Through exploiting temporal, spectral, time-frequency representations, and spatial properties of mismatch negativity (MMN) simultaneously, this study extracts a multi-domain feature of MMN mainly using non-negative tensor factorization. In our experiment, the peak amplitude of MMN between children with reading disability and children with attention deficit was not significantly different; whereas the new feature of MMN significantly discriminated the two groups of children. This is because the feature was derived from multi-domain information with significant reduction of the heterogeneous effect of datasets.

Keywords: EEG; event-related potential; mismatch negativity; multi-domain feature; non-negative tensor factorization.
1. Introduction

Mismatch negativity (MMN) identified by Näätänen and his colleagues is one of the small event-related potentials (ERPs) with the peak amplitude 0.5–5 μV, and its peak latency can be within a range of 100–250 ms due to different elicitation paradigms. Like the spontaneous EEG used in the cognitive and clinical research, MMN has also been generally used in the research of cognitive studies, clinical neuroscience, and neuro-pharmacology. Since MMN was found, data processing plays one of the important roles in the research of MMN because MMN is a relatively small ERP and the signal to noise ratio is very low in the corresponding electroencephalography (EEG) data. This study is devoted to extracting a new multi-domain feature from EEG data to reliably represent MMN for robustly analyzing MMN related brain activity.

1.1. Conventional feature to represent MMN

Specifically, most studies use the mean value across a short interval of EEG data centered on the MMN peak as the amplitude of MMN (or simplicity, it is called peak amplitude of MMN hereinafter). MMN has been acknowledged to be an endophenotype to study normal subjects and subjects with disorders, owing to the smaller peak amplitude of MMN generated by the latter group. The peak amplitude at one electrode only reveals the information of MMN in the time-domain. This feature of MMN is probably not robust since it is very sensitive to many conditions including the criteria to label subjects, and the diversity of ages of subjects, and so on. For example, in the study of Huttunen et al. in 2007, where each group consisted of 21 children, the peak amplitudes of MMNs of children with reading disability (RD) were not significantly different from those of children with attention deficit (AD). In the study of Huttunen-Scott et al. in 2008, where each group consisted of 11 children, the peak amplitudes of the MMNs of children with RD were evidently different from those of AD children. Furthermore, those 11 children from each group in the latter study were covered by the former study. The difference in the criteria to categorize the children made the different sizes of groups in those two studies. In the latter study, the criteria were stricter, hence, the number of children in a group decreased, and consequently, the range of the ages of children in the latter study was also much narrower than that in the former one. These are main reasons why magnitudes of MMN peaks yielded different results in the two studies of children with RD and children with AD.

To signify an ERP, there is more information to be used besides measurements of a peak in the time-domain, such as information derived from an ERP’s spectrum, time-frequency representation (TFR) or topography. These kinds of information have already been exploited in brain computer interfaces (BCIs). However, regarding MMN, its peak amplitude in the time-domain or its TFR in the time-frequency domain is often first calculated. Then, the corresponding topography is produced. Hence, the information in different domains is exploited sequentially. To the best knowledge of authors, it is rare to see the study which simultaneously exploits more information of MMN in more domains mentioned above for MMN research.

1.2. Multi-domain feature to represent MMN

Indeed, the peak amplitude representing an ERP measured from the time-domain waveform at one electrode has been mostly analyzed in the study of ERPs and the information only in one domain, i.e. time-domain, is retrieved. Recently, the region of interest (ROI) of the TFR of an ERP has been used for ERP research. In the latter case, the information in two domains, i.e. time and frequency domains, is studied; in order to facilitate the statistical analysis of an ERP related brain activity, the value of the average over the time and the frequency in ROI is calculated to represent the TFR of the ERP. In contrast to the peak amplitude, the TFR can reveal the brain activity in the time and frequency domains simultaneously. Moreover, the topography of an ERP is also important to study the ERP in the spatial domain. Consequently, it is natural to represent an ERP in the time, frequency and spatial domains concurrently. This is that the data can be represented in the multi-way array (tensor), and the extracted feature from the data can also reflect the multi-linear structure of the data. To achieve this goal, we apply non-negative tensor factorization (NTF) in terms of canonical polyadic (CP) model to extract the discriminative multi-domain...
Multi-Domain Feature of MMN

between children with RD and children with AD here.

