EVALUATION OF MULTI-DIMENSIONAL DECOMPOSITION MODELS USING SYNTHETIC MOVING EEG POTENTIALS

Judith Mengelkamp

Biomedical Engineering Group
Ilmenau University of Technology
P. O. Box 100565, D-98684 Ilmenau, Germany
judith.mengelkamp@tu-ilmenau.de

Martin Weis, Peter Husar

Biosignal Processing Group
Ilmenau University of Technology
P. O. Box 100565, D-98684 Ilmenau, Germany
{martin.weis, peter.husar}@tu-ilmenau.de

ABSTRACT

To identify the scalp projections of the underlying sources of neural activity based on recorded electroencephalographic (EEG) signals, the multi-dimensional decomposition models Parallel Factor Analysis (PARAFAC) and Parallel Factor Analysis 2 (PARAFAC2) have recently attained interest. We evaluate the models based on synthetic EEG data, because this allows an objective assessment by comparing the estimated projections to the parameters of the sources. We simulate EEG data using the EEG forward solution and focus on dynamic sources that change their spatial projection over time. Recently, this type of signal has been identified as the dominant type of signal, e. g. in measurements of visual evoked potentials. Further, we develop a method to objectively evaluate the decomposition models. We show that the decomposition models reconstruct the scalp projections successfully from data with low signal-to-noise ratio (SNR). They perform best if the number of calculated components (model order) equals the number of sources.

***Index Terms*—** Multi-dimensional signal processing, Synthetic EEG data, Moving scalp projections, Forward solution, PARAFAC2

1. INTRODUCTION

Multi-channel EEG data is a summation of potentials that originate from specific and unspecific brain activity of the subject. Specific neural activity refers to specific tasks of the subject. If we assume that two different centers of activity are activated successively (dynamic sources), specific neural activity results into temporally overlapping activity. This can yield a virtual movement of the potential pattern in the measured EEG signals giving rise to the concept of a visually moving neural source. These movements have recently been investigated in the analysis of visual evoked potentials (VEP) (see Figure 1) [11]. Furthermore, unspecific neural fluctuations result into background EEG that superimposes the specific EEG signals. The signal-to-noise ratio (SNR) of measured EEG data is very low.

In neural signal processing it is of major interest to identify the scalp projections of specific EEG signals and to separate them from background EEG. To decompose EEG data into its components the multi-dimensional decomposition models PARAFAC and PARAFAC2 have been introduced. Possible applications are the detection of scalp projections of epileptic seizures and of brain regions that control specific tasks. Based on the analysis of measured EEG data it is known that the PARAFAC model can only analyze the scalp projections of static sources that vary over the channels by a scalar factor [2, 5, 10]. The PARAFAC2 model can also resolve the temporal variation of moving scalp potentials [11]. This renders the PARAFAC2 model a promising approach in the analysis of dynamic EEG sources.

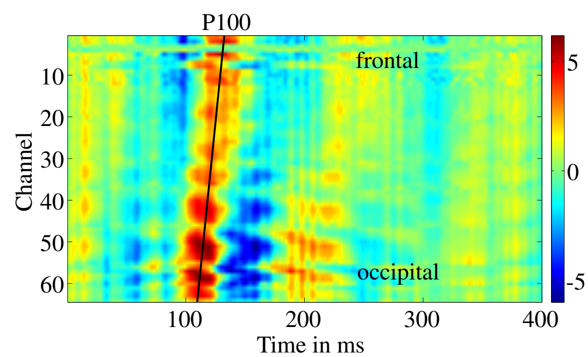


Fig. 1: Time course of a visual evoked potential. The subject is stimulated with a flash light. The typical P100 response wave occurs earlier at the occipital than at the frontal channels [11].

For all applications it is of great interest to analyze how effective the multi-dimensional decomposition models estimate the scalp projections of the underlying sources. However, an objective evaluation requires to compare the estimated components to the parameters of the underlying sources. Such a verification is difficult on measured EEG data, because the original EEG sources are unknown. Therefore, no reference data for a comparison with the calculated components is available and no objective measure for the performance of the decomposition methods can be defined.

