Robust Tensor Decomposition of Resting Brain Networks in Stereotactic EEG

Jian Li, *Student Member IEEE*, John C. Mosher, *Senior Member IEEE*, Dileep R. Nair,
Jorge Gonzalez-Martinez, Richard M. Leahy, *Fellow IEEE*

Abstract—Stereotactically implanted Electro-Encephalography (SEEG) in patients with epilepsy provides a unique insight into spontaneous human brain activity. Exploring dynamic functional connectivity in spontaneous SEEG signals provides a rich framework for studying brain networks. Tensor decomposition is a powerful tool for decoding dynamic networks, capturing the intrinsic interactions between multiple dimensions with less restrictive constraints than traditional 2D matrix decomposition methods such as PCA and ICA. Tensor decomposition, however, is seldom used for decoding large resting brain datasets due to its high computational complexity and poor robustness. In this paper, we describe a Scalable and Robust Sequential Canonical Polyadic Decomposition (SRSCPD) framework that can sequentially and robustly identify tensor models of successively higher rank. We demonstrate that SRSCPD is not only more robust than the popular Alternating Least Square (ALS) algorithm, but can also be extended to large-scale problems.

Index Terms—Tensor decomposition, dynamic functional connectivity, stereotactic EEG, optimization

I. INTRODUCTION

Tensor decomposition, including both the Canonical Polyadic (CP) and Tucker models, is a natural model for data that can be represented as a multi-dimensional array. Standard (2D) matrix decomposition methods such as principal component analysis (PCA) and independent component analysis (ICA) applied to "matricized" or unfolded tensor data [1] are not able to capture the intrinsic interactions and couplings across dimensions. Moreover, the CP decomposition has a unique solution under milder conditions than the orthogonality or independence assumptions implicit in PCA or ICA [2], [3]. Therefore, when analyzing brain networks, the structure inherent in tensor analysis allows us to avoid the perhaps unrealistic assumption that networks behave independently of each other.

The CP decomposition can be computed via several algorithms. Among these, alternating least square (ALS) is widely used because of its relative simplicity [4]. ALS-based CP decomposition has been widely used in EEG analysis by transforming the raw EEG recordings to a time-frequency representation using either short-time Fourier transforms or wavelets, and applying a 3-way CP decomposition (channel by time by frequency) [5]. Möck [6] applied CP to event-related potentials (ERP). Miwakeichi *et al.* [7] analyzed both spontaneous and evoked EEG recordings and showed that theta activity was predominant during a task condition, while alpha activity was observed continuously during both rest and task

conditions. CP decomposition has also been applied to ictal EEG recordings from patients with epilepsy. The extracted components have been used to localize the seizure onset zone [8] as well as to remove artifacts [9].

Exploring functional connectivity (FC) in spontaneous or resting brain signals offers a rich framework for studying brain networks [10]. Of particular recent interest is the dynamic nature of functional connectivity [11]. The most commonly used strategy for decoding dynamic FC (DFC) is to compute the correlation or coherence using a sliding window. Using sufficient samples to obtain robust FC estimates inevitably leads to over-smoothing of dynamic changes. Alternative methods, including various ICA-based approaches are reviewed in [11]. A limitation of ICA is that either the time series of each network are required to be independent (temporal ICA) or the spatial modes of the networks are disjoint (spatial ICA), whereas real networks can overlap and be correlated in both space and time [12]. In contrast, CP decomposition does not impose any specific constraint on any domain, and since the decomposition is performed directly on the tensor representation of the raw data, the temporal smoothing associated with sliding window methods is avoided. Thus CP decomposition is a potentially powerful tool for exploring DFC, although we are unaware of any literature explicitly combining CP and DFC. Two main issues, however, have limited the use of the CP model for studying DFC.

Scalability: Previous studies, such as [7]–[9], [13], [14], almost always truncate the data into short temporal segments (up to a few thousand time samples) in EEG ictal and event related potential recordings, in order to make the CP decomposition tractable. However, a typical 10-minute stereotactic EEG recording consists of at least 120,000 time samples, even after decimating to a sampling rate of 200 Hz. In order to compute CP decompositions on data of this size, a fast and efficient algorithm is required.

Robustness: It is well known that the ALS algorithm is not guaranteed to converge to a global minimum or a stationary point, even when multi-start is applied during the optimization [1], [4]. The local minimum problem becomes more severe as the number of components increases. Performance is further compromised when a larger number of components than necessary are fit to the data, resulting in over-factoring, i.e. splitting rank-one components into two or more factors.

