

# Décomposition canonique polyadique robuste de tenseurs de faible rang basée sur la parcimonie de groupe

Xu HAN<sup>1,2</sup>, Laurent ALBERA<sup>1,2,4</sup>, Amar KACHENOURA<sup>1,2,4</sup>, Huazhong SHU<sup>3,4</sup>, Lotfi SENHADJI<sup>1,2,4</sup>

<sup>1</sup>LTSI Laboratory, University of Rennes 1, Rennes, F-35000, France

<sup>2</sup>Inserm, UMR 1099, Rennes, F-35000, France

<sup>3</sup>LIST, Southeast University, 2 Sipailou, 210096, Nanjing, China

<sup>4</sup>Centre de Recherche en Information Biomédicale Sino-Français (CRIBs), Rennes, France

**Résumé** – La plupart des algorithmes existants de décomposition Canonique Polyadique (CP) nécessitent de connaître préalablement le rang du tenseur et sont sensibles à une surestimation de ce dernier. La particularité de la méthode proposée dans cet article est d’estimer simultanément le rang et les matrices de facteurs du tenseur. Alors que l’optimisation de la norme nucléaire engendre généralement un coût de calcul prohibitif, la méthode proposée repose sur la minimisation d’une métrique moins coûteuse exploitant la parcimonie de groupe dans un domaine transformé. Une évaluation de notre méthode sur données simulées et réelles (issues de spectroscopie par fluorescence) montrent l’intérêt de l’approche vis à vis des algorithmes classiques de décomposition CP.

**Abstract** – A new and robust method for low rank Canonical Polyadic (CP) decomposition of tensors is introduced in this paper. The proposed method imposes the Group Sparsity of the coefficients of each Loading (GSL) matrix under orthonormal subspace. By this way, the low rank CP decomposition problem is solved without any knowledge of the true rank and without using any nuclear norm regularization term, which generally leads to computationally prohibitive iterative optimization for large-scale data. Our GSL-CP technique can be then implemented using only an upper bound of the rank. It is compared in terms of performance with classical methods, which require to know exactly the rank of the tensor. Numerical simulated experiments with noisy tensors and results on fluorescence data show the advantages of the proposed GSL-CP method in comparison with classical algorithms.

## 1 Introduction

Le développement de méthodes de décomposition de tableaux multi-dimensionnels suscite toujours autant d’attention, notamment d’un point de vue applicatif. Le modèle de décomposition Canonique Polyadique (CP) de rang faible est sans nul doute le plus célèbre [1]. Il est grandement utilisé en traitement du signal [2] et en traitement d’image [3]. Depuis les travaux pionniers d’Harshman [4], beaucoup d’algorithmes ont été proposés afin de calculer la décomposition CP. Parmi eux, l’ALS (Alternating Least Squares), présenté simultanément par Carroll et Chang [5] et par Harshman [4], est le plus célèbre. Plusieurs modifications ont été apportées afin d’améliorer son comportement. Par exemple, la procédure ELS (Enhanced Line Search), basée sur une extrapolation sophistiquée permet d’éviter certains minima locaux et d’accélérer la convergence [6]. Malgré les bons résultats pratiques de la technique ELS-ALS, aucune minimisation globale de la fonction objective d’attache aux données n’est assurée. Plus récemment, une méthode semi-algébrique, à savoir la méthode DIAG (DIrect ALGORITHM for CP decomposition), a été proposée [7]. Au lieu de minimiser la fonction d’attache aux données de l’ALS, la décomposition CP est reformulée comme un problème de diagonalisation conjointe par similarité de matrices [8], suivi d’une étape

d’approximations tensorielles de rang 1.

