

Réduction de dimension tensorielle parcimonieuse: Application au clustering de connectivité fonctionnelle

Gaëtan FRUSQUE¹, Julien JUNG³, Pierre BORGNAT², Paulo GONÇALVES¹

¹Univ Lyon, Inria, CNRS, ENS de Lyon, UCB Lyon 1, LIP UMR 5668, F-69342, Lyon, France

²Univ Lyon, ENS de Lyon, UCB Lyon 1, CNRS, Laboratoire de Physique, F-69342 Lyon, France

³HCL, Neuro. Hosp., Functional Neurology and Epileptology Dept & Lyon Neurosc. Res. Cent., INSERM, CNRS
gaetan.frusque@ens-lyon.fr, julien.jung@chu-lyon.fr, pierre.borgnat@ens-lyon.fr, paulo.goncalves@ens-lyon.fr

Résumé – k -means est un algorithme célèbre pour le clustering de données, mais ses performances se dégradent sur des données de grandes dimensions. Nous proposons des décompositions tensorielles parcimonieuses pour réduire la dimension des données avant d’appliquer k -means. Nous illustrons notre méthode sur des mesures de connectivité fonctionnelle d’EEG de crises épileptiques.

Abstract – k -means is famous to cluster a dataset, however it is known to perform badly on high dimensional data. To apply it on EEG functional connectivity measures, as function of the time and for different seizures of a same patient, we develop a new sparse tensorial decomposition to reduce the dimensions of the data before applying k -means.

1 Introduction

Contexte et objectif¹ : L’épilepsie est l’un des troubles neurologiques les plus fréquents dans la population mondiale. Environ 40 % des patients sont résistants aux médicaments et une opération chirurgicale pour extraire la région épileptogène peut être réalisée. Pour localiser cette zone et comprendre l’évolution de l’épilepsie, les praticiens utilisent souvent des enregistrements d’électroencéphalographie intracrânienne (SEEG) [5]. Le patient reste environ une semaine à l’hôpital avec des électrodes implantées afin d’enregistrer plusieurs crises d’épilepsie. Les étapes de la crise se distinguent par des évolutions similaires des signaux SEEG enregistrés en différentes zones du cerveau. Pour quantifier ces similitudes, on calcule les connectivités fonctionnelles (FC) au cours du temps entre toutes les paires de signaux. Rappelons que des mesures de FC classiques sont la cohérence spectrale ou la Phase Locking Value [8]. L’analyse des FC est une tâche complexe et coûteuse, une expérience classique comprenant en moyenne 100 électrodes, soient ≈ 5000 FC à analyser. L’objectif ici est d’aider à regrouper les différentes connectivités fonctionnelles, en identifiant lesquelles ont des comportements communs au cours du temps et se répétant à chaque crise.

Démarche proposée : Partant des séries temporelles d’une ou plusieurs crises, nous calculons les FC par Phase Locking Value. Cela conduit à des données sous forme d’un tenseur d’ordre 3 noté $\mathcal{X} \in \mathbb{R}^{L \times T \times S}$, où L désigne le nombre de FC, T le nombre de pas de temps par crise et S

le nombre de crises. La figure 1 montre les 6 FC calculées entre 4 signaux simulés. Pour regrouper les FC, on utilise l’algorithme k -means [1], où la FC d’indice l a comme attribut $\mathbf{X}_{l,:}$ (matrice correspondant à l’indice l dans le tenseur \mathcal{X}). Cependant, on sait que les performances de cet algorithme se dégradent sur des données de grandes dimensions, ici $T \times S$. Il est alors préférable de réduire au préalable \mathcal{X} à une matrice $\mathbf{F} \in \mathbb{R}^{L \times K}$ de K facteurs représentatifs des FC à travers le temps et les crises. k -means est alors appliqué sur \mathbf{F} plutôt que sur \mathcal{X} . La figure 1 illustre le passage du tenseur \mathcal{X} à une matrice de facteurs réduits (ici $K = 2$). La technique de réduction dimensionnelle la plus connue est la décomposition en valeurs singulières (SVD) [9], cependant, elle ne permet pas systématiquement d’extraire des facteurs débruités communs à toutes les crises. Nous proposons donc une décomposition d’ordre supérieure parcimonieuse : (i) sur le mode temporel afin d’éliminer les périodes communes sans activation et (ii) sur le mode des FC car ceux-ci groupés sur la base des facteurs extraits, s’interprètent plus facilement sur une décomposition parcimonieuse [7].

