

A Comparison of Non-negative Tucker Decomposition and Parallel Factor Analysis for Identification and Measurement of Human EEG Rhythms

Zuzana Rošťáková¹, Roman Rosipal^{1,2}, Saman Seifpour¹, Leonardo Jose Trejo²

¹*Department of Theoretical Methods, Institute of Measurement Science, Slovak Academy of Sciences, Dúbravská cesta 9, 841 04, Bratislava, Slovakia, zuzana.rostakova@savba.sk*

²*Pacific Development and Technology, LLC, 999 Commercial Street, Suite 205 94303, Palo Alto, CA, USA*

Analysis of changes in the brain neural electrical activity measured by the electroencephalogram (EEG) plays a crucial role in the area of brain disorder diagnostics. The elementary latent sources of the brain neural activity can be extracted by a tensor decomposition of continuously recorded multichannel EEG. Parallel factor analysis (PARAFAC) is a powerful approach for this purpose. However, the assumption of the same number of factors in each dimension of the PARAFAC model may be restrictive when applied to EEG data. In this article we discuss the potential benefits of an alternative tensor decomposition method – the Tucker model. We analyze situations, where in comparison to the PARAFAC solution, the Tucker model provides a more parsimonious representation of the EEG data decomposition. We show that this more parsimonious representation of EEG is achieved without reducing the ability to explain variance. We analyze EEG records of two patients after ischemic stroke and we focus on the extraction of specific sensorimotor oscillatory sources associated with motor imagery during neurorehabilitation training. Both models provided consistent results. The advantage of the Tucker model was a compact structure with only two spatial signatures reflecting the expected lateralized activation of the detected subject-specific sensorimotor rhythms.

Keywords: multichannel electroencephalogram, sensorimotor oscillatory brain activity, parallel factor analysis, Tucker model.

1. INTRODUCTION

In the area of brain research, multichannel electroencephalographic (EEG) data provide a rich and complex representation of changes of the brain's neural activity over time, either during rest or cognitive tasks. However, the interpretation of the raw multichannel EEG data for quantitative applications is not straightforward. A long-term interest in electrophysiology is therefore focused on developing a methodology for extracting representation of multichannel EEG data that provides compact quantitative metrics for research and clinical applications [1].

One way to achieve a more compact and easily interpretable form of data is the dimensionality reduction approach, which can be carried out by the detection of latent sources of data variability. If data can be rearranged into the form of a matrix; that is the conventional two-dimensional representation, a latent structure can be detected by the principal component analysis (PCA), factor analysis (FA), or independent component analysis (ICA) methods [2], [3]. However, due to the multichannel character of EEG data, presence of repeated trials and often used focus on EEG signal characteristics in the frequency domain, the three-way or multi-way array (tensor) data representation has advantages over a ma-

trix form, where unfolding of selected dimensions into the two-way representation is needed.

A generalization of PCA into higher dimensions was simultaneously developed by Harshman [4] as the parallel factor analysis (PARAFAC), and by Carroll and Chang [5] as the canonical decomposition (CANDECOMP) models. A straightforward implementation and natural interpretation of the results makes the method a popular multi-way approach for multichannel EEG data decomposition [1], [6].

Our long-term research interest is focused on a decomposition of EEG data recorded during the neurorehabilitation training of patients suffering from hemiplegia. The hemiplegia or a paralysis of one side of the body occurs frequently as a consequence of an ischemic stroke [7]. A visual-feedback training procedure in which a patient is asked to imagine a movement of a paretic limb, but without real movements of the limb, represents a new approach, which we have studied intensively [7], [8]. The impact of such rehabilitation is reflected in changes of latent sources of the neural brain activity, as represented by multichannel EEG data.

We performed neurorehabilitation training with a proprietary robotic splint [9], [10]. During the procedure the patient sits in a chair with the paretic hand fixed in the robotic

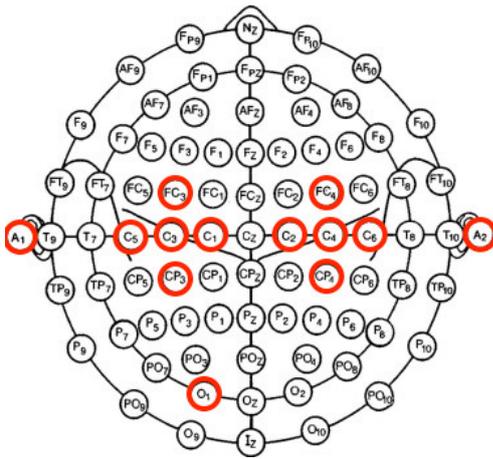


Fig. 1. EEG electrodes placement. The electrode array used in this study is highlighted by red circles. The original image comes from [18].

splint. After several seconds of eyes open or eyes closed resting state, the patient is asked to imagine movement of his or her paretic hand. If the targeted sensorimotor oscillatory neural activity recorded through the scalp EEG meets a given threshold, the patient is rewarded by a real movement of the hand executed by the splint.

The PARAFAC method has been used successfully to extract subject-specific sensorimotor EEG rhythms [9], [11]; however, the assumption of the same number of latent factors within each dimension and simple structure linking factors of the PARAFAC solution may be too restrictive. In the following sections we denote these latent factors as *atoms* [1].

In psychometrics, chemometrics and related areas, a generalized and more flexible version of PARAFAC, called the Tucker model, was developed and often used [12], [13]. There are a few studies applying the Tucker model to EEG data. Moreover, the focus of these studies is either directed to feature extraction or EEG artifact removal [14], [15], [16], [17]. However, a closer focus on the interpretation of the obtained decomposition within the expected neurophysiological changes driven by the applied experimental design is missing in these studies.

To fill this gap and to better understand the applicability of the Tucker model to EEG data decomposition, in this study

- we focus on atoms of the Tucker model decomposition with clear neurophysiological interpretations,
- we aim to answer the question of whether these atoms are in agreement or complimentary to the PARAFAC decomposition and whether EEG decomposition can be obtained with a more parsimonious Tucker model structure,
- finally, as a consequence of this analysis we focus on a constrained non-negative version of the Tucker model in order to improve stability and interpretability of the obtained decomposition when applied to EEG data.

2. DATA

In the study we analyzed multichannel EEG of two chronic hemiplegic patients recorded over several training sessions with the robotic splint [10].

The EEG was recorded with five pairs of symmetrically placed left- and right-hemisphere electrodes (Fig.1.) and referenced to averaged earlobe electrodes. See [11] for details.

For the first subject, five days, each consisting of three three-minute long blocks of the rehabilitation training were used. The total length of the EEG recording was 43.2 minutes. During the resting state, as well as during the motor imagery part, the subject kept his eyes closed. Subject 2 took part in the robotic splint training over 38 weeks with a frequency of two trainings per week. During these training sessions, the conditions of closed and opened eyes during the resting or motor imagery periods were alternating. To match the EEG records of the first subject, in the study only training blocks with the closed eyes were selected from 13 recording days. The total length of the selected EEG recordings for the second subject was 40.9 minutes.

2.1. Data preprocessing

In the first step, the EEG recordings of both patients were segmented into two-second long time windows with 250 ms of overlap. EEG artifacts detection was performed in the semi-automatic and manual mode by a trained technician using the BrainVision Analyzer 2 software [19]; details can be found in [11]. Only artifact-free time windows were considered in the analysis.

Next, for every two-second time window we applied the irregular resampling auto-spectral analysis (IRASA) [20]. IRASA decomposes the amplitude spectrum of each window into a fractal (scale-free) and an oscillatory component, which is the difference between the total spectrum and the fractal component. The oscillatory and fractal components of EEG may be generated by different mechanisms so it is important to estimate them separately, especially when the focus of the measurement is on localized oscillatory sources, as is the case here [21], [22]. The oscillatory part of the spectrum was obtained by subtracting the fractal part from the total spectrum estimate. Negative values of the oscillatory spectrum were set to zero and the transformation $10 * \log_{10}(x + 1)$ was applied. Transformed oscillatory spectrum data were arranged into a three-dimensional tensor $\tilde{X} \in \mathbb{R}_+^{I \times J \times K}$ (*time* \times *electrodes* \times *frequencies*), where I represents the number of all two-second time windows, $J = 10$ is the number of electrodes, and $K = 43$ is the number of frequencies selected. In this study we consider the frequency range between 4 and 25 Hz with a step size of 0.5 Hz. Lower and higher bands were excluded on the assumption that the retained band contained most or all of the EEG signals to be controlled by subjects. Finally, the tensor \tilde{X} was centered across the first dimension

$$X_{ijk} = \tilde{X}_{ijk} - \frac{1}{I} \sum_{l=1}^I \tilde{X}_{ljk}, \quad i = 1, \dots, I; j = 1, \dots, J; k = 1, \dots, K.$$

In order to detect a subject-specific oscillatory activity, the tensor X was constructed separately for Subject 1 and 2.