NTF is a multi-way decomposition method and is the extension of non-negative matrix factorization (NMF). It estimates the dataset’s basis functions based on the linear transform model under multi-way representation of data, assuming elements in the model are all non-negative. EEG data can be modeled by the linear transform of latent variables, and in terms of this model, the principal component analysis (PCA) and independent component analysis (ICA) based blind source separation (BSS) have been applied to reject artifacts and extract the desired ERP components. Furthermore, although EEG data are not non-negative, the TFR of the EEG recordings through the wavelet transform can facilitate NMF and NTF.

1.3. Problems in extracting multi-domain feature to represent MMN

Recently, NMF, NTF, and their derivatives have been used to extract features from spontaneous EEG data and single-trial EEG data in BICs through decomposing data in different domains concurrently and from multispectral images. However, for the feature extraction of MMN, there are still three open and fundamental problems, including what type of the TFR of MMN should be fed to NTF, how many features, i.e. components, should be extracted from the multi-way representation of MMN recordings, and how to determine which multi-domain feature extracted by NTF really corresponds to the desired MMN. This study will try to answer these questions and the proposed method will be performed to extract the multi-domain feature of MMN from EEG data of children with RD and children with AD which have been used in Ref. 11.

2. Method

2.1. Data description

The data was collected at the University of Jyväskylä, Finland. The participants were 114 children aged 8 to 14 years. Three groups were formed: 21 children with RD, 21 children with AD, and 72 normal children without problems in attention or reading. For detailed information on how to categorize the children, please see Ref. 11. This study only discusses the clinical groups. The RD group included 16 boys and 5 girls, and their mean age was 11 years 9 months (age range: 8 years 8 months to 14 years 2 months); and the AD group had 18 boys and 3 girls, and the mean age was 11 years (age range: 8 years 4 months to 13 years 5 months). In the study, an uninterrupted sound under the oddball paradigm was used to elicit MMN. It was invented by Pihko and colleagues in 1995, and such a continuous sound to elicit MMN was used to investigate subject’s ability in detecting temporal changes of the sound. Another advantage of this paradigm is that collecting hundreds of trials of EEG recordings may only take tens of minutes. In order to elicit MMN it is often desired that the recording session should be as short as possible. As shown in Fig. 1, this paradigm consists of the uninterrupted sound, alternating 100 ms sine tones of 600 Hz and 800 Hz (repeated stimuli, see Fig. 1). There was no pause between the alternating tones and their amplitudes did not change. During the experiment 15% of the 600 Hz tones were randomly replaced by shorter ones of 50ms and 30ms duration (called as dev50 and dev30 hereinafter). The deviants consisted of 350 dev50 and 350 dev30. There were at least six repetitions of the alternating 100 ms tones between two deviants.

EEG recordings at nine locations (frontal: F3, Fz, F4; central: C3, Cz, C4; parietal: Pz and mastoids: M1, M2) were collected with Electro-Cap International 20-electrode cap using the standard 10–20 system. Impedances were less than 10 kΩ and in most cases less than 5 kΩ. The potentials were referenced to the tip of nose. After a band-pass filter of 0.1–30 Hz was applied, EEG was down sampled with the rate of 200 Hz. Recording started 300 ms before the

Fig. 1. Stimulus sequence.
onset of a deviant stimulus and lasted 350 ms after the onset of a deviant.

