrustes problem. Applied to $2q$-th order cumulants, the $2q$-CanDep approach allows for a well-performed independent component analysis. Indeed, the proposed approach enjoys i) a good identification resolution, ii) the ability of processing underdetermined mixtures of sources and iii) a good robustness to a Gaussian noise with unknown spatial covariance.

1 Introduction

L'algèbre multilinéaire est défini comme l'algèbre des tableaux d'ordre q ($q > 2$), i.e., les tableaux à plus de deux entrées. Si les premiers travaux en matière d'algèbre multilinéaire remontent au XIXème siècle, il faut attendre la seconde moitié du XXème siècle pour voir ce domaine prendre une toute nouvelle dimension au travers de l'analyse de données et assister au sacre du modèle PARAFAC (PARallel FACtor analysis) [6], que nous désignerons par CAND pour être également référencé sous le nom de Décomposition CANonique.

Le modèle CAND n'est autre que l'écriture d'un tableau d'ordre q sous forme d'une combinaison linéaire de tableaux d'ordre q et de rang 1. L'utilisation d'un tel modèle ne cesse de se répandre, en touchant à présent des domaines aussi riches et variés que ceux de l'ingénierie biomédicale [7] et du traitement d'antennes [10]. Parmi les méthodes de décomposition les plus utilisées, on compte les approches itératives telles que la méthode du gradient conjugué, la méthode de Levenberg-Marquardt [11] mais aussi la non moins populaire méthode ALS (Alternating Least Square) [6] et ses dérivées [9]. Parmi les approches semi-algébrique, on trouve la méthode FOOBI proposée dans [8] pour des tableaux d'ordre 3 et 4, ainsi que l'algorithme BIOME proposée dans [1] pour des tableaux hermitiens semi-définis positifs d'ordre pair. Si les approches algébriques évitent les problèmes de convergence propres aux méthodes itératives, elles sont la plupart du temps très voire trop gourmandes en termes de coût de calcul.

Aussi, présentons-nous ici une nouvelle famille de méthodes itératives, nommée CanDeP (multi-way array Canonical Decomposition based on Procrustes rotation), afin de calculer la décomposition canonique d'un tableau hermitien semi-défini positif d'ordre pair. Cette approche alterne jusqu'à convergence la résolution du problème de Procrustes et la décomposition de tableaux de rang 1. Appliquée aux cumulants d'ordre $2q$, la méthode $2q$-CanDeP permet de réaliser une analyse en composantes indépendantes performante, offrant i) un gain important en résolution, ii) la capacité d'identifier un mélange sous-déterminé de sources et iii) une robustesse à un bruit Gaussien de covariance spatiale inconnue.

2 Outils préliminaires et présentation du problème

Avant tout, rappelons quelques notions d'algèbre multilinéaire [2].

Définition 1 *Soit* $\boldsymbol{\mathcal{T}} \in \mathbb{C}^{N_1 \times \cdots \times N_q}$ *un tableau d'ordre* q $(q \geq 2)$*. On dira que* $\boldsymbol{\mathcal{T}}$ *est de rang 1 s'il peut s'écrire comme le produit externe* $\boldsymbol{u}^{(1)} \circ \cdots \circ \boldsymbol{u}^{(q)}$ *de* q *vecteurs* $\boldsymbol{u}^{(i)} \in \mathbb{C}^{N_i}$ $(1 \leq i \leq q)$ *où chaque composante de* $\boldsymbol{\mathcal{T}}$ *est donnée par* $\mathcal{T}_{i_1,\cdots,i_q} = u^{(1)}_{i_1} \cdots u^{(q)}_{i_q}$.

De même que pour les matrices, il est possible de définir le rang d'un tableau d'ordre supérieur :

Définition 2 *Soit* $\boldsymbol{\mathcal{T}} \in \mathbb{C}^{N_1 \times \cdots \times N_q}$ *un tableau d'ordre* q $(q \geq 3)$*. Le rang de* $\boldsymbol{\mathcal{T}}$*, noté* $\mathrm{rk}(\boldsymbol{\mathcal{T}})$*, est le nombre minimal de tableaux d'ordre* q*, et de rang 1 décrivant exactement* $\boldsymbol{\mathcal{T}}$ *par combinaison linéaire. Cette combinaison linéaire, minimale en son nombre de terme, définit la CAND du tableau* $\boldsymbol{\mathcal{T}}$*.*