Les méthodes mentionnées ci-dessus requièrent une estimation préalable du rang. Toutefois, une telle estimation peut être difficile par exemple pour des Rapports Signal à Bruit (RSB) faibles. D’un autre côté, estimer simultanément le rang et les matrices de facteurs du tenseur n’est pas tâche facile tant les problèmes de minimisation de rang sont généralement NP-difficiles. Certains chercheurs ont alors proposé d’utiliser la norme nucléaire [9], définie comme la somme des valeurs singulières, afin de servir d’enveloppe convexe de la fonction de rang [10]. Cependant, la minimisation de la norme nucléaire engendre généralement un coût de calcul prohibitif pour l’analyse de données de grande taille telles que les vidéos [11].

Dans ce papier, nous proposons la méthode R-CPD (Robust low rank CP Decomposition) pour le calcul d’une décomposition CP de rang faible. Nous utilisons une mesure du rang proposée par Shu et al. [11] pour des matrices de rang faible. Cette métrique exploite la parcimonie de groupe dans un domaine transformé. Par ailleurs, elle admet pour borne inférieure la norme nucléaire et possède le même minimum global que cette dernière [11]. Nous proposons donc de reformuler le problème de décomposition CP de rang faible à l’aide de cette métrique. Ceci nous permet de résoudre le dit problème sans aucune connais-

sance a priori du vrai rang. Des simulations numériques et des résultats sur données issues de spectroscopie par fluorescence [1] montrent les avantages de la méthode R-CPD sur les algorithmes ALS [4], ELS-ALS [6] et DIAG [7]. Bien que l'approche R-CPD est ici présentée pour des tenseurs d'ordre 3, elle peut être facilement généralisée aux ordres supérieurs.

## 2 Notations et préliminaires

Tout au long du papier, nous représenterons les tenseurs par une lettre calligraphique en gras, e.g.  $\mathcal{A}$ , les matrices par une lettre majuscule en gras, e.g.  $\mathbf{A}$ , les vecteurs par une lettre minuscule en gras, e.g.  $\mathbf{a}$ , et les scalaires par une lettre minuscule, e.g.  $a$ . Par ailleurs,  $\odot$  et  $\otimes$  désigneront respectivement les opérateurs du produit de Khatri-Rao et de Kronecker. La  $i$ -ième matrice de dépliement du tenseur  $\mathcal{A}$  sera représentée par  $\mathbf{A}^{(i)}$ . Le produit scalaire de  $\mathcal{X}$  et  $\mathcal{Y}$ , supposant que les deux tenseurs ont mêmes dimensions, est défini par :

$$\langle \mathcal{X}, \mathcal{Y} \rangle = \sum_{i_1=1}^{I_1} \sum_{i_2=2}^{I_2} \cdots \sum_{i_N=1}^{I_N} \mathcal{X}_{i_1 i_2 \dots i_N} \mathcal{Y}_{i_1 i_2 \dots i_N}$$

avec  $(\mathcal{X}, \mathcal{Y}) \in \mathbb{R}^{I_1 \times I_2 \times \dots \times I_N} \times \mathbb{R}^{I_1 \times I_2 \times \dots \times I_N}$ . La norme de Frobenius de  $\mathcal{X} \in \mathbb{R}^{I_1 \times I_2 \times \dots \times I_N}$  est définie par :

$$\|\mathcal{X}\|_F = \sqrt{\sum_{i_1 i_2 \dots i_N} (\mathcal{X}_{i_1 i_2 \dots i_N})^2}$$

Soit  $\mathbf{X}$  une matrice de  $\mathbb{R}^{m \times n}$ ,  $\mathbf{X}_{i,\cdot}$  désignera la  $i$ -ième ligne de  $\mathbf{X}$ . La vectorisation de  $\mathbf{X}$  est notée  $\text{vec}(\mathbf{X}) \in \mathbb{R}^{mn \times 1}$ . Les normes nucléaire et mixte de  $\mathbf{X}$  sont respectivement représentées par  $\|\mathbf{X}\|_*$  et  $\|\mathbf{X}\|_{2,1}$ , avec :