Après un rappel des notations, la section 2 décrit la SVD et ses extensions, puis introduit la SVD d’ordre supérieure parcimonieuse proposée. En section 3, nous évaluons les performances de la méthode sur un modèle, puis nous l’appliquons à des données réelles.

Notations : Valeur, vecteur, matrice et tenseur sont respectivement notés x , \mathbf{x} , \mathbf{X} et \mathcal{X} . Les colonnes (resp. les lignes) d’une matrice \mathbf{X} sont notées $\mathbf{x}_{:,t}$ (resp. $\mathbf{x}_{l,:}$). Une tranche d’un tenseur, e.g. la crise s de \mathcal{X} , est notée $\mathbf{X}_{::s}$.

1. **Remerciements** : travail du projet ACADEMICS de IDEX-LYON de l’Université de Lyon, PIA et ANR-16-IDEX-0005.

$\| \cdot \|_F$ est la norme de Frobenius, et $\| \cdot \|_1$ la norme ℓ_1 .

Enfin, $\mathbf{X}_{(L)} = [\mathbf{X}_{::1}, \dots, \mathbf{X}_{::S}] \in \mathbb{R}^{L \times TS}$ est la représentation matricielle de \mathcal{X} selon le mode FC (de même pour $\mathbf{X}_{(T)} \in \mathbb{R}^{T \times LS}$ et $\mathbf{X}_{(S)} \in \mathbb{R}^{S \times LT}$ selon les deux autres modes) [3].

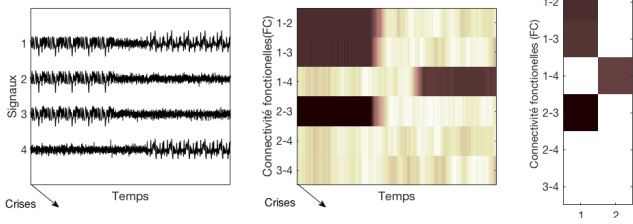


FIGURE 1 – Gauche : 4 signaux SEEG simulés (une crise). Centre : Tenseur \mathcal{X} : FC calculées entre les 6 paires de signaux (PLV). Droite : Réduction du tenseur en une matrice $\mathbf{F} \in \mathbb{R}^{L \times 2}$.

2 Méthodes

2.1 Méthode de regroupement des FC

La méthode consiste tout d'abord à obtenir \mathbf{F} , une réduction dimensionnelle de \mathcal{X} selon K facteurs pertinents. On applique ensuite k -means aux lignes de \mathbf{F} pour identifier N vecteurs $\theta_n \in \mathbb{R}^K$, appelés centroïdes, et solutions de [1] :

$$\operatorname{argmin}_{\theta_1, \dots, \theta_N} \sum_{n=1}^N \sum_{l \in C_n} \| \mathbf{f}_l - \theta_n \|_F^2 \quad (1)$$

La FC d'indice l est affectée au groupe n si le centroïde le plus proche au sens de la norme de Frobenius, est θ_n .

Nous décrivons maintenant, différents types de réductions dimensionnelles pour obtenir \mathbf{F} . Notons, qu'il est possible de traiter directement \mathcal{X} , en appliquant k -means sur sa représentation matricielle $\mathbf{F}_{Direct} = \mathbf{X}_{(L)} = [\mathbf{X}_{::1}, \dots, \mathbf{X}_{::S}]$, correspondant à la concaténation de toutes les crises.

