TEMPORALLY RESOLVED MULTI-WAY COMPONENT ANALYSIS OF DYNAMIC SOURCES IN EVENT-RELATED EEG DATA USING PARAFAC2

Martin Weis, Dunja Jannek, Thomas Guenther, Peter Husar

Biosignal Processing Group
Ilmenau University of Technology
Gustav-Kirchhoff Str. 2, D-98684 Ilmenau, Germany
{martin.weis, peter.husar}@tu-ilmenau.de
{dunja.jannek, thomas.guenther}@tu-ilmenau.de
www.tu-ilmenau.de/bsv

Florian Roemer, Martin Haardt

Communications Research Laboratory
Ilmenau University of Technology
Helmholzplatz 2, D-98684 Ilmenau, Germany
{florian.roemer, martin.haardt}@tu-ilmenau.de
www.tu-ilmenau.de/crl

ABSTRACT

The identification of signal components in electroencephalographic (EEG) data is a major task in neuroscience. The interest to this area has regained new interest due to the possibilities of multi-dimensional signal processing. In this contribution we analyze event-related multi-channel EEG recordings on the basis of the time-varying spectrum for each channel. To identify the signal components it is a common approach to use parallel factor (PARAFAC) analysis. However, the PARAFAC model cannot cope with components appearing time-shifted over the different channels. Furthermore, it is not possible to track PARAFAC components over time. We show how to overcome these problems by using the PARAFAC2 decomposition, which renders it an attractive approach for processing EEG data with highly dynamic (moving) sources. Additionally, we introduce the concept of PARAFAC2 component amplitudes, which resolve the scaling ambiguity in the PARAFAC2 model and can be used to judge the relevance of the components.

1. INTRODUCTION

In this contribution we focus on analyzing measured electroencephalographic (EEG) data to identify the components of activity. This analysis can also be used to detect and localize epileptic seizure onset zones on the scalp as well as projections of cognitive processing like speech or auditory handling. Unfortunately, different sources in the brain can produce the same EEG pattern, which renders them in general non-separable. Source localization algorithms, such as LORETA [15] or dipole fitting methods can resolve this ambiguity by imposing additional assumptions. For further improvements of these methods, preprocessing in form of subspace decompositions, e.g., principle component analysis (PCA), independent component analysis (ICA), singular value decomposition (SVD), or beamforming algorithms [10] have been applied. However, these methods cannot exploit the multi-dimensional nature of the EEG data. Moreover, to obtain matrix decompositions like PCA or ICA, physically unsatisfiable assumptions like orthogonality or independence have to be imposed. Therefore, tensor decompositions are a more promising approach to handle EEG signals. Especially the well known parallel factor (PARAFAC) analysis is widely used in recent literature, because it is essentially unique under mild conditions [2] without any artificial constraints, such as orthogonality. In the last years PARAFAC was applied to EEG signals, e.g., for estimating sources of cognitive processing [13], for the analysis of event-related potentials (ERP) [14], and for epileptic seizure localization [18].

In order to resolve the temporal evolution as well as the frequency content of the EEG recordings, a time-frequency analysis (TFA) is applied for each channel. Therefore, the data is analyzed over three dimensions, i.e., time, frequency, and space (channels). Different TFA algorithms have been studied for the analysis of EEG signals [7]. The most common method is the continuous wavelet transformation (CWT). However, wavelet analysis may not provide adequate time and frequency resolution for EEG data. In [19] it was shown that the reduced interference distribution (RID) [5] is particularly useful for the TFA of EEG data and its subsequent multi-way component analysis, since it provides an improved time-frequency resolution.

The common approaches for the three-way component analysis of EEG data to date are based on the PARAFAC model. However, this model is not able to resolve moving EEG components which appear time-shifted over the different channels. Therefore, the PARAFAC component analysis is only useful in case of static sources. In this contribution we introduce the PARAFAC2 decomposition [9] for the space-time-frequency analysis of EEG data. The PARAFAC2 model supports time-shifted component signals. Furthermore, we show how the PARAFAC2 model can be adopted in order to track the different EEG components over time. The PARAFAC2 model is rarely used up to now, i.e., in [3] it is applied to chemometric data including retention time-shifts, and [1] uses it for the time-space-window analysis of EEG and electrocardiographic (ECG) recordings. The PARAFAC2 model is essentially unique up to scaling and permutation. In order to resolve the scaling ambiguity, we introduce the least squares PARAFAC2 component amplitudes which can be used to judge the influence of the individual components.

This paper is organized as follows: In Section 2 we clarify the notation and define the operators and symbols that are used. In Section 3 we discuss the signal processing steps to analyze EEG signals. Thereby, the Sections 3.1 and 3.2 present the methods for the measurement preprocessing and the time-frequency analysis. Subsequently, Section 3.3 describes the three-way component analysis of the different time-frequency distributions using PARAFAC2. In Section 4 we present the results of the event-related EEG analysis based on measurements, before drawing the conclusions in Section 5.