$$\|\mathbf{X}\|_{2,1} = \sum_{i=1}^m \sqrt{\sum_{j=1}^n (\mathbf{X}_{i,j})^2} = 2\text{Tr}[\mathbf{X}^\top \Phi \mathbf{X}]$$

où  $\text{Tr}[\cdot]$  désigne l'opérateur de trace et  $\Phi$  une matrice diagonale avec  $\Phi_{i,i} = 1/\sqrt{\sum_{j=1}^n (\mathbf{X}_{i,j})^2}$  représentant la  $(i, i)$ -ième composante de  $\Phi$ . Comme le dénominateur de  $\Phi_{i,i}$  peut être égal ou proche de zéro, il est nécessaire d'ajouter une petite valeur  $\varepsilon$  telle que  $\Phi_{i,i} = 1/(\sqrt{\sum_{j=1}^n (\mathbf{X}_{i,j})^2} + \varepsilon)$ .

## 3 La méthode R-CPD

La décomposition CP de rang  $R$  d'un tenseur  $\mathcal{T} \in \mathbb{R}^{N_1 \times N_2 \times N_3}$  est définie par :

$$\mathcal{T} = \mathcal{A} + \mathcal{B} = \sum_{r=1}^R \mathbf{f}_r^{(1)} \circ \mathbf{f}_r^{(2)} \circ \mathbf{f}_r^{(3)} + \mathcal{B} \quad (1)$$

où  $\mathcal{B} \in \mathbb{R}^{N_1 \times N_2 \times N_3}$  est un tenseur de bruit. Soient  $\mathbf{F}^{(1)} = [\mathbf{f}_1^{(1)}, \dots, \mathbf{f}_R^{(1)}]$ ,  $\mathbf{F}^{(2)} = [\mathbf{f}_1^{(2)}, \dots, \mathbf{f}_R^{(2)}]$  et  $\mathbf{F}^{(3)} = [\mathbf{f}_1^{(3)}, \dots, \mathbf{f}_R^{(3)}]$  les trois matrices de facteur du tenseur  $\mathcal{A}$ . Les matrices  $\mathbf{A}^{(1)} = \mathbf{F}^{(1)} (\mathbf{F}^{(3)} \odot \mathbf{F}^{(2)})^\top$ ,  $\mathbf{A}^{(2)} = \mathbf{F}^{(2)} (\mathbf{F}^{(3)} \odot \mathbf{F}^{(1)})^\top$  et  $\mathbf{A}^{(3)} = \mathbf{F}^{(3)} (\mathbf{F}^{(2)} \odot \mathbf{F}^{(1)})^\top$  désignent alors les trois matrices de dépliement de  $\mathcal{A}$ . Sans perte de généralité, nous supposons dans la suite que  $N_1 \leq N_2 \leq N_3$ , et que  $R \ll N_1$  puisque nous traitons de tenseurs de rang faible.

### 3.1 Formulation du problème

Une caractéristique distinctive de la méthode R-CPD est d'imposer la contrainte de rang faible directement sur les matrices de facteur  $\mathbf{F}^{(i)}$ . D'où la fonction de coût que nous proposons de minimiser dans ce papier :

$$\min_{\mathbf{F}^{(1)}, \mathbf{F}^{(2)}, \mathbf{F}^{(3)}} \sum_{i=1}^3 \text{rang}(\mathbf{F}^{(i)}) \quad \text{s.t.} \quad \mathcal{T} = \mathcal{A} + \mathcal{B} \quad (2)$$

D'après [12, theorem 1], la norme nucléaire  $\|\mathbf{X}\|_*$  peut être utilisée comme enveloppe convexe de  $\text{rang}(\mathbf{X})$ . Néanmoins, elle ne peut pas garantir à la fois un rang faible et la parcimonie de  $\mathbf{X}$  si cette matrice satisfait les deux propriétés. Voyons alors quelle autre métrique utiliser [11] :