It should be noted that only the data under dev30 was chosen for analysis in this study. This is because under dev30, the MMN peak amplitude in Ref. 12 discriminated 11 children with RD and 11 children with AD, and it was not able to reveal any significant difference between the 21 children with RD and 21 children with AD in Ref. 11. Furthermore, MMN data may usually possess larger peak amplitudes, i.e., higher signal-to-noise ratio, under the larger magnitude of the deviance to the repeated stimuli. In this study, the repeated stimuli lasted 100 ms, thus, the dev30 lasting 30 ms was with the greater magnitude of the deviance.

2.2. Feature extraction

2.2.1. NTF algorithm

The NTF model can be formulated as follows. For a given Nth-order tensor \( Y \in \mathbb{R}^{I_1 \times I_2 \times \cdots \times I_N} \), performing a factorization into a set of \( N \) unknown non-negative matrices whose elements are non-negative, \( U = [u_1^{(1)}, u_2^{(1)}, \ldots, u_J^{(1)}] \in \mathbb{R}^{I_1 \times J} \) can be described as

\[
\sum_{j=1}^{J} u_j^{(1)} \circ u_2^{(2)} \circ \cdots \circ u_J^{(N)} = Y, \tag{1}
\]

where \( J \) is the number of extracted components, the symbol \( \circ \) denotes the outer product of vectors, \( [u_n^{(i)}]_{ij} = 1 \) for \( n = 1, 2, \ldots, N - 1 \) and \( j = 1, 2, \ldots, J \). Equation (1) does really conform to the CP model.

The target of NTF is to obtain the suitable \( U^{(n)} \), and one \( J \) here is defined to correspond to one NTF model. In the form of tensor products, the NTF model can also be written as

\[
Y \approx \mathbf{1}_{I_1} \times U^{(1)} \times U^{(2)} \cdots \times U^{(N)} = \sum_{j=1}^{J} u_j^{(1)} \circ u_2^{(2)} \circ \cdots \circ u_J^{(N)}, \tag{2}
\]

where \( \mathbf{1} \) is an approximation of the tensor \( Y \), and \( U \) is an identity tensor. Each factor \( U^{(n)} \) explains the data tensor along a corresponding mode. Hence, one factor can be considered as features of the data onto the subspace spanned by the others. The outer product of vectors and the product of a tensor and matrices/vectors are defined in the Appendices A, B and C. Most algorithms for NTF are to minimize a squared Euclidean distance as the following cost function

\[
D(Y, \hat{Y}) = \frac{1}{2} \| Y - \hat{Y} \|_F^2 = \frac{1}{2} \| Y - \mathbf{1}_{I_1} \times U^{(1)} \times U^{(2)} \cdots \times U^{(N)} \|_F^2. \tag{3}
\]

In this study, we applied the hierarchical alternating least squares (HALS) algorithm whose simplified version for NMF has been proved to be superior to the multiplicative algorithms. The HALS is related to the column-wise version of the ALS algorithm for 3-D data. The HALS algorithm used in this study sequentially updates components \( u_j^{(n)} \) by a simple update rule

\[
U_j^{(n)} = \frac{Y \circ (U_{-j}^{(n)})_j - \sum_{j=1}^{J} u_j^{(n)} (U_j^{(n)})_j} {\| \sum_{j=1}^{J} u_j^{(n)} (U_j^{(n)})_j \|_2 - \sum_{j=1}^{J} \| u_j^{(n)} \|_2}. \tag{4}
\]

2.2.2. Multi-domain feature

In detail, to study MMN with NTF, we formulated a fourth-order tensor \( Y \) (including modes of the frequency by time by channel by subject-experimental condition. The number of frequency bins \( I_f \), time-frames/samples \( I_s \), channels \( I_c \), and subjects \( I_s \)) compose the dimensions of the tensor \( Y \). Hence, each tensor actually includes the TFR of the MMN data of all participants at all channels under one or more experimental conditions. For example, if there are two groups of subjects with \( K \) participants in each, \( I_s \) is equal to \( 2K \) under one experimental condition. Furthermore, if there are \( L \) experimental conditions, \( I_c \) is equal to \( 2K \times L \). This study only discusses the case with one experimental condition in the tensor \( Y \). Decomposition of \( Y \) results in four matrices. The first three are the basis matrices \( U^{(1)} (I_f \times J_f) \), \( U^{(2)} (I_f \times J_c) \), and \( U^{(3)} (I_f \times J_s) \) which reflect the spectral, temporal, and spatial factors, respectively. They are common to all subjects. The last factor matrix \( I_f \times J_s \) consists of different compressed multi-domain features of different brain responses elicited in the experiment for \( I_s \) subjects.
Multi-Domain Feature