2. NOTATION

To facilitate the distinction between scalars, vectors, matrices, and higher-order tensors, we use the following notation: scalars are denoted by lower-case italic letters $(a, b, ...)$, vectors by boldface lower-case italic letters $(\boldsymbol{a}, \boldsymbol{b}, ...)$, matrices by boldface upper-case letters $(\boldsymbol{A}, \boldsymbol{B}, ...)$, and tensors are denoted as upper-case, boldface, calligraphic letters $(\boldsymbol{\mathcal{A}}, \boldsymbol{\mathcal{B}}, ...)$. This notation is consistently used for lower-order parts of a given structure, unless stated otherwise. For example $\boldsymbol{\mathcal{A}} \in \mathbb{R}^{I_1 \times I_2 \times \cdots \times I_N}$ represents an N-dimensional tensor of size I_n along mode n. Its elements are referenced by $a_{i_1, i_2, \ldots, i_N}$ for $i_n = 1, 2, \ldots I_n$ and $n = 1, 2, \ldots, N$. For matrices we use the superscripts $^{\mathrm{T}}$, $^{\mathrm{H}}$, $^{-1}$, $^{+}$ for transposition, Hermitian transposition, matrix inverse, and Moore-Penrose pseudo-inverse, respectively.

The tensor operations we use are consistent with [12]. The higher-order norm of a tensor $\boldsymbol{\mathcal{A}}$, symbolized by $\|\boldsymbol{\mathcal{A}}\|_{\mathrm{H}}$, is defined as the square root of the sum of the squared magnitude of all elements in $\boldsymbol{\mathcal{A}}$. The n-mode vectors of a tensor $\boldsymbol{\mathcal{A}}$ are obtained

 ISSN 2076-1465

by varying the n-th index i_n of the tensor elements $a_{i_1,i_2,\ldots,i_N}$ within its range $(1,2,\ldots,I_n)$ while keeping all the other indices fixed. The matrix unfolding of the tensor $\boldsymbol{\mathcal{A}}$, denoted by $[\boldsymbol{\mathcal{A}}]_{(n)} \in \mathbb{R}^{I_n \times I_1 \cdot \ldots \cdot I_{n-1} \cdot I_{n+1} \cdot \ldots \cdot I_N}$ contains all the n-mode vectors of the tensor $\boldsymbol{\mathcal{A}}$. The k-th frontal slice of a third order tensor $\boldsymbol{\mathcal{X}} \in \mathbb{R}^{I_1 \times I_2 \times I_3}$ is addressed by $[\boldsymbol{\mathcal{X}}]_{:,:,k} \in \mathbb{R}^{I_1 \times I_2}$, where k can reach the values $1 \ldots I_3$. The operator $\mathrm{vec}\{\boldsymbol{\mathcal{X}}\}$ aligns all elements of the tensor $\boldsymbol{\mathcal{X}}$ into a column vector. The n-mode product of a tensor $\boldsymbol{\mathcal{A}} \in \mathbb{R}^{I_1 \times \cdots \times I_N}$ and a matrix $\boldsymbol{U} \in \mathbb{R}^{J_n \times I_n}$ is denoted as $(\boldsymbol{\mathcal{A}} \times_n \boldsymbol{U}) \in \mathbb{R}^{I_1 \times \cdots \times J_n \times \cdots \times I_N}$. It is obtained by multiplying all n-mode vectors of $\boldsymbol{\mathcal{A}}$ from the left hand side by the matrix $\boldsymbol{U}$. The outer product of an N-dimensional tensor $\boldsymbol{\mathcal{A}}$ and a K-dimensional tensor $\boldsymbol{\mathcal{B}}$, denoted by $(\boldsymbol{\mathcal{A}} \circ \boldsymbol{\mathcal{B}})$, is a $(N+K)$-dimensional tensor whose elements are given by $(\boldsymbol{\mathcal{A}} \circ \boldsymbol{\mathcal{B}})_{i_1,\ldots,i_N,j_1,\ldots j_K} = a_{i_1,\ldots,i_N} \cdot b_{j_1,\ldots j_K}$. An N-dimensional tensor $\boldsymbol{\mathcal{A}} \in \mathbb{R}^{I_1 \times \cdots \times I_N}$ is of rank one if and only if it can be written as the outer product between N non-zero vectors $\boldsymbol{c}^{(n)} \in \mathbb{R}^{M_n}$, such that $\boldsymbol{\mathcal{A}} = \boldsymbol{c}^{(1)} \circ \ldots \circ \boldsymbol{c}^{(N)}$. The three-dimensional identity tensor $\boldsymbol{\mathcal{I}}_{3,d}$ is defined as

$$\boldsymbol{\mathcal{I}}_{3,d} = \sum_{n=1}^{d} \boldsymbol{e}_{n,d} \circ \boldsymbol{e}_{n,d} \circ \boldsymbol{e}_{n,d} \in \mathbb{R}^{d \times d \times d} , \qquad (1)$$

where $\boldsymbol{e}_{n,d}$ represents the n-th column of a $d \times d$ identity matrix (also termed the n-th pinning vector of size d).

3. SIGNAL PROCESSING STEPS

The processing of EEG data is a very challenging task due to the difficult nature of these signals, e.g., they are non-stationary and suffer from very low signal to noise ratios. Moreover, they are affected by correlated noise with unknown distribution and artifacts originating from eye blinks, eye movements, and muscle activity as well as from diverse technical distortions. Therefore, a suitable preprocessing has to be applied in the form of filters, reference EEG channels, and averaging over several trials. Afterwards, the time-frequency analysis is applied to each channel individually, in order to resolve the temporal evolution as well as the frequency content of the EEG data. This is done by applying the Reduced Interference Distribution (RID) [5], since it provides an improved time and frequency resolution. The components of the resulting three-dimensional signal, which changes in frequency, space (channels), and time, are extracted via PARAFAC2 analysis (see Figure 1).