**Lemme 1.** *Considérons une matrice  $\mathbf{A} \in \mathbb{R}^{m \times n}$  telle que  $m \geq n$ , sa décomposition en valeurs singulières et sa décomposition dans une base orthonormée sont désignées par  $\mathbf{A} = \mathbf{U}\mathbf{S}\mathbf{V}^\top$  et  $\mathbf{A} = \mathbf{D}\boldsymbol{\alpha}$ , où  $\mathbf{D} \in \mathbb{R}^{m \times n}$ ,  $\boldsymbol{\alpha} \in \mathbb{R}^{n \times n}$  et  $\mathbf{D}^\top \mathbf{D} = \mathbf{I}_n$  sans perte de généralité. Les minima respectifs des mesures de parcimonie de groupe, à savoir row-0 and row-1, de  $\mathbf{A}$  par rapport à  $\boldsymbol{\alpha}$  telles que  $\mathbf{A} = \mathbf{D}\boldsymbol{\alpha}$  et  $\mathbf{D}^\top \mathbf{D} = \mathbf{I}_n$  correspondent à  $\text{rang}(\mathbf{A})$  et  $\|\mathbf{A}\|_*$ .*

Rappelons que la norme row-0 de  $\boldsymbol{\alpha}$  est le nombre de lignes non nulles de  $\boldsymbol{\alpha}$  tandis que row-1 désigne la norme mixte, i.e.  $\|\boldsymbol{\alpha}\|_{\text{row-1}} = \|\boldsymbol{\alpha}\|_{2,1}$ . Nous pouvons alors facilement démontrer la proposition suivante :

**Proposition 1.** *Soit  $\mathbf{F} \in \mathbb{R}^{m \times n}$  une matrice de rang  $R$  telle que  $m \geq n$ . On a alors :*

$$\|\mathbf{D}\mathbf{D}^\top \mathbf{F}\|_* \leq \|\mathbf{D}^\top \mathbf{F}\|_{2,1} \quad (3)$$

où  $\mathbf{D} \in \mathbb{R}^{m \times n}$  est une matrice orthogonale quelconque.

Si  $n = m$ , alors on a  $\mathbf{D}\mathbf{D}^\top = \mathbf{I}$  et la proposition 1 implique  $\|\mathbf{F}\|_* \leq \|\mathbf{D}^\top \mathbf{F}\|_{2,1}$ . D'autre part, quand  $\mathbf{D}$  est la matrice identité, l'inégalité (3) devient  $\|\mathbf{F}\|_* \leq \|\mathbf{F}\|_{2,1}$ . D'après le lemme 1 et la proposition 1, il est alors possible de minimiser la norme nucléaire en minimisant la norme mixte, qui assurera les propriétés de rang faible et de parcimonie de groupe dans la direction des lignes. Nous pouvons de ce fait reformuler le problème de minimisation (2) comme :

$$\min_{\mathbf{F}^{(1)}, \mathbf{F}^{(2)}, \mathbf{F}^{(3)}} \sum_{i=1}^3 \|\mathbf{D}^{(i)\top} \mathbf{F}^{(i)}\|_{2,1} \quad \text{t.q.} \quad \mathcal{T} = \mathcal{A} + \mathcal{B} \quad (4)$$

où les matrices  $\mathbf{D}^{(i)}$  sont carrées, orthogonales et quelconques.

### 3.2 Optimization scheme

Ecrivons le lagrangien augmenté correspondant à (4) :

$$\begin{aligned} \mathcal{L}(\{\mathbf{F}^{(i)}\}, \mathcal{Y}) &= \sum_{i=1}^3 2\lambda_i \text{Tr}[(\mathbf{D}^{(i)\top} \mathbf{F}^{(i)})^\top \Phi^{(i)} (\mathbf{D}^{(i)\top} \mathbf{F}^{(i)})] \\ &\quad + \langle \mathcal{Y}, \mathcal{T} - \mathcal{A} \rangle + \frac{u}{2} \|\mathcal{T} - \mathcal{A}\|_F^2 \end{aligned} \quad (5)$$