To overcome this drawback we apply the decomposition models to synthetic EEG data. Thereby, we generate synthetic EEG signals that reflect all relevant characteristics of physiological EEG signals. We simulate EEG data with the help of the EEG forward solution. Thus, we calculate scalp potentials, that derive from intracranial neural activity, at given electrode positions based on electromagnetic field theory. We model the neural activity with equivalent current dipoles and the human head with an analytic volume conductor model. In particular, we introduce synthetic EEG signals with moving scalp potentials as observed in measured EEG data. To simulate such EEG signals we use the moving dipole model [8]. In the last years, the forward solution with static dipoles has been applied to evaluate decomposition models, e. g. in [1, 12]. Further, we simulate background EEG with 1/f noise, because we obtain the same characteristics from measured EEG data. We constrain our evaluation to synthetic EEG data based on one dipole source.

Based on the synthetic EEG data we develop a method to compare the estimated components and the scalp projections of the dipole sources objectively. We analyze in detail if the models are able to separate EEG signals from data with low SNR. Further, we assess the influence of the model order (cf. Section 2). This is of special interest, because in measured EEG data the number of underlying sources is usually unknown and the decomposition is likely to be calculated with an incorrect model order.

The paper is structured as follows. In Section 2, we outline the multi-dimensional models PARAFAC and PARAFAC2. In Section 3, we explain the synthetic EEG data based on the EEG forward solution. In this context, we describe the applied volume conductor model and the simulation of background EEG. In Section 4, we explain our method to evaluate the performance of the multi-dimensional models objectively. Finally, in Section 5 we evaluate how the SNR and the model order influence the performance of the decomposition models before drawing the conclusions in Section 6.

2. THREE-DIMENSIONAL COMPONENT ANALYSIS USING PARAFAC AND PARAFAC2

In order to apply the multi-dimensional models to EEG data we transform the EEG data matrix into a three-dimensional tensor by applying a time-frequency analysis (TFA) to each channel. For the TFA we apply the Reduced Interference Distribution with a Choi-Williams kernel [3]. This approach provides very useful results in the analysis of EEG data especially in combination with tensor decomposition models [10]. The resulting EEG data tensor is denoted by

$$\boldsymbol{\mathcal{Y}} \in \mathbb{R}^{N_\mathrm{F} \times N_\mathrm{T} \times N_\mathrm{C}}, \tag{1}$$

where N_F is the number of frequency samples, N_T the number of time samples and N_C the number of channels.

The PARAFAC model decomposes the tensor $\boldsymbol{\mathcal{Y}}$ into a minimum number of rank-one component tensors $\boldsymbol{\mathcal{C}}_\mathrm{P}^{(r)}$. If the EEG data is superimposed with noise, the PARAFAC decomposition is expressed as

$$\boldsymbol{\mathcal{Y}} = \sum_{r=1}^{R_\mathrm{P}} \boldsymbol{\mathcal{C}}_\mathrm{P}^{(r)} + \boldsymbol{\mathcal{E}} = \sum_{r=1}^{R_\mathrm{P}} \boldsymbol{a}_r \circ \boldsymbol{b}_r \circ \boldsymbol{c}_r + \boldsymbol{\mathcal{E}}. \tag{2}$$

The PARAFAC component vectors $\boldsymbol{a}_r \in \mathbb{R}^{N_\mathrm{F}}$, $\boldsymbol{b}_r \in \mathbb{R}^{N_\mathrm{T}}$ and $\boldsymbol{c}_r \in \mathbb{R}^{N_\mathrm{C}}$ with $r = 1, ..., R_\mathrm{P}$ represent the frequency signature, the time signature and the channel signature of the r-th PARAFAC component. The error tensor $\boldsymbol{\mathcal{E}}$ contains the residuals that are not estimated by the PARAFAC components. Further, R_P is the number of PARAFAC components (PARAFAC model order). The matrix $\boldsymbol{b}_r \circ \boldsymbol{c}_r \in \mathbb{R}^{N_\mathrm{T} \times N_\mathrm{C}}$ describes how the channel signatures vary over time. Since it is restricted to rank one, the PARAFAC model can only analyze static components that vary over the channels by a scalar factor.