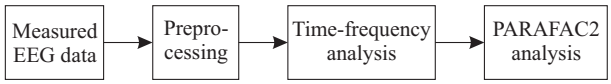

Figure 1: Signal processing steps for the identification of signal components in event-related EEG data. After the measurements and an appropriate preprocessing, the time-frequency analysis is performed. The resulting three-way data is then analyzed using the PARAFAC2 decomposition.

3.1 Measurement Description and Preprocessing

The EEG signal is recorded from a 23 year old, healthy and right-handed woman. The position of the 64 EEG electrodes is based on the international 10-10-system [4] with earlobe references $[(A1 + A2)/2]$. The sampling frequency is chosen to 1000 sps (samples per second). For the preprocessing of the raw signal, the following off-line, digital, zero-phase filters are applied: a 7 Hz high-pass, a 135 Hz low-pass, and a band-stop filter between 45 and 55 Hz. Thereby, all filters showed a stop-band suppression of at least 60 dB. For the investigation of event-related potentials, we record EEG data triggered by a visual stimulus. The subject sits in front of a hemispherical perimeter. The stimulus is a 20 ms central light flash from a white LED to the right eye. The triggered EEG responses to this stimulus are averaged over 1600 trials for all channels (see Figure 2).

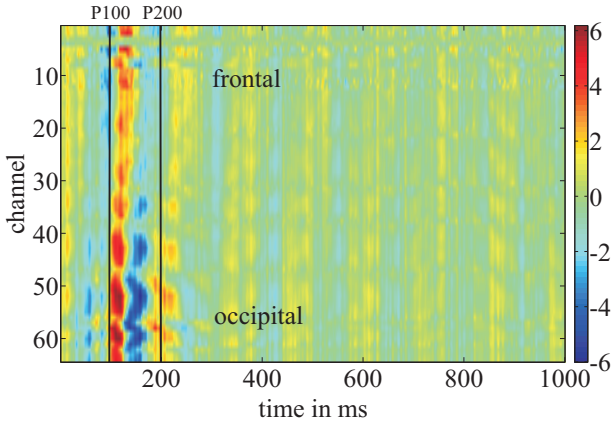

Figure 2: The course of the visual evoked potentials, averaged over 1600 trials and distributed over 64 channels according to the 10-10-system [4]. The visual stimulus is a 20 ms light flash coming from a white LED which is mounted in the central field of the right eye. The occipital channels show the response (P100 and P200) earlier than the frontal ones.

3.2 Time-Frequency Analysis

A powerful approach to time-frequency analysis is given by the family of quadratic time-frequency distributions (TFD), which are based on the temporal correlation function (TCF) $q_x(t,\tau)$ of the signal $x(t)$ defined as [7]

$$q_x(t,\tau) = x(t + \frac{\tau}{2}) x^*(t - \frac{\tau}{2}) . \qquad (2)$$

The Wigner-Ville distribution (WVD) $W_x(t,f)$ of $x(t)$ is defined as the Fourier transform of the TCF with respect to the lag variable τ [20, 17]

$$W_x(t,f) = \int_{-\infty}^{\infty} q_x(t,\tau) \mathrm{e}^{-\mathrm{j}2\pi f \tau} \mathrm{d}\tau . \qquad (3)$$

The ambiguity function $A_x(\theta,\tau)$ is defined as the inverse Fourier transform of the TCF with respect to the time t [7]

$$A_x(\theta,\tau) = \int_{-\infty}^{\infty} q_x(t,\tau) \mathrm{e}^{\mathrm{j}2\pi\theta t} \mathrm{d}t . \qquad (4)$$

Thus, the ambiguity function and the WVD are related by the two-dimensional Fourier transform. The main drawback of the time-frequency analysis based on the TCF is that it produces cross terms in $W_x(t,f)$ as well as in $A_x(\theta,\tau)$. However, the time and frequency resolution can be adjusted separately. In 1966 Cohen introduced an overall class of TFDs based on the WVD which allow the use of kernel functions for reducing cross terms [6]. This group of TFDs $P_x(t,f)$ is defined as

$$P_x(t,f) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} A_x(\theta,\tau) \Theta(\theta,\tau) \mathrm{e}^{-\mathrm{j}2\pi\theta t - \mathrm{j}2\pi\tau f} \mathrm{d}\theta \mathrm{d}\tau , \qquad (5)$$

where $\Theta(\theta,\tau)$ is the kernel function. A large number of TFDs have been proposed, each differing only in the choice of $\Theta(\theta,\tau)$. These kernel functions can be used to suppress the effect of the cross terms on the TFD. Choi and Williams [5] introduced the reduced interference distribution (RID), which is a TFD based on the exponential kernel function

$$\Theta(\theta,\tau) = \mathrm{e}^{-\frac{\theta^2\tau^2}{\sigma}} , \qquad (6)$$

where $\sigma > 0$ is a scaling factor which influences the cross term suppression. The RID has been proven to be especially useful for the analysis of EEG data [19], also in connection with the subsequent tensor decomposition (Figure 1).