où  $\lambda_i$  et  $\mathcal{Y}$  désignent respectivement les paramètres de pénalité et le multiplicateur tensoriel de Lagrange. La fonction  $\mathcal{L}$  est ensuite minimisée à l'aide de la méthode décrite dans [13], adaptée de l'ADMM (Alternating Direction Method of Multipliers) [14]. Plus particulièrement, l'estimée de la première matrice de facteurs obtenue à la  $(k+1)$ -ième itération,  $\mathbf{F}_{(k+1)}^{(1)}$ , est calculée en annulant tout d'abord le gradient de  $\mathcal{L}$  par rapport à  $\mathbf{F}^{(1)}$  :

$$\frac{\partial \mathcal{L}}{\partial \mathbf{F}^{(1)}} = 2\lambda_1 \mathbf{D}^{(1)} \Phi^{(1)} \mathbf{D}^{(1)\top} \mathbf{F}^{(1)} + u \mathbf{F}^{(1)} (\mathbf{F}^{(3)} \odot \mathbf{F}^{(2)})^\top (\mathbf{F}^{(3)} \odot \mathbf{F}^{(2)}) - (\mathbf{Y}^{(1)} + u \mathbf{T}^{(1)}) (\mathbf{F}^{(3)} \odot \mathbf{F}^{(2)}) = \mathbf{0} \quad (6)$$

On utilise ensuite le fait que  $\mathbf{A}\mathbf{X} + \mathbf{X}\mathbf{B} = \mathbf{C}$  implique, d'après l'équation de Lyapunov :

$$\text{vec}(\mathbf{X}) = (\mathbf{I} \otimes \mathbf{A} + \mathbf{B}^\top \otimes \mathbf{I})^{-1} \text{vec}(\mathbf{C}) \quad (7)$$

D'où :

$$\begin{aligned} \text{vec}(\mathbf{F}_{(k+1)}^{(1)}) &= \left( \mathbf{I} \otimes (2\lambda_1 \mathbf{D}^{(1)} \Phi_k^{(1)} \mathbf{D}^{(1)\top}) \right. \\ &\quad \left. + [u_k (\mathbf{F}_k^{(3)} \odot \mathbf{F}_k^{(2)})^\top (\mathbf{F}_k^{(3)} \odot \mathbf{F}_k^{(2)})] \otimes \mathbf{I} \right)^{-1} \\ &\quad \text{vec} \left( (\mathbf{Y}_k^{(1)} + u_k \mathbf{T}^{(1)}) (\mathbf{F}_k^{(3)} \odot \mathbf{F}_k^{(2)}) \right) \end{aligned} \quad (8)$$

Les vecteurs  $\text{vec}(\mathbf{F}_{(k+1)}^{(2)})$  et  $\text{vec}(\mathbf{F}_{(k+1)}^{(3)})$  se déduisent de la même manière :

$$\begin{aligned} \text{vec}(\mathbf{F}_{(k+1)}^{(2)}) &= \left( \mathbf{I} \otimes (2\lambda_2 \mathbf{D}^{(2)} \Phi_k^{(2)} \mathbf{D}^{(2)\top}) \right. \\ &\quad \left. + [u_k (\mathbf{F}_k^{(3)} \odot \mathbf{F}_{(k+1)}^{(1)})^\top (\mathbf{F}_k^{(3)} \odot \mathbf{F}_{(k+1)}^{(1)})] \otimes \mathbf{I} \right)^{-1} \\ &\quad \text{vec} \left( (\mathbf{Y}_k^{(2)} + u_k \mathbf{T}^{(2)}) (\mathbf{F}_k^{(3)} \odot \mathbf{F}_{(k+1)}^{(1)}) \right) \end{aligned} \quad (9)$$

et :

