



**HAL**  
open science

# Dynamic brain effective connectivity analysis based on low-rank canonical polyadic decomposition: application to epilepsy

Pierre-Antoine Chantal, Ahmad Karfoul, Anca Nica, Régine Le Bouquin Jeannès

## ► To cite this version:

Pierre-Antoine Chantal, Ahmad Karfoul, Anca Nica, Régine Le Bouquin Jeannès. Dynamic brain effective connectivity analysis based on low-rank canonical polyadic decomposition: application to epilepsy. *Medical and Biological Engineering and Computing*, 2021, 59 (5), pp.1081-1098. 10.1007/s11517-021-02325-x . hal-03222625

**HAL Id: hal-03222625**

**<https://hal.science/hal-03222625>**

Submitted on 20 Sep 2021

**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.

---

# DYNAMIC BRAIN EFFECTIVE CONNECTIVITY ANALYSIS BASED ON LOW-RANK CANONICAL POLYADIC DECOMPOSITION: APPLICATION TO EPILEPSY

---

A PREPRINT

**Pierre-Antoine Chantal**  
Univ Rennes, INSERM  
LTSI-UMR 1099  
F-35000 Rennes, France  
pierreantchantal@gmail.com

**Ahmad Karfoul**  
Univ Rennes, INSERM  
LTSI-UMR 1099  
F-35000 Rennes, France  
ahmad.karfoul@univ-rennes1.fr

**Anca Nica**  
Univ Rennes, CHU Rennes, INSERM  
LTSI-UMR 1099  
F-35000 Rennes, France  
anca.nica@chu-rennes.fr

**Régine Le Bouquin Jeannès**  
Univ Rennes, INSERM  
LTSI-UMR 1099  
F-35000 Rennes, France  
regine.le-bouquin-jeannes@univ-rennes1.fr

April 22, 2021

## ABSTRACT

In this paper, a new method to track brain effective connectivity networks in the context of epilepsy is proposed. It relies on the combination of partial directed coherence with a constrained low-rank canonical polyadic tensor decomposition. With such combination being established, the most dominating directed graph structures underlying each time window of intracerebral electroencephalographic signals are optimally inferred. Obtained time and frequency signatures of inferred brain networks allow respectively to track the time evolution of these networks and to define frequency bands on which they are operating. Besides, the proposed method allows also to track brain connectivity networks through several epileptic seizures of the same patient. Understanding the most dominating directed graph structures over epileptic seizures and investigating their behavior over time and frequency plans are henceforth possible. Since only few but the the most important directed connections in the graph structure are of interest and also for a meaningful interpretation of obtained signatures to be guaranteed, the low-rank canonical polyadic tensor decomposition is prompted respectively by the sparsity and the non-negativity constraints on the tensor loading matrices. The main objective of this contribution is to propose a new way of tracking brain networks in the context of epileptic iEEG data by identifying the most dominant effective connectivity patterns underlying the observed iEEG signals at each time window. The performance of the proposed method is firstly evaluated on simulated data imitating brain activities and secondly on real intracerebral electroencephalographic signals obtained from an epileptic patient. The partial directed coherence based tensor has been decomposed into space, time and frequency signatures in accordance with the expected ground truth for each consecutive sequence of the simulated data. The method is also in accordance with the clinical expertise for iEEG epileptic signals, where the signatures were investigated through a simultaneous multi-seizure analysis.

Accepted manuscript

**Keywords** Effective connectivity · Canonical polyadic decomposition · Partial directed coherence

## 1 Introduction

Epilepsy is a neurologic network disease affecting between 0.5 and 1 percent of the world population. Epileptic seizures arise from abnormal hypersynchronous electrical activity within a neuronal network. Around 25% of the epileptic patients suffer from drug-resistant epilepsy and cannot be rendered free of seizures using anti-epileptic drugs. Consequently, different options have to be considered, such as epilepsy surgery guided, when needed, by invasive intracerebral investigations [7], vagus nerve stimulation [8, 18], cortical [54] and brain stimulation [57] or epilepsy surgical procedures [9]. The choice of these procedures is decided by a medical team with the help of a pre-surgical evaluation that can benefit from the combination of different modalities. If the surgery solution is chosen by the medical expert, two different procedures may sometimes be adopted, either completely removing the epileptogenic focus (resective surgery) or cutting the nervebundles connected to the epileptogenic focus in order to prevent the seizure spreading without damaging underlying functionalities (disconnective surgery). Choosing surgery act requires sometimes intracerebral investigation of interactions between brain regions in order to identify the source(s) of the seizure onset and potentially its propagation. To this end, intracerebral electroencephalographic (iEEG) signals are used to measure the cerebral activity before and during the seizure. This invasive technique is widely used for epilepsy understanding [6] as it ensures a relatively high signal to noise ratio (SNR) compared to surface electroencephalographic recordings and is able to highlight 3D network brain activity that cannot be detected with other techniques. Intracerebral EEG signals allow to identify the epileptogenic network without facing the ill-posed inverse source localization problem that is common with scalp EEG.

The understanding of the electrical neural information obtained with iEEG recordings relies on the inference of a connectivity graph as a mean to establish the connections between different brain regions. Brain connectivity describes how brain regions or neuronal populations are connected and how the information between these regions is interchanged. In order to identify the information flow between brain regions during the seizure, it is essential to measure and analyze this brain connectivity which can be distinguished in three types. Structural connectivity refers to anatomical connections in the neuronal network and can be measured using only invasive tracing techniques such as Diffusion Tensor Imaging [3]. Functional connectivity describes statistical dependency between electrical activities of different brain regions whatever the anatomical connections. This statistical dependency can be inferred for instance by measuring correlation, covariance, phase locking or coherence [25] and consequently can be described through a symmetric connectivity matrix. This matrix reflects in its turn an underlying undirected connectivity graph. Finally, effective connectivity describes the causal influence that one neuronal region exerts on another one [25] and can be seen as complementary to both functional and structural connectivities. The effective relationships respect the physical causality principle, *i.e.* the causes must precede the effects. Then, it seems natural to use time series to infer this phenomenon as suggested by Granger [27]. In the context of epileptic network identification, the interest is to estimate an asymmetric connectivity matrix describing the causal connections between the different regions of neuronal populations resulting in a directed graph. This paper addresses the inference of a directed graph  $\mathcal{G}(N_v, E)$  composed of  $N_v$  nodes (neurons or neuronal populations) and a set of  $E$  edges or links (corresponding to synapses or pathways).

Different approaches for effective connectivity estimation have been considered in the literature [5] and can be divided into two classes: model-based and model-free (*i.e.* data-driven) ones. Regarding the model-based approaches which include physiological knowledge, they rely on the choice of the best generative model underlying the observed data among predefined model candidates. The most well-known techniques belonging to this model-based class are the structural Equation Modeling (SEM) [59] and the Dynamic Causal Modeling (DCM) [23, 24] (see [48] for details). As far as the data-driven approaches are considered, they are often based on Wiener-Granger causality [27] (see [25, 26] for a complete review). Based on a MultiVariate AutoRegressive (MVAR) modeling of the observed signals, the goal of the Granger method is to determine whether one time series is useful in forecasting another one. A common linear approach to

estimate effective connectivity is the Directed Transfer Function (DTF) [2] but the limitation of this technique comes from its inability to distinguish between direct and indirect connections. Another often used method is the Directed COherence (DCOH) introduced in [49]. This method not only infers the temporal pattern of EEG signals but also their spatial relationship by modeling the generations of different signals with common sets of white noise. The Partial Directed Coherence (PDC) initially proposed in [50] is commonly used as an efficient frequency-based technique to describe causal relationships among time series together with a good robustness to indirect causal relations [4, 15]. Unlike DCOH, PDC is useful to understand the propagation of the seizure when taking into account more than two channels and has been implemented in this study.

Tensor Canonical Polyadic Decomposition (CPD) initially proposed by Hitchcock in 1927 [31] and known later as PARAllel FACtor analysis (PARAFAC) [11, 29, 14, 35] has attracted significant interest in widespread applications such as telecommunications [52, 53], image-completion [60] and biomedical signal processing [44]. Since its introduction, properties of CPD such as uniqueness, stability have been extensively studied. Furthermore, a variety of methods to fit the CPD model have also been developed and various constraints have been explored to obtain meaningful solutions and to avoid degeneracy. CPD has already proven its interest in the analysis of electroencephalographic signals [17, 32, 40, 58, 41]. Following an MVAR modeling, a general framework for tensor analysis of model inversion and multimodal data fusion was presented [32]. Moreover, tensor decomposition has already been used to linearly separate multivariate time-frequency data into mode channels, time and frequency [40, 58, 41]. Interesting performance in tracking brain connectivity is based on coupling the CPD model with a sample-wise time-variant PDC computed on surface EEG signals [45]. However, in addition to its high numerical complexity, this approach is subject to the volume conduction [43] effect and relies on four classical rank estimation methods [1, 51, 12] to insure reliable rank estimation of the constructed PDC-based tensor. In the present study, the PDC measure computed from iEEG signals is combined with a low rank tensor decomposition to provide an efficient way to track the evolution of brain connectivity along the epileptic seizure. More precisely, the PDC measure is combined with a CPD model but with a rank that, contrary to the strategy adopted in [45], is optimally computed as recently suggested in [28, 55]. The PDC is computed on sliding time windows leading to an easy implementation and reasonable numerical complexity. In addition to this single seizure analysis, the proposed approach allows for a simultaneous space, time and frequency analysis of brain connectivity through several recorded epileptic seizures.

This paper is organized as follows: Section 2 is devoted to the presentation of the mathematical background and the useful tools used in the proposed method described in Section 3. Experiments and results are reported in Section 4. Our findings are discussed in Section 5 before giving some concluding remarks in Section 6.

## 2 Background

### 2.1 Notations and mathematical background

Throughout this paper, vectors, matrices and tensors are denoted respectively by bold lower case  $\mathbf{a}$ , bold upper case  $\mathbf{A}$  and bold calligraphic  $\mathcal{A}$  letters. Entries of vectors, matrices or tensors are denoted with the same notation as the structure they belong to but with a normal case ( $a_i$ ,  $A_{i,j}$ , or  $\mathcal{A}_{i,j,k}$ ). In addition,  $\circ$ ,  $\otimes$  and  $\odot$  stand respectively for the outer product, the Kronecker product and the Khatri-Rao product (column-wise Kronecker product) [33]. The operator  $vec(\cdot)$  stands for the matrix-to-vector transformation where, for a matrix  $\mathbf{A}(I \times J)$ , the  $i + (j - 1)I$ -th component of  $vec(\mathbf{A})$  corresponds to the  $(i, j)$ -th entry,  $A_{i,j}$ , of  $\mathbf{A}$ . Furthermore,  $\mathbf{I}_M$  is the identity matrix of size  $(M \times M)$ ,  $\mathbf{0}$  is a vector of zeros and  $\mathbf{0}_{I \times J}$  denotes a matrix of zeros of size  $(I \times J)$ . The operators  $\top$  and  $\dagger$  denote respectively for the transpose and the conjugate transpose operators and  $Tr(\cdot)$  denotes the matrix trace. For  $\mathbf{A}(M \times N)$  and  $\mathbf{B}(M \times N)$ ,  $\mathbf{X} = \max(\mathbf{A}, \mathbf{B})$  is a matrix

whose  $(i, j)$ -th entry is computed as  $X_{i,j} = \max(A_{i,j}, B_{i,j})$  where the max operator returns the value of its highest entry. The  $\ell_{2,1}$  norm of a matrix  $\mathbf{A} (I \times J)$  is defined by:

$$\|\mathbf{A}\|_{2,1} = \sum_{i=1}^I \sqrt{\sum_{j=1}^J A_{i,j}^2} = \text{Tr}(\mathbf{A}^\top \Gamma_{\mathbf{A}} \mathbf{A}) \quad (1)$$

where  $\Gamma_{\mathbf{A}}$  is a diagonal matrix such that

$$\Gamma_{\mathbf{A},i,i} = \frac{1}{\sqrt{\sum_{j=1}^J A_{i,j}^2}} \quad (2)$$

## 2.2 Basic definitions

**Definition 1.** The outer product  $\mathbf{u}^{(1)} \circ \dots \circ \mathbf{u}^{(N)}$  of  $N$  vectors  $(\mathbf{u}^{(n)})_{1 \leq n \leq N} \in \mathbb{R}^{I_n}$  is a rank-1 tensor of size  $(I_1 \times \dots \times I_N)$  whose  $(i_1, \dots, i_N)$ -th element in  $I_1 \times \dots \times I_N$  is given by  $\mathbf{u}_{i_1}^{(1)} \dots \mathbf{u}_{i_N}^{(N)}$ .

**Definition 2.** The CPD of an  $N$ -dimensional tensor  $\mathcal{X} (I_1 \times \dots \times I_N)$ , denoted by  $\llbracket \mathbf{A}^{(1)}, \mathbf{A}^{(2)}, \dots, \mathbf{A}^{(N)} \rrbracket$ , is defined as the minimal sum of  $N$ -dimensional rank-1 tensors such that:

$$\mathcal{X} = \llbracket \mathbf{A}^{(1)}, \mathbf{A}^{(2)}, \dots, \mathbf{A}^{(N)} \rrbracket = \sum_{r=1}^R \mathbf{a}_r^{(1)} \circ \dots \circ \mathbf{a}_r^{(N)} \quad (3)$$

where  $\mathbf{a}_r^{(n)}$  is the  $r$ -th column of the  $n$ -th loading matrix  $\mathbf{A}^{(n)}$  associated with the  $n$ -th direction of  $\mathcal{X}$  and  $R$  is the tensor rank.

**Definition 3.** Let  $\mathcal{X} (I_1 \times I_2 \times \dots \times I_N)$  be an  $N$ -dimensional tensor. Then, unfolding this tensor according to its  $n$ -th direction results in an unfolding matrix, denoted here by  $\mathbf{X}_{(n)}$  of size  $(I_n \times I_{n+1} I_{n+2} \dots I_N I_1 I_2 \dots I_{n-1})$ , such that [33]:

$$\mathbf{X}_{(n),i_n,j} = \mathcal{X}_{i_1,i_2,\dots,i_N} \quad (4)$$

with  $j = 1 + \sum_{\substack{k=1 \\ k \neq n}}^N (i_k - 1) \prod_{\substack{m=1 \\ m \neq n}}^{k-1} I_m$ .