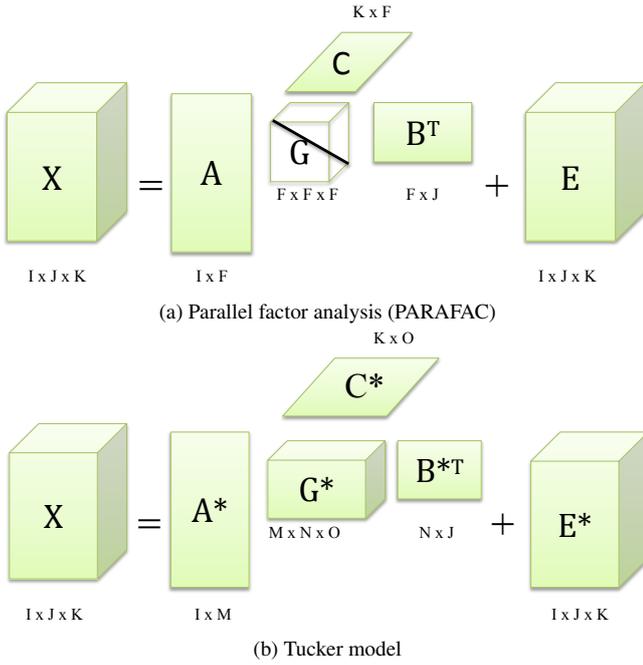


Fig.2. Graphical schemes of the tensor decomposition models used in the study.

3. METHODS

First, we would like to highlight the terminology and notations used in this study. The PARAFAC model decomposes EEG into unique space-time-frequency factors which, as mentioned above, we denote as *atoms*. In our setting atoms represent narrow-band EEG rhythms. The strength of the presence of each EEG atom over the time is measured by the time loadings obtained either by the PARAFAC or the Tucker decomposition for which we use the term *time scores* (TS).

The columns of the decomposition matrices in the PARAFAC or the Tucker model (Sections 3.1 and 3.2), corresponding to the space and frequency parts of the decomposition, represent loading vectors. Following [1], we use the term spatial and frequency *signatures* (SS, FS) to denote these loading vectors. Note that while in the PARAFAC model each atom is represented by a single space signature and a single frequency signature, in the Tucker model each atom can be linked to several frequency and space signatures as well as time score vectors.

3.1. Parallel factor analysis

The PARAFAC method decomposes the tensor $X \in \mathbb{R}^{I \times J \times K}$ into matrices $A \in \mathbb{R}^{I \times F}$, $B \in \mathbb{R}^{J \times F}$, and $C \in \mathbb{R}^{K \times F}$ and a core tensor $G \in \mathbb{R}^{F \times F \times F}$

$$X_{ijk} = \sum_{f=1}^F g_{fff} a_{if} b_{jf} c_{kf} + e_{ijk}, \quad (1)$$

$$i = 1, \dots, I; j = 1, \dots, J; k = 1, \dots, K,$$

by minimizing the sum of squared residuals

$$\sum_{i=1}^I \sum_{j=1}^J \sum_{k=1}^K \left(X_{ijk} - \sum_{f=1}^F g_{fff} a_{if} b_{jf} c_{kf} \right)^2$$

under the constraints

$$\|a_f\|^2 = \sum_{i=1}^I a_{if}^2 = 1, \quad \|b_f\|^2 = \sum_{j=1}^J b_{jf}^2 = 1,$$

$$\|c_f\|^2 = \sum_{k=1}^K c_{kf}^2 = 1, \quad f = 1, \dots, F.$$

The tensor $E = (e_{ijk}) \in \mathbb{R}^{I \times J \times K}$ represents the noise or error term. As depicted in the graphical scheme of the model (Fig.2a.), the core tensor G has non-zero elements only on its main super-diagonal. Consequently, each column of A is related with only one column of B and one column of C . In the case of our data, the f^{th} , $f = 1, \dots, F$ column of A , B , and C represent the vectors of time scores, and space, and frequency signatures of the f^{th} atom.

In contrast to PCA, ICA or FA, the PARAFAC decomposition was proved to be unique under mild conditions [23].

To improve the physiological interpretation of the results, the loading matrices are often constrained to be orthogonal, non-negative, sparse, or smooth [6]. However, the orthogonality constraint does not follow a priori knowledge and expected physiological interpretation of our data solution. The time activation of different atoms is expected not to be independent, but correlated over short time subintervals, and the atoms can also share similar space locations.

Moreover, note that our input data are positive \log_{10} values representing the oscillatory part of the amplitude spectrum. Therefore, we applied the non-negativity constraint on the matrices A, B, C .¹ Similar constraints were considered in [15], and these constraints are also often used in a tensor decomposition of spectral datasets in chemometrics [24]. Finally, we focused on extracting atoms representing separate narrow-band EEG rhythms, therefore we applied the unimodality constraint on the columns of C .

3.2. Tucker model

A generalized version of PARAFAC for the tensor $X \in \mathbb{R}^{I \times J \times K}$ – the Tucker model – follows the formula

$$X_{ijk} = \sum_{m=1}^M \sum_{n=1}^N \sum_{o=1}^O g_{mno}^* a_{im}^* b_{jn}^* c_{ko}^* + e_{ijk}, \quad (2)$$

$$i = 1, \dots, I; j = 1, \dots, J; k = 1, \dots, K,$$

$$\|a_m^*\|^2 = 1, \quad m = 1, \dots, M,$$

$$\|b_n^*\|^2 = 1, \quad n = 1, \dots, N,$$

$$\|c_o^*\|^2 = 1, \quad o = 1, \dots, O.$$

The loading matrices A^*, B^*, C^* are not assumed to have the same number of columns and the core tensor G^* has an

¹We would like to highlight that a combination of orthogonality and non-negativity constraints is not possible.

arbitrary structure (Fig.2b.). The contribution of each loading vector to a given atom is defined by the elements of the core tensor G^* , which can be understood as a tensor of weights.

In the following text, the model with M time score vectors, N spatial signatures, and O frequency signatures is denoted as the (M, N, O) -Tucker model.

In general, and in contrast to PARAFAC, the Tucker decomposition is not unique [25], [6]. Any orthogonal rotation of a loading matrix A^* , B^* or C^* and the following inverse rotation of G^* produces a decomposition with the same data fit as the original solution. Another problem with the Tucker model is the need for careful interpretation of the extracted atoms due to the arbitrary structure of G^* , which links the loading vectors of the decomposition to each atom [25].

In chemometrics, these problems – rotation freedom of the solution and interpretation of G^* – are solved either by constraining chosen elements of G^* to fixed values (usually zeros) using a priori information about the investigated data [26], [27], or by rotating the solution to achieve G^* with sparse or special structure (super-diagonality, diagonal slices, etc.) [25]. Another option for improving the stability of the solution is restricting the loading matrices [26].

In the following analysis the Tucker model with i) unrestricted structure of G^* , and ii) non-negative G^* was taken into account. The same non-negativity and unimodality constraints were applied to the loading matrices A^* , B^* , C^* as in PARAFAC.

The restriction ii) was proposed due to the difficult interpretation of negative elements of G^* . Together with the non-negativity of the loading matrices, the model is known as the non-negative Tucker decomposition (NTD) [26], [6].

In our analysis, we did not want to constrain chosen elements of G^* to zeros in order to investigate and compare the obtained solution with the constrained NTD model. Therefore, to avoid instability of the solution, the algorithm was run several times and then the best model was chosen. The criteria of the model quality are described in the following section. We also tried to rotate the G^* core tensor in order to obtain a clear neurophysiological interpretation of the decomposition.

3.3. Measures of model quality

To formally evaluate the performance of the Tucker and PARAFAC models we considered two measures – the proportion of variance explained (*VarExpl*) and the core consistency diagnostic (*CorConDiag*) [28], [29].

The proportion of variance explained follows the formula

$$\text{VarExpl} = 1 - \frac{\sum_{i=1}^I \sum_{j=1}^J \sum_{k=1}^K (X_{ijk} - \hat{X}_{ijk})^2}{\sum_{i=1}^I \sum_{j=1}^J \sum_{k=1}^K X_{ijk}^2} \leq 1,$$

where

$$\hat{X}_{ijk} = \sum_{f=1}^F \hat{g}_{fff} \hat{a}_{if} \hat{b}_{jf} \hat{c}_{kf} \quad (\text{PARAFAC})$$

$$\hat{X}_{ijk} = \sum_{m=1}^M \sum_{n=1}^N \sum_{o=1}^O \hat{g}_{mno}^* \hat{a}_{im}^* \hat{b}_{jn}^* \hat{c}_{ko}^* \quad (\text{Tucker model}).$$

Estimates of vectors or matrices are labelled by “^”.