In contrast to the PARAFAC model the components of the PARAFAC2 model have channel-dependent time signatures. Thus, the PARAFAC time signature $\boldsymbol{b}_r$ is extended to a matrix $\boldsymbol{F}_r \in \mathbb{R}^{N_\mathrm{T} \times N_\mathrm{C}}$ that contains the time signatures of the r-th PARAFAC2 component $\boldsymbol{\mathcal{C}}_\mathrm{P2}^{(r)}$. The PARAFAC2 model can be formulated as

$$\begin{aligned} \boldsymbol{\mathcal{Y}} &= \sum_{r=1}^{R_\mathrm{P2}} \boldsymbol{\mathcal{C}}_\mathrm{P2}^{(r)} + \boldsymbol{\mathcal{E}} = \sum_{r=1}^{R_\mathrm{P2}} \boldsymbol{a}_r \circ (\boldsymbol{F}_r \cdot \mathrm{diag}\{\boldsymbol{c}_r\}) + \boldsymbol{\mathcal{E}} \\ &= \sum_{r=1}^{R_\mathrm{P2}} \boldsymbol{a}_r \circ \boldsymbol{G}_r + \boldsymbol{\mathcal{E}}, \end{aligned} \tag{3}$$

with R_P2 being the PARAFAC2 model order. The matrix $\boldsymbol{G}_r$ contains the time-varying channel signatures of the r-th PARAFAC2 component. It is the equivalent to the matrix $\boldsymbol{b}_r \circ \boldsymbol{c}_r$ of the PARAFAC model, however it can obey full rank. Therefore, the PARAFAC2 components can exceed tensor rank one and allow to analyze time shifts between the channels. An analysis of the matrix $\boldsymbol{G}_r$ enables us to resolve the scalp projections of the different components with the same resolution as the original EEG data.

Both models are unique up to a permutation and a scaling ambiguity. The permutation ambiguity refers to the fact that the order of the components is arbitrary. The scaling ambiguity refers to the fact that the component vectors can be multiplied with arbitrary scalar values α_r, β_r and γ_r without changing the decomposition, i. e. for the PARAFAC2 model,

$$\boldsymbol{\mathcal{Y}} = \sum_{r=1}^{R_\mathrm{P2}} (\alpha_r \boldsymbol{a}_r) \circ (\beta_r \boldsymbol{F}_r \cdot \mathrm{diag}\{\gamma_r \boldsymbol{c}_r\}) + \boldsymbol{\mathcal{E}}, \tag{4}$$

as long as $\alpha_r \beta_r \gamma_r = 1$ for $r = 1, ..., R_\mathrm{P2}$ [4, 11].

3. SIMULATION OF EEG DATA USING THE EEG FORWARD SOLUTION

3.1. The Volume Conductor Model

The applied volume conductor model is an analytic four-layer shell model. Starting from the inside the shells represent the brain, the cerebrospinal fluid, the skull and the scalp. The conductivities of the shells are $0.33\,\mathrm{S/m}$, $1\,\mathrm{S/m}$, $0.0042\,\mathrm{S/m}$ and $0.33\,\mathrm{S/m}$, and the radii are 63 mm, 65 mm, 71 mm and 75 mm, respectively. This model has been applied in a large number of source localization algorithms, e. g. in [9]. Furthermore, for the EEG measurements we apply the international 10-10 electrode system [7].