Notons que, contrairement au cas matriciel, le rang d'un tableau d'ordre strictement supérieur à 2 n'est pas obligatoirement majoré par la plus petite de ses dimensions. Par ailleurs, la définition 2 montre que l'ordre et la pondération des différents tableaux de rang 1 définissant la CAND d'un tableau $\boldsymbol{\mathcal{T}}$ peuvent être modifiés sans changer la valeur de $\boldsymbol{\mathcal{T}}$. Si l'existence d'une telle décomposition est toujours assurée, son unicité quant à elle ne sera garantie, aux deux indéterminations citées précédemmment, que sous certaines conditions (voir les références dans [10] pour des conditions suffisantes d'unicité). Considérons à présent la transformation tableau-à-matrice suivante :

Définition 3 *Soit* $\boldsymbol{\mathcal{T}}$ *un tableau carré d'ordre* $2q$ $(q \geq 2)$ *de dimensions* N*. Soient respectivement* $\lfloor q/2 \rfloor$ *et* $\lceil q/2 \rceil$ *les parties entières inférieure et supérieure de* $q/2$*. La* (i,j)*ème composante de la matrice* $\boldsymbol{T} = \mathrm{mat}_1(\boldsymbol{\mathcal{T}})$ *de taille* $(N^q \times N^q)$ *est définie par :*

$$(\mathrm{mat}_1(\boldsymbol{\mathcal{T}}))_{i,j} = \mathcal{T}_{n_1,\cdots,n_{\lceil q/2 \rceil},n_{\lceil q/2 \rceil+1},\cdots,n_q,n_{q+1},\cdots,n_{q+\lceil q/2 \rceil}n_{2q+\lceil q/2 \rceil+1},\cdots,n_{2q}}$$

où

$$\begin{cases} i = (n_1 - 1)N^{q-1} + \cdots + (n_{\lceil q/2 \rceil} - 1)N^{\lfloor q/2 \rfloor} \\ \quad + (n_{q+\lceil q/2 \rceil+1} - 1)N^{\lfloor q/2 \rfloor - 1} + \cdots + (n_{2q} - 1)N + n_{2q} \\ j = (n_{q+1} - 1)N^{q-1} + \cdots + (n_{q+\lceil q/2 \rceil} - 1)N^{\lfloor q/2 \rfloor} \\ \quad + (n_{\lceil q/2 \rceil+1} - 1)N^{\lfloor q/2 \rfloor - 1} + \cdots + (n_q - 1)N + n_q \end{cases}$$

Prenons à titre d'exemple le cas d'un tableau $\boldsymbol{\mathcal{T}}$ hermitien d'ordre $2q$ $(q \geq 2)$ de dimensions N tel que $\xi(\omega) = 1$ pour $\omega \in \{1, \ldots, q\}$ et $\xi(\omega) = -1$ pour $\omega \in \{q+1, \ldots, 2q\}$.

Définition 4 *Un tableau* $\boldsymbol{\mathcal{T}}$ *d'ordre* q $(q \geq 2)$ *à valeurs complexes sera dit hermitien si sa CAND prend la forme suivante :*

$$\boldsymbol{\mathcal{T}} = \sum_{p=1}^{P} \lambda_p \, \boldsymbol{a}_p^{\xi(1)} \circ \boldsymbol{a}_p^{\xi(2)} \circ \cdots \circ \boldsymbol{a}_p^{\xi(q)} \qquad (1)$$

où les P *valeurs* λ_p *sont réelles et l'application* ξ *est définie par* $\xi(\omega) = \pm 1$ *avec pour conventions* $x^1 = x$ *et* $x^{-1} = x^*$*,* x^* *désignant la conjuguée de* x*.*

Supposons également que $\boldsymbol{\mathcal{T}}$ soit semi-défini positif, les P valeurs λ_p de la CAND (1) de $\boldsymbol{\mathcal{T}}$ sont alors toutes positives. Construisons la matrice $\boldsymbol{T} = \mathrm{mat}_1(\boldsymbol{\mathcal{T}})$ de taille $(N^q \times N^q)$, on peut alors démontrer d'après les définitions 3 et 4 l'égalité $\boldsymbol{T} = \boldsymbol{A}_q \boldsymbol{\Lambda} \boldsymbol{A}_q{}^{\mathsf{H}}$, où $\boldsymbol{A}_q = \boldsymbol{A}^{\oslash \lceil q/2 \rceil} \oslash \boldsymbol{A}^{* \lfloor q/2 \rfloor}$, $\oslash$ est l'opérateur du produit de *Khatri-Rao product* [1], $^{\oslash}$ désigne l'opérateur de *puissance de Khatri-Rao* [1], $\boldsymbol{A}$ est la matrice de taille $(N \times P)$ égale à $[\boldsymbol{a}_1, \ldots, \boldsymbol{a}_P]$ et $\boldsymbol{\Lambda}$ est la matrice diagonale $\mathrm{diag}\{[\lambda_1, \cdots, \lambda_P]\}$. De cette égalité découle le caractère hermitien semi-défini positif de $\boldsymbol{T}$ garantissant l'existence d'une racine carré de $\boldsymbol{T}$. Enfin, introduisons la transformation tableau-à-vecteur suivante :