CorConDiag is a measure originally developed for the PARAFAC model and represents the appropriateness of the constraints applied to G (super-diagonal structure). In the first step, the matrices A, B, C and the tensor G are estimated by PARAFAC with F factors. Then, using these matrices, the G^{unconstr} core tensor of the (F, F, F) -Tucker model is estimated in the least-squares sense.

$$G^{\text{unconstr}} \in \arg \min_{H \in \mathbb{R}^{F \times F \times F}} \sum_{i=1}^I \sum_{j=1}^J \sum_{k=1}^K \left(X_{ijk} - \sum_{m=1}^F \sum_{n=1}^F \sum_{o=1}^F h_{mno} a_{im} b_{jn} c_{ko} \right)^2. \quad (3)$$

Finally

$$\text{CorConDiag}_{\text{PARAFAC}} = 1 - \frac{\sum_{m=1}^F \sum_{n=1}^F \sum_{o=1}^F (g_{mno} - g_{mno}^{\text{unconstr}})^2}{g_{mno}^2} \leq 1. \quad (4)$$

Values close to 1 indicate that the constrained structure of G is appropriate. *CorConDiag* < 0 reflects the need to relax the constraints and to use a more general model.

The methodology can be generalized also for the (M, N, O) -Tucker model with a restricted G^* structure [29]. In this case

$$\text{CorConDiag}_{\text{Tucker}} = 1 - \frac{\sum_{m=1}^M \sum_{n=1}^N \sum_{o=1}^O (g_{mno}^* - g_{mno}^{\text{unconstr}})^2}{g_{mno}^{*2}} \leq 1. \quad (5)$$

Similarly to (4), G^* represents the restricted core tensor estimator and G^{unconstr} of the unconstrained (M, N, O) -Tucker model is estimated in the least-squares sense similarly as in (3).

Moreover, the two measures can be used as an indicator of the appropriate selection of the number of atoms either in PARAFAC or in the Tucker model. However, Bro [24] claims that it is difficult to decide about the optimal value of this number and there is no general approach like the cross-validation or proportion of variance explained in PCA. For the Tucker model, this represents an even more challenging task, because the number of time, space, and frequency loading vectors of the decomposition can be different and their links to each extracted atom vary.

Therefore, these formal measures have to be considered with caution and they should serve more like an indicator about the model fit than strict measures upon which the final model should be selected. Plausible interpretation of the final model components needs to be considered.

3.4. Detection of subject-specific atoms

The \log_{10} -transformed oscillatory amplitude EEG spectrum and the non-negativity and unimodality constraints in

both investigated models are motivated by the aim to detect subject-specific oscillatory sources. Following our previous study [11], we hypothesized that the subject-specific PARAFAC atoms should be present in models with a sufficient number of factors F . With this in mind we applied several PARAFAC models with a varying number of atoms instead of a single model with a fixed size. A similar procedure was applied in the case of the Tucker model.

More specifically, we varied the number of factors F in PARAFAC between six and 20. In the Tucker model, the number of frequency signatures O varied between six and 12. The number of time score vectors M in the model with O frequency signatures varied between O and $2 * O$. According to the set of electrodes used, the number of spatial signatures N was set to 2, 3, or 4.

The subject-specific atoms were then selected in the next step, where a cluster analysis was applied to a set of all extracted atoms from all models. We applied the nonparametric density-based clustering; the DBSCAN algorithm [30]. The minimum number of “neighbors” in the method was set to two and the diameter of the eps-neighborhood of a point parameter was varied between 0.2 and 0.5 [30]. A distance between the i^{th} and the j^{th} atom was measured as the weighted sum of distances between their corresponding time scores, space and frequency signatures.

All analyses were performed in MATLAB®. We used APECSgui, a set of proprietary MATLAB codes developed by Pacific Development and Technology (2012) and the routines *parafac.m* and *tucker.m* with related subroutines from the public N -way toolbox for MATLAB [31]. For NTD, we replaced the standard least squares criterion for the core tensor estimation in the *tucker.m* routine by a non-negative least squares algorithm (*fastnnls.m* from the N -way toolbox). For all models, the algorithm stops either when the maximum number of iterations (set to 200) is reached or the relative change of the fit is smaller than 10^{-6} . The other parameters in *parafac.m* and *tucker.m* were kept at their default values.

4. RESULTS

4.1. Tucker model with unrestricted G^*

In the first step, a set of the (M, N, O) -Tucker models, $O \in \{6, \dots, 12\}$, $M \in \{O, O + 1, \dots, 2 * O\}$, $N \in \{2, 3, 4\}$ with no restrictions to the structure of G^* were applied to the EEG data. In order to avoid convergence to a local minimum, the algorithm was run five times and the solution with the highest *VarExpl* and *CorConDiag* was chosen.

The obtained decompositions showed several weaknesses. To graphically demonstrate some of the listed problems we selected the $(11, 2, 6)$ -Tucker model (Fig.3.). We observed:

- **interpretation of frequency signatures**

The unimodality of frequency signatures (FS) was violated (Fig.3., first row). Moreover, the estimated FS represent either low (≈ 4.5 Hz) oscillatory activity or noise (> 20 Hz) and therefore are not of interest in this study.

- **duplicate components**

Duplicate frequency signatures and time scores were observed in the solution in the majority of models.

- **relationship between loading vectors**

Difficult interpretation of the negative elements in G^* .

- **rotation of solution**

It was difficult to find an orthogonal rotation which would lead to G^* with simplified structure and at the same time not violating the non-negativity and unimodality constraints applied to the loading matrices.

- **problems to find a stable solution**

Regardless of the choice of M, N, O , the algorithm reached the maximum number of 200 iterations, but the relative change of fit was still above the set threshold of 10^{-6} . The number of iterations needed for PARAFAC and NTD to find a stable solution with the same stopping criterion was lower than 100.

4.2. Non-negative Tucker decomposition

In the second step, we restricted the core tensor G^* to be non-negative, that is the NTD model was applied. We observed that this restriction solved the problems described in Section 4.1. The convergence problem as well as the duplicity in time scores and frequency signatures diminished. The solution of the $(10, 2, 6)$ -NTD model is depicted in Fig.4., where frequency signatures with a clear unimodal structure can be observed. Moreover, the sparse structure of G^* is easier to interpret (Fig.4., second row) and no orthogonal rotation of the core tensor is needed.

Spatial signatures Following Fig.4., the estimated spatial signatures represent scalp EEG activation over the right (SS 1) and left (SS 2) hemisphere. The same spatial patterns were present in all NTD models with the fixed number of spatial signatures $N = 2$, a varied number of frequency signatures $O = 6, \dots, 12$, and time scores $M = O, \dots, 2 * O$. We would like to highlight that neither NTD nor PARAFAC have any prior information about laterality of the solution. Therefore, we hypothesized that two space components are enough for description of our data variability in the space domain.

A detailed inspection of models with $N > 2$ has shown that the increased number of space components provides the solution where the two lateralized spatial patterns are horizontally divided and represent scalp EEG activity over the right and left hemisphere. Importantly, the cluster analysis of the space signatures from all NTD models (Fig.5.) shows that these two dominant spatial patterns were present in the majority of models. Therefore, in the following analysis only the NTD models with two space signatures were further investigated. Note that the inference about the corresponding cortical activation in the left and right hemisphere is beyond the scope of this paper. For example, EEG source localization using an inverse mapping method applied to high-density EEG could be considered.