3.2. The EEG Forward Solution for Static Dipoles

The EEG forward solution considering static dipoles is defined as

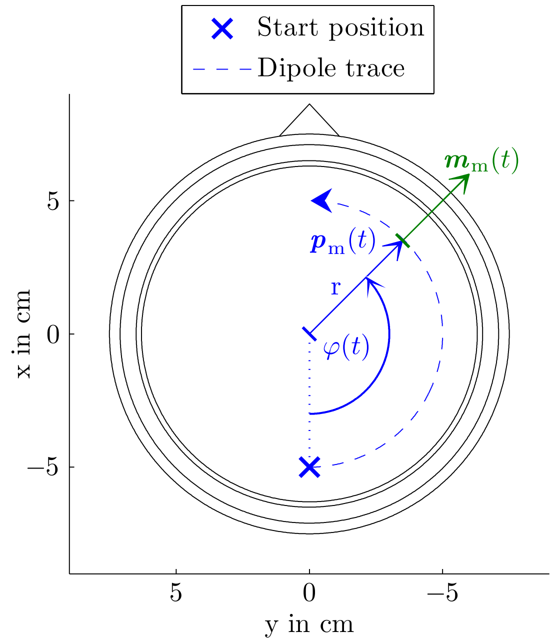
$$\boldsymbol{y}(t) = \boldsymbol{L} \cdot \boldsymbol{x}(t) + \boldsymbol{e}(t), \tag{5}$$

where $\boldsymbol{x}(t) \in \mathbb{R}^{3 \cdot N_\mathrm{D}}$ contains the moments of the N_D dipole sources. It is $\boldsymbol{y}(t) \in \mathbb{R}^{N_\mathrm{C}}$ the vector of the synthetic EEG signals in the N_C channels and $\boldsymbol{e}(t) \in \mathbb{R}^{N_\mathrm{C}}$ the noise vector. The lead field matrix $\boldsymbol{L} \in \mathbb{R}^{N_\mathrm{C} \times 3 \cdot N_\mathrm{D}}$ defines the projections from the dipole sources at discrete positions in the volume conductor model to the measurement channels. The superscript $3 \cdot N_\mathrm{D}$ indicates that each source can be assigned with specific 3-D spatial information (x,y,z). We perform the forward simulations using the Matlab toolbox FieldTrip [6]. For the PARAFAC analysis (see Figure 5) we simulate EEG alpha waves (10-Hz sinus) by placing one static dipole into the brain model.

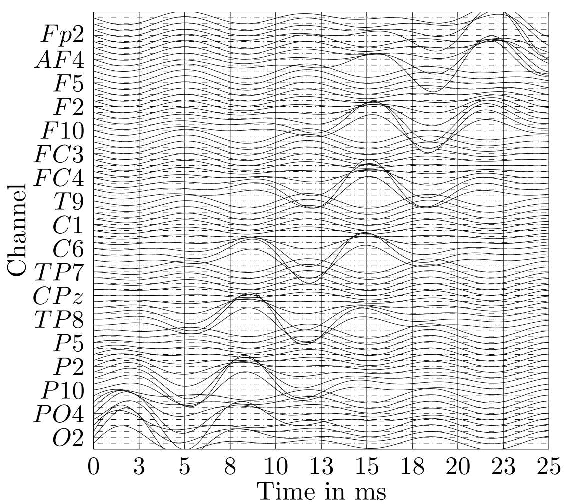
3.3. Synthetic EEG Data Including Moving Scalp Potentials

We generate EEG data with moving scalp projections by placing a moving dipole into the volume conductor model. Thus, the lead field matrix is time-dependent, $\boldsymbol{L} = \boldsymbol{L}(t)$, and the discrete positions of the dipole sources are allowed to vary over time.

Figure 2a shows the trace of the moving dipole. The position is denoted by $p_\mathrm{m}(t)$ and the moment by $m_\mathrm{m}(t)$. The dipole moves in the xy-plane of the volume conductor model from the occipital to the frontal cortex on a half circle with the radius $r = 50$ mm. The starting position is in the occipital cortex of the volume conductor model, since a real VEP signal is generated in the visual center located in the occipital cortex. The angular velocity of the moving dipole is constant over time: $d\varphi/dt = 40\,\pi\,\mathrm{rad/s}$. Further, the dipole moment is always perpendicular to the surface of the volume conductor model.



(a) Trace of the moving dipole source. The axis labels correspond to the orientation of the volume conductor model. The x-coordinate points towards the noise. The dipole position and moment are denoted by $p_{\mathrm{m}}(t)$ and $m_{\mathrm{m}}(t)$, respectively. The dipole moves from the starting position indicated by the blue cross to the end position indicated by the blue arrow. The trace is a half circle with constant radius. The moment is always perpendicular to the surface of the volume conductor model.