Définition 5 *Soit* $\boldsymbol{\mathcal{T}} \in \mathbb{C}^{N_1 \times \cdots \times N_q}$ *avec* $q \geq 2$*. Les composantes du vecteur colonne* $\boldsymbol{t} = \mathrm{vec}(\boldsymbol{\mathcal{T}})$ *de dimension* $N_1 \cdots N_q$ *sont définies par :*

$$t_{(n_1-1)(N_2 \cdots N_q)+(n_2-1)(N_3 \cdots N_q)+\cdots+n_q} = \mathcal{T}_{n_1,\cdots,n_q} \qquad (2)$$

On peut alors définir l'opérateur réciproque unvec *tel que l'on ait* $\mathrm{unvec}\,(\mathrm{vec}(\boldsymbol{\mathcal{T}})) = \boldsymbol{\mathcal{T}}$.

Le problème traité dans ce papier peut être résumé de la manière suivante :

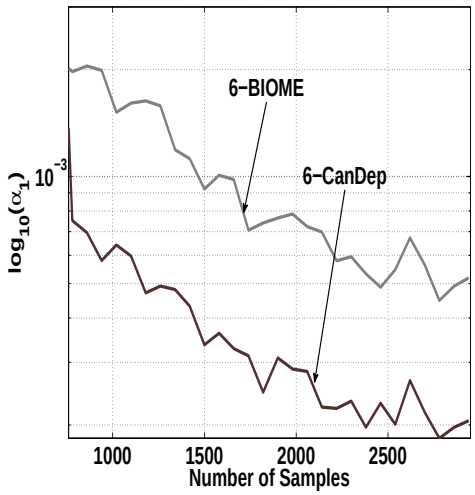
Problème 1 *Calculer la CAND d'un tableau* $\boldsymbol{\mathcal{T}}$ *hermitien semi-défini positif d'ordre pair.*

3 La famille de méthodes CanDeP

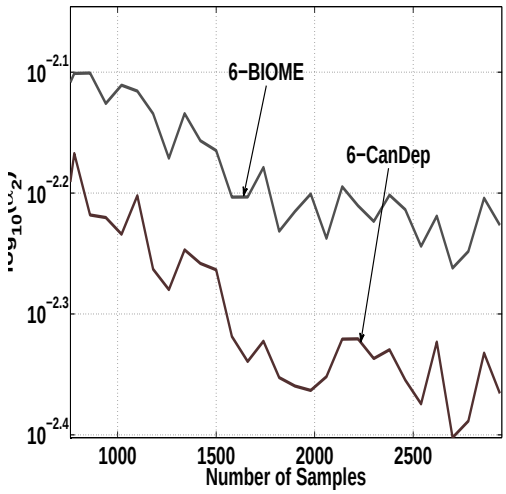
Nous proposons dans cette section une méthode permettant de résoudre le problème 1 évoqué ci-dessus. Soit $\boldsymbol{\mathcal{T}}$ un tableau hermitien semi-défini positif d'ordre $2q$ $(q \geq 2)$. La méthode proposée ici, baptisée $2q$-CanDeP, calcule la CAND de $\boldsymbol{\mathcal{T}}$ en alternant jusqu'à convergence deux étapes de minimisation. Présentons tout d'abord la première étape. Pour ce faire, notons $\boldsymbol{T}^{1/2}$ une racine carrée de $\boldsymbol{T} = \mathrm{mat}_1(\boldsymbol{\mathcal{T}})$ obtenue par diagonalisation de $\boldsymbol{T}$. En effet, $\boldsymbol{T}$ est diagonalisable sous la forme $\boldsymbol{E}\boldsymbol{L}\boldsymbol{E}^{\mathsf{H}}$, du fait du caractère hermitien de $\boldsymbol{T}$, où $\boldsymbol{E}$ est la matrice des vecteurs propres orthonormés associés aux P valeurs propres non nulles de $\boldsymbol{T}$ rangées dans la matrice diagonale $\boldsymbol{L}$. On a alors $\boldsymbol{T}^{1/2} = \boldsymbol{E}\boldsymbol{L}^{1/2}$ où $\boldsymbol{L}^{1/2}$ est la racine carrée de $\boldsymbol{L}$. On a alors la proposition suivante :