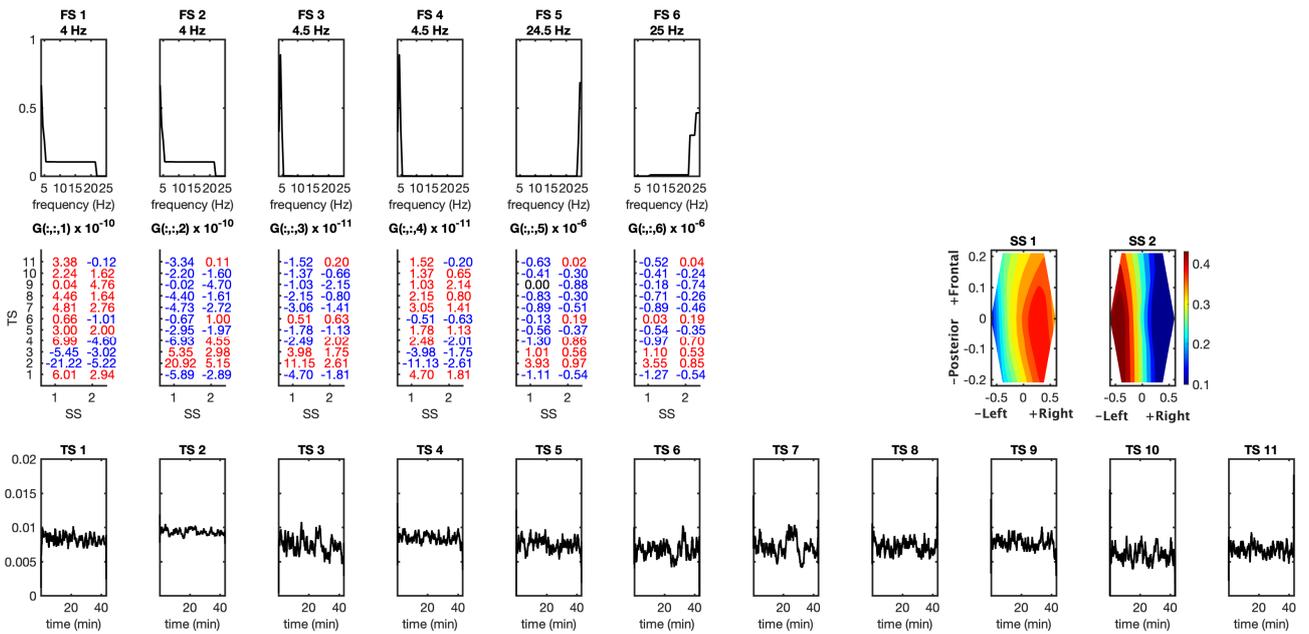


Fig. 3. Subject 1. The unconstrained (11, 2, 6)-Tucker decomposition. The frequency signatures (FS) (loading vectors) are depicted in the first row. Slices of G^* with the fixed frequency dimension are in the second row of the plot on the left. These slices show links between each frequency signature and two spatial signatures (SS, second row on the right) and 11 time score (TS, third row) vectors.

G^* structure The sparse structure of G^* allows to find clear links between the time scores, frequency and spatial signatures.

By fixing a given frequency signature, the corresponding slice of G^* (depicted in the second row of Fig.4.) is a matrix where each row relates to the set of extracted time scores and each of the two columns to the one of the spatial signatures. Elements of this matrix represent weights (links) between time scores and spatial signatures. By inspecting Fig.4., it can be observed that frequency signatures FS1, FS2, FS3, and FS5 show links of about the same numerical value for two different time scores related either to the left or right spatial signature. In other words, this solution indicates different activation of the given oscillatory component over the time recorded over the left and right hemisphere. In contrast, this is not the case for FS4 and FS6, where the links in the corresponding G^* slices indicate the same time activation over the left and right hemisphere. By inspecting raw EEG data and their spectral decomposition we identified that FS6 represents artifact oscillation centered at 22 Hz and affecting several central and centro-parietal EEG channels. The amplitude of this artifact varied across individual recording sessions and was strongest during one of the recording days. This period can be clearly identified in the trace of TS7 as a short burst with increased time score values.

At this point we investigated whether the PARAFAC model would provide a similar lateralized solution for the spatial signatures and size of the model. Using the PARAFAC model with nine atoms we observed spatial laterality for a subset of atoms (Fig.6.). After increasing the number of atoms to 14, we could observe lateralized spatial decomposition for the majority of atoms (Fig.7.). However, this is at the expense

of increasing the number of frequency components, which makes the physiological interpretation of the solution challenging.

Fig.8. compares in detail the left and right time scores of the 8-8.5 Hz atom (dominant sensorimotor mu rhythm) of NTD and PARAFAC for Subject 1. First, differences in the left and right time scores indicate different lateralized activation of this 8-8.5 Hz rhythm over some time periods. Second, the time scores of the PARAFAC and NTD models closely match, indicating that both methods estimate the same EEG sources.

In summary, these examples show that the compact form of the NTD model with a variable number of spatial and frequency signatures and time scores can outperform a PARAFAC model with a more rigid structure and an equal number of loading vectors for frequency, space, and time dimensions of the decomposition. In the next section, a more formal comparison of the two models is discussed.

4.3. Comparison of PARAFAC and NTD

In this section, PARAFAC and NTD are compared according to the *VarExpl* values.

As expected, *VarExpl* monotonically increases with the increasing number of components in both the PARAFAC and NTD models (Fig.9a.). In these examples, the value was reaching its maximum at about 0.5. When comparing *VarExpl* for PARAFAC with F and NTD with $N = 2$ spatial and $O = F$ frequency signatures and $M \in \{O, \dots, 2 * O\}$ time scores, the differences were negligible.

Moreover, we analyzed also the *CorConDiag* values to detect appropriateness of the non-negative (the NTD model) or

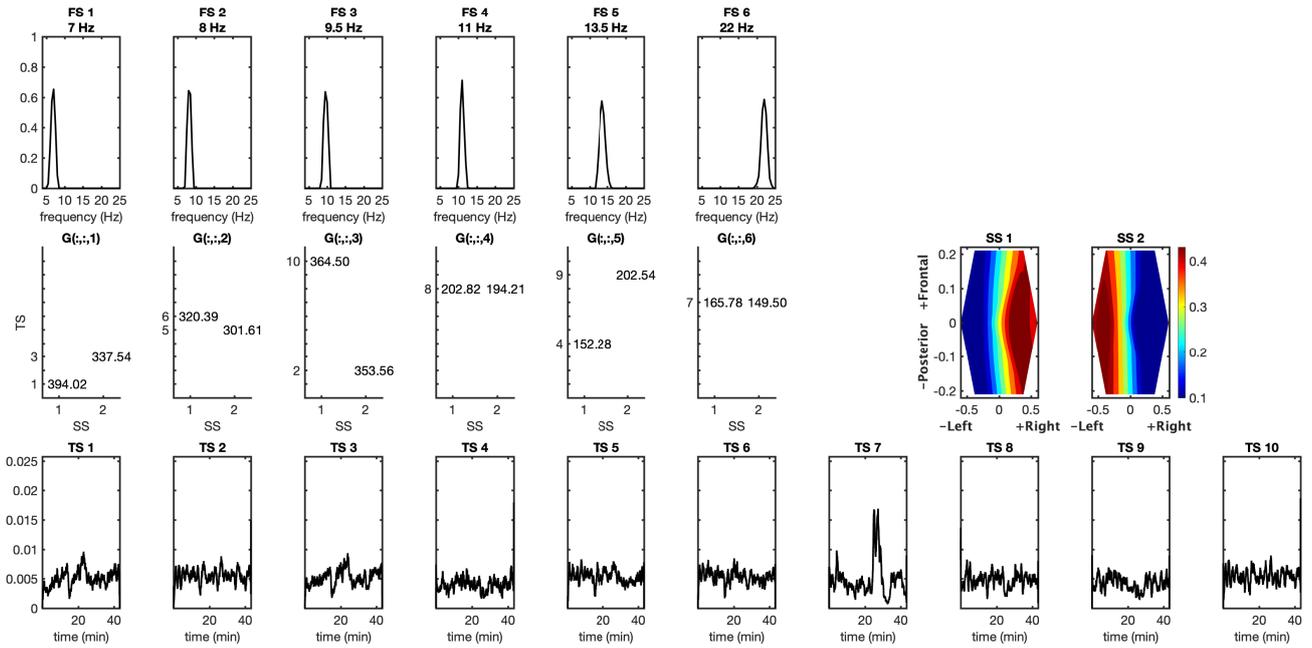


Fig. 4. Subject 1. The $(10, 2, 6)$ -Tucker decomposition with non-negative G^* (the NTD model). The frequency signatures (FS) are depicted in the first row. Slices of the core tensor G^* with the fixed frequency dimension are in the second row of the plot on the left. These slices depicts non-zero elements of G^* representing weights between each frequency signature and two spatial signatures (SS, second row on the right) and 10 time score (TS, third row) vectors.

super-diagonal structure (the PARAFAC model) of the core tensor. We would like to highlight that our aim is not a direct comparison of the *CorConDiag* values between the PARAFAC and NTD models, which would not be meaningful because (4) and (5) are computed using the unconstrained Tucker models with a different number of factors. Therefore, the analysis was performed separately for the PARAFAC and NTD models.

CorConDiag values for PARAFAC are stable across the models with different number of atoms, but in general small (Fig.9b.). The values increase with the number of components, but do not exceed the value of 0.5. This was true for both subjects. According to [28] this indicates that the super-diagonal structure of the core tensor is probably too restrictive for our data and should be relaxed.

CorConDiag for NTD was high regardless of the selected size of the model. In the majority of models it exceeds the value of 0.8 (Fig.9b.). This indicates that the NTD model is appropriate for the description of our data structure. The only exception in this trend were the $(8, 2, 8)$ - and $(16, 2, 10)$ -NTD models for Subject 1 with *CorConDiag* ≈ 0.12 , but we consider this as a random effect.

However, as mentioned before, both *VarExpl* and *CorConDiag* should be considered with caution when selecting the final PARAFAC or NTD model. For example, for NTD we may observe that all models seem to be of equal quality according to Fig.9b., therefore a further exploratory step to interpret the obtained solution needs to be applied. One of the solutions we used in this study was the cluster analysis applied to a set of solutions obtained by varying the model size.