(b) Synthetic EEG data based on one moving dipole. The simulated wave form is a sinus signal. The scalp projections of the moving dipole source appear earlier at the occipital electrodes (O2) than at the frontal ones (Fp2). The time shifts are in the same magnitude as in measured EEG data (see Figure 1). Please note that only every third electrode is labeled.

Fig. 2: Simulation of EEG signals with moving scalp projections. The dynamic neural activity is modeled with one moving dipole.

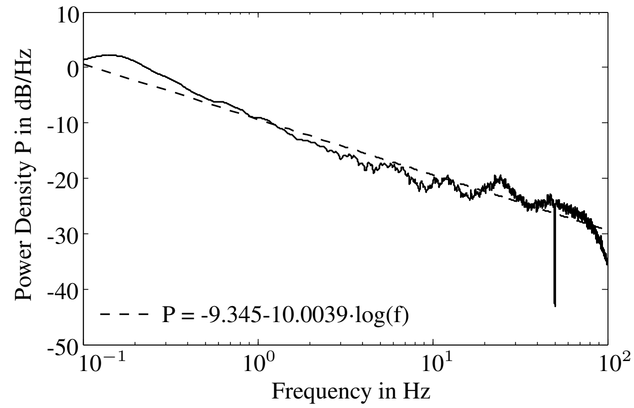


Fig. 3: Spectrum of measured background EEG. The EEG data is recorded from a 23-year old healthy male, who sits in front of a white wall, keeps the eyes open and does not perform any task. The data is filtered and eye blinks are reduced. The spectrum is averaged over all channels.

This parameter setting results from the fact that EEG measurement systems are only sensitive to radial sources.

The resulting synthetic EEG data is depicted in Figure 2b. The simulated wave form is a 150-Hz sinus signal. It is clearly visible that the spatial projection of the moving dipole changes over time. Comparing the synthetic EEG data to the measured EEG data (see Figure 1), we see that the time shifts are in the same magnitude. Thus, the synthetic EEG data reflects well the virtual movement of physiological potential patterns, which we intend to analyze with the decomposition models.

3.4. Recording and Simulation of Background EEG

We simulate background EEG with 1/f noise, because we obtain the same characteristics from measured EEG data. We record human resting EEG from a 23-year old healthy male who sits in front of a white wall, keeps the eyes open and does not perform any task. The sampling frequency is set to 512 samples per second. The international 10-10 electrode system with common reference is applied [7]. During measurements we ensure that there is no crosstalk between the channels.

For the preprocessing of the raw data we apply the following filters: a 100 Hz low-pass, a 0.1 Hz high-pass and a 50-Hz notch filter. Further, we apply the FastICA to reduce eye blinks. Then, we estimate the power spectrum of each channel using the Welch method. Since the spectra of the channels are similar, we average over the channels (see Figure 3). We calculate a line of best fit and obtain a gradient of approximately -10 Decibel per Decade. This gradient is characteristic for 1/f noise. Moreover, we calculate the correlations between the channels. The results show that there are high correlations between neighboring channels. This proves that the measured EEG data represents a neurological process and not only amplifier noise. We add the 1/f noise to the EEG data of the dipoles (see Equation (5)).

4. EVALUATION OF ESTIMATED COMPONENTS

To evaluate how effective the multi-dimensional models estimate specific EEG signals in noisy EEG data, we compare the estimated

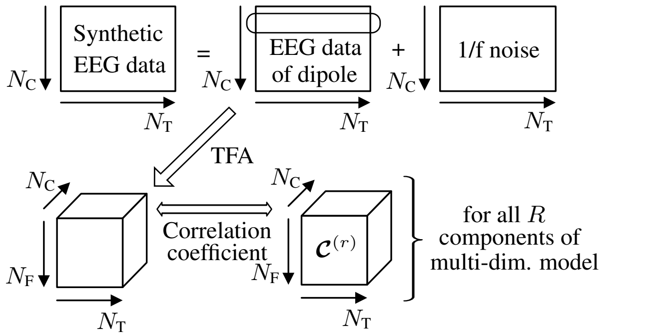


Fig. 4: Evaluation of the estimated components. The EEG data of the dipole is transformed into a tensor of the same size as the component tensors by applying a time-frequency analysis (TFA) to each channel. We calculate the Pearson correlation coefficient between the dipole tensor and each component tensor. It is $r = 1, ..., R$ with R being the model order.