Proposition 1 *Sous les hypothèses précédentes, la matrice* $\boldsymbol{B} = (\boldsymbol{L}^{1/2})^{-1} \boldsymbol{E}^{\mathsf{H}} \boldsymbol{A}_q \boldsymbol{\Lambda}^{1/2}$ *est une matrice orthonormée (unitaire) de taille* $(P \times P)$*, à valeurs réelles lorsque* q *est pair.*

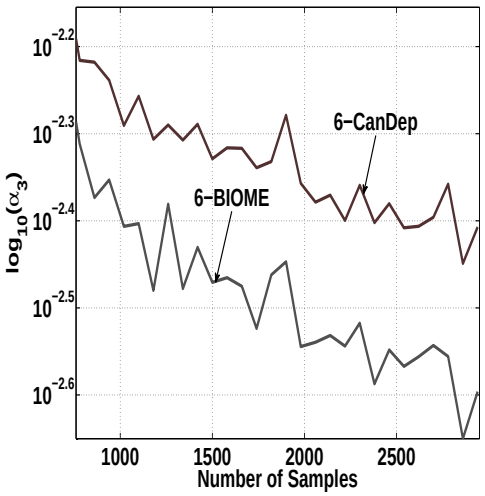
Ainsi, la première étape de $2q$-CanDeP consiste à minimiser la norme de Frobenius de la différence entre $\boldsymbol{T}^{1/2}$ et le produit matriciel $\boldsymbol{A}_q \boldsymbol{\Lambda}^{1/2} \boldsymbol{B}^{\mathsf{H}}$, ie, sans perte de généralité au vu des indéterminations de la CAND, entre $\boldsymbol{T}^{1/2}$ et le produit matriciel $\boldsymbol{A}_q \boldsymbol{B}^{\mathsf{H}}$ sous la contrainte d'avoir $\boldsymbol{B}$ unitaire. En supposant connue la matrice $\boldsymbol{A}_q$, ce problème d'optimisation est connu sous le nom de problème de Procrustes [5]. La solution optimale au sens des moindres carrés consiste à choisir $\boldsymbol{B} = \boldsymbol{U}\boldsymbol{V}^{\mathsf{H}}$, où $\boldsymbol{U}$ et $\boldsymbol{V}$ sont respectivement les matrices singulières gauche et droite de la matrice $\boldsymbol{A}_q^{\mathsf{H}} \boldsymbol{T}^{1/2}$ (la démonstration proposée dans [5] dans le cas de matrices réelles peut facilement être étendue au cas complexe).



(a) source 1



(b) source 2



(c) source 3

FIGURE 1 – Pseudo-distance D calculée pour 6-CanDeP et 6-BIOME en fonction du nombre d'échantillons

La seconde étape de la méthode $2q$-CanDeP consiste à identifier la matrice $\boldsymbol{A}_q$, et plus exactement les P vecteurs colonnes $\boldsymbol{a}_p$ de $\boldsymbol{A}$. Supposant connue la matrice $\boldsymbol{B}$, la matrice $\boldsymbol{A}_q$ peut être identifiée, à une matrice diagonale près, en calculant le produit $\boldsymbol{T}^{1/2}\boldsymbol{B}$. On peut alors montrer qu'appliquer l'opérateur unvec au p-ème vecteur colonne de dimension N^q de la matrice $\boldsymbol{A}_q$ identifiée permet d'obtenir le tableau carré $\boldsymbol{\mathcal{A}}^{(p)}$ d'ordre q et de rang 1, défini par $\boldsymbol{\mathcal{A}}^{(p)} = \boldsymbol{a}_p^{\circ\lceil q/2\rceil} \circ \boldsymbol{a}_p^{*\circ\lfloor q/2\rfloor}$. La seconde étape consiste donc à calculer la CAND de rang 1 de chaque tableau $\boldsymbol{\mathcal{A}}^{(p)}$ à l'aide de la méthode des puissances d'ordre supérieure [2] afin d'identifier à un scalaire près les P vecteurs colonnes $\boldsymbol{a}_p$. Notons que la décomposition est proposée dans [2] pour des tableaux de rang 1 est également réalisée au sens des moindres carrés.