2.2 Réductions dimensionnelles par SVD

SVD : Appliquée à $\mathbf{X}_{(L)}$, la SVD donne $\mathbf{X}_{(L)} = \mathbf{U}\mathbf{\Lambda}\mathbf{Z}^t$, avec $\mathbf{U} \in \mathbb{R}^{L \times L}$, $\mathbf{Z} \in \mathbb{R}^{TS \times TS}$ matrices unitaires et $\mathbf{\Lambda} \in \mathbb{R}^{L \times TS}$ une matrice diagonale positive telle que, si $L \leq TS$, $\lambda_{11} \geq \lambda_{ll} \geq \lambda_{LL}$. La meilleure approximation de rang faible selon la norme de Frobenius de $\mathbf{X}_{(L)}$ est $\mathbf{U}^{(K)}\mathbf{\Lambda}^{(K)}\mathbf{Z}^{(K)t}$ avec $\mathbf{U}^{(K)} = \mathbf{U}_{:[1, \dots, K]}$, $\mathbf{Z}^{(K)} = \mathbf{Z}_{:[1, \dots, K]}$ ainsi que $\mathbf{\Lambda}^{(K)} = \mathbf{U}^{(K)t}\mathbf{X}\mathbf{Z}^{(K)}$. La matrice $\mathbf{X}_{(L)}$ est alors réduite en une matrice $\mathbf{F}_{SVD} = \mathbf{U}^{(K)}\mathbf{\Sigma}^{(K)}$. Le résultat optimal de k -means sur \mathbf{F}_{SVD} est une bonne approximation du résultat optimal sur $\mathbf{X}_{(L)}$ [4]. \mathbf{F}_{SVD} étant de faible dimension, il est facile de répéter k -means pour s'assurer d'atteindre le résultat optimal.

SVD parcimonieuse : Le calcul des matrices $\mathbf{U} \in \mathbb{R}^{L \times K}$ et $\mathbf{Z} \in \mathbb{R}^{TS \times K}$ cherche ici, à satisfaire un compromis entre approximation de rang faible et parcimonie. Les colonnes de \mathbf{U} et \mathbf{Z} ne sont plus orthogonales et il n'est donc plus pertinent d'utiliser une matrice diagonale $\mathbf{\Lambda}^{(K)}$. Les colonnes de \mathbf{U} et de \mathbf{Z} sont renormalisées de telle sorte que

$\| \mathbf{z}_{:,k} \|_F^2 = 1 \forall k \in 1, \dots, K$. La version parcimonieuse de la SVD (notée sSVD) minimise alors le critère :

$$\operatorname{argmin}_{\mathbf{U}, \mathbf{Z}} \| \mathbf{X}_{(L)} - \mathbf{U}\mathbf{Z}^t \|_F^2 + \gamma_1 \| \mathbf{U} \|_1 + \gamma_2 \| \mathbf{Z} \|_1 \quad (2)$$

La réduction est ici obtenue en prenant $\mathbf{F}_{sSVD} = \mathbf{U}$. On peut voir (2) comme l'apprentissage d'un dictionnaire \mathbf{Z} , où la parcimonie est imposée sur le dictionnaire et sur les coefficients. Pour résoudre ce problème, nous utilisons la bibliothèque SPAMS sur matlab [6].

SVD d'ordre supérieur : Les données étant *naturellement* sous la forme d'un tenseur \mathcal{X} , il peut être préférable d'utiliser une version d'ordre supérieur de la SVD plutôt que de concaténer les crises. Il s'agit alors de trouver trois matrices $\mathbf{U} \in \mathbb{R}^{L \times L}$, $\mathbf{V} \in \mathbb{R}^{T \times T}$ et $\mathbf{W} \in \mathbb{R}^{S \times S}$ unitaires et solutions de :

$$\operatorname{argmin}_{\mathbf{U}, \mathbf{V}, \mathbf{W}} \| \mathbf{X}_{(L)} - \mathbf{U}\mathbf{G}_{(L)}(\mathbf{W} \otimes \mathbf{V})^t \|_F^2 \quad (3)$$

où $\mathbf{G}_{(L)} \in \mathbb{R}^{L \times T \times S}$ est une matrice dense (contrairement au cas diagonal de la SVD) et \otimes symbolise le produit de Kronecker. L'algorithme HOSVD [3] permet de résoudre le problème d'optimisation (3). Pour trouver les matrices réduites $\mathbf{U}^{(K_L)} \in \mathbb{R}^{L \times K_L}$, $\mathbf{V}^{(K_T)} \in \mathbb{R}^{T \times K_T}$ et $\mathbf{W}^{(K_S)} \in \mathbb{R}^{S \times K_S}$ minimisant (3) (avec $\mathbf{G}_{(L)} \in \mathbb{R}^{K_L \times K_T \times K_S}$), l'algorithme HOOI [14] permet d'approcher une solution optimale.