Several techniques have been explored to improve the robustness and efficiency of the ALS algorithm. For example,

Rajih *et al.* [15] added a line search after each major ALS iteration. Navasca *et al.* [16] applied Tikhonov regularization on each sub-problem in the ALS iteration. These modifications result in significantly higher computational cost at each iteration, limiting their practical utility, particularly for large scale problems.

II. NOTATION AND PRELIMINARIES

A. CP Decomposition

CP decomposes a tensor into a sum of rank-one tensors or components. For a three-order tensor $\mathcal{X} \in \mathbb{R}^{I \times J \times K}$

$$\mathcal{X} = \sum_{r=1}^{R} a_r \circ b_r \circ c_r + \mathcal{E} \tag{6}$$

where $a_r \in \mathbb{R}^I$, $b_r \in \mathbb{R}^J$, $c_r \in \mathbb{R}^K$, R is the *rank* or the number of components, $a \circ b$ denotes the outer product between a and b and $\mathcal{E}$ is the error tensor. If we group the components in each mode into a matrix, i.e. let $A = [a_1 \, a_2 \cdots a_R] \in \mathbb{R}^{I \times R}$ and similar for $B \in \mathbb{R}^{J \times R}$ and $C \in \mathbb{R}^{K \times R}$, then the CP decomposition can be expressed as

$$X_{(1)} = A(C \odot B)^T + E_{(1)} \tag{7}$$

or

$$X_{(2)} = B(C \odot A)^T + E_{(2)} \tag{8}$$

or

$$X_{(3)} = C(B \odot A)^T + E_{(3)} \tag{9}$$

where A, B, C are called the *loading matrices* for the three modes respectively and $A \odot B$ represents the Khatri-Rao product of matrix A and B.

B. Computation of CP decomposition and the ALS algorithm

Suppose we want to find the best rank R approximation of $\mathcal{X} \in \mathbb{R}^{I \times J \times K}$ via

$$\min_{\widehat{\mathcal{X}}} \|\mathcal{X} - \widehat{\mathcal{X}}\| + g(\widehat{\mathcal{X}}) \tag{10}$$

where $\widehat{\mathcal{X}} = \sum_{r=1}^{R} \lambda_r \, a_r \circ b_r \circ c_r$, λ_r represents the scale of component r, $g(\widehat{\mathcal{X}}) = \mu_1 g_1(A) + \mu_2 g_2(B) + \mu_3 g_3(C)$ is a data-dependent regularizer with (μ_1, μ_2, μ_3) the corresponding regularization parameters. The ALS algorithm solves this problem in an iterative fashion. We first solve for A with B and C fixed, then solve for B with A and C fixed, and so on. This procedure is repeated until some convergence criterion is satisfied. Note that each sub-problem reduces to an ordinary least square problem. Specifically, let's assume B and C are fixed and we are solving for A. Using the equivalent matrix expression discussed above, we can write the optimization problem as

$$\widehat{A} = \underset{A}{\operatorname{argmin}} \|X_{(1)} - A(C \odot B)^T\|_F + \mu_1 g_1(A) \tag{11}$$

The solution with $\mu_1 = 0$ (without regularization) reduces to a regular least square solution:

$$\widehat{A} = X_{(1)}[(C \odot B)^T]^\dagger \tag{12}$$

where $Y^\dagger$ is the Moore-Penrose pseudo inverse of Y. Using a property of the Khatri-Rao product [4], we can rewrite as

$$\widehat{A} = X_{(1)}(C \odot B)(C^T C * B^T B)^\dagger \tag{13}$$

where $A * B$ represents the Hadamard product of matrix A and B. This expression is almost always preferable to (12) because it achieves a much lower computational complexity by only calculating the pseudo-inverse of an $R \times R$ matrix. Finally, we normalize each component and store the norm in λ to avoid scale ambiguity. For the case $\mu_1 \neq 0$, the solution in (12) is replaced by the solution to (11), which will be closed form if $g_1(A)$ is quadratic, but may require iterative solution in other cases. The full ALS algorithm is shown in Algorithm I.