3.3 Three-Way Component Analysis

After the time-frequency analysis the EEG data is represented by a time-varying frequency distribution for every channel. This three-way data can be expressed in form of a tensor

$$\boldsymbol{\mathcal{X}} \in \mathbb{R}^{N_\mathrm{F} \times N_\mathrm{T} \times N_\mathrm{C}} \,, \tag{7}$$

where N_F and N_T are the number of samples in frequency and time, and N_C is the number of channels, respectively. In order to separate the signal components in this tensor, it is common to use a multi-dimensional extension of the singular value decomposition that is known as the PARAFAC decomposition [8]. Thereby, the tensor is decomposed into a minimal sum of rank one components. In the absence of noise, the PARAFAC model for the tensor (7) can be represented as

$$\boldsymbol{\mathcal{X}} = \sum_{n=1}^{d} \boldsymbol{\mathcal{Y}}^{(n)} = \sum_{n=1}^{d} \boldsymbol{a}_n \circ \boldsymbol{b}_n \circ \boldsymbol{c}_n \,, \tag{8}$$

where the vectors $\boldsymbol{a}_n \in \mathbb{R}^{N_\mathrm{F}}$, $\boldsymbol{b}_n \in \mathbb{R}^{N_\mathrm{T}}$, and $\boldsymbol{c}_n \in \mathbb{R}^{N_\mathrm{C}}$, represent the frequency, time, and channel signatures of the n-th PARAFAC component. Moreover, d represents the number of signal components (PARAFAC model order). Since each PARAFAC component $\boldsymbol{\mathcal{Y}}^{(n)}$ is constructed from the outer product of the channel, time and frequency signature, it represents a component signal with a rank-one time-frequency distribution. Furthermore, the component signal can vary over the different channels only by a scalar factor $c_{k,n}$, which is the k-th element of the channel signature $\boldsymbol{c}_n$. Therefore, the k-th frontal slice $[\boldsymbol{\mathcal{Y}}^{(n)}]_{:,:,k} \in \mathbb{R}^{N_\mathrm{F} \times N_\mathrm{T}}$ for $k = 1 \ldots N_\mathrm{C}$ of each component tensor $\boldsymbol{\mathcal{Y}}^{(n)}$ is given by

$$[\boldsymbol{\mathcal{Y}}^{(n)}]_{:,:,k} = \boldsymbol{a}_n \cdot c_{k,n} \cdot \boldsymbol{b}_n^\mathrm{T} \,. \tag{9}$$

For the analysis of moving EEG sources, it is crucial to allow that the component signals can appear time-shifted over the channels. Therefore, we have to adopt the PARAFAC components $\boldsymbol{\mathcal{Y}}^{(n)}$ such that the time signature $\boldsymbol{b}_n$ can vary over the channel indices $k = 1 \ldots N_\mathrm{C}$. This yields the following PARAFAC2 [9] component

$$[\boldsymbol{\mathcal{Z}}^{(n)}]_{:,:,k} = \boldsymbol{a}_n \cdot c_{k,n} \cdot \boldsymbol{t}_{n,k}^\mathrm{T} \,, \tag{10}$$

where the vector $\boldsymbol{t}_{k,n} \in \mathbb{R}^{N_\mathrm{T}}$ is the time signature for the k-th channel ($k = 1 \ldots N_\mathrm{C}$). By introducing the component matrices $\boldsymbol{A} = [\boldsymbol{a}_1, \boldsymbol{a}_2, \ldots, \boldsymbol{a}_r] \in \mathbb{R}^{N_\mathrm{F} \times r}$, $\boldsymbol{T}_k = [\boldsymbol{t}_{k,1}, \boldsymbol{t}_{k,2}, \ldots, \boldsymbol{t}_{k,r}] \in \mathbb{R}^{N_\mathrm{T} \times r}$, as well as the vector of channel signatures $\boldsymbol{s}_k = [c_{k,1}, c_{k,2}, \ldots, c_{k,r}]^\mathrm{T}$ the decomposition of the tensor $\boldsymbol{\mathcal{X}}$ reads as

$$[\boldsymbol{\mathcal{X}}]_{:,:,k} = \boldsymbol{A} \cdot \mathrm{diag}\{\boldsymbol{s}_k\} \cdot \boldsymbol{T}_k \,. \tag{11}$$

Here, r is the number of PARAFAC2 components (PARAFAC2 model order). The decomposition model (11) is not essentially unique without an additional constraint introduced by [9]

$$\boldsymbol{T}_k^\mathrm{T} \cdot \boldsymbol{T}_k = \boldsymbol{H} \in \mathbb{R}^{r \times r} \,. \tag{12}$$

This constraint together with the model equation (11) yields the PARAFAC2 decomposition. Please note that equation (12) constrains the sample cross-correlation matrix $\boldsymbol{H}$ between the time signatures of the different PARAFAC2 components to be constant over the channel index $k = 1 \ldots N_\mathrm{C}$. Thereby, relative time-shifts from channel to channel between the time signatures of each PARAFAC2 component are allowed. However, these relative time-shifts have to remain constant for all components.

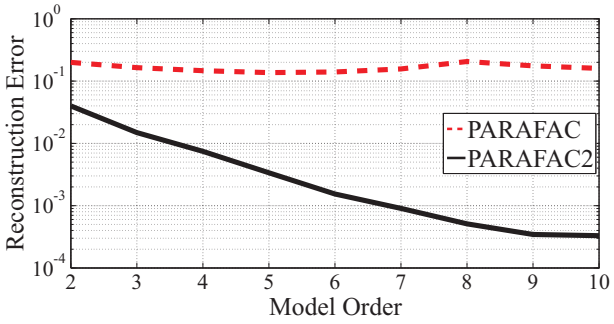

Figure 3: Reconstruction error for the PARAFAC and the PARAFAC2 model for different number of sources in a time window from 121 ms to 240 ms. The PARAFAC reconstruction error does not decrease with increasing model order, which indicates that the model does not fit to the data. The PARAFAC2 model is able to explain the VEP data with adequate number of signal components.