2.3 Nouvelles réductions proposées

HOOI modifiée (mHOOI) : On suppose que pour un même patient, la dynamique des FCs est similaire d'une crise à l'autre. Nous proposons alors de fixer $K_S = 1$ pour ne retenir que les activations communes à toutes les crises. Par commodité, on fixe $K_L = K_T = K$. De plus, nous proposons ici, de simplifier l'algorithme HOOI, en alternant à chaque itération i les deux étapes suivantes : (A) le calcul simultané de $\mathbf{U}_{i+1}^{(K)}$, $\mathbf{V}_{i+1}^{(K)}$ connaissant $\mathbf{W}_i^{(1)}$. On calcule pour cela les K premières composantes singulières de $\mathbf{X}_{(L)}(\mathbf{W}_i^{(1)} \otimes \mathbf{I})$ ($\mathbf{I} \in \mathbb{R}^{T \times T}$ est la matrice identité); (B) le calcul de $\mathbf{W}_{i+1}^{(1)}$ connaissant $\mathbf{U}_{i+1}^{(K)}$ et $\mathbf{V}_{i+1}^{(K)}$ en calculant le vecteur singulier gauche dominant de $\mathbf{X}_{(S)}(\mathbf{V}_{i+1}^{(K)} \otimes \mathbf{U}_{i+1}^{(K)})$. Nous avons observé empiriquement, une convergence vers la solution optimale plus rapide avec notre algorithme (si le calcul des SVD ne dégénère pas [13]), qu'avec HOOI. L'algorithme 1 détaille le calcul de \mathbf{F}_{mHOOI} issu de la réduction mHOOI proposée.

HOOI modifiée parcimonieuse (HOsSVD) : Nous proposons également une version parcimonieuse de mHOOI, en remplaçant dans l'étape (A), le calcul de la SVD par la SVD parcimonieuse (2). La matrice \mathbf{U}_{i+1} n'étant pas forcément orthogonale, on procède à sa décomposition QR et seuls les K premiers vecteurs colonnes de \mathbf{Q} sont retenus. On procède de même pour \mathbf{V}_{i+1} et on note $qr(\mathbf{V}_{i+1}, K)$ cette opération. L'algorithme 2 présente le calcul de \mathbf{F}_{HOsSVD} obtenu avec cette nouvelle réduction HOsSVD. Notons que cette décomposition est proche d'une *Block Term Decomposition* à un terme [2] avec contrainte de parcimonie. Il s'agit

Algorithm 1 Estimation de \mathbf{F}_{mHOOI}

Require: \mathcal{X} , K , tolérance angulaire $\epsilon > 0$, et i_{max} .

```
 $i = 0$   
 $\mathbf{W}_0^{(1)} = \text{left-1-SVD}(\mathbf{X}_{(3)})$   
while  $i < i_{max}$  or  $\text{acos}(\langle \mathbf{W}_i^{(1)}, \mathbf{W}_{i-1}^{(1)} \rangle) > \epsilon$  do  
  (A).  $[\mathbf{U}_{i+1}^{(K)}, \mathbf{\Lambda}_{i+1}, \mathbf{V}_{i+1}^{(K)}] = K\text{-SVD}(\mathbf{X}_{(S)}(\mathbf{W}_i^{(1)} \otimes \mathbf{I}))$   
  (B).  $\mathbf{W}_{i+1}^{(1)} = \text{left-1-SVD}(\mathbf{X}_{(L)}(\mathbf{V}_{i+1}^{(K)} \otimes \mathbf{U}_{i+1}^{(K)}))$   
   $i = i + 1$   
end while  
 $\mathbf{F}_{mHOOI} = \mathbf{U}_i^{(K)} \mathbf{\Lambda}_i$ 
```

aussi d'une variante de la décomposition de Tucker parcimonieuse avec une matrice $\mathbf{G}_{(L)}$ diagonale [12]. Bien que nous n'ayons pas de preuve de convergence, les résultats empiriques sont encourageants (voir section 3).

Algorithm 2 Estimation de \mathbf{F}_{HOsSVD}

Require: \mathcal{X} , K , γ_1 , γ_2 , tolérance angulaire $\epsilon > 0$, et i_{max} .

```
 $i = 0$   
 $\mathbf{W}_0^{(1)} = \text{left-1-SVD}(\mathbf{X}_{(3)})$   
while  $i < i_{max}$  or  $\text{acos}(\langle \mathbf{W}_i^{(1)}, \mathbf{W}_{i-1}^{(1)} \rangle) > \epsilon$  do  
  (A).  $\mathbf{U}_{i+1}, \mathbf{V}_{i+1}$  minimisent (2) selon  $\mathbf{X}_{(S)}(\mathbf{W}_i^{(1)} \otimes \mathbf{I})$ .  
   $\mathbf{A} = \text{qr}(\mathbf{U}_{i+1}, K)$ ,  $\mathbf{B} = \text{qr}(\mathbf{V}_{i+1}, K)$   
  (B).  $\mathbf{W}_{i+1}^{(1)} = \text{left-1-SVD}(\mathbf{X}_{(L)}(\mathbf{B} \otimes \mathbf{A}))$   
   $i = i + 1$   
end while  
 $\mathbf{F}_{HOsSVD} = \mathbf{U}_i^{(K)}$ 
```