4.4. Subject-specific sensorimotor oscillatory activity

In this section we focus on the interpretation of the obtained results, with the aim to detect the subject-specific EEG rhythms. For this purpose the DBSCAN cluster analysis was applied separately on extracted spatial and frequency signatures and time scores of the PARAFAC model with $F = 6, \dots, 20$ atoms and NTD model with $O = 6, \dots, 12$ frequency signatures, $M = O, \dots, 2 * O$ time scores, and $N = 2$ spatial signatures. The cluster-specific patterns (computed as cluster averages) for both subjects are depicted in Fig.10. (left hemisphere) and Fig.11. (right hemisphere).

Both subjects show activity at around 8 Hz, which represents the well-known mu rhythm associated with motor imagery. Synchronization and desynchronization changes of this rhythm were used to control the robotic splint during the training. Two other motor-related rhythms were identified and located at 11-11.5 Hz and at 13.5 Hz. This was true for both subjects. Following [32], we interpret them as the lower and higher sensorimotor rhythms (SMR). In addition, for Subject 2 we could identify two rhythms in the beta band, which we denote as the lower and higher beta. These beta rhythms were not clearly identified for Subject 1. Note, all these motor imagery related rhythms were interpreted following their change during the periods of motor imagery and resting conditions, their estimated spectral decomposition, as well as their spatial and frequency signatures and our experience with prior experiments with mirror box training of these two subjects [11]. In addition, we identified the posterior alpha rhythm (9.5 to 10 Hz) with its source in the occipital region of the brain as well as the frontal theta rhythm (6 to 7

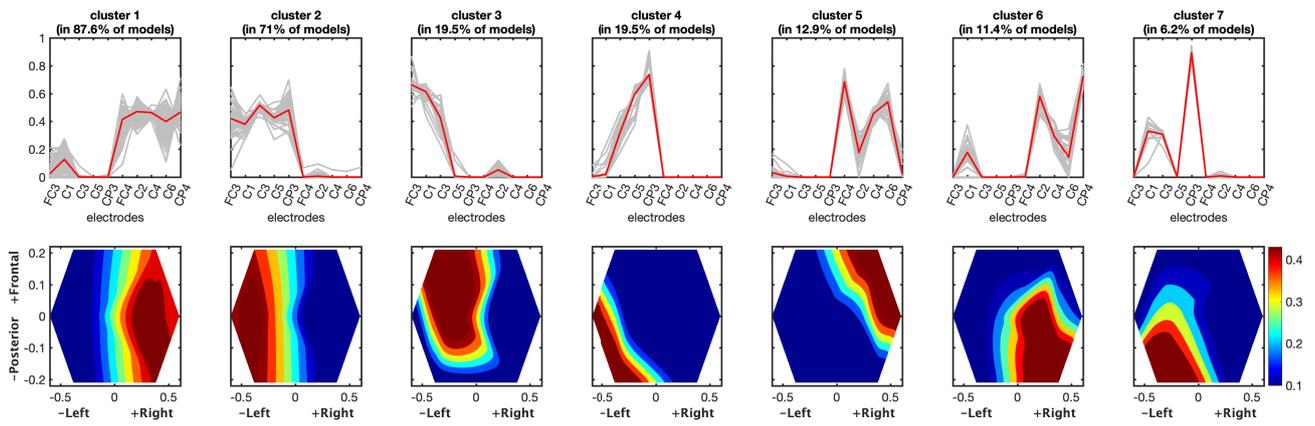


Fig. 5. Cluster analysis of the space components from all non-negative Tucker models (NTD). The patterns from the first and second cluster were present in at least 70% of models and they represent the location of scalp EEG activation over the right and left hemisphere.

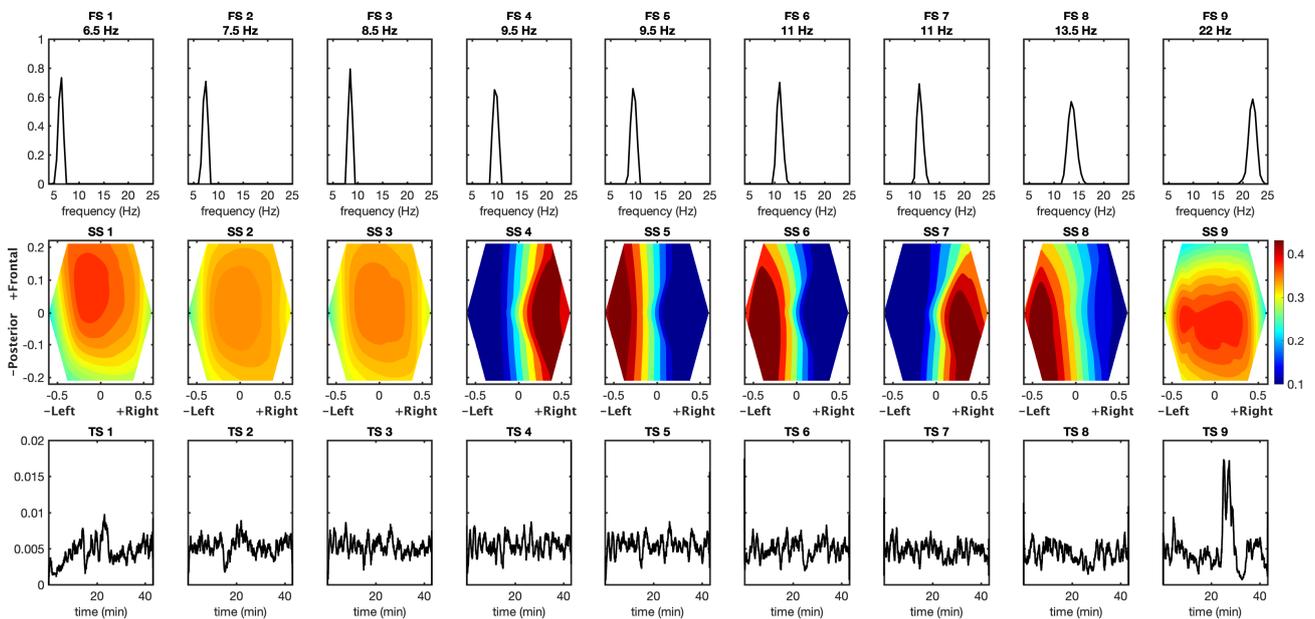


Fig. 6. Subject 1. The PARAFAC decomposition of the EEG data with 9 atoms (factors). The frequency signatures (FS), time scores (TS) and spatial signatures (SS) of the atoms are depicted in the first, second and third row, respectively.

Hz). To confirm visual alpha we also used spectral decomposition of the O1 EEG electrode signal recorded during the resting state eyes-opened and eyes-closed conditions before and after the training. We observed no systematic changes of the theta rhythm related to motor imagery.

5. DISCUSSION AND CONCLUSIONS

Identification and estimation of sources of neural activity in EEG recordings with the Tucker model is very rare in the literature. In this article, we discussed potential advantages of the Tucker model in comparison to the better-known PARAFAC model.

In the first part of the analysis we observed that when the unconstrained Tucker model was applied to multichannel EEG recordings, the results were difficult to interpret due to a complex structure of the core tensor elements. This was

specifically true for the task where, based on results of our previous studies, we expected a clear detection of a set of sensorimotor EEG oscillatory sources. We observed that by restricting the core tensor G^* of the model to be non-negative (non-negative Tucker decomposition, NTD), the interpretation of the obtained results was improved. A core tensor with a sparse structure was obtained and for which the detected atoms closely resembled those estimated by the PARAFAC model.