The direct fitting algorithm for the computation of the PARAFAC2 model used within this contribution was introduced by [11]. It is based on the fact that the model equations (11) and (12) can be transformed into a PARAFAC decomposition of the tensor $\boldsymbol{\mathcal{W}} \in \mathbb{R}^{N_\mathrm{F} \times r \times N_\mathrm{C}}$, which is given by

$$[\boldsymbol{\mathcal{W}}]_{:,:,k} = [\boldsymbol{\mathcal{X}}]_{:,:,k} \cdot \boldsymbol{P}_k = \boldsymbol{A} \cdot \mathrm{diag}\{\boldsymbol{s}_k\} \cdot \boldsymbol{D}^\mathrm{T} \,, \tag{13}$$

where $\boldsymbol{P}_k$ is a matrix of size $N_\mathrm{T} \times r$ with orthogonal columns, and $\boldsymbol{D}$ is of size $r \times r$. The matrix $\boldsymbol{P}_k$ can be computed via alternating least squares (ALS) iterations using the singular value decomposition (SVD) of

$$\boldsymbol{D} \cdot \mathrm{diag}\{\boldsymbol{s}_k\} \cdot \boldsymbol{A}^\mathrm{T} \cdot [\boldsymbol{\mathcal{X}}]_{:,:,k} = \boldsymbol{U}_k \cdot \boldsymbol{\Sigma}_k \cdot \boldsymbol{V}_k^\mathrm{T} \tag{14}$$

$$\boldsymbol{P}_k = \boldsymbol{V}_k \cdot \boldsymbol{U}_k^\mathrm{T} \tag{15}$$

The time signatures of the PARAFAC2 decomposition of $\boldsymbol{\mathcal{X}}$ are then given by

$$\boldsymbol{T}_k = \boldsymbol{P}_k \cdot \boldsymbol{D} \,. \tag{16}$$

Please note that for the solution of (13) an arbitrary PARAFAC algorithm can be used. For the computation of the PARAFAC decomposition the most common methods to date are based on iterative alternating least squares algorithms. However, these algorithms may require many iterations and are not guaranteed to converge to the global minimum. In this contribution we use the recently developed closed-form PARAFAC algorithm [16], which outperforms the iterative approaches.

3.3.1 Closed-Form PARAFAC

The closed-form PARAFAC algorithm is based on the higher order singular value decomposition (HOSVD) of the tensor $\boldsymbol{\mathcal{W}}$ which is defined as [12]

$$\boldsymbol{\mathcal{W}} = \boldsymbol{\mathcal{Q}} \times_1 \boldsymbol{U}_1 \times_2 \boldsymbol{U}_2 \times_3 \boldsymbol{U}_3 \,, \tag{17}$$

where $\boldsymbol{\mathcal{Q}} \in \mathbb{R}^{N_\mathrm{F} \times r \times N_\mathrm{C}}$ is the full core tensor of same size as $\boldsymbol{\mathcal{W}}$. The unitary matrices $\boldsymbol{U}_1 \in \mathbb{R}^{N_\mathrm{F} \times N_\mathrm{F}}$, $\boldsymbol{U}_2 \in \mathbb{R}^{r \times r}$ and $\boldsymbol{U}_3 \in \mathbb{R}^{N_\mathrm{C} \times N_\mathrm{C}}$ provide an orthonormal basis for the 1-mode, 2-mode, and 3-mode vector spaces of $\boldsymbol{\mathcal{W}}$, respectively. Thus, the HOSVD can easily be obtained from the matrix singular value decomposition of the n-mode matrix unfoldings of $\boldsymbol{\mathcal{W}}$ [12]. In the non-degenerate case ($r \leq \min\{N_\mathrm{F}, N_\mathrm{C}\}$) the HOSVD of the tensor $\boldsymbol{\mathcal{W}}$ can be truncated to

$$\boldsymbol{\mathcal{W}} = \boldsymbol{\mathcal{Q}}^{[\mathrm{s}]} \times_1 \boldsymbol{U}_1^{[\mathrm{s}]} \times_2 \boldsymbol{U}_2^{[\mathrm{s}]} \times_3 \boldsymbol{U}_3^{[\mathrm{s}]} \,, \tag{18}$$

where $\mathcal{Q}^{[s]} \in \mathbb{R}^{r\times r\times r}$ and where $\boldsymbol{U}_1^{[s]}$, $\boldsymbol{U}_2^{[s]}$ and $\boldsymbol{U}_3^{[s]}$ are of size $(N_F \times r)$, $(r \times r)$, and $(N_C \times r)$, respectively. By defining the matrix $\boldsymbol{C} = [\boldsymbol{c}_1, \ldots \boldsymbol{c}_r] \in \mathbb{R}^{N_C \times r}$ we can rewrite the PARAFAC model (8) in terms of the identity tensor $\mathcal{I}_{3,r}$

$$\mathcal{W} = \mathcal{I}_{3,r} \times_1 \boldsymbol{A} \times_2 \boldsymbol{D} \times_3 \boldsymbol{C} \,. \tag{19}$$

Comparing the equations (18) and (19) indicates that there is a link between the PARAFAC model and the HOSVD. To exploit this connection we define the transformation matrices $\boldsymbol{T}_1 \in \mathbb{R}^{r\times r}$, $\boldsymbol{T}_2 \in \mathbb{R}^{r\times r}$, and $\boldsymbol{T}_3 \in \mathbb{R}^{r\times r}$ such that

$$\boldsymbol{A} = \boldsymbol{U}_1^{[s]} \cdot \boldsymbol{T}_1 \,, \boldsymbol{D} = \boldsymbol{U}_2^{[s]} \cdot \boldsymbol{T}_2 \,, \boldsymbol{C} = \boldsymbol{U}_3^{[s]} \cdot \boldsymbol{T}_3 \,. \tag{20}$$