3 Evaluation et application

3.1 Evaluation sur un modèle

Modèle : Afin d'évaluer et de comparer nos différentes méthodes de réduction, nous proposons un modèle simple de la dynamiques caractéristiques des FC d'une crise épileptique. Dans la matrice $\mathbf{X} \in \mathbb{R}^{L \times T}$ obtenue, $x_{lt} = 0.75$ si la FC $l \in \{1, \dots, L\}$ est active au temps $t \in \{1, \dots, T\}$, et $x_{lt} = 0.25$ si elle est inactive. Toutes les FC d'un même groupe $n \in 1, \dots, N$ s'activent sur la même période de temps T_n . A ce comportement déterministe, nous rajoutons trois sources d'aléa : (i) les durées d'activation T_n sont uniformément distribuées entre 1 et $\frac{T}{N}$ (la variable $\alpha = \{0, 1\}$ permet de choisir le mode certain ou aléatoire); (ii) chaque FC active d'un groupe donné a une probabilité β d'être remplacée par une autre FC choisie aléatoirement; (iii), on rajoute aux coefficients x_{lt} un bruit blanc gaussien de puissance σ^2 (SNR= σ^{-2}). Enfin, S réalisations i.i.d. avec le même jeu de paramètres $b = [\alpha, \beta, \sigma]$, simulent les crises et forment le tenseur $\mathcal{X} \in \mathbb{R}^{L \times T \times S}$. La figure 2a correspond à la réalisation $\mathbf{X}_{(L)}$ de 4 crises simulées.

Expériences et résultats : On compare la méthode appliquée directement au tenseur, i.e. sur $\mathbf{F}_{Direct} = \mathbf{X}_{(L)}$, à ses différentes réductions \mathbf{F}_{SVD} , \mathbf{F}_{sSVD} , \mathbf{F}_{mHOOI} , \mathbf{F}_{HOsSVD} et aux crises moyennée $\mathbf{F}_{Mean} = \sum_{s=1}^S \mathbf{X}_{::s}$ (les hyperparamètres $\epsilon = 10^{-3}$, $\gamma_1 = 0.1$ et $\gamma_2 = 4$ sont optimisés

empiriquement). Pour les réductions, K varie de 1 à 4 et on ne retient que la valeur de K offrant le meilleur regroupement. L'algorithme k -means++ [1] est stoppé au bout de 1000 itérations et relancé sur 120 réalisations de \mathbf{F} . Le nombre de groupes est fixé à $N = 4$. Le score ARI (Adjusted Rand Index) [11] calculé entre chaque regroupement obtenu et la vérité-terrain, vaut 1 si les deux regroupements sont identiques, et 0 s'ils ne se correspondent pas plus qu'un groupement aléatoire. La fig. 2b présente la moyenne et l'écart type des scores ARI pour chaque méthode appliquée à 120 réalisations. On considère 6 combinaisons possibles du jeu de paramètres d'aléas b . On observe que les méthodes avec réduction dimensionnelle ont de meilleures performances que les méthodes standards. Les méthodes tensorielles ont de meilleures performances que leurs homologues matriciels, de même que les méthodes parcimonieuses surpassent globalement les méthodes classiques. Enfin, HOsSVD se distingue systématiquement, surtout pour des SNR faibles.