Then, equation (4) can be written in a matrix form as follows:

$$\mathbf{X}_{(n)} = \mathbf{A}^{(n)} (\mathbf{A}^{(n-1)} \circ \mathbf{A}^{(n-2)} \circ \dots \circ \mathbf{A}^{(1)} \circ \mathbf{A}^{(N)} \circ \dots \circ \mathbf{A}^{(n+2)} \circ \mathbf{A}^{(n+1)})^\top \quad (5)$$

## 2.3 Partial Directed Coherence

Let  $\mathbf{X}^w (N_v \times M)$  be a spatio-temporal matrix of time series (epileptic iEEG recordings) observed on  $N_v$  sensors (iEEG electrodes) over a window of  $M$  samples. As brain connectivity patterns are changing over time, inferring its related connectivity graph is generally realized on sufficiently small epochs obtained from the observed time series where the connectivity graph is supposed to be stationary. These time epochs are defined using a sliding window of wisely chosen width,  $M$ , such that the stationarity assumption of graph connectivity is verified. Note that the choice of an adequate width  $M$  is a real challenge since this parameter is highly related to the data and to the application context (epilepsy, cognitive task, ...). Now, let  $\mathbf{x}^w(m)$  be the  $m$ -th ( $m \in \{1, \dots, M\}$ )  $N_v$ -dimensional column vector of the  $w$ -th ( $w \in \{1, \dots, N_w\}$ ) spatio-temporal matrix obtained at the  $w$ -th sliding window where  $M$  and  $N_w$  denote the width of the sliding window and the number of overlapped windows used for the entire time series, respectively. An MVAR modeling of  $\mathbf{x}^w(m)$  can then be written as:

$$\mathbf{x}^w(m) = \sum_{l=1}^{L_w} \mathbf{D}_l^w \mathbf{x}^w(m-l) + \mathbf{e}^w(m) \quad (6)$$

where  $\mathbf{D}_l^w(N_v \times N_v)$  is the matrix of the MVAR model coefficients associated with the time lag  $l$ ,  $\mathbf{e}^w(m)$  is an  $N_v$ -dimensional column vector of white Gaussian noise whose  $n$ -th component,  $e_n^w(m)$ , is a standardized Gaussian random variable and  $L_w$  stands for the model order. The latter parameter is estimated for each time window using the Akaike Information Criterion (AIC) [1]. Other techniques for the model order estimation can be used instead (see [19] for more details). The cross power spectral density  $\mathbf{S}^w(f)(N_v \times N_v)$  of  $\mathbf{x}^w(m)$  defined in equation (6) at the frequency bin  $f$  is then defined by the following equation [47]:

$$\mathbf{S}^w(f) = \mathbf{H}^w(f)\mathbf{H}^{w\dagger}(f) \quad (7)$$

where

$$\mathbf{H}^w(f) = (\overline{\mathbf{D}}^w(f))^{-1} = (\mathbf{I}_{N_v^2} - \mathbf{D}^w(f))^{-1} \quad (8)$$

with

$$\mathbf{D}^w(f) = \sum_{l=1}^{L_w} \mathbf{D}_l^w e^{-j2\pi fl}. \quad (9)$$

Let  $\overline{D}_{i,j}^w(f)$  be the  $(i, j)$ -th entry of  $\overline{\mathbf{D}}^w(f) = [\overline{\mathbf{d}}_1^w(f), \overline{\mathbf{d}}_2^w(f), \dots, \overline{\mathbf{d}}_{N_v}^w(f)]$ . Then, the PDC measured between the  $i$ -th and  $j$ -th nodes (sensors)  $1 \leq i \neq j \leq N_v$  at the frequency bin  $f$  is given by:

$$\tilde{\mathcal{P}}_{i,j}^w(f) = \frac{\overline{D}_{i,j}^w(f)}{\sqrt{(\overline{\mathbf{d}}_i^w(f))^\dagger \overline{\mathbf{d}}_j^w(f)}}, \quad 1 \leq w \leq N_v \quad (10)$$

where  $1 \leq f \leq N_f$  with  $N_f$  standing for the number of frequency bins.

### 3 Method

#### 3.1 Construction of the PDC-based tensor

It is easy to note from equation (10) that the module of each measured PDC value  $\tilde{\mathcal{P}}_{i,j}^w(f)$  can be considered as the  $(k, w, f)$ -th entry of a  $3^{rd}$  order space $\times$ time $\times$ frequency PDC-based tensor  $\mathcal{P}$  of size  $((N_v^2 - N_v) \times N_w \times N_f)$ , such that:

$$\mathcal{P}_{k,w,f} = |\tilde{\mathcal{P}}_{i,j}^w(f)| \quad (11)$$

where:

$$k = \begin{cases} i + (j - 1)(N_v - 1) - 1 & \text{if } i > j \\ i + (j - 2)(N_v - 1) - 1 & \text{if } i < j \end{cases} \quad (12)$$

According to equation (11), the PDC-based tensor captures the variability of the PDC measure over time and frequency directions. It is noteworthy that this tensor can be further extended to capture the potential PDC variability over several epileptic seizures of the same patient. In this case, the  $3^{rd}$  order PDC tensor can be extended to a  $4^{th}$  order space $\times$ time $\times$ frequency $\times$ seizure PDC-based one. More precisely, let  $\mathcal{P}^c$  be the space $\times$ time $\times$ frequency PDC-based tensor defined in equation (11) for the  $c$ -th seizure ( $c \in \{1, \dots, N_c\}$ ) where  $N_c$  is the number of recorded epileptic seizures, then,  $\mathcal{P}^c$  can be considered as the  $c$ -th sub-tensor of the  $4^{th}$  order space $\times$ time $\times$ frequency $\times$ seizure PDC-based one  $\mathcal{Q}$  of size  $((N_v^2 - N_v) \times N_w \times N_f \times N_c)$  such that  $\mathcal{Q}_{k,w,f,c} = \mathcal{P}_{k,w,f}^c$ .

### 3.2 Tensor decomposition

As our main objective is to track the temporal variation of brain effective connectivity patterns that underlie the observed iEEG signals and specify which one is acting on which time epoch and for which frequency band and possibly analyze these patterns over epileptic seizures, the tensor decomposition of the constructed PDC-based tensor is quite appropriate for such a task. In fact, based on a tensor decomposition of the constructed  $3^{rd}$  order/ $4^{th}$  order PDC-based tensor, one can easily identify spatial, temporal and frequency signatures of each connectivity pattern contributing to the observed iEEG signals. To this end, the CPD of the  $3^{rd}$  order/ $4^{th}$  order PDC-based tensor is investigated in this paper. According to definition 2 and for meaningful CPD, the tensor rank,  $R$ , should be optimally determined. Finding the exact rank of a tensor is an NP-hard problem [22, 30]. It can be very complex especially for low SNR. This may lead to the well-known over-factoring problem. Therefore, the rank estimation of the tensor could be part of the overall optimization problem to fit the CPD model. Several techniques have been proposed to solve the rank estimation problem in the case of CPD such as the CORCONDIA [12], the minimum description length (MDL) [38], the Laplace Method [39], the cross-validation based method [13], the method for simultaneously estimating the rank and noise level [34], the quotient of differences in additional values [42] and the group-sparsity of the over-estimated loading matrices technique recently proposed in [28, 55] which showed higher performance over the above mentioned techniques. Indeed, recent works [28, 55] suggested to use a new group-sparsity of the over-estimated loading matrices of the considered tensor as a powerful way to optimally estimate the tensor rank. More precisely, authors in [28, 55] showed that the mixed  $\ell_{2,1}$ -norm, as a mean to describe the group sparsity constraint, is a tighter complex envelop of the matrix rank function than the nuclear norm commonly used in this context. Furthermore, the  $\ell_{2,1}$ -norm defined in equation (1) is attractive from a numerical complexity point of view compared to the nuclear norm whose minimization relies essentially on an iterative computation of the singular value decomposition of the considered matrix. Therefore, in this paper, the group sparsity-based technique [28, 55] is employed for an optimal estimation of the PDC-based tensor rank. Moreover, as generally few but the most significant connections are of interest, the tensor loading matrix associated with the space direction is also a sparse matrix. Therefore, the  $\ell_1$ -norm on this spatial loading matrix as a way to describe this sparsity constraint is used in the optimization problem at hand. Besides, for easy and straightforward spatial, temporal and frequency analysis of inferred directed graph connectivity over one or several seizures, all loading matrices are assumed to be non-negative. Therefore, non-negativity constraints of all loading matrices are imposed in the optimization problem as shown in the subsequent subsections.

#### 3.2.1 Single-seizure analysis

In the case of a single seizure analysis, the simultaneous space, time and frequency analysis of brain effective connectivity is possible, as mentioned previously, using a robust low-rank CPD of the  $3^{rd}$  order space $\times$ time $\times$ frequency PDC-based tensor,  $\mathcal{P}$  obtained by solving the following optimization problem:

$$\begin{aligned} \min_{\mathbf{S}, \mathbf{T}, \mathbf{F}} \alpha_{\mathbf{S}} \|\mathbf{S}\|_{2,1} + \delta \|\mathbf{S}\|_1 + \alpha_{\mathbf{T}} \|\mathbf{T}\|_{2,1} + \alpha_{\mathbf{F}} \|\mathbf{F}\|_{2,1} \\ \text{s.t. } \mathbf{S} \geq 0, \mathbf{T} \geq 0, \mathbf{F} \geq 0 \ \& \ \mathcal{P} = \sum_{r=1}^R \mathbf{s}_r \circ \mathbf{t}_r \circ \mathbf{f}_r \end{aligned} \quad (13)$$

where  $\mathbf{S} = [\mathbf{s}_1, \dots, \mathbf{s}_R]$ ,  $\mathbf{T} = [\mathbf{t}_1, \dots, \mathbf{t}_R]$  and  $\mathbf{F} = [\mathbf{f}_1, \dots, \mathbf{f}_R]$  are the loading matrices of the tensor  $\mathcal{P}$  associated with its space, time and frequency directions, respectively.  $\alpha_{\mathbf{S}}$ ,  $\alpha_{\mathbf{T}}$ ,  $\alpha_{\mathbf{F}}$  and  $\delta$  stand for the penalty parameters that control the balance between the different norms in the minimization problem. The above problem can be solved using the well-known Alternating Direction Method of Multipliers (ADMM) [10], which is based on the minimization of the augmented Lagrangian function  $\mathcal{L}_1$  associated with equation (13). Therefore, the above minimization problem defined in equation (13) can be rewritten as:

$$\begin{aligned} \min_{\mathbf{S}, \mathbf{T}, \mathbf{F}, \mathbf{Y}} \mathcal{L}_1 \\ \text{s.t. } \mathbf{S} \geq 0, \mathbf{T} \geq 0, \mathbf{F} \geq 0, \mathbf{Y} = \mathbf{S} \end{aligned} \quad (14)$$

with

$$\mathcal{L}_1 = \alpha_{\mathbf{S}} \|\mathbf{S}\|_{2,1} + \delta \|\mathbf{Y}\|_1 + \alpha_{\mathbf{T}} \|\mathbf{T}\|_{2,1} + \alpha_{\mathbf{F}} \|\mathbf{F}\|_{2,1} + \text{vec}(\mathbf{W})^\top \text{vec}(\mathbf{Y} - \mathbf{S}) + \frac{\rho}{2} \|\mathbf{Y} - \mathbf{S}\|_F^2 + \lambda \left\| \mathcal{P} - \sum_{r=1}^R \mathbf{s}_r \circ \mathbf{t}_r \circ \mathbf{f}_r \right\|_F^2$$

where  $\rho$  and  $\lambda$  are, respectively, the penalty and regularization parameters and  $\mathbf{W}$  denotes the Lagrange multiplier.

The update rules of  $\mathbf{S}$ ,  $\mathbf{T}$  and  $\mathbf{F}$  are computed by looking for the stationary points of  $\mathcal{L}_1$  in  $\mathbf{S}$ ,  $\mathbf{T}$ , and  $\mathbf{F}$ , respectively. These update rules are given by 27, 30, and 33 in Appendix A.1. Technical details regarding the derivation of these update rules are given in appendices A.1.1, A.1.2 and A.1.3. Regarding the dual variable  $\mathbf{Y}$ , and the Lagrange multiplier  $\mathbf{W}$ , their update rules are given in Appendix A.1 (see equations 34 and 35).

Regarding the non-negativity constraints on  $\mathbf{S}$ ,  $\mathbf{T}$  and  $\mathbf{F}$ , they are dealt with at each iteration by setting to zero all their respective possible negative entries. Regarding the initialization of  $\mathbf{S}$ ,  $\mathbf{T}$  and  $\mathbf{F}$ , it is performed using a few iterations of a non-negative CPD of the tensor  $\mathcal{P}$ . At each iteration of the algorithm, each matrix is updated by fixing the other matrices to their last estimate. The algorithm stops either when the number of maximum iterations is reached or when the CPD quality is higher than 0.92 or when the relative error on the spatial loading matrix,  $\mathbf{S}$ , exhibits between two successive iterations a value that is smaller than  $\epsilon = 10^{-5}$ . A pseudo-code summarizing the main steps to solve the optimization problem in equation (14) is given in Algorithm 1 hereafter.

---

### Algorithm 1

---

1. Initialization of  $\mathbf{S}$ ,  $\mathbf{T}$ ,  $\mathbf{F}$  and  $R_{init}$
  2. While the no. max. of iterations is not reached &  $\frac{\|\mathbf{S}^{it} - \mathbf{S}^{it-1}\|^2}{\|\mathbf{S}^{it}\|^2} > \epsilon$  &  $\frac{\|\mathcal{P} - \sum_{r=1}^R \mathbf{s}_r \circ \mathbf{t}_r \circ \mathbf{f}_r\|^2}{\|\mathcal{P}\|^2} > \nu$ 
    - Update  $\mathbf{S}$  using equation (27)
    - $\mathbf{S} \leftarrow \max(\mathbf{S}, \mathbf{0}_{(N_v^2 - N_v) \times R})$
    - Update  $\mathbf{Y}$  using equation (34)
    - Update  $\mathbf{W}$  using equation (35)
    - Update  $\mathbf{T}$  using equation (30)
    - $\mathbf{T} \leftarrow \max(\mathbf{T}, \mathbf{0}_{N_w \times R})$
    - Update  $\mathbf{F}$  using equation (33)
    - $\mathbf{F} \leftarrow \max(\mathbf{F}, \mathbf{0}_{N_c \times R})$
- 

where  $it$  stands for the iteration number.