Inserting these equations into (19) and comparing it with (18) yields

$$\mathcal{Q}^{[s]} \times_1 \boldsymbol{T}_1^{-1} \times_2 \boldsymbol{T}_2^{-1} \times_3 \boldsymbol{T}_3^{-1} = \mathcal{I}_{3,r} \,. \tag{21}$$

Therefore, the closed-form PARAFAC algorithm estimates the transformation matrices that diagonalize the truncated core tensor $\mathcal{Q}^{[s]}$ to the identity tensor $\mathcal{I}_{3,r}$. In [16] it is shown that this can be accomplished very efficiently by means of joint matrix diagonalizations, also in the degenerate case. The resulting closed-form algorithm outperforms iterative approaches especially in critical scenarios, since it does not require alternating least squares iterations. By considering different criteria for the choice of the final component estimates, it provides the opportunity to obtain a tradeoff between accuracy and computational complexity.

3.3.2 Scaling ambiguity in PARAFAC2

The PARAFAC2 model equation (11) can be reformulated in terms of component tensors as

$$\mathcal{X} = \sum_{n=1}^{r} \mathcal{Z}^{(n)} = \sum_{n=1}^{r} \boldsymbol{a}_n \circ (\boldsymbol{F}_n \cdot \mathrm{diag}\{\boldsymbol{c}_n\}) \,, \tag{22}$$

where $\boldsymbol{F}_n = [\boldsymbol{t}_{1,n}, \boldsymbol{t}_{2,n}, \ldots, \boldsymbol{t}_{N_C,n}]$ is the matrix of time signatures for the n-th PARAFAC2 component of size $N_T \times N_C$. Similarly to the PARAFAC decomposition, the PARAFAC2 decomposition in (22) together with (12) is essentially unique up to scaling and permutation. In order to resolve the scaling ambiguity and to judge the influence of the components, we extend the idea of least squares PARAFAC component amplitudes [19] to the PARAFAC2 model. To this end we introduce the following normalization.

$$\boldsymbol{a}'_n = \frac{\boldsymbol{a}_n}{\|\boldsymbol{a}_n\|_F} \,, \boldsymbol{t}'_{k,n} = \frac{\boldsymbol{t}_{k,n}}{\|\boldsymbol{t}_{k,n}\|_F} \,, \boldsymbol{c}'_n = \frac{\boldsymbol{c}_n}{\|\boldsymbol{c}_n\|_F} \,, \tag{23}$$

for $n = 1 \ldots r$ and $k = 1 \ldots N_C$. Here $\|\cdot\|_F$ denotes the Frobenius norm. Please note that these normalization leads to $\|\boldsymbol{a}'_n \circ (\boldsymbol{F}'_n \cdot \mathrm{diag}\{\boldsymbol{c}'_n\})\|_H = 1$ with $\boldsymbol{F}'_n = [\boldsymbol{t}'_{1,n}, \boldsymbol{t}'_{2,n}, \ldots, \boldsymbol{t}'_{N_C,n}]$. Next we introduce the PARAFAC2 component amplitudes γ_n by

$$\mathcal{X} \approx \sum_{n=1}^{r} \gamma_n \cdot \boldsymbol{a}'_n \circ (\boldsymbol{F}'_n \cdot \mathrm{diag}\{\boldsymbol{c}'_n\}) = \sum_{n=1}^{r} \gamma_n \cdot \mathcal{Z}^{(n)\prime} \,. \tag{24}$$

To determine all amplitudes jointly we rewrite this equation by applying the vec operator to

$$\mathrm{vec}\{\mathcal{X}\} = [\mathrm{vec}\{\mathcal{Z}^{(1)\prime}\}, \mathrm{vec}\{\mathcal{Z}^{(2)\prime}\}, \ldots, \mathrm{vec}\{\mathcal{Z}^{(r)\prime}\}] \cdot \boldsymbol{\gamma} \,, \tag{25}$$

where the vector $\boldsymbol{\gamma} = [\gamma_1, \ldots, \gamma_r]^T$ contains all component amplitudes. The least squares solution for the set of linear equations (25) is given by

$$\boldsymbol{\gamma} = \left[\mathrm{vec}\{\mathcal{Z}^{(1)\prime}\}, \mathrm{vec}\{\mathcal{Z}^{(2)\prime}\}, \ldots, \mathrm{vec}\{\mathcal{Z}^{(r)\prime}\}\right]^{+} \mathrm{vec}\{\mathcal{X}\} \,. \tag{26}$$

In practical applications the PARAFAC2 model often does not exactly fit the data, and no apriori knowledge can be used to resolve the scaling and permutation ambiguity. In these cases we suggest to judge the influence of the components based on the magnitudes of the component amplitudes γ_n.