$$\begin{aligned} \text{vec}(\mathbf{F}_{(k+1)}^{(3)}) &= \left( \mathbf{I} \otimes (2\lambda_3 \mathbf{D}^{(3)} \Phi_k^{(3)} \mathbf{D}^{(3)\top}) \right. \\ &\quad \left. + [u_k (\mathbf{F}_{(k+1)}^{(2)} \odot \mathbf{F}_{(k+1)}^{(1)})^\top (\mathbf{F}_{(k+1)}^{(2)} \odot \mathbf{F}_{(k+1)}^{(1)})] \otimes \mathbf{I} \right)^{-1} \\ &\quad \text{vec} \left( (\mathbf{Y}_k^{(3)} + u_k \mathbf{T}^{(3)}) (\mathbf{F}_{(k+1)}^{(2)} \odot \mathbf{F}_{(k+1)}^{(1)}) \right) \end{aligned} \quad (10)$$

La mise à jour du multiplicateur de Lagrange est quant-à-elle donnée par :

$$\mathcal{Y}_{(k+1)} = \mathcal{Y}_k + u_k (\mathcal{T} - \mathcal{A}_{(k+1)}) \quad (11)$$

## 4 Résultats

Deux types d'expériences sont considérées afin de montrer l'intérêt de l'approche R-CPD. Alors que la seconde exploite des données réelles de spectroscopie par fluorescence d'acides aminés [1], la première repose sur l'utilisation de données simulées de la manière suivante :

$$\mathcal{T} = \frac{\mathcal{A}}{\|\mathcal{A}\|_F} + \sigma \frac{\mathcal{B}}{\|\mathcal{B}\|_F} \quad (12)$$

où le paramètre  $\sigma$  contrôle le RSB défini par  $\text{RSB} = -20 \log_{10}(\sigma)$ . À noter que le tenseur  $\mathcal{A} \in \mathbb{R}^{N_1 \times N_2 \times N_3}$  de rang  $R$  est généré à partir de matrices de facteurs  $\mathbf{F}^{(i)} \in \mathbb{R}^{N_i \times R}$ , tirées aléatoirement suivant la loi normale. Le tenseur de bruit  $\mathcal{B}$  est également gaussien. La qualité d'estimation des trois matrices de facteurs  $\mathbf{F}^{(i)}$  du tenseur  $\mathcal{A}$  est étudiée par le biais de la mesure  $\alpha$  définie par les équations [15, (41)] et [15, (42)].

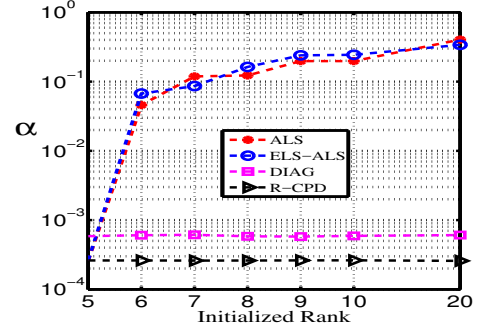


FIGURE 1 – Décomposition CP en fonction de  $\hat{R} \geq R$ .

### 4.1 Expérience sur des tenseurs simulés

Le comportement de R-CPD est étudié sur des tenseurs simulés de taille  $(100 \times 100 \times 100)$  et comparé à celui de l'ALS [4], de l'ELS-ALS [6] et de DIAG [7]. Cette expérience vise à analyser l'influence d'une surestimation  $\hat{R}$  du rang  $R$  utilisée comme initialisation des quatre algorithmes. Ces derniers sont arrêtés dès que l'erreur relative normalisée de la fonction de coût est inférieure à  $10^{-6}$  ou que le nombre d'itérations dépasse 1000. La loi normale est utilisée pour initialiser les matrices de facteurs.