As generally rank-1 terms in the CPD are not necessarily ordered according to their power, performing such ordering is helpful for an easy identification of the right minimal number of rank-1 tensors contributing to the tensor  $\mathcal{P}$ . Therefore, the rank-1 tensors are ordered in descending way according to their rank-1 power score proposed in this paper and defined by:

$$h(r) = \frac{\sum_{n=1}^{N_v^2 - N_v} \mathbf{S}_{n,r} \sum_{w=1}^{N_w} \mathbf{T}_{w,r} \sum_{f=1}^{N_f} \mathbf{F}_{f,r}}{\sum_{n=1}^{N_v^2 - N_v} \sum_{r=1}^R \mathbf{S}_{n,r} \sum_{w=1}^{N_w} \sum_{r=1}^R \mathbf{T}_{w,r} \sum_{f=1}^{N_f} \sum_{r=1}^R \mathbf{F}_{f,r}}, \quad \forall r \in \{1, \dots, R_{init}\} \quad (15)$$

with  $R_{init}$  denoting the initial rank value. Once the rank-1 tensors are ordered, the retained tensor rank corresponds to the number of rank-1 tensors whose  $h$  value exceeds a predefined threshold.

### 3.2.2 Multi-seizure analysis

In the case where several epileptic seizures are available, the variability of effective connectivity through these seizures can be further evaluated following the same idea as in the above single-seizure analysis 3.2.1. More

precisely, a simultaneous space, time, frequency and seizure analysis is also possible and can be performed using a constrained low rank CPD of the 4<sup>th</sup> order space×time×frequency×seizure PDC-based tensor  $\mathcal{Q}$ . This is henceforth possible by solving the following optimization problem:

$$\begin{aligned} \min_{\mathbf{S}, \mathbf{T}, \mathbf{F}, \mathbf{C}} \quad & \alpha_{\mathbf{S}} \|\mathbf{S}\|_{2,1} + \delta \|\mathbf{S}\|_1 + \alpha_{\mathbf{T}} \|\mathbf{T}\|_{2,1} + \alpha_{\mathbf{F}} \|\mathbf{F}\|_{2,1} + \alpha_{\mathbf{C}} \|\mathbf{C}\|_{2,1} \\ \text{s.t.} \quad & \mathbf{S} \geq 0, \mathbf{T} \geq 0, \mathbf{F} \geq 0, \mathbf{C} \geq 0 \ \& \ \mathcal{Q} = \sum_{r=1}^R \mathbf{s}_r \circ \mathbf{t}_r \circ \mathbf{f}_r \circ \mathbf{c}_r \end{aligned} \quad (16)$$

Similarly, the ADMM method [10] is employed to solve this optimization problem by minimizing its associated Lagrangian function  $\mathcal{L}_2$ . This leads to:

$$\begin{aligned} \min_{\mathbf{S}, \mathbf{T}, \mathbf{F}, \mathbf{C}, \mathbf{Y}} \quad & \mathcal{L}_2 \\ \text{s.t.} \quad & \mathbf{S} \geq 0, \mathbf{T} \geq 0, \mathbf{F} \geq 0, \mathbf{C} \geq 0, \mathbf{Y} = \mathbf{S} \end{aligned} \quad (17)$$

with

$$\begin{aligned} \mathcal{L}_2 = \quad & \alpha_{\mathbf{S}} \|\mathbf{S}\|_{2,1} + \delta \|\mathbf{Y}\|_1 + \alpha_{\mathbf{T}} \|\mathbf{T}\|_{2,1} + \alpha_{\mathbf{F}} \|\mathbf{F}\|_{2,1} + \alpha_{\mathbf{C}} \|\mathbf{C}\|_{2,1} + \text{vec}(\mathbf{W})^\top \text{vec}(\mathbf{Y} - \mathbf{S}) \\ & + \frac{\rho}{2} \|\mathbf{Y} - \mathbf{S}\|_F^2 + \lambda \left\| \mathcal{Q} - \sum_{r=1}^R \mathbf{s}_r \circ \mathbf{t}_r \circ \mathbf{f}_r \circ \mathbf{c}_r \right\|_F^2 \end{aligned}$$

where  $\mathbf{s}_r, \mathbf{t}_r, \mathbf{f}_r, \mathbf{c}_r$  stand respectively for the  $r$ -th column of the loading matrices  $\mathbf{S}, \mathbf{T}, \mathbf{F}$  and  $\mathbf{C}$ .

The update rules of  $\mathbf{S}, \mathbf{T}, \mathbf{F}$  and  $\mathbf{C}$ , are computed by looking for the stationary points of  $\mathcal{L}_2$  in  $\mathbf{S}, \mathbf{T}$ , and  $\mathbf{F}$ , respectively. These update rules are given by 40, 43, 46 and 49 in Appendix A.2. Technical details regarding the derivation of these update rules are given in appendices A.2.1, A.2.2, A.2.3 and A.2.4. Regarding the dual variable  $\mathbf{Y}$ , and the Lagrange multiplier  $\mathbf{W}$ , their update rules are given in Appendix A.2 (see equations 50 and 51).

Similarly to the case of single-seizure analysis 3.2.1, the non-negativity constraints on  $\mathbf{S}, \mathbf{T}$  and  $\mathbf{F}$  are dealt with at each iteration by setting to zero all their respective possible negative entries. Furthermore, matrices  $\mathbf{S}, \mathbf{T}, \mathbf{F}$  and  $\mathbf{C}$ , are initialized using few non-negative CPD iterations of the 4<sup>th</sup> order tensor  $\mathcal{Q}$ .

At each iteration of the algorithm, each matrix is updated by fixing the other matrices to their last estimate. Similarly, the algorithm stops either when the number of maximum iterations is reached or when the CPD quality is higher than 0.92 or when the relative error on the spatial loading matrix,  $\mathbf{S}$ , exhibits between two successive iterations a value that is smaller than  $\epsilon = 10^{-5}$ . A pseudo-code summarizing the main steps to solve the above optimization problem defined in equation (17) is given in Algorithm 2 hereafter.

---

### Algorithm 2

---

1. Initialization of  $\mathbf{S}, \mathbf{T}, \mathbf{F}, \mathbf{C}$  and  $R_{init}$
  2. While the no. max. of iterations is not reached  $\& \frac{\|\mathbf{S}^{it} - \mathbf{S}^{it-1}\|^2}{\|\mathbf{S}^{it}\|^2} > \epsilon \ \& \ \frac{\|\mathcal{Q} - \sum_{r=1}^R \mathbf{s}_r \circ \mathbf{t}_r \circ \mathbf{f}_r \circ \mathbf{c}_r\|^2}{\|\mathcal{Q}\|^2} > \nu$ 
    - Update  $\mathbf{S}$  using equation (40)
    - $\mathbf{S} \leftarrow \max(\mathbf{S}, \mathbf{0}_{(N_v^2 - N_v) \times R})$
    - Update  $\mathbf{Y}$  using equation (50)
    - Update  $\mathbf{W}$  using equation (51)
    - Update  $\mathbf{T}$  using equation (43)
    - $\mathbf{T} \leftarrow \max(\mathbf{T}, \mathbf{0}_{N_w \times R})$
    - Update  $\mathbf{F}$  using equation (46)
    - $\mathbf{F} \leftarrow \max(\mathbf{F}, \mathbf{0}_{N_f \times R})$
    - Update  $\mathbf{C}$  using equation (49)
    - $\mathbf{C} \leftarrow \max(\mathbf{C}, \mathbf{0}_{N_c \times R})$
-

Table 1: Numerical complexity of both methods

	Numerical complexity (flops)
single seizure analysis	$N_{it}[\frac{2R^3}{3}(N_f^3 + N_w^3 + N_v^3(N_v - 1)^3) + R^2(2(N_w^2 + N_v^2(N_v - 1) + N_f^2) + N_v(N_v - 1)N_w + N_v(N_v - 1)N_f + N_fN_w) + R(3(N_f + N_w + N_v(N_v - 1)) + 5N_v(N_v - 1)N_w + N_v(N_v - 1)N_f + N_fN_w)]$
multi-seizure analysis	$N_{it}[\frac{2R^3}{3}(N_f^3 + N_w^3 + N_v^3(N_v - 1)^3 + N_c^3) + R^2(2(N_w^2 + N_v^2(N_v - 1) + N_f^2 + N_c^2) + N_v(N_v - 1)N_wN_c) + R^2(N_v(N_v - 1)N_fN_c + N_fN_wN_c + N_v(N_v - 1)N_fN_w) + R(3(N_f + N_w + N_v(N_v - 1) + N_c) + 5N_v(N_v - 1)N_wN_c) + R(N_v(N_v - 1)N_fN_c + N_fN_wN_c + N_v(N_v - 1)N_fN_c)]$

where  $it$  stands for the iteration number.

In order to identify the right tensor rank, the estimated rank-1 tensors are ordered in a descending way according to the proposed rank-1 power score :

$$h(r) = \frac{\sum_{n=1}^{N_v^2 - N_v} \mathbf{S}_{n,r} \sum_{v=1}^{N_w} \mathbf{T}_{v,r} \sum_{f=1}^{N_f} \mathbf{F}_{f,r} \sum_{c=1}^{N_c} \mathbf{C}_{c,r}}{\sum_{n=1}^{N_v^2 - N_v} \sum_{r=1}^R \mathbf{S}_{n,r} \sum_{v=1}^{N_w} \sum_{r=1}^R \mathbf{T}_{v,r} \sum_{f=1}^{N_f} \sum_{r=1}^R \mathbf{F}_{f,r} \sum_{c=1}^{N_c} \sum_{r=1}^R \mathbf{C}_{c,r}}, \quad \forall r \in \{1, \dots, R_{init}\} \quad (18)$$

Once the rank-1 tensors are ordered, the retained tensor rank corresponds to the number of rank-1 tensors whose  $h$  value exceeds a predefined threshold.

### 3.3 Numerical Complexity

The numerical complexity of the two proposed strategies (single-seizure and multi-seizure analysis) are expressed here in the number of numerical flops required for each method to reach the final solution. A numerical flop is defined as a multiplication followed by an addition. However, in practice more multiplications than additions are encountered. Therefore, in Table 1 only the number of multiplications is considered to compute the numerical complexity. Note that  $N_{it}$  stands for the total number of iterations used, by each scheme, to reach the final solution.

## 4 Numerical Experiments

The performance of the proposed method has been evaluated on simulated data and on real iEEG epileptic signals. To do so, the order of the MVAR model and PDC values were estimated over overlapping sliding windows of the same size ( $M_w = 2048$  points) and sliding at 75% of this size. Note that, for each sliding window, the model order was estimated using the AIC [1].

### 4.1 Simulated data

A realistic dynamic network is considered in this section to evaluate the proposed method. This dynamic network was built by concatenating four sequences of simulated brain activities. Each sequence was generated using a directed graph with specific effective connectivity patterns. The four sequences were generated based on the MVAR model introduced in [21] to simulate brain activity. For this network to be dynamic over both time and space, some connections (edges) in this MVAR model [21] were activated/deactivated and permuted across time and space, respectively (see equations (19)-(22)). The number of active edges in sequences 1, 2, 3 and 4, were set respectively to 4, 6, 10 and 11 as shown in Figure 1. The first three sequences correspond to graphs with limited numbers of nodes and edges whereas the fourth sequence was extracted from [21].

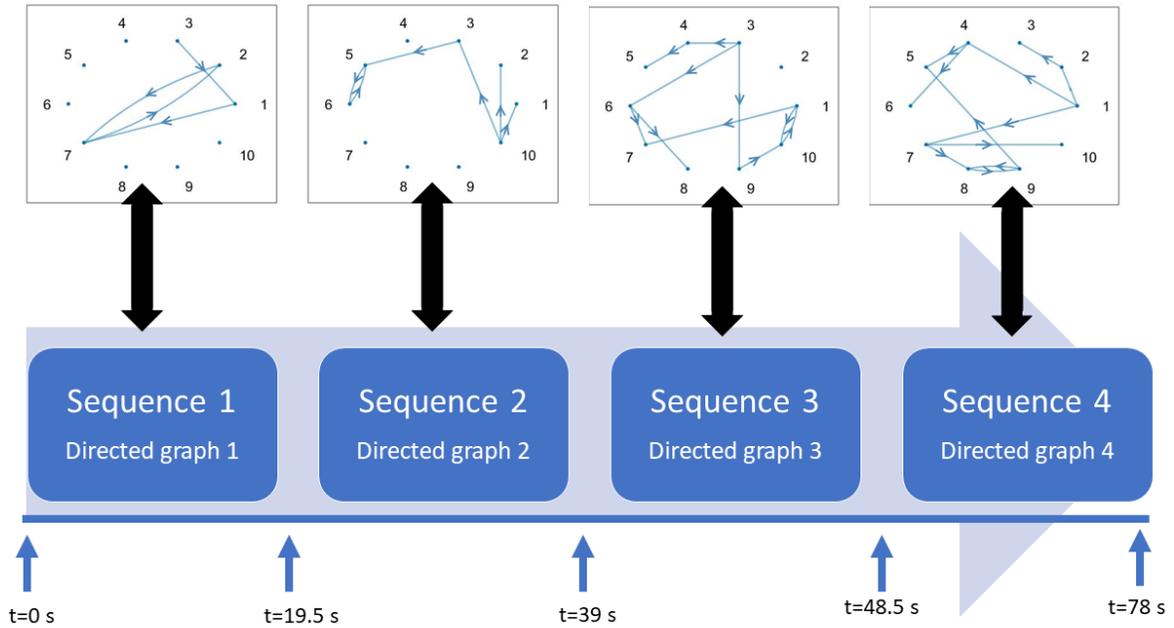


Figure 1: The four directed graphs generated using an MVAR modelling according to equations (19)-(22) and used to define sequences 1, 2, 3 and 4.

Simulated brain activity signals consisted of 20000-point signals, each sequence being composed of 5000 samples. The time-frequency representations of these signals are represented in Figure 4. The generated data were sampled at  $F_e = 256$  Hz with the same time windowing than the one used for the PDC measure in section 2.3. For each channel, the four different sequences are concatenated in a single time signal. Indeed, it is easy to foresee the cuts between sequences around 19, 38 and 57 seconds. Moreover, the frequential informative content is mainly around 35 Hz and some secondary activity can also be noticed between 70 Hz and 120 Hz.