ALGORITHM I: CP-ALS

Algorithm *CP-ALS* $(\mathcal{X}, R, \{A^*, B^*, C^*, \lambda^*\})$
  *Initialize $A \in \mathbb{R}^{I \times R}, B \in \mathbb{R}^{J \times R}, C \in \mathbb{R}^{K \times R}, \lambda \in \mathbb{R}^R$
  While not converged†
  $\quad A \leftarrow \underset{A}{\operatorname{argmin}} \|X_{(1)} - A(C \odot B)^T\|_F + \mu_1 g_1(A)$
  $\quad B \leftarrow \underset{B}{\operatorname{argmin}} \|X_{(2)} - B(C \odot A)^T\|_F + \mu_2 g_2(B)$
  $\quad C \leftarrow \underset{C}{\operatorname{argmin}} \|X_{(3)} - C(B \odot A)^T\|_F + \mu_3 g_3(C)$
  Normalize A, B, C and store the norms in λ
  End While
  Return A, B, C and λ
End Algorithm

---

* The initialization is typically performed using either random matrices or the R leading singular vectors of the matricized $\mathcal{X}$. We define specific initializations below for our SRSCPD algorithm.
† Algorithm convergence is determined when the sum of the l_1 norm of the difference between two adjacent iterations in all modes is less than some small constant, e.g. 10^{-5}.

III. METHODS

Eckart and Young [17] showed that the best rank-r approximation of a matrix with respect to the Frobenius norm is given by the leading r factors of the SVD. This is not the case for CP decomposition of a higher-order tensor. Kolda *et al.* [18] showed an example where the best rank-1 approximation is not part of the best rank-2 approximation of a tensor. As a result, all components in the CP decomposition for a given desired rank have to be found simultaneously.

The determination of tensor rank is NP-hard [19]. Many metrics have been proposed to help us find the correct rank, e.g. the core consistency diagnostic (CORCONDIA) [20], difference in fit (DIFFIT) [21] and automatic relevance determination (ARD) [22]. However, all these metrics require a set of decomposition results for all ranks up to the maximum rank R. Obtaining such a set of solutions using CP decomposition is quadratically ($O(R^2)$) more complex than finding a rank-1 approximation. This represents a significant challenge to use of higher rank tensor models and was a primary motivation for our development of the SRSCPD framework.

The SRSCPD framework is built on the original ALS algorithm. Our goal is to compute a rank-recursive set of

decompositions from rank 1 to rank R. Our approach uses the result for rank r to initialize the decomposition for rank $r + 1$. Initialization for the additional component is found by fitting a rank one tensor to the residual from the rank r fit. While this is a relatively straightforward approach, it has not previously been described and the "warm start" greatly improves the convergence speed of the ALS algorithm relative to alternatives as well as helps to avoid local minima in this non-convex optimization problem. As a result, we are able to address the problems with robustness and scalability for large-scale datasets.

The full SRSCPD framework is shown in Algorithm II, for a three-order tensor example. The inputs of the algorithm are a tensor $\mathcal{X} \in \mathbb{R}^{I \times J \times K}$ and the desired maximum rank R. For each iteration r, a rank-r approximation is calculated using the original *CP-ALS* algorithm with initializations $\{A^*, B^*, C^*, \lambda^*\}$. The initializations are formed by concatenating the solutions $\{A^{r-1}, B^{r-1}, C^{r-1}, \lambda^{r-1}\}$ from the previous $(r-1)$ recursion with the rank-1 approximation $\{a', b', c', \lambda'\}$ of the residue tensor $\mathcal{X}_{res}$, where $\mathcal{X}_{res}$ is obtained by subtracting the reconstructed tensor using $\{A^{r-1}, B^{r-1}, C^{r-1}, \lambda^{r-1}\}$ from the original data tensor $\mathcal{X}$.

SRSCPD is flexible in the sense that techniques that has been proposed to improve the ALS algorithm can be directly incorporated. For example, one can add a line search at the end of each major iteration of ALS. Moreover, data-dependent constraints and regularization terms can be applied to each of the ALS sub-problems, e.g. non-negativity, sparsity, and smoothness.