La figure 1 montre la mesure  $\alpha$  moyennée sur 50 réalisations de Monte Carlo en sortie des quatre algorithmes, pour différentes valeurs de  $\hat{R} \geq 5$  alors que la vraie valeur du rang  $R$  est de 5. Le RSB est choisi égal à 10 dB. Toutes les méthodes semblent bien fonctionner si  $\hat{R} = R$  est utilisé comme initialisation. Par contre, seuls DIAG et R-CPD se montrent robustes à une surestimation de  $R$ .

### 4.2 Expérience sur données réelles

Dans cette expérience, les quatre algorithmes sont appliqués à des données de spectroscopie par fluorescence (excitation 240-300 nm, émission 250-450 nm, intervalles de 1nm) mesurées sur cinq échantillons faits main à partir d'un spectrofluoromètre PE LS50B [1]. Chaque échantillon contient différentes quantités de tyrosine, tryptophan et phenylalanine dissouts dans un tampon phosphate. Le tableau d'ordre trois résultant est de taille  $(5 \times 61 \times 201)$ .

La figure 2 montre les facteurs du mode émission calculés pour  $\hat{R} = 3$  et  $\hat{R} = 5$  utilisés comme initialisation alors que le véritable rang  $R$  du tenseur est 3. La première ligne montre un bon comportement des quatre méthodes lorsque le rang est parfaitement estimé. Par contre, seuls DIAG et R-CPD permettent d'estimer avec précision les facteurs attendus pour un rang surestimé égal à 5 comme le montre la seconde ligne. Enfin, seule R-CPD n'estime pas de facteurs fantômes.

## 5 Conclusion

Ce papier présente une méthode de décomposition CP de tenseur, nommée R-CPD, qui permet d'estimer simultanément le rang et les matrices de facteurs du dit tenseur.

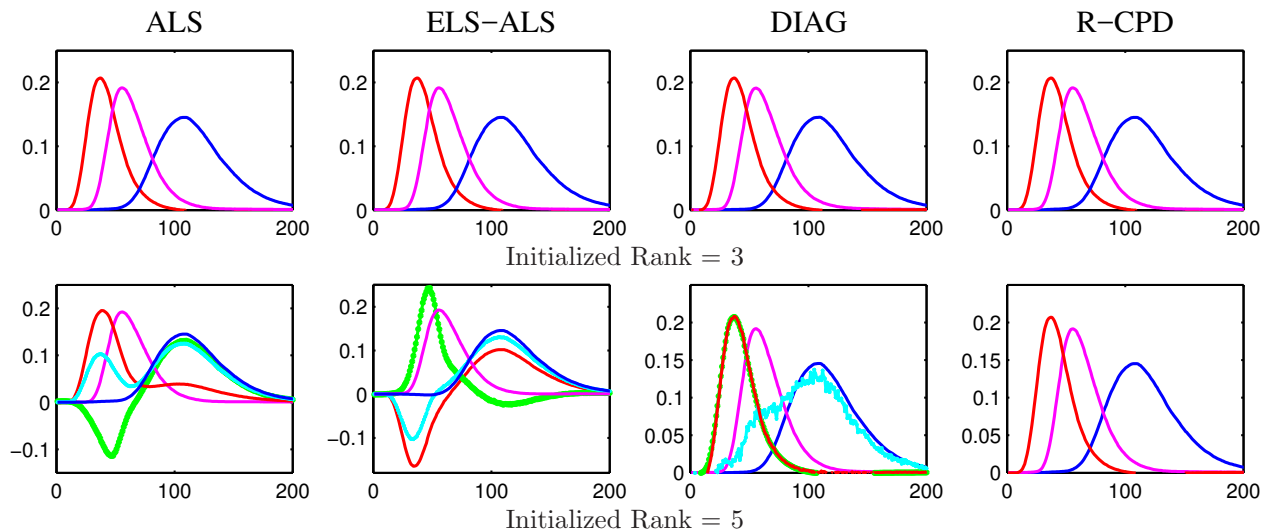


FIGURE 2 – Estimation of emission factors for different initialized ranks.