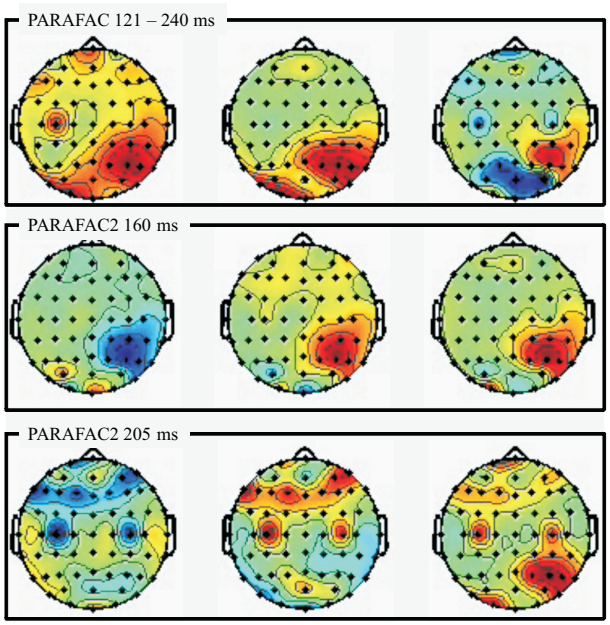

Figure 4: PARAFAC and PARAFAC2 signal components for the TFA based on the reduced interference distribution. The analysis window reaches from 121 to 240 ms. The PARAFAC2 analysis clearly indicates that there is an occipital component over the right visual cortex as well as a short component at 205 ms over the motoric center (electrodes C3 / C4). The temporal location of the components is resolved only by the PARAFAC2 analysis.

4. EXPERIMENTAL RESULTS

For the experimental validation of the component analysis algorithm, we apply both the PARAFAC and the PARAFAC2 decomposition on the measured visual evoked potentials (VEP) presented in Figure 2. In this data-set a strong positive wave is observed around 100 ms and 200 ms. This P100 and P200 component is well known in literature for this kind of VEP. However, both components appear slightly time-shifted on the different channels, e.g., they are observed slightly earlier on the occipital parts. This clearly indicates the presence of moving sources, which cannot be analyzed via the PARAFAC decomposition. This is also reflected by the relative reconstruction error (27) of the PARAFAC model, which is depicted in the upper part of Figure 3.

$$E_P(d) = \frac{\|\mathcal{X} - \sum_{n=1}^{d} \boldsymbol{a}_n \circ \boldsymbol{b}_n \circ \boldsymbol{c}_n\|_H}{\|\mathcal{X}\|_H} \,. \tag{27}$$

It is clearly recognized that the relative reconstruction error does not decrease with increasing number of PARAFAC components (note that for each model order the decompositions (8) and (11) are recomputed). Therefore, the PARAFAC model is not able to explain the measured data, since the time-shifts lead to a dramatic growth of the model order. On the other hand, the PARAFAC2 model shows a significant degradation in the relative reconstruction error (28) for increasing number of PARAFAC2 components.

$$E_{P2}(r) = \frac{\|\mathcal{X} - \sum_{n=1}^{r} \boldsymbol{a}_n \circ (\boldsymbol{F}_n \cdot \mathrm{diag}\{\boldsymbol{c}_n\})\|_H}{\|\mathcal{X}\|_H} \,. \tag{28}$$

Using a model order of $r = 3$, PARAFAC2 is able to explain more than 98% of the measured data. This clearly indicates that the PARAFAC2 decomposition is able to resolve the signal components of moving (time-shifted) EEG sources. Moreover, the PARAFAC2

model allows us to track the signal components over the time. This is achieved by combining the PARAFAC2 time signatures $\boldsymbol{F}_n$ and the channel signatures $\boldsymbol{c}_n$ to a matrix $\boldsymbol{G}_n = \boldsymbol{F}_n \cdot \mathrm{vec}\{\boldsymbol{c}_n\}$ of size $N_\mathrm{T} \times N_\mathrm{F}$ such that the PARAFAC2 model (22) becomes

$$\boldsymbol{\mathcal{X}} = \sum_{n=1}^{r} \boldsymbol{a}_n \circ (\boldsymbol{F}_n \cdot \mathrm{diag}\{\boldsymbol{c}_n\}) = \sum_{n=1}^{r} \boldsymbol{a}_n \circ \boldsymbol{G}_n \,. \qquad (29)$$

Therefore, each PARAFAC2 component can be interpreted as outer product of a constant frequency signature $\boldsymbol{a}_n$ and a time-varying channel signature $\boldsymbol{G}_n$. Please note that this enables us to track the channel signatures over time with the same resolution as the original signal $\boldsymbol{\mathcal{X}}$ (note that it is not necessary to compute the PARAFAC2 model in overlapping time windows to achieve the time tracking.). The time tracking of channel signatures is not possible with the PARAFAC component analysis. The time-varying PARAFAC2 channel signature provide new insights into the component analysis of EEG data, as demonstrated in Figure 4. The upper part of this Figure shows three PARAFAC channel signatures within a time window ranging from 120 ms to 240 ms. They show a strong component on the right occipital parts of the scalp, as well as some activities on the motoric center (C3 / C4 electrode). However, it is not possible to extract the exact temporal location of these components. Moreover, we observed that distinct temporally separated sources mix up to one PARAFAC channel signature, since the PARAFAC model has to describe the whole data window with rank-one components. This effect can be seen in comparison with the temporally exactly located PARAFAC2 channel signatures in the middle (160 ms) and the lower part (205 ms) of Figure 4. At 160 ms the PARAFAC2 components indicate a clean component originating from the visual cortex on the right occipital part of the scalp, whereas the motoric components appear very shortly 45 ms later over the motoric center (C3 / C4 electrode).