However, the formal testing reveals that in comparison to PARAFAC, a smaller and compact NTD model is usually sufficient to describe the same amount of data variability. *CorConDiag*, a formal measure of the model fit, was also favorable for NTD. We observed that the NTD models, where the number of time scores M was approximately two times higher than the number of frequency signatures O , *CorConDiag* reached the value of 0.8. This indicates, that this com-

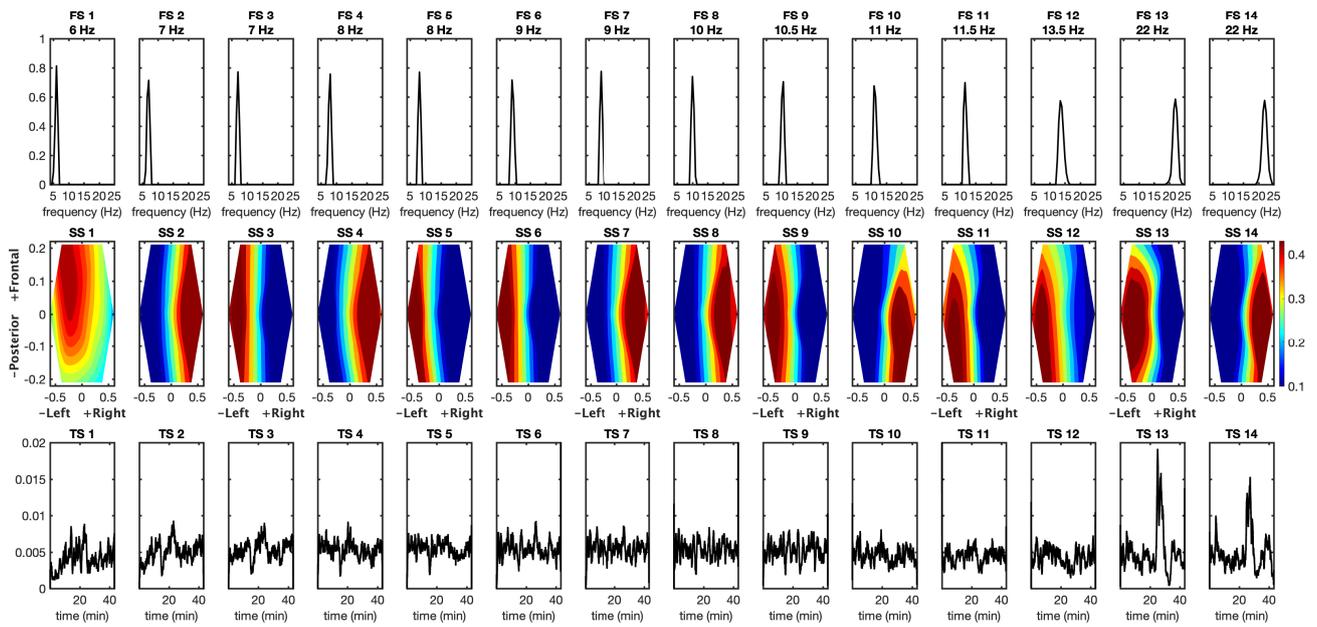


Fig. 7. Subject 1. The PARAFAC decomposition of the EEG data with 14 components. The frequency signatures (FS), time scores (TS) and spatial signatures (SS) of the atoms are depicted in the first, second and third row, respectively.

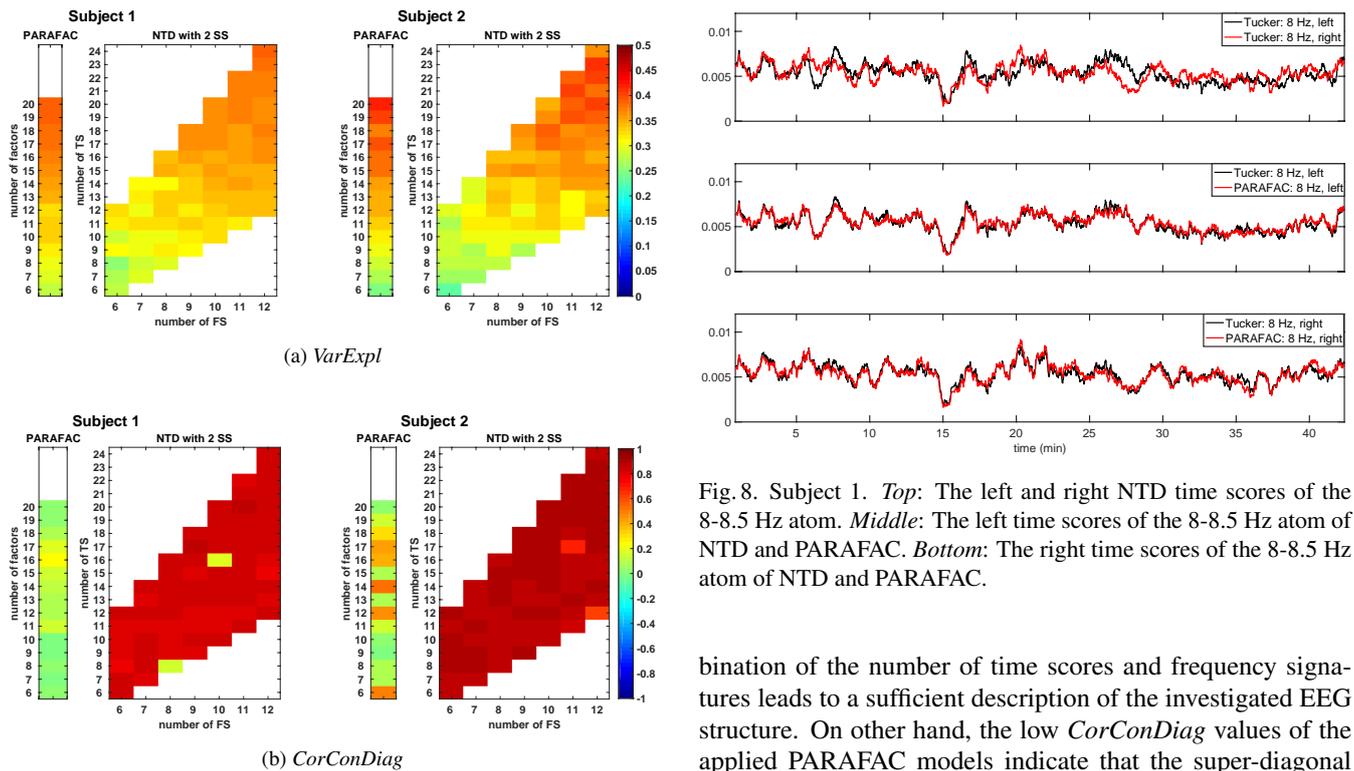
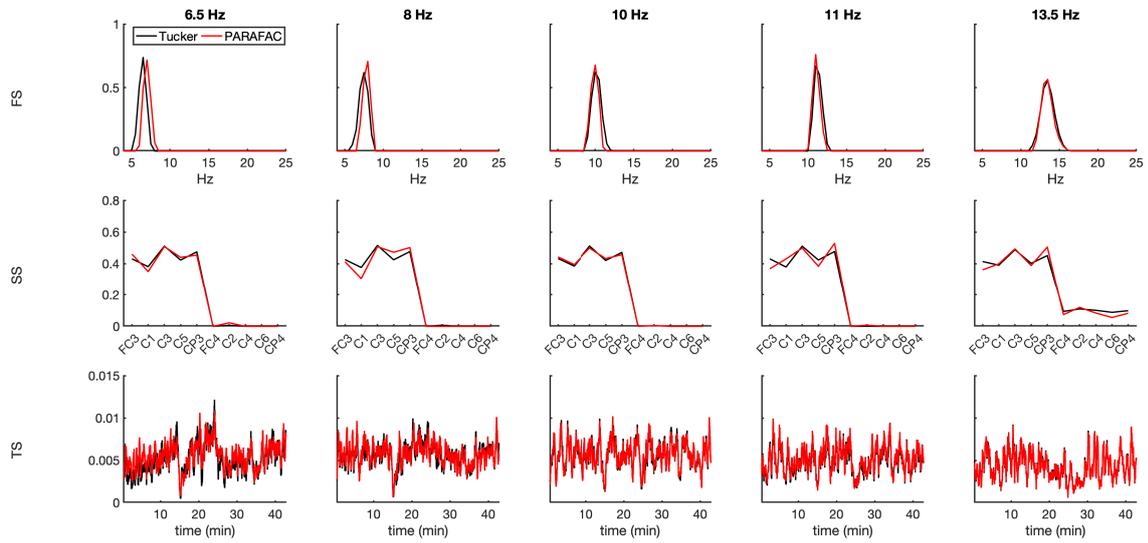