Une mesure du rang promouvant la parcimonie de groupe dans un domaine transformé est préférée à la norme nucléaire, assurant un coût de calcul moindre. La supériorité de R-CPD sur des algorithmes classiques de la littérature est illustrée à la fois sur données simulées et données réelles dans le contexte de surestimation du rang.

## Références

- [1] R. Bro, “PARAFAC, tutorial and applications,” *Chemometrics and Intelligent Laboratory Systems*, vol. 38, no. 2, pp. 149–171, 1997.
- [2] H. Becker, L. Albera, P. Comon, R. Gribonval, F. Wendling, and I. Merlet, “Brain source imaging : from sparse to tensor models,” *IEEE Signal Processing Magazine, special issue on Brain-Computer Interfaces*, vol. 32, no. 6, pp. 100–112, November 2015.
- [3] A. Fargeas, L. Albera, A. Kachenoura, G. Drean, J.-D. Ospina, J. Coloigner, C. Lafond, J.-B. Delobel, R. De Crevoisier, and O. Acosta, “On feature extraction and classification in prostate cancer radiotherapy using tensor decompositions,” *Elsevier Medical Engineering & Physics*, vol. 37, no. 1, pp. 126–131, January 2015.
- [4] R. A. 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.
- [5] J. Carroll and J. Chang, “Analysis of individual differences in multidimensional scaling via an n-way generalization of eckart-young decomposition,” *Psychometrika*, no. 9, pp. 267–283, 1970.
- [6] M. Rajih, P. Comon, and R. Harshman, “Enhanced line search : A novel method to accelerate PARAFAC,” *SIAM Journal in Matrix Analysis and Applications*, vol. 30, no. 3, pp. 1128–1147, 2008.
- [7] X. Luciani and L. Albera, “Canonical polyadic decomposition based on joint eigenvalue decomposition,” *Chemometrics and Intelligent Laboratory Systems*, vol. 132, pp. 152–167, March 2014.
- [8] X. Luciani and L. Albera, “Joint eigenvalue decomposition of non-defective matrices based on the lu factorization with application to ICA,” *IEEE Transactions On Signal Processing*, vol. 63, no. 17, pp. 4594–4608, September 2015.
- [9] J. Liu, P. Musialski, P. Wonka, and J. Ye, “Tensor completion for estimating missing values in visual data,” *IEEE Transactions Pattern Analysis and Machine Intelligence*, vol. 35, no. 1, pp. 208–220, January 2013.
- [10] B. Recht, M. Fazel, and P. A. Parillo, “Guaranteed minimum-rank solutions of linear matrix equations via nuclear norm minimization,” *SIAM Review*, vol. 52, no. 3, pp. 471–501, 2010.
- [11] X. Shu, F. Porikli, and N. Ahuja, “Robust orthonormal subspace learning : Efficient recovery of corrupted low-rank matrices,” in *2014 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2014, pp. 3874–3881.
- [12] M. Fazel, *Matrix rank minimization with applications*, Ph.D. thesis, Stanford University, 2002.
- [13] Z. Lin, M. Chen, and Y. Ma, “The augmented lagrange multiplier method for exact recovery of corrupted low-rank matrices,” *Technical report*.
- [14] S. Boyd, N. Parikh, E. Chu, B. Peleato, and J. Eckstein, “Distributed optimization and statistical learning via the alternating direction method of multipliers,” *Foundations and Trends in Machine Learning*, vol. 3, no. 1, pp. 1–122, 2011.
- [15] L. Wang, L. Albera, A. Kachenoura, H. Z. Shu, and L. Senhadji, “Canonical polyadic decomposition of 3rd order semi-nonnegative semi-symmetric tensors using LU and QR matrix factorizations,” *EURASIP Journal on Advances in Signal Processing*, vol. 150, October 2014, Special Issue on Recent Advances in Tensor Based Signal and Image Processing.