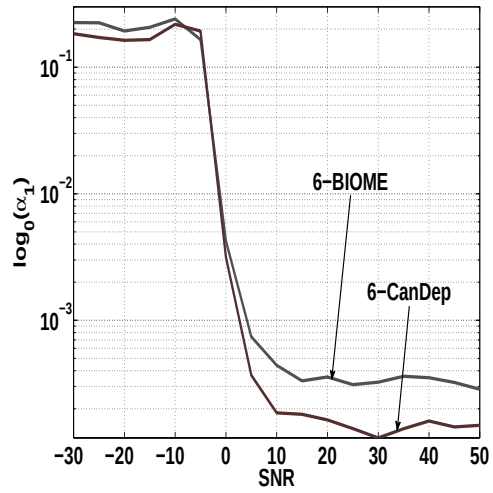
4 Simulations

Une étude comparative en termes de simulations numériques est présentée ci-dessous dans le contexte de l'identification de mélange sous-déterminé de sources. Cette étude permet de comparer les performances de la méthode CanDeP et d'une autre méthode plus connue sous le nom de BIOME [1]. Notons que BIOME a pour caractéristique de proposer une solution semi-algébrique au problème de CAND. En fait, les méthodes 6-CanDeP et 6-BIOME ont été utilisées afin d'obtenir la CAND d'un tableau cumulant $\boldsymbol{\mathcal{C}}_{6,\boldsymbol{x}}$ estimé à partir de réalisations d'un vecteur aléatoire $\boldsymbol{x}$ suivant le modèle $\boldsymbol{x} = \boldsymbol{A}\boldsymbol{s} + \boldsymbol{\nu}$ où $\boldsymbol{s}$ est un vecteur aléatoire dont les composantes sont mutuellement indépendantes et $\boldsymbol{\nu}$ un vecteur Gaussien. La CAND de $\boldsymbol{\mathcal{C}}_{6,\boldsymbol{x}}$ conduit alors à l'identification du mélange $\boldsymbol{A}$ de taille $(N \times P)$ par propriété des cumulants [3]. A titre d'exemple, nous nous sommes placés dans le cadre de communications satellitaires où $P = 3$ modulations QPSK filtrées Nyquist (roll-off de 0.3) ont été générées et dont la réception sur un réseau de $N = 2$ capteurs a été simulée. Les porteuses des QPSK ont été respectivement prises égales à 0, 0.35 et 0.7 (au facteur près de la fréquence d'échantillonnage). Le critère d'évaluation de performance utilisé ici est la pseudo-distance, $D(\boldsymbol{A}, \widehat{\boldsymbol{A}}) = (\alpha_1, \alpha_2, \cdots, \alpha_P)$, entre $\boldsymbol{A}$ et son estimée $\widehat{\boldsymbol{A}}$ où $\alpha_p = \min_{1 \le i \le P}\{d(\boldsymbol{a}_p, \widehat{\boldsymbol{a}}_i)\}$ avec :

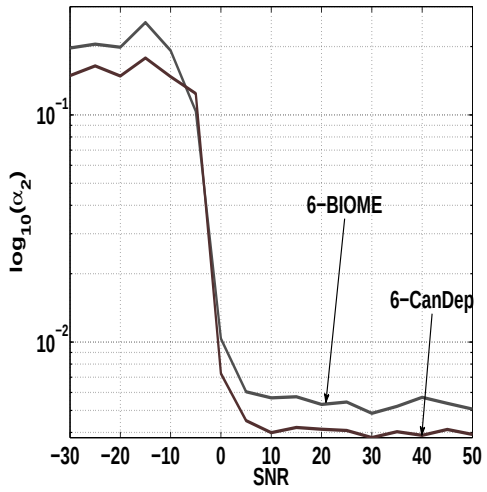
$$d(\boldsymbol{u}, \boldsymbol{v}) = 1 - |\boldsymbol{u}^{\mathrm{H}}\boldsymbol{v}|^2 / (||\boldsymbol{u}||\,||\boldsymbol{v}||) \qquad (3)$$

Les résultats ont été moyennés sur 100 expériences indépendantes.