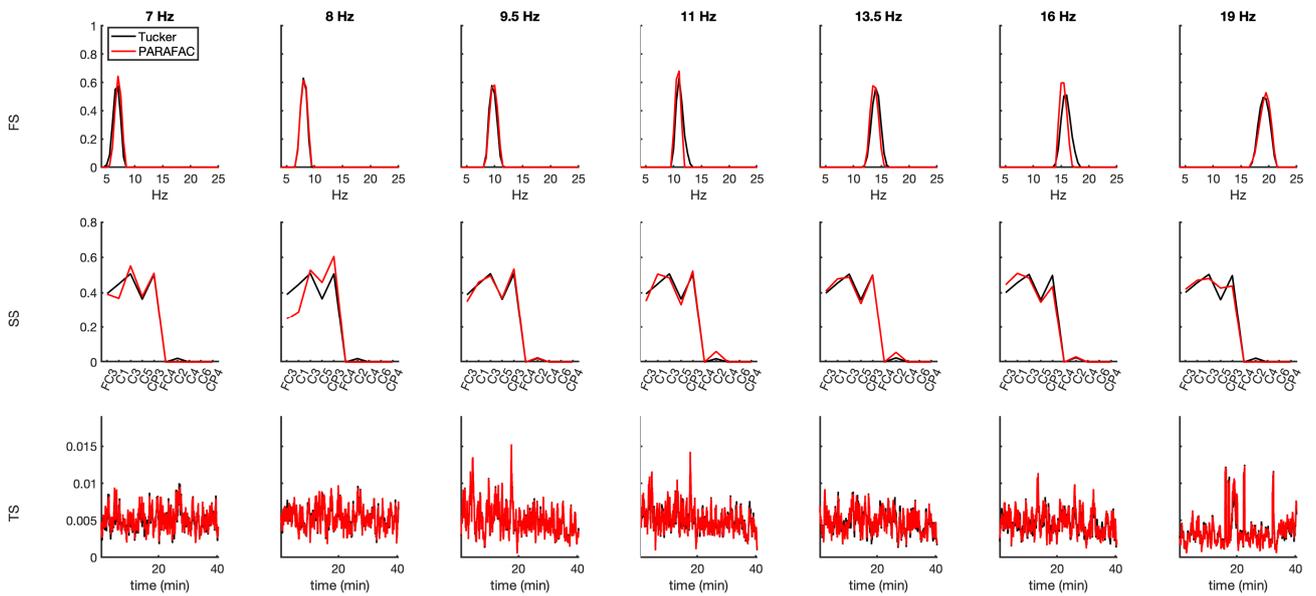
Fig. 9. The variance explained (*VarExpl*, top) and the core consistency diagnostics (*CorConDiag*, bottom) for the PARAFAC and non-negative Tucker decomposition (NTD) models with varying number of factors. The number of factors for the PARAFAC model was considered between six and 20. For NTD, the number of space signatures (SS) was fixed and set to two. The number O of frequency signatures (FS) varied between six and 12 and number of time scores (TS) was chosen from the range $[O, 2 * O]$.

Fig. 8. Subject 1. *Top*: The left and right NTD time scores of the 8-8.5 Hz atom. *Middle*: The left time scores of the 8-8.5 Hz atom of NTD and PARAFAC. *Bottom*: The right time scores of the 8-8.5 Hz atom of NTD and PARAFAC.

combination of the number of time scores and frequency signatures leads to a sufficient description of the investigated EEG structure. On other hand, the low *CorConDiag* values of the applied PARAFAC models indicate that the super-diagonal structure of the core tensor and the same number of factors in each dimension may be too restrictive for our data. Finally, we can also speculate that by focusing on the sensorimotor rhythms activation through motor imagery and by using the scalp EEG recorded over the sensorimotor cortical areas, the observed left and right pairs of identified rhythms reflect expected laterality of sensorimotor cortical activation.

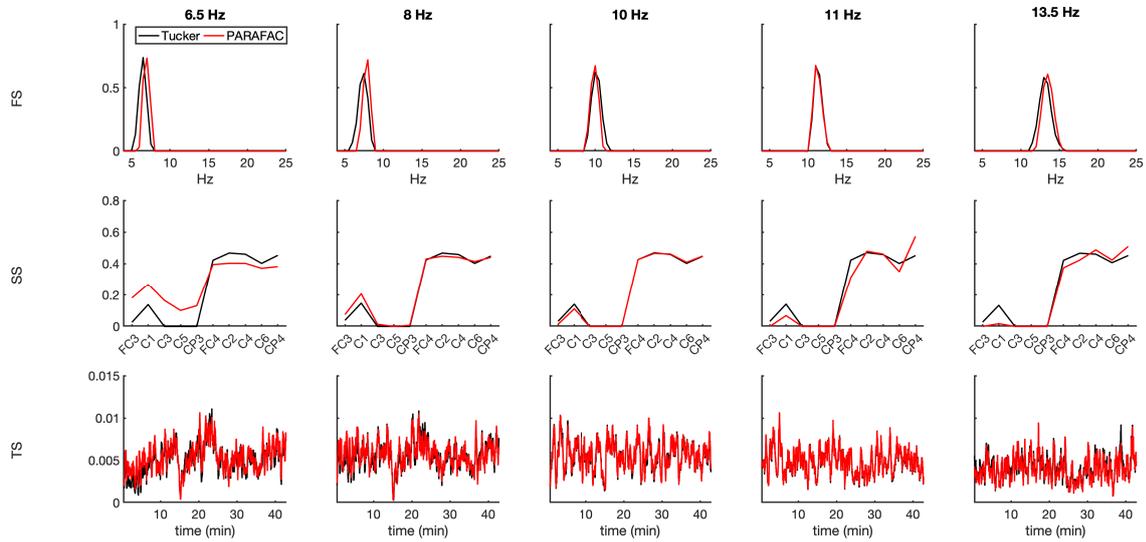


(a) Subject 1

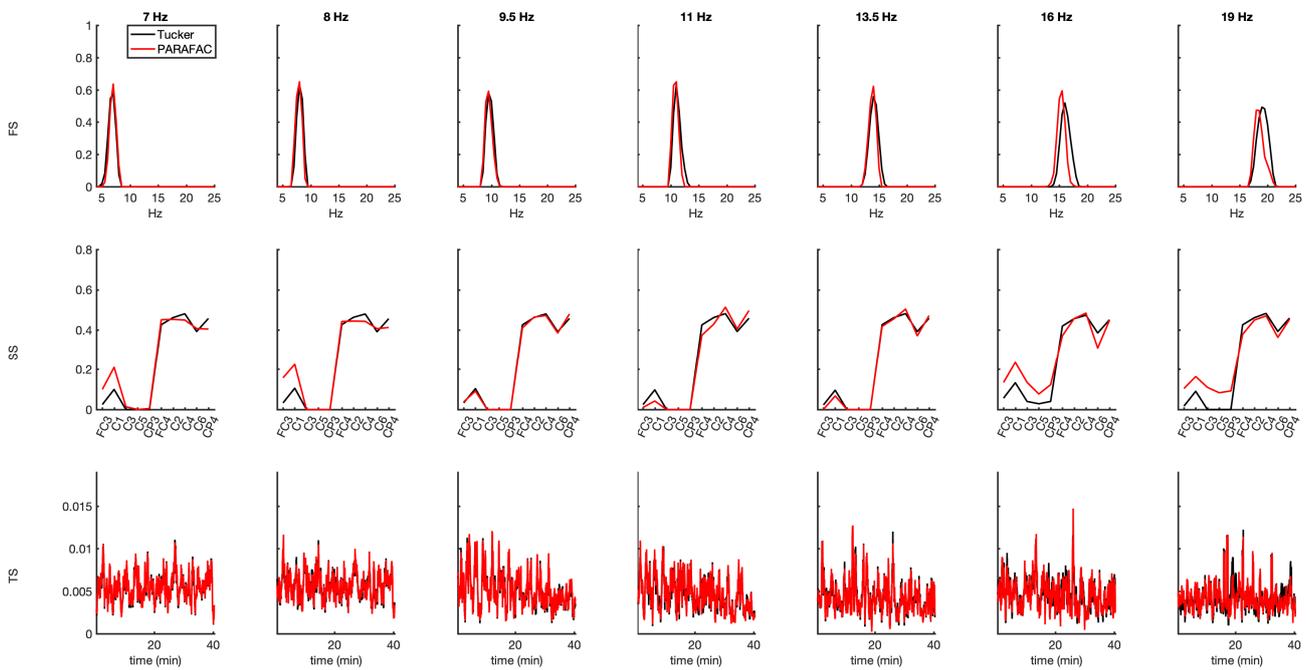


(b) Subject 2

Fig. 10. Subject-specific atoms detected by PARAFAC (red) and NTD (black) located in the left hemisphere.



(a) Subject 1



(b) Subject 2

Fig. 11. Subject-specific atoms detected by PARAFAC (red) and NTD (black) located in the right hemisphere.

An interesting feature of the results was the symmetric and lateralized structure of some of the obtained spatial signatures. Although the placement of EEG electrodes used in our study was restricted to the scalp area covering the sensorimotor cortical region, neither PARAFAC nor NTD has any prior information about laterality of the solution. A compact NTD model with two spatial signatures representing the left- and right-hemisphere neural activity was observed to adequately describe the spatial distribution of atoms in our data.

This was also supported by the cluster analysis applied to a whole set of the NTD solutions obtained by varying the number of spatial and frequency signatures and time scores. This clustering approach allowed us to identify dominant subject-specific atoms, because we expected them to occur in the majority of models and represent dominant clusters in comparison to atoms occurring at random. The lateralized spatial distribution of individual EEG rhythms suggests the presence of at least two equivalent current dipoles with a polarity reversal near the midline, and which may arise from pools of neurons related to motor activity and imagery. It is natural to expect that the temporal activation of these sources may show a different level of synchrony. Therefore, we varied the number of time scores M between O and $2 * O$ with O equal to the number of frequency signatures used in a particular NTD model. This choice turned out to be reasonable and for several rhythms we observed differences in time scores for the two symmetric sources.