components to the corresponding scalp projections of the dipoles. To achieve this, we transform the two-dimensional synthetic EEG data of the dipole into a three-dimensional tensor by applying a time-frequency analysis to each channel (Figure 4). For the time-frequency analysis we apply the Reduced Interference Distribution (RID) with the same parameters as in the preprocessing of the decomposition. As a result, the dipole tensor and the component tensors are of the same size. Please note that an inverse transformation of the component tensors is not possible, because the RID does not preserve the phase information.

Then, we calculate the Pearson correlation coefficient between the dipole tensor and each component tensor. Thus, the number of calculated correlation coefficients for one decomposition equals the model order. The model component that has the highest absolute correlation estimates best the scalp projections of the dipole. The correlation coefficient proves to be a good measure, because it does not depend on the absolute amplitude but instead on the relative shape of the signal.

5. RESULTS OF MULTI-DIMENSIONAL COMPONENT ANALYSIS

For an objective assessment of the decomposition models we apply both the PARAFAC and PARAFAC2 model to synthetic EEG data. We evaluate the influence of the SNR and of the model order by varying both parameters. Additionally, we analyze the temporal variation of the scalp projections of a moving dipole source with the PARAFAC2 model.

We fit the PARAFAC model to the synthetic EEG data based on one static dipole and superimposed with 1/f noise (see Sections 3.2 and 3.4). The SNR of the EEG data ranges from -40 dB to 40 dB. Further, the PARAFAC model order is varied: $R_\mathrm{P} \in \{1, 2, 3\}$. To evaluate the results objectively we calculate the correlation coefficient (see Section 4).

Figure 5 shows the correlation coefficient in dependence on the SNR of the EEG data and on the PARAFAC model order. If the SNR is above -6 dB, the correlation coefficient of all model orders exceeds 0.8. If the SNR exceeds 0 dB, the correlation coefficient for $R_\mathrm{P} = 1$ is close to one and constant over the SNR. This shows that the PARAFAC model is able to reconstruct the scalp projections of a static dipole in EEG data with low SNR. The reconstruction is

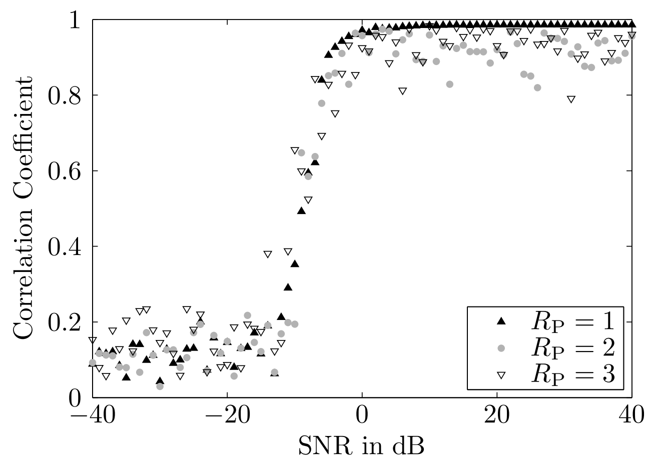


Fig. 5: Correlation coefficients of the PARAFAC model. The EEG data is based on one static dipole and superimposed with 1/f noise. The PARAFAC model order is varied: $R_\mathrm{P} \in \{1, 2, 3\}$.

most accurate if the model order equals the number of sources. An overfitting of the PARAFAC model, $R_\mathrm{P} \in \{2, 3\}$, results in a less accurate decomposition. Further, we fit the PARAFAC model to the EEG data with one moving source. We obtain that the PARAFAC components do not allow to analyze time-varying EEG signals, since the components are rank-one tensors.

We show that the PARAFAC2 model can analyze the time-varying scalp projections of a moving dipole by calculating the PARAFAC2 model on the synthetic EEG data based on one moving dipole not superimposed with noise (see Figure 2b in Section 3.3). The PARAFAC2 model order is set to one.