Sequence 1:

$$\begin{aligned}
 x_1(m) &= 0.4x_3(m-1) - 0.4x_3(m-2) + e_1(m) \\
 x_2(m) &= -0.3x_7(m-3) + 0.4x_2(m-3) + e_2(m) \\
 x_3(m) &= 0.95\sqrt{2}x_3(m-1) - 0.9025x_3(m-2) + e_3(m) \\
 x_4(m) &= 0.5x_4(m-1) + e_4(m) \\
 x_5(m) &= 0.5x_5(m-1) + e_5(m) \\
 x_6(m) &= 0.5x_6(m-1) + e_6(m) \\
 x_7(m) &= -0.9x_1(m-2) + 0.4x_7(m-3) + 0.3x_2(m-3) + e_7(m) \\
 x_8(m) &= 0.5x_8(m-1) + e_8(m) \\
 x_9(m) &= 0.5x_9(m-1) + e_9(m) \\
 x_{10}(m) &= 0.5x_{10}(m-1) + e_{10}(m)
 \end{aligned} \tag{19}$$

Sequence 2:

$$\begin{aligned}
x_1(m) &= 0.5x_{10}(m-2) + e_1(m) \\
x_2(m) &= -0.5x_{10}(m-2) + e_2(m) \\
x_3(m) &= 0.4x_{10}(m-1) - 0.4x_{10}(m-2) + e_3(m) \\
x_4(m) &= 0.5x_4(m-1) + e_4(m) \\
x_5(m) &= -0.9x_3(m-2) + 0.4x_5(m-3) + 0.3x_6(m-3) + e_5(m) \\
x_6(m) &= -0.3x_5(m-1) + 0.4x_6(m-3) + e_6(m) \\
x_7(m) &= 0.5x_7(m-1) + 0.4x_7(m-3) + 0.3x_2(m-3) + e_7(m) \\
x_8(m) &= 0.5x_8(m-1) + e_8(m) \\
x_9(m) &= 0.5x_9(m-1) + e_9(m) \\
x_{10}(m) &= 0.95\sqrt{2}x_{10}(m-1) - 0.9025x_{10}(m-2) + e_{10}(m)
\end{aligned} \tag{20}$$

Sequence 3:

$$\begin{aligned}
x_1(m) &= -0.3x_{10}(m-3) + 0.4x_1(m-3) + e_1(m) \\
x_2(m) &= -0.5x_2(m-1) + e_2(m) \\
x_3(m) &= 0.95\sqrt{2}x_3(m-1) - 0.9025x_3(m-2) + e_3(m) \\
x_4(m) &= 0.5x_3(m-2) + e_4(m) \\
x_5(m) &= 0.9x_4(m-3) + e_5(m) \\
x_6(m) &= -0.5x_3(m-2) + e_6(m) \\
x_7(m) &= 0.8x_6(m-4) - 0.4x_1(m-2) + e_7(m) \\
x_8(m) &= -0.8x_6(m-4) + e_8(m) \\
x_9(m) &= 0.4x_3(m-1) - 0.4x_3(m-2) + e_9(m) \\
x_{10}(m) &= -0.9x_9(m-2) + 0.4x_{10}(m-3) + 0.3x_1(m-3) + e_{10}(m)
\end{aligned} \tag{21}$$

Sequence 4:

$$\begin{aligned}
x_1(m) &= 0.95\sqrt{2}x_1(m-1) - 0.9025x_1(m-2) + e_1(m) \\
x_2(m) &= 0.5x_1(m-2) + e_2(m) \\
x_3(m) &= 0.9x_2(m-3) + e_3(m) \\
x_4(m) &= -0.5x_1(m-2) + e_4(m) \\
x_5(m) &= 0.8x_4(m-4) - 0.4x_9(m-2) + e_5(m) \\
x_6(m) &= -0.8x_4(m-4) + e_6(m) \\
x_7(m) &= 0.4x_1(m-1) - 0.4x_1(m-2) + e_7(m) \\
x_8(m) &= -0.9x_7(m-2) + 0.4x_8(m-3) + 0.4x_9(m-3) + e_8(m) \\
x_9(m) &= -0.3x_8(m-3) + 0.4x_9(m-3) + e_9(m) \\
x_{10}(m) &= -0.75x_7(m-4) + e_{10}(m)
\end{aligned} \tag{22}$$

where noise terms  $e_i(t)$ ,  $1 \leq i \leq 10$ , are realizations of independent and standardized Gaussian random variables  $E_i$ .

Regarding the penalty parameters  $\alpha_S$ ,  $\alpha_T$ ,  $\alpha_F$ ,  $\delta$  and  $\lambda$  in equation (13), they were set to 3, 1, 3, 1 and 1, respectively. The choice of these values was made in a trial-and-error manner such that the best solution to the minimization problem given in equation (13) was obtained. An optimal choice of these parameters is still an open question and requires further and deeper investigation which is beyond the scope of this paper. As the rank estimation method is based on the group sparsity of the overestimated loading matrices,  $\mathbf{S}$ ,  $\mathbf{T}$  and  $\mathbf{F}$ , (see [28] for more details), the initial rank was set to  $R_{init} = 10$ . Figure 2 shows the rank-1 tensor power

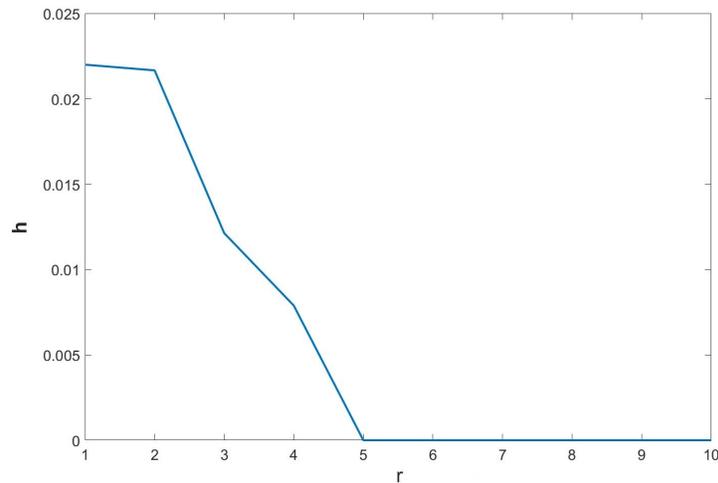


Figure 2: Normalized power of the 10 rank-1 tensors.

score of the 10 rank-1 tensors ordered in descending order. We can notice from this figure that the rank-1 tensor power becomes close to zero beyond a rank equal to four. Hence, the retained rank was  $R = 4$ .

According to Figure 3 and following the temporal signature of the retained rank-1 terms, the inferred sequence of graph connectivity can be ordered chronologically as  $(d) \rightarrow (c) \rightarrow (a) \rightarrow (b)$ . The contribution of each graph structure is well described and can be easily identified by investigating its corresponding temporal signature. This result is in accordance with the ground truth. Regarding the frequency analysis, according to Figure 3, all estimated directed graph structures act on the same frequency band  $[0, 128]$  Hz but with higher contribution around 30 Hz and between 70 Hz and 120 Hz as suggested by the time-frequency analysis from Figure 4.

## 4.2 iEEG data

The proposed approach is then evaluated on real epileptic iEEG signals recorded during three different epileptic seizures of the same patient. This patient suffered from a drug-resistant focal epilepsy on mesial temporal region. A non-invasive pre-surgical examination has been done by the medical team, including an EEG-video recording. Among the different iEEG recordings, three of them were selected by the expert to be analyzed and were limited to the first 32 seconds of the ictal phase. These signals were sampled at 256 Hz and the amplitudes of these real data varied from  $-3000$  to  $2000 \mu\text{V}$  as shown in Figure 5, 6 and 7. The same three channels for which the ground truth was provided by the clinical expert were considered for the three epileptic seizures. These three channels were labeled in two groups according to their involvement along the ictal phase. According to the clinical expert, (i) nodes 1 and 2 were the ones corresponding to the area of the cortex that was responsible for initiating the seizure (*i.e.*, brain areas belonging to the onset group) and (ii) node 3 was considered as a sink receiving the information flow (brain area belonging to the seizure propagation group). Time-Frequency representations of the first, second and third seizures are depicted in decibels in Figures 5, 6 and 7, respectively. These figures highlight a high contribution of low frequencies (below 50 Hz) in the iEEG signals of channels 1 and 2 whereas a wide range of frequency components in the frequency band  $[1, 128]$  Hz are present in the signal of channel 3. This fact holds valid for the three epileptic seizures under study. For each epileptic seizure the model order was estimated using the AIC [1] for each seizure independently and then a PDC-based 3<sup>rd</sup> order tensor was built and consequently analyzed.

The initial rank was set to  $R_{init} = 10$ , and the penalty parameters were set to  $\alpha_S = 3$ ,  $\alpha_T = 1$ ,  $\alpha_F = 3$ ,  $\alpha_C = 0.5$ ,  $\delta = 1$  and  $\lambda = 1$ . According to Figure 3, we note that the inferred directed graphs are in accordance with the expert's opinion since a causal influence from nodes 1 and 2 to node 3 is observed in

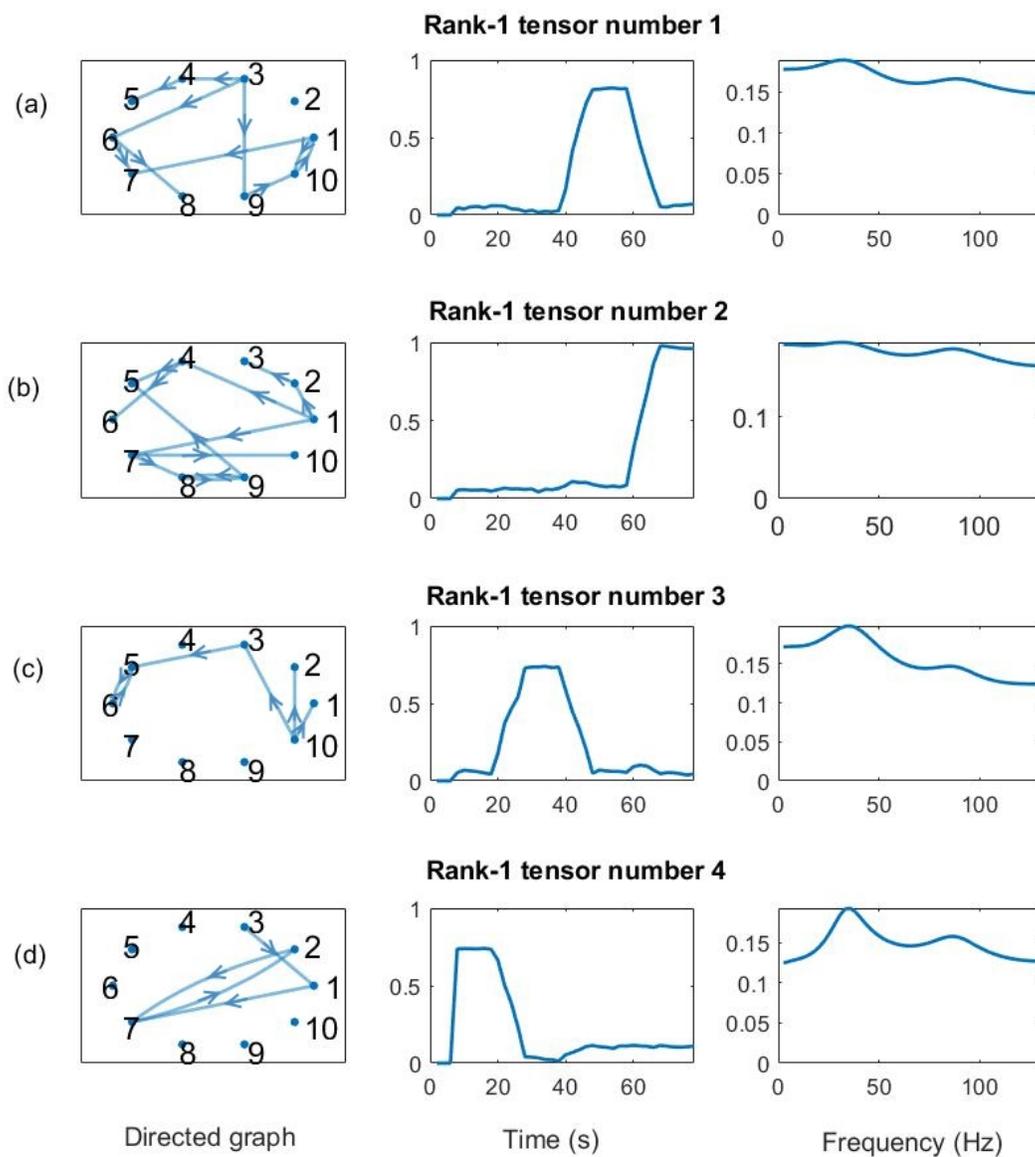


Figure 3: Space, time and frequency signatures of the retained four relevant rank-1 tensors.

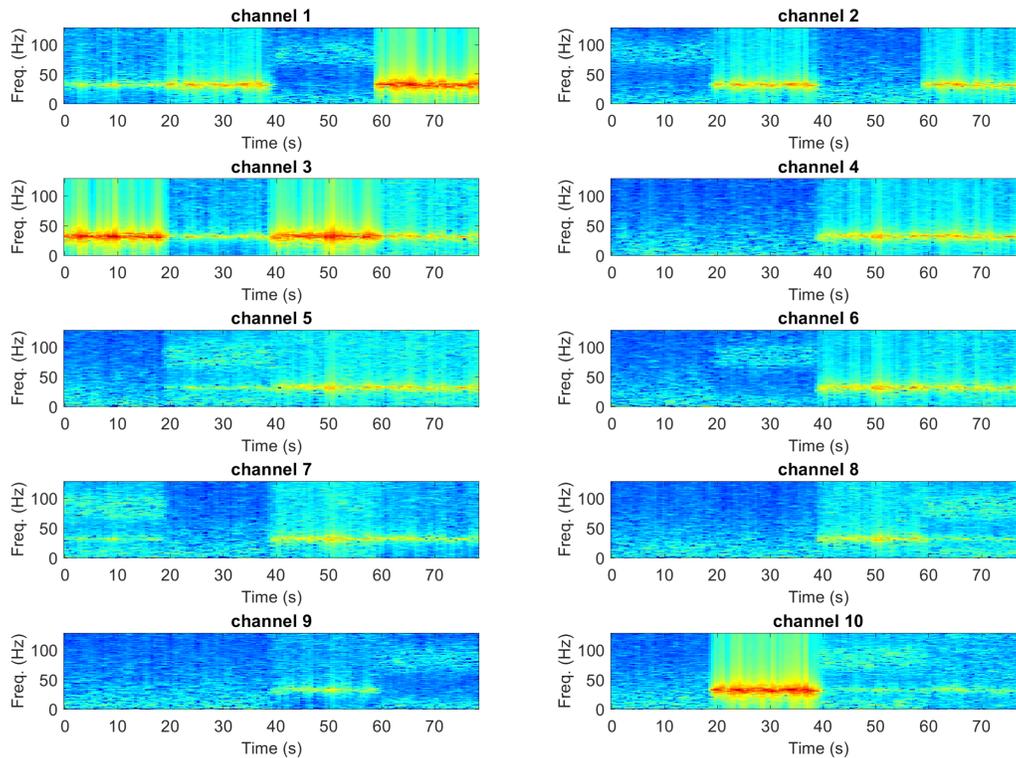


Figure 4: Time-frequency analysis of the 10 channels of simulated data

each seizure. For each seizure, the retained tensor rank is equal to one which means that each seizure can be described by a single directed graph that generates the epileptic iEEG signals at hand.