In PARAFAC, the laterality of the identified dominant rhythms was observed only in models with a higher number of components. While the obtained frequency and spatial signatures and time scores of these atoms closely resemble the ones obtained by the NTD model, their identification is more difficult due to the larger size of the PARAFAC models. Again, the proposed approach of applying cluster analysis to a set of solutions obtained from the PARAFAC models of different size seems to be a useful strategy for selecting a representative set of atoms.

Although the focus of the paper was on general applicability of the Tucker model to EEG data, the obtained results for two post-stroke patients during the neurorehabilitation training with motor imagery is interesting to mention. When considering the NTD solutions, we found five homogeneous clusters with high cardinality (number of elements per cluster) for Subject 1. The average time scores, space, and frequency signatures of the clusters represent characteristics of three sensorimotor related rhythms, and two rhythms not directly related to sensorimotor activation and deactivation states. For Subject 2, the cluster analysis detected seven stable atoms. In addition to Subject 1, low and high beta rhythms were found, which may relate to sensorimotor activity. Although these two rhythms were sporadically detected in Subject 1, they did not create stable dominant clusters. The cluster analysis of the PARAFAC atoms led to similar results, and a high level of similarity between the time scores, space, and frequency signatures of both PARAFAC and NTD models was observed.

Finally, we can conclude, that the NTD model is able to ex-

plain the same amount of EEG variability as PARAFAC, but with fewer components and direct information about the atom lateralization due to the sparse structure of G^* . In PARAFAC, the lateralized versions of atoms were observed only in models with a large number of components. Thus, we hypothesize that the NTD model can be applied also for other EEG studies where the atomic decomposition of the oscillatory spectrum is the focus.

ACKNOWLEDGEMENT

This research was supported by the Slovak Research and Development Agency (grant APVV-16-0202), the VEGA grant 2/0081/19 and additional support from Pacific Development and Technology, LLC.

REFERENCES

- [1] Miwakeichi, F., Martinez-Montes, E., Valdés-Sosa, P. A., Nishiyama, N., Mizuhara, H., Yamaguchi, Y. (2004). Decomposing EEG data into space–time–frequency components using Parallel Factor Analysis. *NeuroImage* 22(3), 1035–1045.
- [2] Jolliffe, I. (2002). *Principal Component Analysis: Second Edition*. Springer.
- [3] Hyvärinen, A., Oja, E. (2000). Independent component analysis: Algorithms and applications. *Neural networks* 13(4-5), 411–430.
- [4] Harshman, R. A. (1970). Foundations of the PARAFAC procedure: Models and conditions for an “explanatory” multimodal factor analysis. *UCLA Working Papers in Phonetics* 16, 1–84.
- [5] Carroll, J. D., Chang, J.-J. (1970). Analysis of individual differences in multidimensional scaling via an N-way generalization of “Eckart-Young” decomposition. *Psychometrika* 35(3), 283–319.
- [6] Cichocki, A., Zdunek, R., Phan, A. H., Amari, S.-I. (2009). *Nonnegative Matrix and Tensor Factorizations: Applications to Exploratory Multi-Way Data Analysis and Blind Source Separation*. John Wiley & Sons.
- [7] Dobkin, B. H. (2004). Strategies for stroke rehabilitation. *The Lancet Neurology* 3(9), 528–536.
- [8] Carrasco, D. G., Cantalapiedra, J. A. (2016). Efectividad de la imaginación o práctica mental en la recuperación funcional tras el ictus: revisión sistemática [Effectiveness of motor imagery or mental practice in functional recovery after stroke: A systematic review]. *Neurología* 31(1), 43–52.
- [9] Rosipal, R., Porubcová, N., Cimrová, B., Farkaš, I. (2018). Mirror-therapy as a way to start BCI robot-assisted rehabilitation: A single case longitudinal study of a patient with hemiparesis. In: *Seventh International BCI Meeting: Abstract Book*. BCI Society.
- [10] Rosipal, R., Porubcová, N., Cimrová, B., Farkaš, I. (2017). Neurorehabilitation training based on mental imagery of movement (using a robotic splint).

http://aiolos.um.savba.sk/~roman/rrLab/video/RoboticArm_EN.mp4.

- [11] Rosipal, R., Porubcová, N., Barančok, P., Cimrová, B., Farkaš, I., Trejo, L. J. (2019). Effects of mirror-box therapy on modulation of sensorimotor EEG oscillatory rhythms: A single-case longitudinal study. *Journal of Neurophysiology* 121(2), 620–633.
- [12] Tucker, L. R. (1966). Some mathematical notes on three-mode factor analysis. *Psychometrika* 31(3), 279–311.
- [13] Geladi, P. (1989). Analysis of multi-way (multi-mode) data. *Chemometrics and Intelligent Laboratory Systems* 7(1), 11–30.
- [14] Estienne, F., Matthijs, N., Massart, D., Ricoux, P., Leibovici, D. (2001). Multi-way modelling of high-dimensionality electroencephalographic data. *Chemometrics and Intelligent Laboratory Systems* 58(1), 59–72.
- [15] Cong, F., Lin, Q.-H., Kuang, L.-D., Gong, X.-F., Astikainen, P., Ristaniemi, T. (2015). Tensor decomposition of EEG signals: A brief review. *Journal of Neuroscience Methods* 248, 59–69.
- [16] Latchoumane, C.-F. V., Vialatte, F.-B., Solé-Casals, J., Maurice, M., Wimalaratna, S. R., Hudson, N., Jeong, J., Cichocki, A. (2012). Multiway array decomposition analysis of EEGs in Alzheimer's disease. *Journal of Neuroscience Methods* 207(1), 41–50.
- [17] Acar, E., Aykut-Bingol, C., Bingol, H., Bro, R., Yener, B. (2007). Multiway analysis of epilepsy tensors. *Bioinformatics* 23(13), i10–i18.
- [18] Seeck, M., Koessler, L., Bast, T., Leijten, F., Michel, C., Baumgartner, C., He, B., Beniczky, S. (2017). The standardized EEG electrode array of the IFCN. *Clinical Neurophysiology* 128(10), 2070–2077.
- [19] Brain Products, GmbH (2013). *BrainVision Analyser 2*.
- [20] Wen, H., Liu, Z. (2016). Separating fractal and oscillatory components in the power spectrum of neurophysiological signal. *Brain Topography* 29(1), 13–26.
- [21] Buzsáki, G., Draguhn, A. (2004). Neuronal oscillations in cortical networks. *Science* 304(5679), 1926–1929.
- [22] He, B. J., Zempel, J. M., Snyder, A. Z., Raichle, M. E. (2010). The temporal structures and functional significance of scale-free brain activity. *Neuron* 66(3), 353–369.
- [23] Kruskal, J. B. (1989). Rank, decomposition, and uniqueness for 3-way and N-way arrays. In: *Multway Data Analysis*. Elsevier Science Publishers B.V. (North-Holland), pp. 7–18.
- [24] Bro, R. (1997). PARAFAC. Tutorial and applications. *Chemometrics and Intelligent Laboratory Systems* 38(2), 149–171.
- [25] Kiers, H. A. L. (1998). Recent developments in three-mode factor analysis: Constrained three-mode factor analysis and core rotations. In: *Data Science, Classification, and Related Methods*. Springer-Verlag Tokyo, pp. 563–574.
- [26] Kiers, H. A. L., Smilde, A. K. (1998). Constrained three-mode factor analysis as a tool for parameter estimation with second-order instrumental data. *Journal of Chemometrics* 12(2), 125–147.
- [27] Smilde, A. K., Tauler, R., Henshaw, J. M., Burgess, L. W., Kowalski, B. R. (1994). Multicomponent determination of chlorinated hydrocarbons using a reaction-based chemical sensor. 3. medium-rank second-order calibration with restricted Tucker models. *Analytical Chemistry* 66(20), 3345–3351.
- [28] Bro, R., Kiers, H. A. L. (2003). A new efficient method for determining the number of components in PARAFAC models. *Journal of Chemometrics* 17(5), 274–286.
- [29] Kompany-Zareh, M., Akhlaghi, Y., Bro, R. (2012). Tucker core consistency for validation of restricted Tucker3 models. *Analytica Chimica Acta* 723, 18–26.
- [30] Ester, M., Kriegel, H.-P., Sander, J., Xu, X. (1996). A density-based algorithm for discovering clusters in large spatial databases with noise. In: *Second International Conference on Knowledge Discovery & Data Mining*. American Association for Artificial Intelligence, pp. 226–231.
- [31] Andersson, C., Bro, R. (2000). The N-way toolbox for Matlab. *Chemometrics and Intelligent Laboratory Systems* 52, 1–4.
- [32] Sterman, M. (1996). Physiological origins and functional correlates of EEG rhythmic activities: Implications for self-regulation. *Biofeedback and Self-Regulation* 21(1), 3–33.

Received January 13, 2020.

Accepted June 16, 2020.