ALGORITHM II: SRSCP-ALS

Algorithm *SRSCP-ALS* $(\mathcal{X}, R)$

$\quad a^1, b^1, c^1, \lambda^1 \leftarrow CP\text{-}ALS\,(\mathcal{X}, 1)$

$\quad \mathcal{X}_{res} \leftarrow \mathcal{X} - Tensor\_Recon\,(a^1, b^1, c^1, \lambda^1)$

$\quad a', b', c', \lambda' \leftarrow CP\text{-}ALS\,(\mathcal{X}_{res}, 1)$

$\quad A^* \leftarrow [a^1\ a']; B^* \leftarrow [b^1\ b']; C^* \leftarrow [c^1\ c']; \lambda^* \leftarrow \begin{bmatrix} \lambda^1 \\ \lambda' \end{bmatrix}$

$\quad$ For $r = 2, 3, \dots, R$

$\quad\quad A^r, B^r, C^r, \lambda^r \leftarrow CP\text{-}ALS\,(\mathcal{X}, r, \{A^*, B^*, C^*, \lambda^*\})$

$\quad\quad \mathcal{X}_{res} \leftarrow \mathcal{X} - Tensor\_Recon\,(A^r, B^r, C^r, \lambda^r)$

$\quad\quad a', b', c', \lambda' \leftarrow CP\text{-}ALS\,(\mathcal{X}_{res}, 1)$

$\quad\quad A^* \leftarrow [A^r\ a']; B^* \leftarrow [B^r\ b']; C^* \leftarrow [C^r\ c']; \lambda^* \leftarrow \begin{bmatrix} \lambda^r \\ \lambda' \end{bmatrix}$

$\quad$ End For

$\quad$ Return a set of solutions

$\quad\quad \{a^1, b^1, c^1, \lambda^1\}, \{A^2, B^2, C^2, \lambda^2\}, \dots, \{A^R, B^R, C^R, \lambda^R\}$

End Algorithm

IV. EXPERIMENT AND RESULTS

We simulated stereotactic EEG data [23] with 100 channels, 200 Hz sampling rate, 2 second duration. We assume that there are R true underlying networks or components. In each network, spatially, a total of N channels are co-activated, where N is chosen randomly between 2 and 10. Temporally, we
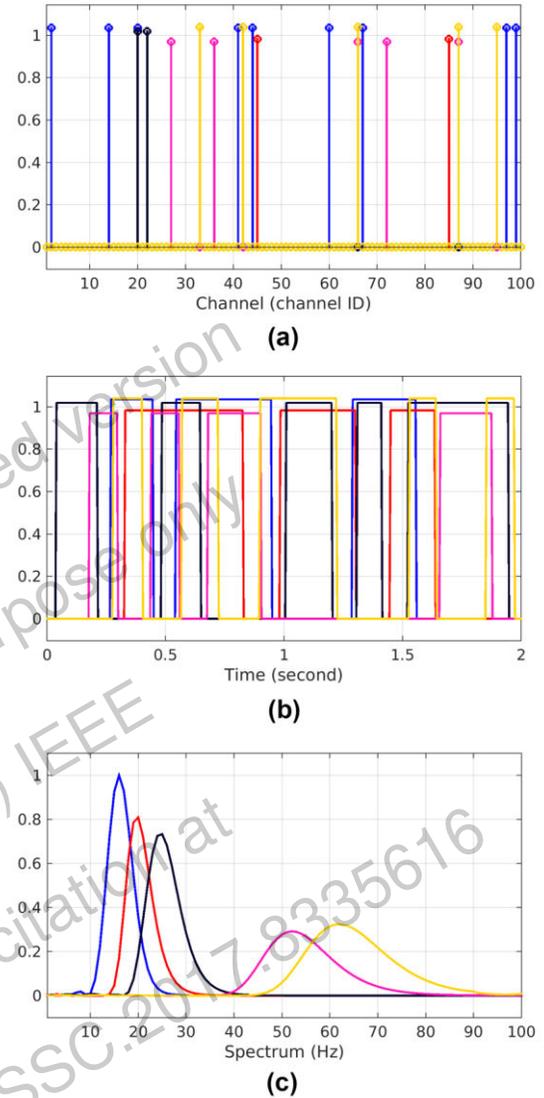

Figure 1: An example of the simulated data with 5 components. Each component is represented by a distinct color in all three modes. (a) The channel (spatial) mode shows the activated channels that participate in each network; (b) The time (temporal) mode shows the block activation pattern for each network; (c) The spectrum (spectral) mode shows the frequency spectrum for each network.

assume the network is either "on" or "off", in an interleaved manner that forms a block activation pattern. The number of activated blocks is selected randomly between 2 and 5 and both the minimum block length and the minimum interval between any adjacent activated blocks are set to be 0.1 second. Spectrally, the true signals in each network are sinusoidal with frequencies chosen randomly between 10 and 80 Hz. Finally, we add white Gaussian noise to the simulated data with a range of signal to noise ratios ($SNRs$).