Regarding the case of multi-seizure analysis (*i.e.* the case of  $4^{th}$  order PDC-based tensor), similar results compared to the case of  $3^{rd}$  order PDC-based tensor were obtained. The retained rank is also equal to one as shown in Figure 8 which indicates that the three seizures have the same structure in terms of spatial graph, temporal evolution and frequency responses. Again, the obtained results are in accordance with the clinical expert. More precisely, obtained frequency signatures are in agreement with the time-frequency analysis obtained with power spectral density shown in Figures 5, 6 and 7. In fact, obtained frequency signatures show clearly that all frequencies in the frequency band  $[0, 128]$  Hz are generally contributing to the observed iEEG spectrum but, as observed in Figures 5, 6 and 7, with relatively higher contribution of low frequencies (below 50 Hz) compared to those beyond 50 Hz.

## 5 Discussion

In this study, the most dominant effective connectivity patterns were inferred based on simultaneous space-time-frequency analysis of epileptic iEEG data. More precisely, the time periods and the frequency bands on which these patterns are acting were identified by the proposed method. Both simulated and real iEEG signals were used to evaluate the performance of this approach. For simulated signals, the method showed good results with 3 to 10 nodes in the epileptic graph of propagation and was able to extract unidirectional and bidirectional links. It should be noticed that the sparsity constraints can be violated when the number of considered nodes is relatively small. However, our method behaved well even for small number of electrodes as demonstrated on real iEEG data. Inferred directed graphs for both simulated and real epileptic iEEG data

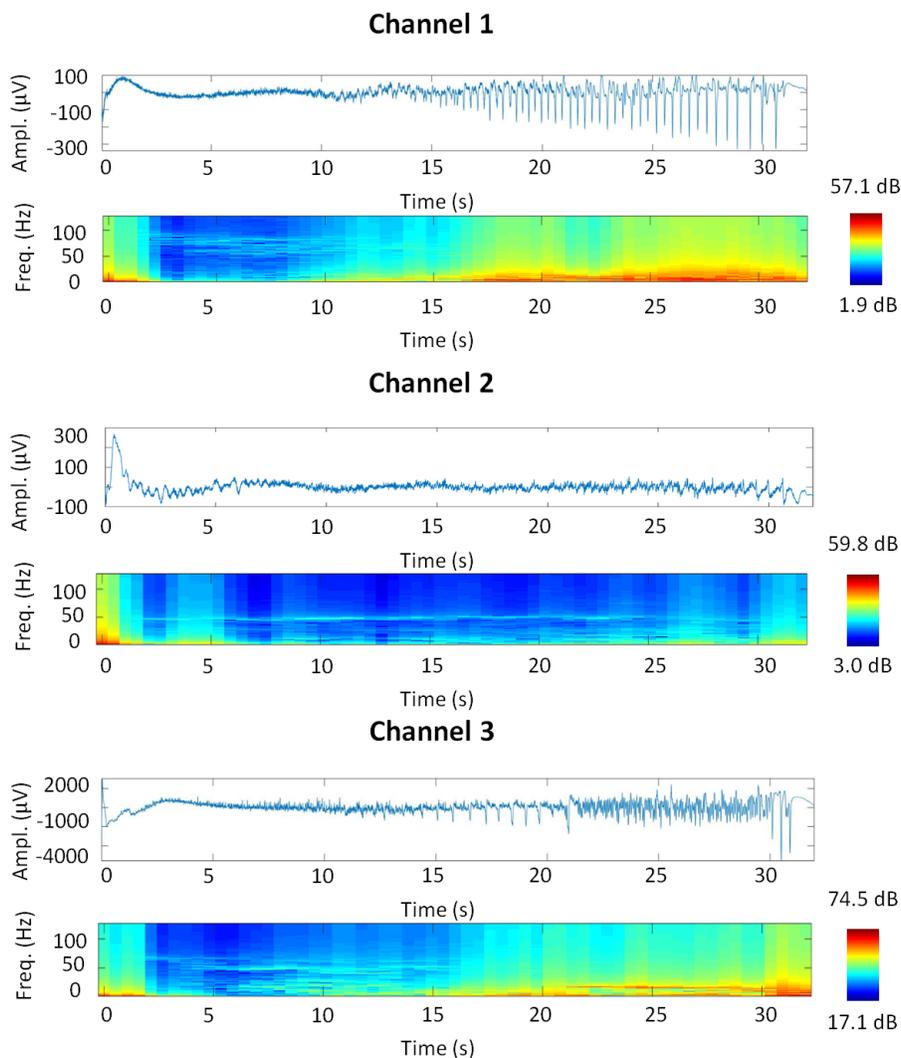


Figure 5: Time profile and time-frequency representation of the three channels from seizure 1

were compatible with the available ground truth. Once the directed graph is established using our method, further graph analysis study could be conducted. For example, one can resort to specific graph measures such as the node inward/outward degree, to distinguish between nodes involved in the seizure initiation and those implied in the seizure propagation. For instance, in our experiment on real data, the outward degrees of node 1, node 2 and node 3 were respectively 1, 1, 0 while their corresponding inward degrees were 0, 0, 2. This shows clearly that nodes 1 and 2 irrigate node 3 which is in accordance with the clinical expertise. In addition, promising results provided by our approach pave the way for an extensive study on a larger database to come up with its applicability in a clinical context. Providing frequency and time signatures of each inferred graph connectivity could help the clinician in his understanding of the seizure initiation and/or evolution. However, in some cases like neocortical or frontal focal epilepsies, the electrical discharges affect simultaneously different structures and then deeper investigations could be needed.

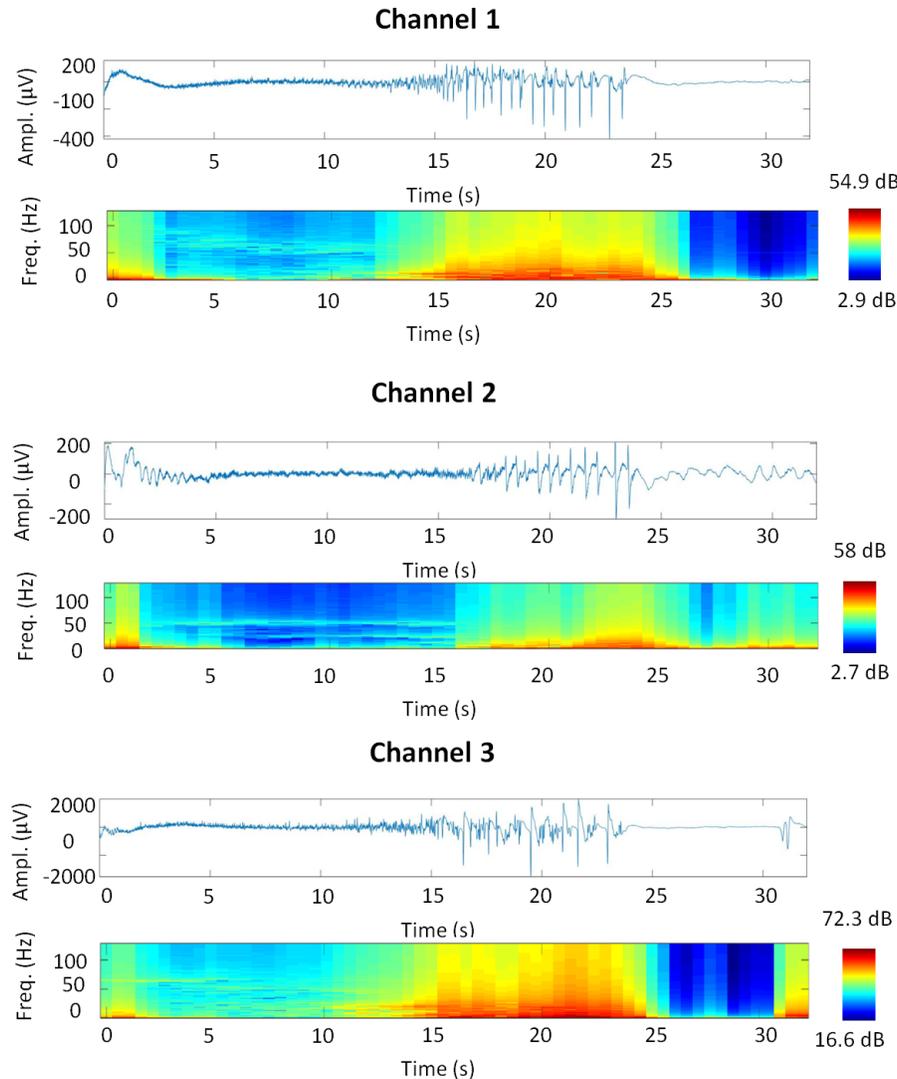


Figure 6: Time profile and time-frequency representation of the three channels from seizure 2

## 6 Conclusion

In this paper, a new method to track brain effective connectivity over the space, time and frequency domains and also over several seizures was proposed. It is based on the low-rank canonical polyadic decomposition of a specific PDC-based tensor built from epileptic iEEG data of a given seizure. Two space, time and frequency analyses have been introduced in this paper: (i) a single-seizure analysis where a robust low rank CPD of a space $\times$ time $\times$ frequency PDC-based tensor was used and (ii) a multi-seizure analysis where several epileptic seizures from the same patient were simultaneously analyzed using a low rank CPD of a space $\times$ time $\times$ frequency $\times$ seizure PDC-based tensor. The good performance has been assessed with simulated signals and real iEEG signals from a patient suffering from mesial temporal focal epilepsy. Obtained time and frequency signatures of inferred directed graphs permitted to efficiently track these graphs over time and frequency for one or several epileptic seizures. In addition to the CPD, other tensor decomposition models could be figured out such as the Block Term Decomposition (BTD) [36, 37] and the Tucker decomposition [56] which could be further investigated to deal with possible interactions between tensor directions.

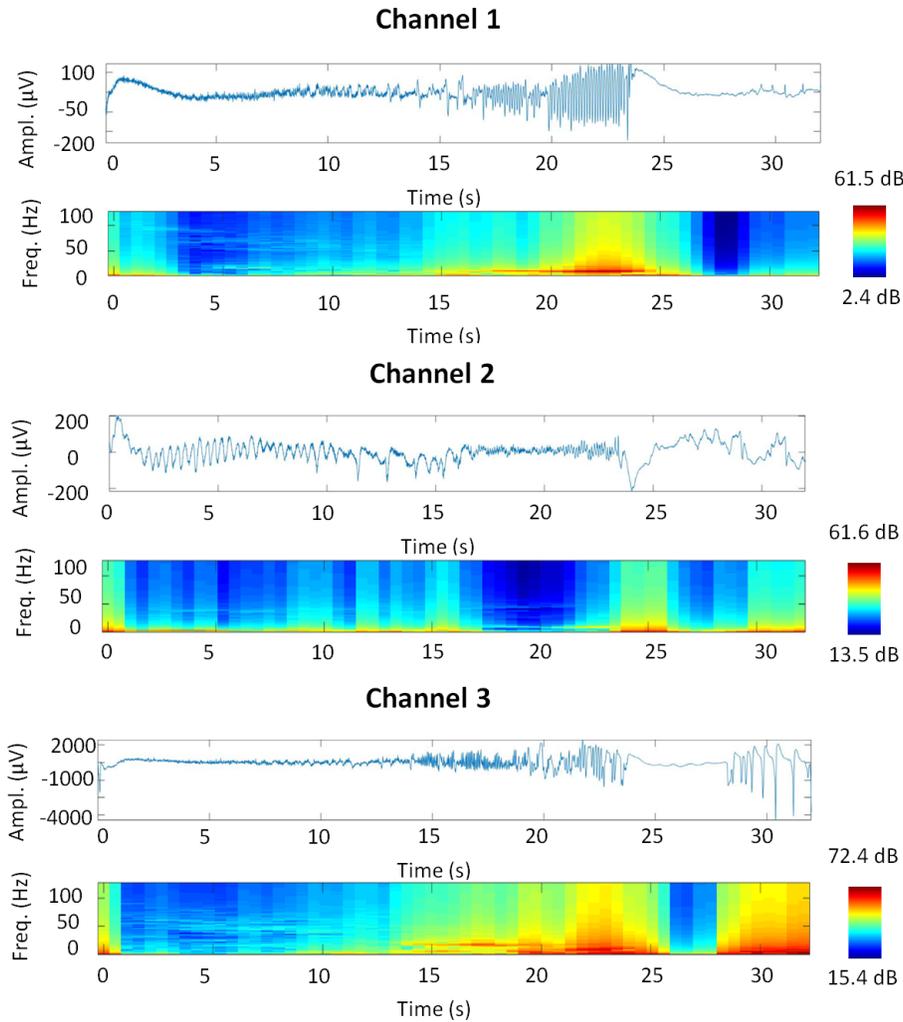


Figure 7: Time profile and time-frequency representation of the three channels from seizure 3

## 7 Acknowledgement

Authors would like to thank the CHU (Centre Hospitalier Universitaire), Rennes, France for the iEEG dataset.

## 8 Ethics approval

All experimental procedures have been conducted at the CHU (Centre Hospitalier Universitaire), Rennes, France, following the ethical and regulatory standards. The patient has signed a consent and has been informed that his iEEG data would be used for clinical research and might serve for publication.