(i = 1, 2, ..., I) is characterized by the ith row of $U^{(i)}$ in this study. In summary, factorizing a fourth-order tensor of MMN data with the dimensions of frequency by time by channel by subject can be expressed as the following,

$$
\hat{Y} = \sum_{j=1}^{J} U^{(f)}_{j} \times U^{(t)}_{j} \times U^{(c)}_{j} \times \mathbf{f}_{j}
$$

(5)

where the last factor matrix consists of the extracted multi-domain features of brain responses in the MMN experiment onto the subspaces spanned by the spectral (i.e. $U^{(f)}$), temporal (i.e. $U^{(t)}$) and spatial (i.e. $U^{(c)}$) factors. In the vector form, Eq. (5) can read

$$
\hat{Y} = \sum_{j=1}^{J} u^{(f)}_{j} \circ u^{(t)}_{j} \circ u^{(c)}_{j} \circ \mathbf{f}_{j},
$$

(6)

Hence, one new feature reveals the multi-domain information of the corresponding brain response. For example, the jth multi-domain feature $f_{j}$ corresponds to the jth temporal ($u^{(t)}_{j}$), spatial ($u^{(c)}_{j}$) and spectral ($u^{(f)}_{j}$) components. These components reveal the properties of the jth multi-domain feature in the time, spatial and frequency domains.

After the $J$ multi-domain features of EEG data are extracted, the next step is to select the multi-domain feature of MMN activity for cognitive research

2.2.3. Type of TFR of MMN

For the time-frequency analysis of ERPs two types of TFR are the event-related spectral perturbation (ERSP) representing TFR averaged across single trials, and the TFR of the averaged EEG. Indeed, the former one contains the induced and evoked brain activities and the latter mainly reflects the evoked one. As mentioned earlier, MMN is one of the small ERPs. Hence, hundreds of single trials are usually collected during the MMN experiment for averaging to produce the evident MMN component. After averaging, the induced brain activity is severely reduced; therefore, most of the well-known knowledge of MMN results from the evoked brain activity. For example, the previously mentioned knowledge of the peak amplitude of MMN between children with RD and children with AD was actually derived from the average over single trials, i.e. from the evoked brain activity.

As a result, this study performed the wavelet transform on the average of EEG data to obtain the TFR of MMN. Then, NTF was used to extract the multi-domain feature of MMN from this type of TFR.

2.2.4. Number of features to be extracted

In the application of NTF to extract features of brain responses, the number of features is the number of extracted components in each factor of tensor factorization in terms of CP model. Determining this number is very important to NTF because different numbers in different quantitative levels may probably correspond to very different decompositions. Actually, the choice of the number of components to extract is an inherent problem of model order selection which is usually for the underdetermined linear transform model, i.e. the number of electrodes to collect data is greater than the number of sources in the model under the context of EEG and fMRI. However, in our study, we used only nine electrodes to collect EEG data, hence, the model of our data is highly probably underdetermined, i.e. the number of sensors to collect data is smaller than the number of sources in the model. This really does bring the difficulty in estimating the number of components to be extracted by NTF using the general model order selection methods. In such a condition, regarding the tensor factorization under the CP model, another two methods have been developed and they are known as DIFFIT and CORCONDIA. Indeed, DIFFIT and CORCONDIA respectively measure the change of the fit (explained variance of the raw data by the proposed model) and the core tensor of the decomposition among a number of models. For simplicity, DIFFIT was used in this study. The fit of a NTF model is defined as below

$$
\text{fit}(m) = 1 - \frac{\|Y - \hat{Y}_m\|_F}{\|Y\|_F},
$$

where $\hat{Y}_m$ is the approximation for the raw data tensor $Y$ by a rank-$m$ NTF model ($m$ plays the same role as $J$ in Eq. (1)). $\| \cdot \|_F$ is the Frobenius norm, $m = 1, \ldots, M$, and fit($m$) is monotonely rising.