The order-three tensor $\mathcal{X}$ is generated by calculating the power of the Morlet wavelet transform (MWT) of the simulated data matrix with center frequency 1 Hz and a time-frequency index of two in a linearly spaced frequency range of 1 to 100 Hz with interval 1 Hz. Thus, the final tensor $\mathcal{X}$ has the dimensions of $\mathbb{R}^{I \times J \times K}$, where $I = 100, J = 400, K = 100$. An example of the model used to simulate the data is shown in Fig.
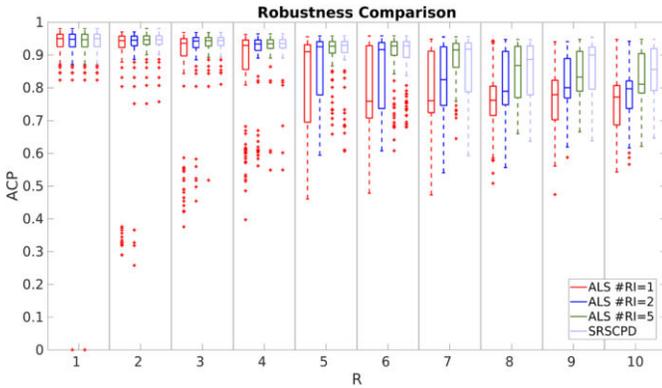
Figure 2: Simulation results. Boxplots of ACP over 100 Monte Carlo trials are shown as a function of R. "#RI" denotes the number of random initializations when using original ALS algorithm.

1. Note that overlaps between components may occur in any of the three modes.

We first compared the robustness of the decomposition using the SRSCPD framework against ALS using 1, 2 and 5 randomly selected initializations. The same convergence criterion was used for both algorithms and in all cases we computed solutions from rank 1 to R. In both algorithms we used a non-negativity constraints on all modes, because the power of Morlet wavelet coefficients are naturally non-negative and the constraint helps avoid degeneracy [24]. Let $A \in \mathbb{R}^{I \times R}$, $B \in \mathbb{R}^{J \times R}$, $C \in \mathbb{R}^{K \times R}$ be the loading matrix in each of the three modes as described in Section II. Then in each sub-problem of the ALS, we used the following cost function for A (likewise for B and C)

$$\widehat{A} = \underset{A}{\mathrm{argmin}} \left\| X_{(1)} - A(C \odot B)^T \right\|_F \quad s.t.\, A \geqslant 0 \quad (14)$$

where "$\geqslant$" denotes the element-wise inequality.

Since we know the ground truth under the simulated settings, we assessed the quality of the solutions using the averaged congruence product [25]. Let A, B, C be the column-wise normalized ground truth loading matrices and $\widehat{A}, \widehat{B}, \widehat{C}$ their estimated counterparts. Then the averaged congruence product is defined as

$$ACP = \max_{P} \mathrm{tr}((A^T\widehat{A} * B^T\widehat{B} * C^T\widehat{C})\, P) \quad (15)$$

where P is a permutation matrix accounting for the ambiguity of the ordering of the solutions [26] and $\mathrm{tr}(X)$ indicates the
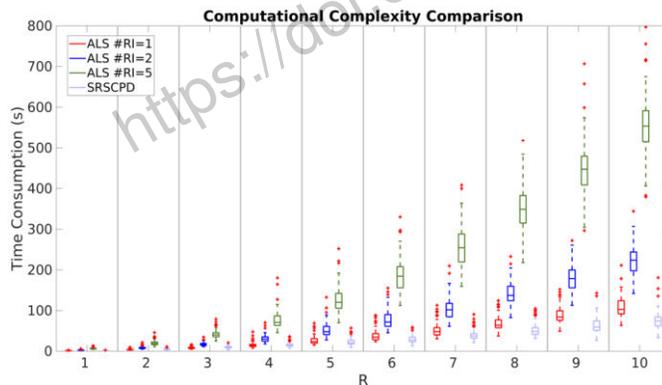