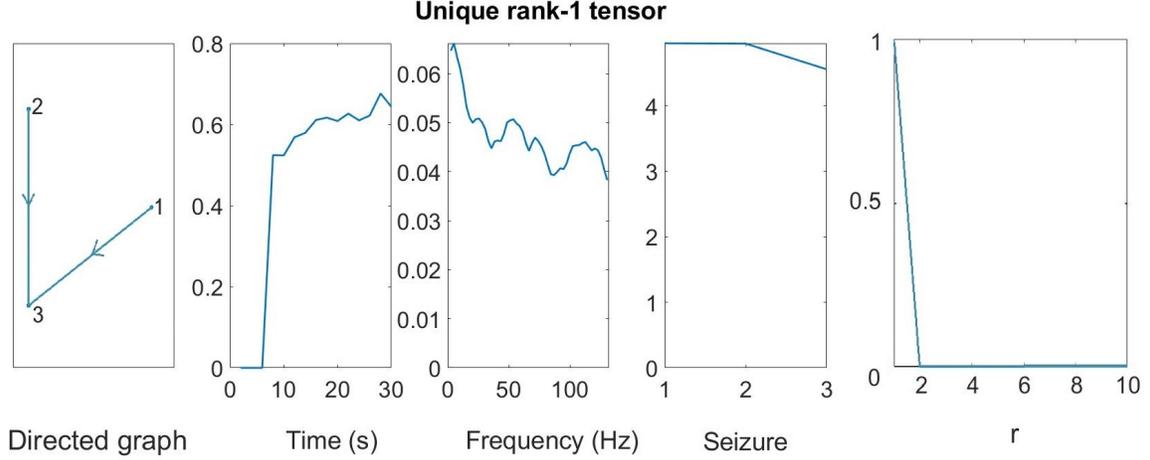


Figure 8: Space, time and frequency signatures of the single retained rank-1 tensor (case of multi-seizure analysis of real epileptic iEEG signals).

## A Appendices

### A.1 Technical materials on the minimization of $\mathcal{L}_1$ defined in equation (14)

As mentioned in section 3.2.1, solving the minimization problem defined in equation (13) is possible by minimizing its associated augmented Lagrangian function  $\mathcal{L}_1$  given by:

$$\begin{aligned} \mathcal{L}_1 = & \alpha_{\mathbf{S}} \|\mathbf{S}\|_{2,1} + \delta \|\mathbf{Y}\|_1 + \alpha_{\mathbf{T}} \|\mathbf{T}\|_{2,1} + \alpha_{\mathbf{F}} \|\mathbf{F}\|_{2,1} + \text{vec}(\mathbf{W})^\top \text{vec}(\mathbf{Y} - \mathbf{S}) + \frac{\rho}{2} \|\mathbf{Y} - \mathbf{S}\|_F^2 \\ & + \lambda \left\| \mathcal{P} - \sum_{r=1}^R \mathbf{s}_r \circ \mathbf{t}_r \circ \mathbf{f}_r \right\|_F^2 \end{aligned} \quad (23)$$

$$\begin{aligned} = & \alpha_{\mathbf{S}} \text{Tr}(\mathbf{S}^\top \mathbf{\Gamma}_{\mathbf{S}} \mathbf{S}) + \delta \|\mathbf{Y}\|_1 + \alpha_{\mathbf{T}} \text{Tr}(\mathbf{T}^\top \mathbf{\Gamma}_{\mathbf{T}} \mathbf{T}) + \alpha_{\mathbf{F}} \text{Tr}(\mathbf{F}^\top \mathbf{\Gamma}_{\mathbf{F}} \mathbf{F}) \\ & + \text{vec}(\mathbf{W})^\top \text{vec}(\mathbf{Y} - \mathbf{S}) + \frac{\rho}{2} \|\mathbf{Y} - \mathbf{S}\|_F^2 + \lambda \left\| \mathcal{P} - \sum_{r=1}^R \mathbf{s}_r \circ \mathbf{t}_r \circ \mathbf{f}_r \right\|_F^2 \end{aligned} \quad (24)$$

with respect to its variable matrices  $\mathbf{S}$ ,  $\mathbf{T}$ ,  $\mathbf{F}$ . This minimization is possible by looking for the stationary points of  $\mathcal{L}_1$  in these variable matrices, as shown hereafter. Note that according to definition 3, we can write  $\mathbf{P}_{(1)} = \mathbf{S}(\mathbf{F} \odot \mathbf{T})^\top$ ,  $\mathbf{P}_{(2)} = \mathbf{T}(\mathbf{S} \odot \mathbf{F})^\top$  and  $\mathbf{P}_{(3)} = \mathbf{F}(\mathbf{T} \odot \mathbf{S})^\top$ . For mathematical convenience, derivation of the update rules will be expressed in a vector form.

#### A.1.1 Update of $\mathbf{S}$

The stationary point of  $\mathcal{L}_1$  defined in equation (14) in  $\mathbf{S}$  is computed by solving the equation  $\frac{\partial \mathcal{L}_1}{\partial \text{vec}(\mathbf{S})} = \mathbf{0}$ . Then according to equation (24) we can write:

$$\frac{\partial \mathcal{L}_1}{\partial \text{vec}(\mathbf{S})} = \alpha_{\mathbf{S}} \frac{\partial \text{Tr}(\mathbf{S}^\top \mathbf{\Gamma}_{\mathbf{S}} \mathbf{S})}{\partial \text{vec}(\mathbf{S})} + \frac{\partial \text{vec}(\mathbf{W})^\top \text{vec}(\mathbf{Y} - \mathbf{S})}{\partial \text{vec}(\mathbf{S})} + \frac{\rho}{2} \frac{\partial \|\mathbf{Y} - \mathbf{S}\|_F^2}{\partial \text{vec}(\mathbf{S})} + \lambda \frac{\partial \|\mathbf{P}_{(1)} - \mathbf{S}(\mathbf{F} \odot \mathbf{T})^\top\|_F^2}{\partial \text{vec}(\mathbf{S})} \quad (25)$$

Now, based on the properties of matrix trace, Kronecker product [46] and on those A.1, A.2 [16], we obtain:

$$\begin{aligned} \frac{\partial \mathcal{L}_1}{\partial \text{vec}(\mathbf{S})} &= -2\lambda((\mathbf{F} \odot \mathbf{T}) \otimes \mathbf{I}_{N_v^2 - N_v})^\top \text{vec}(\mathbf{P}_{(1)}) - \text{vec}(\mathbf{W}) - \rho \text{vec}(\mathbf{Y}) \\ &\quad + (2\alpha_S \mathbf{I}_R \otimes \mathbf{\Gamma}_S + 2\lambda((\mathbf{F} \odot \mathbf{T})^\top (\mathbf{F} \odot \mathbf{T})) \otimes \mathbf{I}_{N_v^2 - N_v} + \rho \mathbf{I}_{R(N_v^2 - N_v)}) \text{vec}(\mathbf{S}) \end{aligned} \quad (26)$$

Then, by solving  $\frac{\partial \mathcal{L}_1}{\partial \text{vec}(\mathbf{S})} = \mathbf{0}$  with respect to  $\text{vec}(\mathbf{S})$ , we get:

$$\begin{aligned} \text{vec}(\mathbf{S}) &= (\alpha_S \mathbf{I}_R \otimes \mathbf{\Gamma}_S + \lambda((\mathbf{F} \odot \mathbf{T})^\top (\mathbf{F} \odot \mathbf{T})) \otimes \mathbf{I}_{N_v^2 - N_v} + \rho \mathbf{I}_{R(N_v^2 - N_v)})^{-1} \\ &\quad \times ((\lambda(\mathbf{F} \odot \mathbf{T})^\top \otimes \mathbf{I}_{N_v^2 - N_v}) \text{vec}(\mathbf{P}_{(1)}) + \frac{\text{vec}(\mathbf{W})}{2} + \frac{\rho}{2} \text{vec}(\mathbf{Y})) \end{aligned} \quad (27)$$

### A.1.2 Update of $\mathbf{T}$

Similarly, the stationary point of  $\mathcal{L}_1$  defined in equation (14) in  $\mathbf{T}$  is computed by solving the equation  $\frac{\partial \mathcal{L}_1}{\partial \text{vec}(\mathbf{T})} = \mathbf{0}$ . Then according to equation (24), we can write:

$$\frac{\partial \mathcal{L}_1}{\partial \text{vec}(\mathbf{T})} = \alpha_T \frac{\partial \text{Tr}(\mathbf{T} \mathbf{\Gamma}_T \mathbf{T}^\top)}{\partial \text{vec}(\mathbf{T})} + \lambda \frac{\partial \|\mathbf{P}_{(2)} - \mathbf{T}(\mathbf{S} \odot \mathbf{F})^\top\|_F^2}{\partial \text{vec}(\mathbf{T})} \quad (28)$$

Based also on the properties of matrix trace, Kronecker product [46] and [16], we obtain:

$$\frac{\partial \mathcal{L}_1}{\partial \text{vec}(\mathbf{T})} = 2\lambda((\mathbf{S} \odot \mathbf{F}) \otimes \mathbf{I}_{N_w})^\top \text{vec}(\mathbf{P}_{(2)}) + (2\alpha_S \mathbf{I}_R \otimes \mathbf{\Gamma}_T + 2\lambda((\mathbf{S} \odot \mathbf{F})(\mathbf{S} \odot \mathbf{F})^\top) \otimes \mathbf{I}_{N_w}) \text{vec}(\mathbf{T}) \quad (29)$$

Then, by solving  $\frac{\partial \mathcal{L}_1}{\partial \text{vec}(\mathbf{T})} = \mathbf{0}$ , we get:

$$\text{vec}(\mathbf{T}) = (\alpha_T \mathbf{I}_R \otimes \mathbf{\Gamma}_T + \lambda((\mathbf{S} \odot \mathbf{F})^\top (\mathbf{S} \odot \mathbf{F})) \otimes \mathbf{I}_{N_w})^{-1} (\lambda(\mathbf{S} \odot \mathbf{F})^\top \otimes \mathbf{I}_{N_w}) \text{vec}(\mathbf{P}_{(2)}) \quad (30)$$

### A.1.3 Update of $\mathbf{F}$

Similarly, the stationary point of  $\mathcal{L}_1$  defined in equation (14) in  $\mathbf{F}$  is computed by solving the equation  $\frac{\partial \mathcal{L}_1}{\partial \text{vec}(\mathbf{F})} = \mathbf{0}$ . Then according to equation (24) we can write:

$$\frac{\partial \mathcal{L}_1}{\partial \text{vec}(\mathbf{F})} = \alpha_F \frac{\partial \text{Tr}(\mathbf{F} \mathbf{\Gamma}_F \mathbf{F}^\top)}{\partial \text{vec}(\mathbf{F})} + \lambda \frac{\partial \|\mathbf{P}_{(3)} - \mathbf{F}(\mathbf{T} \odot \mathbf{S})^\top\|_F^2}{\partial \text{vec}(\mathbf{F})} \quad (31)$$

Based also on the properties of matrix trace, Kronecker product [46] and [16], we obtain:

$$\frac{\partial \mathcal{L}_1}{\partial \text{vec}(\mathbf{F})} = 2\lambda((\mathbf{T} \odot \mathbf{S}) \otimes \mathbf{I}_{N_f})^\top \text{vec}(\mathbf{P}_{(3)}) + (2\alpha_F \mathbf{I}_R \otimes \mathbf{\Gamma}_F + 2\lambda((\mathbf{T} \odot \mathbf{S})(\mathbf{T} \odot \mathbf{S})^\top) \otimes \mathbf{I}_{N_f}) \text{vec}(\mathbf{F}) \quad (32)$$

Then, by solving  $\frac{\partial \mathcal{L}_1}{\partial \text{vec}(\mathbf{F})} = \mathbf{0}$ , we get:

$$\text{vec}(\mathbf{F}) = (\alpha_F \mathbf{I}_R \otimes \mathbf{\Gamma}_F + \lambda((\mathbf{T} \odot \mathbf{S})^\top (\mathbf{T} \odot \mathbf{S})) \otimes \mathbf{I}_{N_f})^{-1} (\lambda(\mathbf{T} \odot \mathbf{S})^\top \otimes \mathbf{I}_{N_f}) \text{vec}(\mathbf{P}_{(3)}) \quad (33)$$

### A.1.4 Update of $\mathbf{Y}$

The update rule of the dual variable  $\mathbf{Y}$  is given by:

$$\mathbf{Y} \leftarrow \mathbf{Y} + \text{prox}_{\phi, \frac{\delta}{\rho}}\left(\mathbf{S} + \frac{\mathbf{W}}{\rho}\right) \quad (34)$$

### A.1.5 Update of $\mathbf{W}$

The update rule of the Lagrange multiplier  $\mathbf{W}$  is given by:

$$\mathbf{W} \leftarrow \mathbf{W} + \rho(\mathbf{S} - \mathbf{Y}) \quad (35)$$

## A.2 Technical materials on the minimization of $\mathcal{L}_2$ defined in equation (17)

As mentioned in section 3.2.2, solving the minimization problem defined in equation (16) is possible by minimizing its associated augmented Lagrangian function  $\mathcal{L}_2$  given by:

$$\begin{aligned} \mathcal{L}_2 = & \alpha_{\mathbf{S}} \|\mathbf{S}\|_{2,1} + \delta \|\mathbf{Y}\|_1 + \alpha_{\mathbf{T}} \|\mathbf{T}\|_{2,1} + \alpha_{\mathbf{F}} \|\mathbf{F}\|_{2,1} + \alpha_{\mathbf{C}} \|\mathbf{C}\|_{2,1} + \text{vec}(\mathbf{W})^\top \text{vec}(\mathbf{Y} - \mathbf{S}) + \frac{\rho}{2} \|\mathbf{Y} - \mathbf{S}\|_F^2 \\ & + \lambda \left\| \mathcal{Q} - \sum_{r=1}^R \mathbf{s}_r \circ \mathbf{t}_r \circ \mathbf{f}_r \circ \mathbf{c}_r \right\|_F^2 \end{aligned} \quad (36)$$

$$\begin{aligned} = & \alpha_{\mathbf{S}} \text{Tr}(\mathbf{S}^\top \mathbf{\Gamma}_S \mathbf{S}) + \delta \|\mathbf{Y}\|_1 + \alpha_{\mathbf{T}} \text{Tr}(\mathbf{T}^\top \mathbf{\Gamma}_T \mathbf{T}) + \alpha_{\mathbf{F}} \text{Tr}(\mathbf{F}^\top \mathbf{\Gamma}_F \mathbf{F}) + \alpha_{\mathbf{C}} \text{Tr}(\mathbf{C}^\top \mathbf{\Gamma}_C \mathbf{C}) \\ & + \text{vec}(\mathbf{W})^\top \text{vec}(\mathbf{Y} - \mathbf{S}) + \frac{\rho}{2} \|\mathbf{Y} - \mathbf{S}\|_F^2 + \lambda \left\| \mathcal{Q} - \sum_{r=1}^R \mathbf{s}_r \circ \mathbf{t}_r \circ \mathbf{f}_r \circ \mathbf{c}_r \right\|_F^2 \end{aligned} \quad (37)$$

with respect to its variable matrices  $\mathbf{S}$ ,  $\mathbf{T}$ ,  $\mathbf{F}$  and  $\mathbf{C}$ . This minimization is possible by looking for the stationary points of  $\mathcal{L}_2$  in these variable matrices, as shown hereafter. Note that according to definition 3, the unfolding matrices associated with the first, the second, the third and the fourth direction of the  $4^{rd}$  order tensor  $\mathcal{Q}$  are given, respectively, by:  $\mathbf{Q}_{(1)} = \mathbf{S}(\mathbf{C} \odot \mathbf{F} \odot \mathbf{T})^\top$ ,  $\mathbf{Q}_{(2)} = \mathbf{T}(\mathbf{S} \odot \mathbf{C} \odot \mathbf{F})^\top$ ,  $\mathbf{Q}_{(3)} = \mathbf{F}(\mathbf{T} \odot \mathbf{S} \odot \mathbf{C})^\top$  and  $\mathbf{Q}_{(4)} = \mathbf{C}(\mathbf{F} \odot \mathbf{T} \odot \mathbf{S})^\top$ . For mathematical convenience, derivation of the update rules will be expressed in a vector form.