3.2 Application à des données réelles

Données : On considère des données SEEG réelles d'un patient souffrant d'épilepsie focale. Les électrodes utilisées sont réparties sur des tiges implantées en profondeur dans le cerveau. L'activité du cerveau est enregistrée via 5 à 10 électrodes par tige. L'espace entre deux électrodes consécutives est de 3.5 mm. Sur une durée d'enregistrement de 15 jours, 4 crises ont été sélectionnées. Chaque crise est délimitée en temps par une fenêtre de 100 secondes centrée sur le début de la crise. Le signal est échantillonné à 256 Hz. Sur les 108 contacts, nous en choisissons 33 équitablement répartis pour éviter les corrélations spatiales trop fortes. La mesure de connectivité fonctionnelle utilisée est la PLV [8]. Une PLV forte entre deux signaux signifie que leurs phases sont similaires. Les 528 FC ont été calculées par pas de 1 seconde sur une fenêtre rectangulaire glissante de 4 secondes. Après élimination des points soumis aux effets de bord, les données peuvent être représentées par un tenseur $\mathcal{X} \in \mathbb{R}^{528 \times 96 \times 4}$.

Application : A partir du tenseur \mathcal{X} , on obtient la matrice \mathbf{F}_{HOsSVD} (algorithme 2), où les paramètres, $K = 4$, $\gamma_1 = 1$, $\gamma_2 = 1$ et $\epsilon = 10^{-3}$ sont fixés empiriquement. La figure 3 (haut) montre les profils d'activation des composantes de \mathbf{F}_{HOsSVD} (correspondant à la matrice $\mathbf{V}^{(K)}$ issue de l'algorithme 2). On observe 4 périodes d'activation que l'on peut aisément associer aux périodes d'avant, de début, de propagation et de fin de crise. Aucun facteur n'est actif autour de 50 secondes, ce qui pourrait correspondre au découplage fonctionnel du début de crise [10], une courte période où les activités SEEG en différentes zones du cerveau sont subitement décorréées. On applique k -means sur \mathbf{F}_{HOsSVD} afin d'identifier les 5 groupes de FC correspondants ($N = 5$). Comme \mathbf{F}_{HOsSVD} est une matrice orthogonale et parcimonieuse, on retient les 4 groupes de plus petites tailles, que l'on associe à chacune des 4 pé-

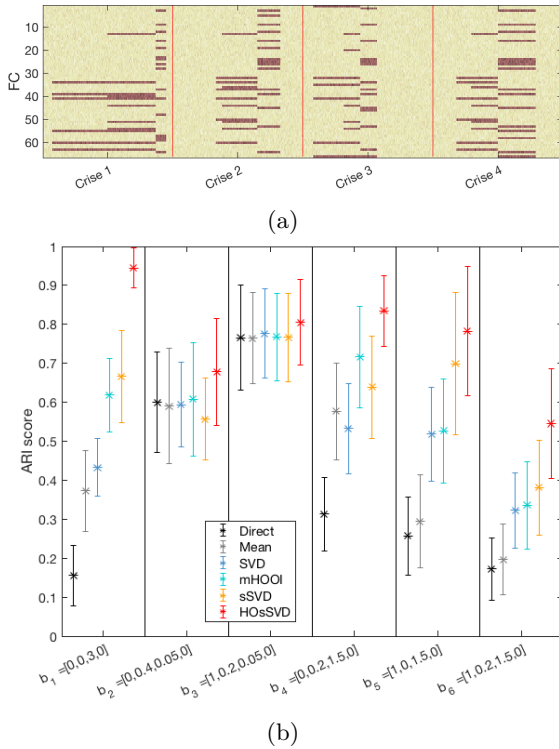


FIGURE 2 – (a) - Réalisation du modèle avec 4 crises simulées pour $b = [1, 0.2, 0.05]$, (b) ARI scores de chaque méthode pour 6 modèles associés à des vecteurs bruits b_i différents.

riodes d'activation; le 5ème groupe correspondant aux FC non synchronisées. La fig. 3 (bas) matérialise les positions des 33 électrodes projetées sur le plan transverse (selon le repaire de Talairach). Une FC est représentée par un lien entre les deux électrodes en phase. Chacun des 4 groupes de FC apparaît alors comme un graphe variant dans le temps : avant et en début de crise, l'activité est localisée dans le demi-plan $y < 0$ (hémisphère gauche du cerveau), puis s'étend aux $y > 0$ (hémisphère droit du cerveau). Les graphes obtenus sont en accord avec les résultats cliniques, le foyer du début de crise se trouvant proche des graphes "Avant crise" et "Début Crise". La propagation des crises dans l'hémisphère droit du cerveau est bien représentée par les graphes "Propagation" et "Fin de crise".