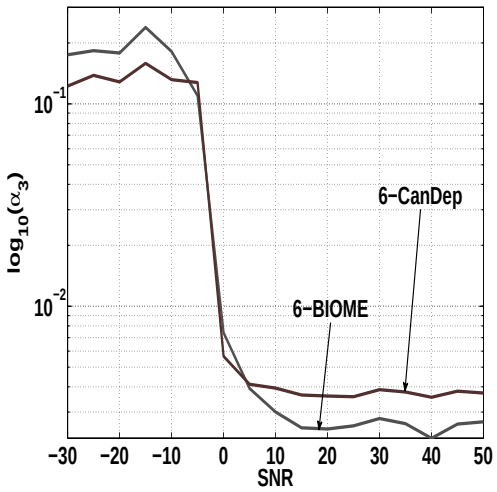
La figure 1 montre les variations de D en fonction du nombre de réalisations du vecteur $\boldsymbol{x}$ pour un Rapport Signal à Bruit (RSB) de 20dB. Les deux méthodes ont des performances très voisines comme en témoigne la figure 1. 750 échantillons suffisent aux deux algorithmes pour assurer des performances tout à fait correctes (i.e. $D < 0.01$) malgré l'usage de cumulants d'ordre supérieur (l'ordre 6 est ici utilisé) réputés pour avoir une grande variance d'estimation. La figure 2 montre quant-à-elle les variations de D en fonction du RSB pour 2500 réalisations de $\boldsymbol{x}$. On note un bon comportement de 6-CanDeP au delà de 0dB, comportement similaire à celui de 6-BIOME.



(a) source 1



(b) source 2



(c) source 3

FIGURE 2 – Pseudo-distance D calculée pour 6-CanDeP et 6-BIOME en fonction du RSB.

Ce travail a été en partie financé par l'ANR au travers de deux contrats ANR blancs, à savoir "Decotes" (ANR-06-BLAN-0074) et "mv-EMD" (ANR-BLAN07-0314-02).

Références

[1] L. ALBERA, A. FERREOL, P. COMON, and P. CHEVALIER, "Blind Identification of Overcomplete Mixtures of sources (BIOME)," *Linear Algebra Applications*, vol. 391C, pp. 3–30, November 2004.

[2] L. De LATHAUWER, B. De MOOR, and J. VANDEWALLE, "On the best rank-1 and rank-(R1,R2,...,RN) approximation of high-order tensors," *SIAM Journal Matrix Analysis and Applications*, vol. 21, no. 4, pp. 1324–1342, 2000.

[3] A. KACHENOURA, L. ALBERA, L. SENHADJI, and P. COMON, "ICA : a potential tool for BCI systems," *IEEE Signal Processing Magazine, special issue on Brain-Computer Interfaces*, vol. 25, no. 1, pp. 57–68, January 2008.

[4] A. FERREOL, L. ALBERA, and P. CHEVALIER, "Fourth order blind identification of underdetermined mixtures of sources (FOBIUM)," *IEEE Transactions On Signal Processing*, vol. 53, pp. 1254–1271, April 2005.

[5] G. H. GOLUB and C. F. V. LOAN, *Matrix computations, second edition*. The Johns Hopkins University Press, Baltimore, MD, 1989.

[6] R. HARSHMAN, "Foundation of PARAFAC procedure : Models and conditions for an 'explanatory' multi-mode factor analysis," *UCLA working papers in Phonetics*, no. 16, pp. 1–84, 1970.

[7] C. S. HERMANN, J.PARNAS, M. MORUP, L. K. HANSEN, and S. M. ARNFRED, "Parallel factor analysis as an exploratory tool for wavelet transformed eventrelated EEG," *NeuroImage*, vol. 29, no. 3, pp. 938–947, 2006.

[8] L. D. LATHAUWER, "A link between the canonical decomposition in multilinear algebra and simultaneous matrix diagonalization," *SIAM Journal in matrix Analysis and Applications*, vol. 28, no. 3, pp. 642–666, 2006.

[9] M. RAJIH, P. COMON, and R. HARSHMAN, "Enhanced line search : A novel method to accelerate PARAFAC," *SIAM Journal in Matrix Analysis and Applications, special issue on Tensor Decomposition and Applications*, 2008.

[10] N. D. SIDIROPOULOS, R. BRO, and G. B. GIANNAKIS, "Parallel factor analysis in sensor array processing," *IEEE Transactions On Signal Processing*, vol. 48, no. 8, pp. 2377–2388, August 2000.

[11] G. TOMASI and R. BRO, "A comparison of algorithms for fitting the parafac model," *Computational Statistics and Data Analysis*, vol. 50, no. 7, pp. 1700–1734, 2006.