Then, the difference fit of two adjacent fits is given by

$$
diff(m) = \text{fit}(m) - \text{fit}(m - 1),
$$
where \( m = 2, \ldots, M \). Next, the ratio of the adjacent difference fits is defined as

\[
diffit(m) = \frac{\text{diff}(m)}{\text{diff}(m+1)},
\]

where \( m = 2, \ldots, M - 1 \). The model with the largest diffit value is regarded as the appropriate NTF model for the raw data tensor \( \sum \).

2.2.5. Determination of feature of MMN based on its prior knowledge

In this study, we hope to obtain the multi-domain feature of MMN which can better discriminate the two groups of children than the peak amplitude of MMN does. Hence, the feature selection includes two steps. The first step is based on the statistical tests and the second is based on prior knowledge of MMN in different domains. First, for each extracted multi-domain feature in this study, statistical tests were used to investigate the difference in the multi-domain feature between two groups, and only the feature with the \( p \)-value smaller than 0.05 was chosen for the further analysis. Subsequently, the temporal, spectral and spatial components associated with each selected feature were examined according to the properties of MMN in different domains to finally obtain the multi-domain feature of MMN.

Indeed, the averaged EEG recordings are still mixtures of sources of electrical brain activities, and NTF may decompose the mixtures into individual sources of respective electrical brain activities presented in different domains. Hence, the components in the three common factors and the multi-domain features extracted by NTF correspond to different sources of brain activities. Then, it is necessary to determine which multi-domain feature corresponds to the MMN in this study. If the feature of interest was not properly chosen, the following analysis would fail. This process is analogous to the application of ICA on EEG. Through ICA, a number of components can be extracted, and then, one or more interesting components are usually chosen for the further analysis based on the prior knowledge of EEG.

Regarding the spontaneous EEG, its spectrum and topography are often used to judge the features of the interesting oscillations. For the spontaneous EEG and ERPs, their spectrums may be distinct. However, for different ERPs, their spectrums are not so discriminative to verify an ERP of interest. ERPs are time-locked, and the temporal characteristics are usually discriminative. Furthermore, the topography of an ERP is also an important property to ascertain the ERP. Thus, in order to seek the proper multi-domain feature of MMN, all the temporal, spectral and spatial information of MMN can be used if they are obtainable prior knowledge.

As illustrated in Fig. 2(a), the majority of brain activities of MMN elicited by the oddball paradigm in our study would appear between 50 ms and 200 ms after the offset of the deviant. Particularly, the latency of MMN in our dataset fell in the range from 100 ms to 160 ms. The spectrum of MMN in our dataset was thoroughly analyzed before, and its optimal range was from 2 to 8.5 Hz and the peak of the spectrum was below 5 Hz. Children with RD may have different topography of MMN from the view of lateralization of MMN peak in contrast to children with AD. For children with RD, the peak amplitude of MMN was more likely pronounced in the left hemisphere, and regarding the children with AD, the peak amplitude was probably greater on the right part in our dataset. Furthermore, as visually inspected by Huttunen et al., the peak amplitude of MMN between the RD group and the AD group was more distinct in the left hemisphere than in the right. As the decomposition by NTF is also implemented in the spatial domain, the component of MMN in the spatial factor extracted by NTF is able to reveal the location where the difference of MMN between the two groups may be most likely to appear along the scalp. Consequently, such a spatial component of MMN should present the difference topography of MMN between the two groups in this study. The difference topography means the subtraction of an ERP’s topography of one group from that of another group, thus, it may be different in contrast to the topography of any group.