### A.2.1 Update of $\mathbf{S}$

The stationary point of  $\mathcal{L}_2$  defined in equation (17) in  $\mathbf{S}$  is computed by solving the equation  $\frac{\partial \mathcal{L}_2}{\partial \text{vec}(\mathbf{S})} = \mathbf{0}$ . Then according to equation (37) we can write:

$$\frac{\partial \mathcal{L}_2}{\partial \text{vec}(\mathbf{S})} = \alpha_{\mathbf{S}} \frac{\partial \text{Tr}(\mathbf{S}^\top \mathbf{\Gamma}_S \mathbf{S})}{\partial \text{vec}(\mathbf{S})} + \frac{\partial \text{vec}(\mathbf{W})^\top \text{vec}(\mathbf{Y} - \mathbf{S})}{\partial \text{vec}(\mathbf{S})} + \frac{\rho}{2} \frac{\partial \|\mathbf{Y} - \mathbf{S}\|_F^2}{\partial \text{vec}(\mathbf{S})} + \lambda \frac{\partial \|\mathbf{Q}_{(1)} - \mathbf{S}(\mathbf{C} \odot \mathbf{F} \odot \mathbf{T})^\top\|_F^2}{\partial \text{vec}(\mathbf{S})} \quad (38)$$

Now based on the properties of matrix trace, Kronecker product [46] and [16], we obtain:

$$\begin{aligned} \frac{\partial \mathcal{L}_2}{\partial \text{vec}(\mathbf{S})} = & -2\lambda((\mathbf{F} \odot \mathbf{T}) \otimes \mathbf{I}_{N_v^2 - N_v})^\top \text{vec}(\mathbf{Q}_{(1)}) - \text{vec}(\mathbf{W}) - \rho \text{vec}(\mathbf{Y}) \\ & + (2\alpha_{\mathbf{S}} \mathbf{I}_R \otimes \mathbf{\Gamma}_S + 2\lambda((\mathbf{F} \odot \mathbf{T})^\top (\mathbf{F} \odot \mathbf{T})) \otimes \mathbf{I}_{N_v^2 - N_v} + \rho \mathbf{I}_{R(N_v^2 - N_v)}) \text{vec}(\mathbf{S}) \end{aligned} \quad (39)$$

Then, by solving  $\frac{\partial \mathcal{L}_2}{\partial \text{vec}(\mathbf{S})} = \mathbf{0}$  with respect to  $\text{vec}(\mathbf{S})$ , we get:

$$\begin{aligned} \text{vec}(\mathbf{S}) = & (\alpha_{\mathbf{S}} \mathbf{I}_R \otimes \Gamma_{\mathbf{S}} + \lambda((\mathbf{C} \otimes \mathbf{F} \odot \mathbf{T})^\top (\mathbf{C} \otimes \mathbf{F} \odot \mathbf{T})) \otimes \mathbf{I}_{N_v^2 - N_v} + \rho \mathbf{I}_{R(N_v^2 - N_v)})^{-1} \\ & \times ((\lambda(\mathbf{C} \otimes \mathbf{F} \odot \mathbf{T})^\top \otimes \mathbf{I}_{N_v^2 - N_v}) \text{vec}(\mathbf{Q}_{(1)}) + \frac{\text{vec}(\mathbf{W})}{2} + \frac{\rho}{2} \text{vec}(\mathbf{Y})) \end{aligned} \quad (40)$$

### A.2.2 Update of $\mathbf{T}$

Similarly, the stationary point of  $\mathcal{L}_2$  defined in equation (17) in  $\mathbf{T}$  is computed by solving the equation  $\frac{\partial \mathcal{L}_2}{\partial \text{vec}(\mathbf{T})} = \mathbf{0}$ . Then according to equation (37) we can write:

$$\frac{\partial \mathcal{L}_2}{\partial \text{vec}(\mathbf{T})} = \alpha_{\mathbf{T}} \frac{\partial \text{Tr}(\mathbf{T} \Gamma_{\mathbf{T}} \mathbf{T}^\top)}{\partial \text{vec}(\mathbf{T})} + \lambda \frac{\partial \|\mathbf{Q}_{(2)} - \mathbf{T}(\mathbf{S} \odot \mathbf{F})^\top\|_F^2}{\partial \text{vec}(\mathbf{T})} \quad (41)$$

Based also on the properties of matrix trace, Kronecker product [46] and [16], we obtain:

$$\frac{\partial \mathcal{L}_2}{\partial \text{vec}(\mathbf{T})} = 2\lambda((\mathbf{S} \odot \mathbf{F}) \otimes \mathbf{I}_{N_w})^\top \text{vec}(\mathbf{Q}_{(2)}) + (2\alpha_{\mathbf{S}} \mathbf{I}_R \otimes \Gamma_{\mathbf{T}} + 2\lambda((\mathbf{S} \odot \mathbf{C} \odot \mathbf{F})(\mathbf{S} \odot \mathbf{C} \odot \mathbf{F})^\top) \otimes \mathbf{I}_{N_w}) \text{vec}(\mathbf{T}) \quad (42)$$

Then, by solving  $\frac{\partial \mathcal{L}_2}{\partial \text{vec}(\mathbf{S})} = \mathbf{0}$ , we get:

$$\text{vec}(\mathbf{T}) = (\alpha_{\mathbf{T}} \mathbf{I}_R \otimes \Gamma_{\mathbf{T}} + \lambda((\mathbf{S} \odot \mathbf{C} \odot \mathbf{F})^\top (\mathbf{S} \odot \mathbf{C} \odot \mathbf{F})) \otimes \mathbf{I}_{N_w})^{-1} (\lambda(\mathbf{S} \odot \mathbf{C} \odot \mathbf{F})^\top \otimes \mathbf{I}_{N_w}) \text{vec}(\mathbf{Q}_{(2)}) \quad (43)$$

### A.2.3 Update of $\mathbf{F}$

The stationary point of  $\mathcal{L}_2$  defined in equation (17) in  $\mathbf{F}$  is computed by solving the equation  $\frac{\partial \mathcal{L}_2}{\partial \text{vec}(\mathbf{F})} = \mathbf{0}$ . Then, according to equation (24) we can write:

$$\frac{\partial \mathcal{L}_1}{\partial \text{vec}(\mathbf{F})} = \alpha_{\mathbf{F}} \frac{\partial \text{Tr}(\mathbf{F} \Gamma_{\mathbf{F}} \mathbf{F}^\top)}{\partial \text{vec}(\mathbf{F})} + \lambda \frac{\partial \|\mathbf{Q}_{(3)} - \mathbf{F}(\mathbf{T} \odot \mathbf{S})^\top\|_F^2}{\partial \text{vec}(\mathbf{F})} \quad (44)$$

Based also on the properties of matrix trace, Kronecker product [46] and [16], we obtain:

$$\begin{aligned} \frac{\partial \mathcal{L}_2}{\partial \text{vec}(\mathbf{F})} = & 2\lambda((\mathbf{T} \odot \mathbf{S} \odot \mathbf{C}) \otimes \mathbf{I}_{N_f})^\top \text{vec}(\mathbf{Q}_{(3)}) \\ & + (2\alpha_{\mathbf{F}} \mathbf{I}_R \otimes \Gamma_{\mathbf{F}} + 2\lambda((\mathbf{T} \odot \mathbf{S} \odot \mathbf{C})(\mathbf{T} \odot \mathbf{S} \odot \mathbf{C})^\top) \otimes \mathbf{I}_{N_f}) \text{vec}(\mathbf{F}) \end{aligned} \quad (45)$$

Then, by solving  $\frac{\partial \mathcal{L}_2}{\partial \text{vec}(\mathbf{F})} = \mathbf{0}$ , we get:

$$\text{vec}(\mathbf{F}) = (\alpha_{\mathbf{F}} \mathbf{I}_R \otimes \Gamma_{\mathbf{F}} + \lambda((\mathbf{T} \odot \mathbf{S} \odot \mathbf{C})^\top (\mathbf{T} \odot \mathbf{S} \odot \mathbf{C})) \otimes \mathbf{I}_{N_f})^{-1} (\lambda(\mathbf{T} \odot \mathbf{S} \odot \mathbf{C})^\top \otimes \mathbf{I}_{N_f}) \text{vec}(\mathbf{Q}_{(3)}) \quad (46)$$

### A.2.4 Update of $\mathbf{C}$

Similarly, the stationary point of  $\mathcal{L}_2$  defined in equation (17) in  $\mathbf{C}$  is computed by solving the equation  $\frac{\partial \mathcal{L}_2}{\partial \text{vec}(\mathbf{C})} = \mathbf{0}$ . Then according to equation (24) we can write:

$$\frac{\partial \mathcal{L}_2}{\partial \text{vec}(\mathbf{C})} = \alpha_{\mathbf{C}} \frac{\partial \text{Tr}(\mathbf{C}\mathbf{\Gamma}_{\mathbf{C}}\mathbf{C}^{\top})}{\partial \text{vec}(\mathbf{C})} + \lambda \frac{\partial \|\mathbf{Q}_{(4)} - \mathbf{C}(\mathbf{F} \odot \mathbf{T} \odot \mathbf{S})^{\top}\|_F^2}{\partial \text{vec}(\mathbf{C})} \quad (47)$$

Based also on the properties of matrix trace, Kronecker product [46] and [16], we obtain:

$$\begin{aligned} \frac{\partial \mathcal{L}_2}{\partial \text{vec}(\mathbf{C})} = & 2\lambda((\mathbf{F} \odot \mathbf{T} \odot \mathbf{S}) \otimes \mathbf{I}_{N_c})^{\top} \text{vec}(\mathbf{Q}_{(4)}) \\ & + (2\alpha_{\mathbf{C}}\mathbf{I}_R \otimes \mathbf{\Gamma}_{\mathbf{C}} + 2\lambda((\mathbf{F} \odot \mathbf{T} \odot \mathbf{S})(\mathbf{F} \odot \mathbf{T} \odot \mathbf{S})^{\top}) \otimes \mathbf{I}_{N_c}) \text{vec}(\mathbf{C}) \end{aligned} \quad (48)$$

Then, by solving  $\frac{\partial \mathcal{L}_2}{\partial \text{vec}(\mathbf{C})} = \mathbf{0}$ , we get:

$$\text{vec}(\mathbf{C}) = (\alpha_{\mathbf{C}}\mathbf{I}_R \otimes \mathbf{\Gamma}_{\mathbf{C}} + \lambda((\mathbf{F} \odot \mathbf{T} \odot \mathbf{S})(\mathbf{F} \odot \mathbf{T} \odot \mathbf{S})^{\top}) \otimes \mathbf{I}_{N_c})^{-1} (\lambda(\mathbf{F} \odot \mathbf{T} \odot \mathbf{S})^{\top} \otimes \mathbf{I}_{N_c}) \text{vec}(\mathbf{Q}_{(4)}) \quad (49)$$

### A.2.5 Update of $\mathbf{Y}$

The update rule of the dual variable  $\mathbf{Y}$  is given by:

$$\mathbf{Y} \leftarrow \mathbf{Y} + \text{prox}_{\phi, \frac{\delta}{\rho}}(\mathbf{S} + \frac{\mathbf{W}}{\rho}) \quad (50)$$

### A.2.6 Update of $\mathbf{W}$

The update rule of the Lagrange multiplier  $\mathbf{W}$  is given by:

$$\mathbf{W} \leftarrow \mathbf{W} + \rho(\mathbf{S} - \mathbf{Y}) \quad (51)$$

## References

- [1] H. Akaike, (1969) Fitting autoregressive models for prediction. In *Annals of the Institute of Statistical Mathematics*, Vol. 21, no 1, pp. 243-247.
- [2] L. Astolfi, F. Cincotti, D. Mattia, S. Salinari, C. Babilioni, A. Basilisco, M. Rossini, L. Ding, Y. Ni, B. He, M. Marciani, F. Babilioni. (2004) Estimation of the effective and functional human cortical connectivity with structural equation modeling and directed transfer function applied to high-resolution EEG. In *Magnetic Resonance Imaging*, Vol. 22, Issue 10, pp.1457-1470.
- [3] P. Basser, J. Mattiello, D. Le Bihan, (1994) MR diffusion tensor spectroscopy and imaging. In *Biophysical Journal*, Vol. 66, Issue 1, pp. 259-267.
- [4] L. Baccalá, K. Sameshima, (2001) Partial directed coherence: a new concept in neural structure determination. In *Biological Cybernetics*, Vol. 84, No. 6, pp. 463-474.
- [5] K. Blinowska, (2011) Review of the methods of determination of directed connectivity from multichannel data. In *Medical & Biological Engineering & Computing*, Vol. 49, Issue 5, pp. 521-529.
- [6] P. Bancaud, R. Angelergues, A. Bonis, M. Bordas-Ferrer, M. Bresson, P. Buser, L. Covello, P. Morel, G. Szikla, A. Takeda, J. Talairach, (1970) Functional stereotaxic exploration (SEEG) of epilepsy. In *Electroencephalographic and Clinical Neurophysiology*, Vol. 8, Issue 1, pp. 85-86.