Figure 6 compares the topographic maps of the time-varying amplitudes of the synthetic EEG data to the time-varying PARAFAC2 channel signatures summarized in the matrix $\boldsymbol{G}_r$ (see Equation (3)). The comparison is shown for seven different time samples that are equally spaced within the time axis. The individual topographic maps are normalized. It is clearly visible that the PARAFAC2 model can reconstruct the time-dependent spatial projection of the moving dipole source. However, there are visible differences especially at the beginning (0 ms) of the signal. These differences are due to the temporal smearing of the time-frequency analysis. Moreover, the PARAFAC2 model cannot reconstruct the sign of the components because of the scaling ambiguity. Further, we fit the PARAFAC2 model to the synthetic EEG data based on one static dipole. The results show, that the PARAFAC2 model estimates the static EEG component with a similar accuracy as the PARAFAC model.

To assess how the performance of the PARAFAC2 model depends on the SNR of the EEG data and the model order, we calculate the PARAFAC2 model on the synthetic EEG data with moving scalp projections and superimposed with 1/f noise (see Sections 3.3 and 3.4). The SNR of the EEG data is varied between -40 dB and 40 dB. Moreover, the PARAFAC2 model is calculated using different model orders: $R_\mathrm{P2} \in \{1, 2, 3\}$. Figure 7 shows the resulting correlation coefficients.

If the SNR is positive and $R_\mathrm{P2} = 1$, the correlation coefficient exceeds 0.8. If the SNR exceeds 10 dB, the correlation is close to one and constant for all models orders. Thus, the PARAFAC2 model reconstructs the scalp projections best from EEG data with low SNR, if the model order equals the number of sources. If the SNR exceeds

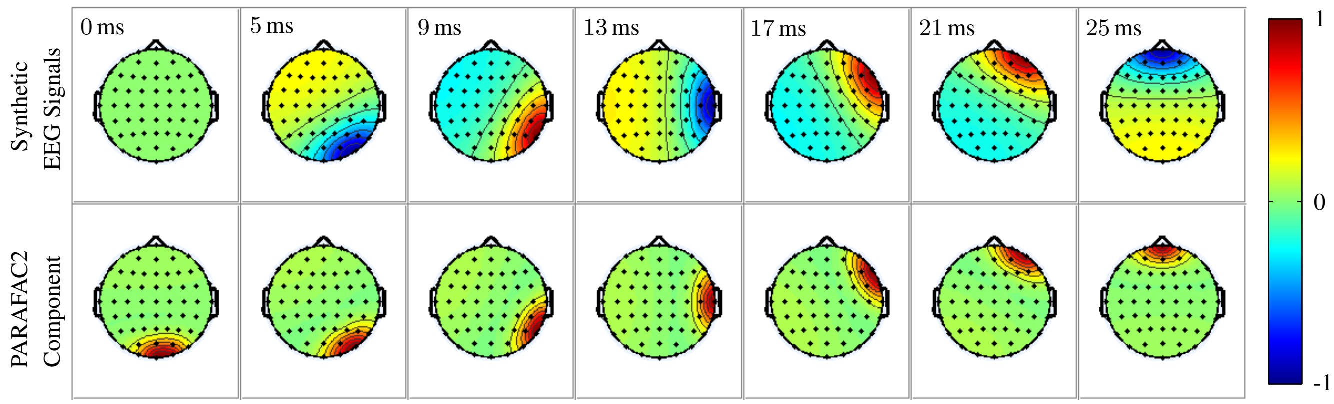


Fig. 6: Comparison of the time-varying amplitudes of the synthetic EEG data (upper row) and the time-varying channel signatures of the PARAFAC2 model (lower row). The amplitudes and channel signatures are shown in topographic maps for seven different time samples. The individual topographic maps are normalized. The analyzed EEG data is based on one moving dipole and depicted in Figure 2b. The PARAFAC2 model order is set to one.