Conclusion : La méthode HOsSVD proposée est une alternative intéressante pour la réduction de données tensorielles de grandes dimensions. Appliquée en pré-traitement des données brutes, elle permet d'améliorer significativement les performances de clustering de k -means. Appliquée à des données réelles de SEEG enregistrés pendant une crise épileptique, notre méthode nous a permis d'identifier les 4 groupes d'activation de FC correspondant aux 4 périodes de temps significatives de l'évolution de la crise. Dans les perspectives immédiates de ce travail, nous envisageons d'automatiser le choix des hyperparamètres pour une analyse plus systématique et plus exhaustive des performances comparées. Cela nous permettra ensuite d'analyser plus rigoureusement la dynamique des connectivités fonctionnelles sur un jeu de données cliniques plus riche.

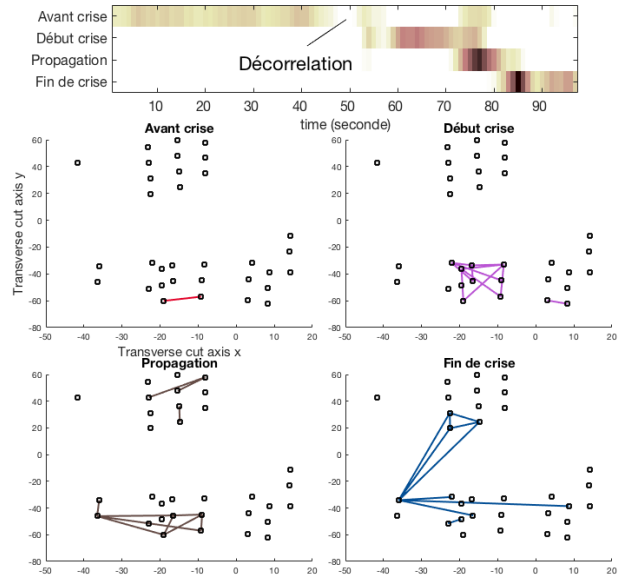


FIGURE 3 – (haut) - Profil d'activation de $HOsSVD$ (variable V de l'algorithme 2), (bas) Groupes de liens correspondant aux 4 périodes d'activations des crises du patient considéré.

Références

- [1] D. Arthur and S. Vassilvitskii. k -means++ : The advantages of careful seeding. In *Proc. of the 18th annual ACM-SIAM symp. on Discrete algo.*, 2007.
- [2] Lieven De Lathauwer. Decompositions of a higher-order tensor in block terms—part ii : Definitions and uniqueness. *SIAM J. on Matrix Analysis and Applications*, 30(3) :1033–1066, 2008.
- [3] De Lathauwer et al. A multilinear singular value decomposition. *SIAM J. on Matrix Anal. and App.*, 21(4), 2000.
- [4] Feldman et al. Turning big data into tiny data : Constant-size coresets for k -means, pca and projective clustering. In *Proc. of the 24th ACM-SIAM symp. on Dis. algo.*, 2013.
- [5] Guenot et al. Neurophysiological monitoring for epilepsy surgery : the Talairach SEEG method. *Stereotactic and functional neurosurgery*, 77(1-4), 2001.
- [6] Mairal et al. SPAMS : A SPARSE Modeling Software. URL <http://spams-devel.gforge.inria.fr/>, 2014.
- [7] Sorzano et al. A survey of dimensionality reduction techniques. *arXiv preprint arXiv :1403.2877*, 2014.
- [8] Van Mierlo et al. Functional brain connectivity from EEG in epilepsy : Seizure prediction and epileptogenic focus localization. *Progress in neurobiology*, 121, 2014.
- [9] Wall et al. Singular value decomposition and principal component analysis. In *A pract. approach to microarray data anal.* 2003.
- [10] Wendling et al. Epileptic fast intracerebral EEG activity : evidence for spatial decorrelation at seizure onset. *Brain*, 126(6), 2003.
- [11] L. Hubert and P. Arabie. Comparing partitions. *Journal of classification*, 2(1), 1985.
- [12] Morten Mørup, Lars Kai Hansen, and Sidse M Arnfred. Algorithms for sparse nonnegative Tucker decompositions. *Neural computation*, 20(8) :2112–2131, 2008.
- [13] Y. Xu. On the convergence of higher-order orthogonal iteration. *Linear and Multilinear Algebra*, 66(11), 2018.
- [14] A. Zhang and D. Xia. Tensor SVD : Statistical and Computational Limits. *IEEE Transactions on Info. Th.*, 2018.