5. CONCLUSION

In this contribution we have derived a novel concept for the component analysis of multi-channel EEG data by applying the PARAFAC2 decomposition on the time-frequency distributions of all channels. We have shown that the PARAFAC2 model is able to cope with time-shifted signal components originating from moving EEG sources. Therefore, the PARAFAC2 decomposition clearly outperforms the commonly used PARAFAC decomposition in case of highly dynamic signals. Additionally, we have demonstrated how the PARAFAC2 model can be used to implement a temporal tracking of the channel signatures for each component. Thereby, the temporal resolution is the same as for the original EEG signal, which provides new insights into the temporal evolution of EEG components. In order to increase the robustness of the direct fitting PARAFAC2 algorithm we use the recently developed closed-form PARAFAC algorithm for the computation of the PARAFAC2 model. Moreover, we introduced the least squares PARAFAC2 amplitudes in order to resolve the scaling ambiguities of the PARAFAC2 model. Furthermore, we suggest to use these amplitudes to judge the influence of the PARAFAC2 components.

ACKNOWLEDGMENTS

The authors gratefully acknowledge the Thüringer Aufbaubank (TAB) for sponsoring this work within the PsychoPort Project (support-code 2007 FE 0150).

REFERENCES

[1] B. M. Abadi, D. Jarchi, and S. Sanei, "Simultaneous localization and separation of biomedical signals by tensor factorization", in *Proc. IEEE/SP 15th Workshop on Statistical Signal Processing*, pp. 61–78, 2009.

[2] R. Bro, "PARAFAC. tutorial and applications", in *Chemometrics and Intelligent Laboratory Systems, 38*, pp. 149 – 171, 1997.

[3] R. Bro, C. A. Andersson, and H. A. L. Kiers, "PARAFAC2 — Part II. Modeling chromatographic data with retention time shifts", *J. Chemometrics*, vol. 11, pp. 295–309, 1997.

[4] G. E. Chatrian, E. Lettich, and P. L. Nelson, "Ten percent electrode system for topographic studies of spontaneous and evoked EEG activities.", in *American Journal of EEG Technology*, vol. 25, pp. 83–92, 1985.

[5] H. Choi and W.J. Williams, "Improved time-frequency representation of multicomponent signals using exponential kernels", in *IEEE Trans. Acoust., Speech, Signal Process.*, vol. 37, no. 6, pp. 862–871, 1989.

[6] L. Cohen, "Time-frequency analysis: Theory and applications", Prentice Hall, Upper Saddle River, NJ, 1995.

[7] C. M. Guerrero, A. M. Triguerosand, and J. I. Franco, "Time-frequency EEG analysis in epilepsy: What is more suitable?", in *Proceedings of the Fifth IEEE International Symposium on Signal Processing and Information Technology*, pp. 202 – 207, Athens, Greece, 2005.

[8] R. A. Harshman, "Foundations of the PARAFAC procedure: Models and conditions for an "explanatory" multi-modal factor analysis", *UCLA Working Papers in Phonetics*, vol. 16, pp. 30 pp., 1970.

[9] R. A. Harshman, "PARAFAC2: Mathematical and technical notes", *UCLA Working Papers in Phonetics*, vol. 22, pp. 30–44, 1972.

[10] P. Husar, S. Berkes, and M. Drozd, "An approach to adaptive beamforming in measurement of EEG", in *Proc. of the European Biomedical Engineering Congress (EMBEC)*, pp. 1438 – 1439, Vienna, Austria, 2002.

[11] H. A. L. Kiers, J. M. F. Ten Berge, and R. Bro, "PARAFAC2 — Part I. A direct fitting algorithm for the PARAFAC2 model", *J. Chemometrics*, vol. 13, pp. 349–374, 1999.

[12] L. De Lathauwer, B. de Moor, and J. Vandewalle, "A multilinear singular value decomposition", *SIAM J. Matrix Anal. Appl.*, vol. 21, no. 4, 2000.

[13] 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", in *Neuroimage 22(3)*, pp. 1035–1045, 2004.

[14] M. Mørup, L. K. Hansen, and C. S. Hermann, "Parallel factor analysis as an exploratory tool for wavelet transformed event-related EEG", in *NeuroImage 29(3)*, pp. 938–947, 2006.

[15] R. D. Pascual-Marqui, C. M. Michel, and D. Lehmann, "Low resolution electromagnetic tomography: A new method for localizing electrical activity in the brain", in *Int. J. Psychophysiol. 18*, pp. 49 – 65, 1994.

[16] F. Roemer and M. Haardt, "A closed-form solution for parallel factor (PARAFAC) analysis", in *Proc. IEEE Int. Conf. Acoust., Speech, and Signal Processing (ICASSP)*, pp. 2365–2368, Las Vegas, NV, April 2008.

[17] J. Ville, "Théorie et applications de la notion de signal analytique", in *Câbles et transmissions*, vol. 2A, pp. 66–74, 1948.

[18] M. De Vos, L. De Lathauwer, B. Vanrumste, S. Van Huffel, and W. Van Paesschen, "Canonical decomposition of ictal scalp EEG and accurate source localisation: Principles and simulation study", in *Comput. Intell. Neurosci.*, 2007.

[19] M. Weis, F. Roemer, M. Haardt, D. Jannek, and P. Husar, "Multi-dimensional space-time-frequency component analysis of event-related EEG data using closed-form PARAFAC", in *Proc. IEEE Int. Conference on Acoustics, Speech, and Signal Processing (ICASSP)*, Taipei, Taiwan, Apr. 2009.

[20] E. P. Wigner, "On the quantum correction for thermodynamic equilibrium", in *Phys. Rev.*, vol. 40, pp. 749–759, 1932.