Figure 3: Simulation results. Boxplots of the run time in seconds over 100 Monte Carlo trials are shown as a function of R. "#RI" denotes the number of random initializations when using original ALS algorithm.
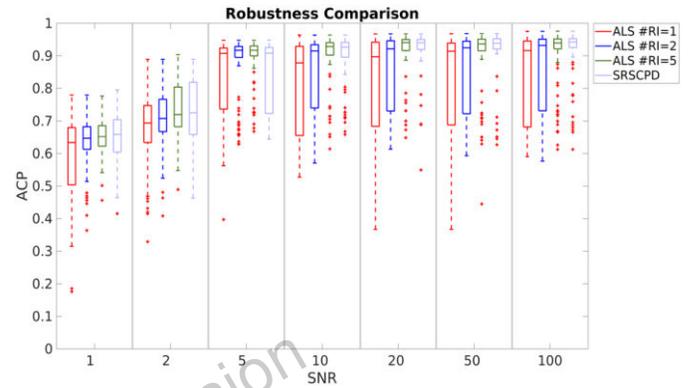

Figure 4: Simulation results. Boxplots of the ACP over 100 Monte Carlo trials are shown as a version of SNR. "#RI" denotes the number of random initializations when using original ALS algorithm.

trace of X.

We evaluated the ACP of the solutions obtained from both ALS and SRSCPD as a function of R for $SNR = 10$. For each R, we ran 100 Monte Carlo trials and boxplots of ACP were generated. For each simulated tensor, we repeated ALS M times, where $M = 1, 2$, and 5, each time using a different random initialization. The final solution was selected as that which has the lowest cost. We also recorded the computational cost for each of the methods. We also repeated the above study, but instead of varying R we conducted the experiment as a function of SNR with $R = 5$.

Fig. 2 shows performance of ALS vs SRSCPD as a function of rank R. For small R all results are similar. However, for larger we see that the ALS results are strongly dependent on initialization, that performance for $RI = 5$ is significantly better than for $RI = 3$ and $RI = 1$. SRSCPD benefits from using the results of the lower rank as an initialization, resulting in overall improved performance relative to all three versions of ALS.

Fig. 3 shows the computation cost as a function of R. As expected, the ratios among the ALS methods are approximately proportional to RI, the number of different initializations. The cost of SRSCPD is significantly lower than that for ALS, $RI = 3$ and 5. As the rank increases, the cost for SRSCPD is even lower than ALS without restart. The reason for this is that the initialization with the result from the lower rank, not only produces improved performance (Fig. 2) but also faster convergence of the ALS sub-problems.

Fig. 4 shows that as the SNR increases, ACP also improves for all methods. SRSCPD shows generally similar performance to ALS with $RI = 5$ restarts and is significantly better than results for $RI = 1$. However, for lower SNRs, performance of ALS with $RI = 5$ restarts is superior to SRSCPD.

V. CONCLUSION

Identification of dynamic functional connectivity in single subjects in resting-brain data is a difficult problem due to low SNR and the spatio-temporal complexity of the data. Traditional sliding window-based correlation approaches cannot achieve optimality in both temporal resolution and robustness of the estimation of the correlation. On the other hand, PCA- or ICA-based methods impose orthogonality or

independence constraints on the data, which may be inconsistent with the underlying physiological processes that govern network dynamics.

In this paper, we have described a SRSCPD framework based on the original ALS algorithm, aiming at robustly identifying dynamic functional connectivity using a tensor decomposition. This framework is scalable to large datasets due to its use of a warm start of the optimization problem for each tensor rank. Using simulations, we have shown that SRSCPD consistently outperforms the multi-start ALS algorithm over a range of ranks and SNRs, with lower computation cost.

REFERENCES

[1] A. Cichocki *et al.*, "Tensor decompositions for signal processing applications: From two-way to multiway component analysis," *IEEE Signal Process. Mag.*, vol. 32, no. 2, pp. 145–163, 2015.

[2] J. B. Kruskal, "Rank, decomposition, and uniqueness for 3-way and N-way arrays," *Multiway Data Analysis*. pp. 7–18, 1989.

[3] N. D. Sidiropoulos and R. Bro, "On the uniqueness of multilinear decomposition of N-way arrays," *J. Chemom.*, vol. 14, no. 3, pp. 229–239, 2000.

[4] T. G. Kolda and B. W. Bader, "Tensor Decompositions and Applications," *SIAM Rev.*, vol. 51, no. 3, pp. 455–500, 2009.

[5] F. Cong, Q.-H. Lin, L.-D. Kuang, X.-F. Gong, P. Astikainen, and T. Ristaniemi, "Tensor decomposition of EEG signals: A brief review," *J. Neurosci. Methods*, vol. 248, pp. 59–69, 2015.