- [7] J. Bancaud, J. Talairach, A. Bonis, G. Schaub, M. Bordas-Ferer, (1965) La stéréo-électro-encéphalographie dans l'épilepsie : informations neurophysiopathologiques apportées par l'investigation fonctionnelle stéréotaxique. Edited by *Masson*.
- [8] P. Boon, K. Vonck, J. De Reuck, J. Caemaert, (2001) Vagus nerve stimulation for refractory epilepsy. In *Seizure*, Vol. 10, Issue 6, pp. 448-455.
- [9] P. Boon, T. Vandekerckhove, E. Achten, E. Thiery, L. Goossens, K. Vonck, M. D'Have, G. van Hoey, B. Vanrumste, B. Legros, L. Defreyne, J. De Reuck, (1999) Epilepsy surgery in Belgium, the experience in Gent. In *Acta Neurologica Belgica*, 99 (4), pp. 256-265.
- [10] S. Bohyd, N. Parikh, E. Chu, B. Peleato, J. Eckstein, (2010) Distributed optimization and statistical learning via the alternating direction method of multipliers. In *Foundations and Trends in Machine Learning*, Vol.3, No. 1, pp. 1-122.
- [11] R. Bro, (1997) PARAFAC, Tutorial and applications. In *Chemometrics and Intelligence Laboratory systems*, Vol. 38, pp. 149-171.
- [12] R. Bro, A. Henk, L. Kiers, (2003) A new efficient method for determining the number of components in PARAFAC models. In *Journal of Chemometrics*, Vol. 17, Issue 5, pp. 274-279.
- [13] R. Bro, K. Kjeldahl, A. Smilde, H. Kiers, (2003) Cross-validation of component models: a critical look at current methods. In *Analytical and Bioanalytical Chemistry*, Vol. 390, No. 5, pp. 1241-1251.
- [14] J. Carroll, J. Chang, (1970) Analysis of individual differences in multidimensional scaling via an n-way generalization of Eckart-Young decomposition. In *Psychometrika*, Vol. 35, pp. 283-319.
- [15] C. Yang, R. Le Bouquin Jeannès, G. Faucon, H. Shu, (2019) Detecting information flow direction in multivariate linear and nonlinear models. In *Signal Processing*, Vol.93, Issue 1, pp. 304-312.
- [16] J. Coloigner, A. Karfoul, L. Albera, P. Common, (2014) Line search and trust region strategies for canonical decomposition of semi-nonnegative semi-symmetric  $3^{rd}$  order tensors. In *Linear Algebra and its Applications*, Volume 450, pp.334-374.
- [17] F. Cong, Q. Lin, L. Kuang, X. Gong, P. Astikainen, T. Ristaniemi, (2015) Tensor decomposition of EEG signals: A brief review. In *Computational Neuroscience*, Vol. 248, pp. 59-69.
- [18] V. De Herdt, P. Boon, B. Ceulemans, H. Hauman, L. Lagae, B. Legros, B. Sadzot, P. Van Bogaert, K. van Rijckevorsel, H. Verhelst, K. Vonck, (2007) Vagus nerve stimulation for refractory epilepsy: A Belgian multicenter study. In *European Journal of Paediatric Neurology*, Volume 11, Issue 5, pp. 261-269.
- [19] J. Ding, V. Tarokh, Y. Yang, (2018) Model selection techniques: An overview. In *IEEE Signal Processing Magazine*, Vol. 35, Issue 6, pp. 16-34.
- [20] L. Ding, (2009) Reconstructing cortical current density by exploring sparseness in the transform domain. In *Physics in Medicine and Biology*, Vol. 54, No. 9, pp. 2683-2697.
- [21] A. Fasoula, Y. Attal, D. Schwartz, (2013) Comparative performance evaluation of data-driven causality measures applied to brain networks. In *Journal of Neuroscience Methods*, Vol. 215, Issue 2, pp. 170-189.
- [22] M. Fazel, (2002) Matrix rank minimization with applications. In PhD thesis, Stanford University.
- [23] K. Friston, J. Kahan, A. Razi, K. Stephan, O. Sporns, (2014) On nodes and modes in resting state fMRI. In *NeuroImage*, Vol. 99 pp. 535-547.
- [24] K. Friston, J. Kahan, B. Biswal, A. Razi, (2014) A DCM for resting state fMRI. In *NeuroImage*, Vol. 94, pp. 396-407.
- [25] K. Friston, (2011) Functional and effective connectivity: a review. In *Brain Connectivity*, Vol. 1, No. 1, pp. 13-36.
- [26] B. Gourevitch, R. Le Bouquin Jeannès, G. Faucon, (2006) Linear and non-linear causality between signals: Methods, examples and neurophysiological applications. In *Biological Cybernetics*, Vol. 95, No. 4, pp. 349-369.
- [27] C. Granger, (1969) Investigating causal relations by econometric models and cross-spectral methods. In *Econometrica*, Vol. 37, No 3, pp. 424-438.

- [28] X. Han, L. Albera, A. Kachenoura, L. Senhadji, H. Shu, (2017) Low rank canonical polyadic decomposition of tensors based on group sparsity. In *25th European Signal Processing Conference*, No. 25, pp. 668-672.
- [29] R. Harshman, (1972) PARAFAC2: Mathematical and technical notes. In *UCLA Working papers in phonetics*, Vol. 22, University of California, Los Angeles.
- [30] C. Hillard, L. Lim, (2013) Most tensor problems are NP-hard. In *journal of the ACM*, Vol. 60, No. 6, pp. 39-45.
- [31] F. Hitchcock, (1927) The expression of a tensor or a polyadic as sum of products. In *Journal Of Mathematics and Physics*, Vol. 6, Issue 1-4 pp. 164-189.
- [32] E. Karahan, P. Rojas-López, M. Bringas-Vega, P. Valdés-Hernández, P. Valdes-Sosa, (2015) Tensor analysis and fusion of multimodal brain images. In *Proceedings of the IEEE*, Vol. 103, Issue 9, pp. 1531-1559.
- [33] T. Kolda, B. Bader, (2009) Tensor decompositions and applications. In *SIAM review*, Vol. 51, Issue 3, pp. 455-500.
- [34] S. Kritchman and B. Nadler (2008) Determining the number of components in a factor model from limited noisy data. In *Chemometrics Intelligent Laboratory Systems*, Vol. 94, Issue 1, pp. 19-32.
- [35] J. Kruskal, (1977) Three-way arrays: rank and uniqueness of trilinear decompositions, with application to arithmetic complexity and statistics. In *Linear algebra and its applications*, Vol. 18, Issue 12, pp. 95-138.
- [36] L. De Lathauwer, (2008) Decompositions of a higher-order tensor in block terms - Part I: Lemmas for partitioned matrices. In *SIAM Journal on Matrix Analysis and Applications*, Vol. 30, No. 3, pp. 1022-1032.
- [37] L. De Lathauwer, (2008) Decompositions of a higher-order tensor in block terms - Part II: Definitions and uniqueness. In *SIAM Journal on Matrix Analysis and Applications*, Vol. 30, No. 3 pp. 1033-1066.
- [38] K. Liu, J. da Costa, H. Cheung So, L. Huang, J. Ye, (2016) Detection of number of components in CANDECOMP/PARAFAC models via minimum description length. In *Digital Signal Processing*, Vol. 51, pp. 110-123.
- [39] T. Minka, (2000) Automatic choice of dimensionality for PCA. In *13th International Conference on Neural Information Processing Systems*, No. 13, pp. 577-583.
- [40] F. Miwakeichi, E. Martinez-Montes, P. Valdés-Sosa, N. Nishiyama, H. Mizuhara, Y. Yamaguchi, (2004) Decomposing EEG data into space-time-frequency components using Parallel Factor Analysis. In *NeuroImage*, Vol. 22, Issue 3, pp. 1035-1045.
- [41] M. Mørup, L. Hansen, C. Herrmann, J. Parnas, S. Arnfred, (2006) Parallel Factor Analysis as an exploratory tool for wavelet transformed event-related EEG. In *NeuroImage*, Vol. 29, Issue 3, pp. 938-947.
- [42] J. Niesing, (1997) Simultaneous component and factor analysis methods for two or more groups: a comparative study. volume 2<sup>nd</sup> ed. Leiden: The Netherlands: DSWO Press, Leiden University.
- [43] P. Nunez, R. Srinivasan, A. Westdrop, R. Wijesinghe, D. Tucker, R. Silberstein, P. Cadusch, (1997) EEG coherency: I: statistics, reference electrode, volume conduction, Laplacians, cortical imaging, and interpretation at multiple scales. In *Electroencephalography and Clinical Neurophysiology*, Vol. 103, Issue 5, pp. 449-515.
- [44] C. Ozcaglar, A. Shabbeer, S. Vandenberg, B. Yener, K. Bennett, (2011), Sublineage structure analysis of *Mycobacterium tuberculosis* complex strains using multiple-biomarker tensors. In *BMC Genomics*. Vol. 12 (Suppl 2)S1:1-26.
- [45] B. Pester, C. Ligges, L. Leistritz, H. Witte, K. Schiecke, (2015) Advanced insights into functional brain connectivity by combining tensor decomposition and partial directed coherence. In *PIOs one*, Vol. 10, Issue 6.
- [46] K. Petersen, M. Pedersen, (2012) The Matrix Cookbook. Published by *Technical University of Denmark*, IMM Group: *Intelligent Signal Processing*.

- [47] M. Priestley, (1981) *Spectral Analysis and Time Series*. By *Academic Press, London*, Vol.1-2.
- [48] P. Valdes-Sosa, A. Roebroeck, J. Daunizeau, K. Friston, (2011) Effective connectivity: influence, causality and biophysical modeling. In *NeuroImage*, Vol. 58, Issue 2, pp. 339-361.
- [49] Y. Saito, H. Harashima, (1981) Tracking of information within multichannel EEG record - Causal analysis in EEG. In *Recent Advances in EEG and EMG Data Processing*. Amsterdam: Elsevier, pp. 133-146.
- [50] K. Sameshima, L. Baccalá, (1999) Using partial directed coherence to describe neuronal ensemble interactions. In *Journal of Neuroscience Methods*, Vol. 94, Issue 1, pp. 93-103.
- [51] G. Schwartz, (1978) Estimating the dimension of a model. In *Annals of Statistics*, Vol. 6, No. 2, pp. 461-464.
- [52] N. Sidiropoulos, R. Bro, G. Giannakis, (2000) Parallel factor analysis in sensor array processing. In *IEEE Transactions on Signal Processing*, Vol. 48, Issue 8, pp. 2377-2388.
- [53] M. Sørensen, L. De Lathauwer, (2013) Blind signal separation via tensor decomposition with Vandermonde factor: Canonical Polyadic Decomposition. In *IEEE Transactions on Signal Processing*, Vol. 61, No. 22, pp. 5507-5519.
- [54] F. Sun, M. Morell, (2014) The RNS System: responsive cortical stimulation for the treatment of refractory partial epilepsy. In *Expert Review of Medical Devices*, Vol. 11, Issue 6, pp. 563-572.
- [55] N. Taheri, X. Han, A. Karfoul, K. Ansari, I. Merlet, L. Senhadji, L. Albera, A. Kachenoura, (2018) Brain source localization using constrained low rank canonical Polyadic decomposition. In *52<sup>nd</sup> IEEE Asilomar Conference on Signals, Systems and Computers*.
- [56] L. Tucker, (1966) Some mathematical notes on three-mode factor analysis. In *Psychometrika*, Vol. 31, Issue 3, pp. 279-311.
- [57] K. Vonck, P. Boon, P. Claeys, S. Dedeurwaerdere, R. Achten, D. van Roost, (2005) Long-term deep brain stimulation for refractory temporal lobe epilepsy. In *Epilepsia*, Vol. 46, Suppl. 5, pp. 98-99.
- [58] M. Weiss, F. Römer, M. Haardt, D. Jannek, P. Husar, (2009) Multi-dimensional space-time-frequency component analysis of event related EEG data using closed-form PARAFAC. In *IEEE International Conference on Acoustics, Speech, and Signal Processing*, pp. 349-352.
- [59] S. Wright, (1921) Correlation and causation. In *Journal of Agricultural research*, Vol. 20, Issue 7, pp. 557-585.
- [60] Q. Zhao, G. Zhou, L. Zhang, A. Cichocki, S. Amari, (2015) Bayesian robust tensor factorization for incomplete multiway data. In *IEEE Transactions on Neural Networks and Learning Systems*, Vol. 13, No. 9, pp. 736-748.

## B Author biographies



**Pierre-Antoine Chantal** received the image processing engineering degree from Telecom Sud Paris, Evry, France in 2017. He is now a Ph.D. student in signal processing in Rennes University. He investigates the help of tensor decompositions for epileptic signals analysis.



**Ahmad Karfoul** received the electronic engineering degree from AL-Baath University, Syria, in 2000. He received his Master degree in Signal processing and his Ph.D. degree in Signal Processing and Telecommunications from Rennes University respectively in 2006 and 2009. From 2010 to 2014, he was associate professor at the Al-Baath University, Syria. From 2014 to 2015 he worked as research engineer with the LTSI (Laboratoire Traitement du Signal et de l'Image, Inserm U1099). He is currently an associate professor in LTSI-Inserm U1099 at Rennes University. His research interests include statistical signal processing, MRI quantification, brain connectivity, and tensor-based brain source imaging.



**Dr Anca Nica** received her MD in adult neurology from the Rennes University, France. She also has a neuropediatrics degree from the Carol Davila University of Medicine of Bucharest, Romania. She is in charge with the noninvasive and intracranial depth electrode neurophysiology of the Epilepsy Unit of the University Rennes Hospital since 2008. In addition to her clinical activity, she is member of LTSI (Laboratoire Traitement du Signal et de l'Image, Inserm U1099) Unit of Rennes University, her research focuses on EEG and SEEG signal processing.



**Régine Le Bouquin Jeannès** received her Ph.D. degree in Signal Processing and Telecommunications from Rennes University, France, in 1991. She is currently Professor in LTSI (Laboratoire Traitement du Signal et de l'Image, Inserm U1099) at Rennes University and belongs to CRIBs (Centre de Recherche en Information Biomédicale sino-français), a French-Chinese laboratory associated with Inserm (Institut national de la santé et de la recherche médicale). Her research activities mainly focus on biomedical signals processing and modeling in the field of epilepsy and brain connectivity as well as on health monitoring and telecare.