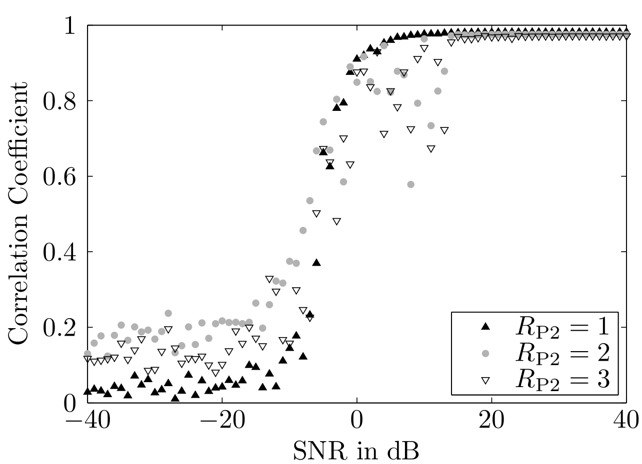


Fig. 7: Correlation coefficients of the PARAFAC2 model. The EEG data is based on one moving dipole and superimposed with 1/f noise. The PARAFAC2 model order is varied: $R_{P2} \in \{1, 2, 3\}$.

10 dB, an overfitting of the PARAFAC2 model does not result into a less accurate decomposition.

6. CONCLUSION

In this contribution we evaluated the multi-dimensional decomposition models PARAFAC and PARAFAC2 based on synthetic EEG data. This enabled us to compare the estimated components to the scalp projections of the underlying sources. We simulated EEG data using the EEG forward solution. Thereby, we focused on simulating scalp projections that appear time shifted over the channels by using the moving dipole model. We verified that the synthetic EEG data reflects the characteristics of measured EEG data. To objectively assess the results of the decomposition models we calculated the correlation coefficient between the estimated components and the scalp projections of the dipole sources. The results show that the decomposition models can reconstruct the scalp projections of the dipole sources from synthetic EEG data with low SNR. Further, they provide best results if the model order equals the number of sources.

7. REFERENCES

[1] B. Abadi, D. Jarchi, and S. Sanei, "Simultaneous localization and separation of biomedical signals by tensor factorization," in *15th Workshop on Statistical Signal Processing*, September 2009, pp. 497 –500.

[2] E. Acar, C. Aykut-Bingol, H. Bingol, R. Bro, and B. Yener, "Multiway analysis of epilepsy tensors," *Bioinformatics*, vol. 23, no. 13, pp. i10–i18, 2007.

[3] L. Cohen, *Time Frequency Analysis: Theory and Applications*. Prentice Hall, December 1994.

[4] T. G. Kolda and W. B. Bader, "Tensor decompositions and applications," *SIAM Review*, vol. 51, June 2009.

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

[6] R. Oostenveld, P. Fries, E. Maris, and J.-M. Schoffelen, "FieldTrip: open source software for advanced analysis of MEG, EEG, and invasive electrophysiological data," *Intell. Neuroscience*, pp. 1:1–1:9, January 2011.

[7] R. Oostenveld and P. Praamstra, "The five percent electrode system for high-resolution EEG and ERP measurements," *Clinical Neurophysiology*, vol. 112, no. 4, pp. 713 – 719, 2001.

[8] M. Scherg, "Fundamentals of dipole source potential analysis," in *Auditory Evoked Magnetic Fields and Electric Potentials*, G. F., M. Hoke, and R. G., Eds. S. Karger, 1990, vol. 6, pp. 40–69.

[9] P. Schimpf, C. Ramon, and J. Haueisen, "Dipole models for the EEG and MEG," *IEEE Transactions on Biomedical Engineering*, vol. 49, no. 5, pp. 409 –418, May 2002.

[10] M. Weis, F. Romer, 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)*, April 2009, pp. 349 –352.

[11] M. Weis, F. Romer, M. Haardt, D. Jannek, P. Husar, and T. Günther, "Temporally resolved multi-way component analysis of dynamic sources in event-related EEG data using PARAFAC2," in *Proc. IEEE 18th European Signal Processing Conference (EUSIPCO)*, August 2010, pp. 696–700.

[12] J. Yao and J. P. Dewald, "Evaluation of different cortical source localization methods using simulated and experimental EEG data," *NeuroImage*, vol. 25, no. 2, pp. 369 – 382, 2005.