[6] J. Möcks, "Topographic components model for event-related potentials and some biophysical considerations," *IEEE Trans. Biomed. Eng.*, vol. 35, no. 6, pp. 482–484, 1988.

[7] F. Miwakeichi, E. Martínez-Montes, P. A. Valdés-Sosa, N. Nishiyama, H. Mizuhara, and Y. Yamaguchi, "Decomposing EEG data into space-time-frequency components using Parallel Factor Analysis," *Neuroimage*, vol. 22, no. 3, pp. 1035–1045, 2004.

[8] E. Acar, C. Aykut-Bingol, H. Bingol, R. Bro, and B. Yener, "Multiway analysis of epilepsy tensors," *Bioinformatics*, vol. 23, no. 13, pp. 10–18, 2007.

[9] W. Deburchgraeve *et al.*, "Neonatal seizure localization using PARAFAC decomposition," *Clin. Neurophysiol.*, vol. 120, no. 10, pp. 1787–1796, 2009.

[10] K. J. Friston, "Functional and Effective Connectivity: A Review," *Brain Connect.*, vol. 1, no. 1, pp. 13–36, 2011.

[11] R. M. Hutchison *et al.*, "Dynamic functional connectivity: Promise, issues, and interpretations," *Neuroimage*, vol. 80, pp. 360–378, 2013.

[12] F. I. Karahanoğlu and D. Van De Ville, "Transient brain activity disentangles fMRI resting-state dynamics in terms of spatially and temporally overlapping networks," *Nat. Commun.*, vol. 6, p. 7751, 2015.

[13] H. Lee, Y.-D. Kim, A. Cichocki, and S. Choi, "Nonnegative Tensor Factorization for Continuous EEG Classification," *Int. J. Neural Syst.*, vol. 17, no. 4, pp. 305–317, 2007.

[14] K. Vanderperren *et al.*, "Single trial ERP reading based on parallel factor analysis," *Psychophysiology*, vol. 50, no. 1, pp. 97–110, 2013.

[15] M. Rajih, P. Comon, and R. A. Harshman, "Enhanced Line Search: A Novel Method to Accelerate PARAFAC," *SIAM J. Matrix Anal. Appl.*, vol. 30, no. 3, pp. 1128–1147, 2008.

[16] C. Navasca, L. De Lathauwer, and S. Kindermann, "Swamp reducing technique for tensor decomposition," in *European Signal Processing Conference*, 2008, no. Eusipco.

[17] C. Eckart and G. Young, "The approximation of one matrix by another of lower rank," *Psychometrika*, vol. 1, no. 3, pp. 211–218, 1936.

[18] T. Kolda, "Orthogonal Tensor Decompositions," *SIAM J. Matrix Anal. Appl.*, vol. 23, no. 1, pp. 243–255, 2001.

[19] J. Håstad, "Tensor rank is NP-complete," *J. Algorithms*, vol. 11, no. 4, pp. 644–654, 1990.

[20] R. Bro and H. A. L. Kiers, "A new efficient method for determining the number of components in PARAFAC models," *J. Chemom.*, vol. 17, no. 5, pp. 274–286, 2003.

[21] M. E. Timmerman and H. a Kiers, "Three-mode principal components analysis: choosing the numbers of components and sensitivity to local optima.," *Br. J. Math. Stat. Psychol.*, vol. 53 (Pt 1), pp. 1–16, 2000.

[22] M. Mørup and L. K. Hansena, "Automatic relevance determination for multi-way models," *J. Chemom.*, vol. 23, no. 7–8, pp. 352–363, 2009.

[23] J. Bancaud and J. Talairach, *La Stéréo-électroencéphalographie dans l'épilepsie : Informations neurophysiopathologiques apportées par l'investigation fonctionnelle stéréotaxique*. Masson, 1965.

[24] L. Lim, "Optimal solutions to non-negative PARAFAC / multilinear NMF always exist," in *Workshop on Tensor Decompositions and Applications, Centre International de rencontres Math{é}matiques, Luminy, France*, 2005.

[25] G. Tomasi and R. Bro, "PARAFAC and missing values," *Chemom. Intell. Lab. Syst.*, vol. 75, no. 2, pp. 163–180, 2005.

[26] R. a Harshman, "Foundations of the PARAFAC procedure: Models and conditions for an 'explanatory' multimodal factor analysis," *UCLA Work. Pap. Phonetics*, vol. 16, no. 10, pp. 1–84, 1970.