In summary, a multi-domain feature extracted by NTF is regarded to be associated with MMN related brain activity in the used datasets, the corresponding temporal component should peak within the range between 100 ms and 160 ms, and the energy of spectral component should maximize below 5 Hz and the frequency band ranges from 2 Hz to 8.5 Hz, and the spatial component should reveal the difference topography between two groups of children.
2.3. Data processing and analysis

2.3.1. Overview

The data were processed and analyzed in MATLAB (v 2010b, The Mathworks, Inc., Natick, MA). In order to remove artifacts, two types of exclusion criteria were applied. First, trials with the amplitude exceeding ±100 μV were rejected. Second, trials with recordings of zero variance were deleted. After the artifacts rejection, the mean number of trials per child was about 331 with the standard deviation of 21.6. In order to obtain a stable MMN, the kept trials were averaged for each subject, and then a difference wave was obtained through subtracting the responses of a standard sweep from those of a deviant sweep, followed by removing the baseline formed by the average of its first 50 ms recordings. This is because a difference wave has been mostly used to observe MMN in its research.

Subsequently, complex Morlet wavelet transform was performed on the difference wave to obtain the TFR of MMN. For the Morlet, the half wavelet length was set to be six for the optimal resolutions of the frequency and the time length was set to be six for the optimal resolutions of the time-frequency domain, the values of TFR of MMN. For the Morlet, the half wavelet length was set to be six for the optimal resolutions of the frequency and the time length was set to be six for the optimal resolutions of the time-frequency domain. To obtain the feature of MMN in the time-frequency domain, the values of TFR were averaged across the time and the frequency for producing one value to represent the feature of MMN in the time-frequency domain. It was calculated channel by channel and subject by subject.

After the features were ready, the statistical tests of one-way ANOVA (analysis of variance) were performed to investigate the difference in MMN between the two groups with 0.05 as the level of significance. Indeed, it is well known that the ANOVA test makes the following assumptions about the analyzed data: all sample populations are normally distributed, all sample populations have equal variance, and all observations are mutually independent.

When the assumptions are not met (for example, there are some outliers in the analyzed data), the assumption of the normal distribution is violated. One reason is that the number of subjects is relatively small and the other is that while features of a group are positive, features of another group tend to be sparse under the framework of NTF. In this case, the Kruskal–Wallis test is better for revealing the difference between two groups. The difference between ANOVA test and Kruskal–Wallis test is that when the data conform the above assumptions ANOVA test is more sensitive to group difference and that when the data include outliers the nonparametric procedure, e.g. Kruskal–Wallis test, is often not severely affected in revealing group difference. Hence, in this study, both ANOVA test and Kruskal–Wallis test were used for statistical analysis. The minimal p-value of the two tests was reported to show the level of significance of group difference in the multi-domain feature of MMN. This is because the multi-domain feature of MMN extracted by NTF may have a few outliers due to constraints of non-negativity. Indeed, the best way for the statistical test is to validate the distribution of the tested data and to use the test whose assumption is met by the data. It is out of the range of this study and is not discussed hereafter.

In summary, the data processing to extract the multi-domain feature of MMN includes five steps:

1. Implementing the conventional data processing approach to produce ERPs.
2. Transferring the time-domain ERPs to the time-frequency domain for producing the tensor including TFRs of ERPs.
3. Implementing the conventional data processing approach to produce ERPs.
4. Transferring the time-domain ERPs to the time-frequency domain for producing the tensor including TFRs of ERPs.
5. Implementing the conventional data processing approach to produce ERPs.
(3) defining the number of extracted components for NTF,
(4) performing NTF on the tensor,
(5) selecting the desired multi-domain feature of an ERP.

2.3.2. Stability analysis of multi-domain feature of MMN
The NTF algorithm used in this study is adaptive. The adaptive algorithm is first initialized; and then, after a number of iterations the predefined stop criterion is met; finally, the results of the last iteration are the outputs of the algorithm.

In this study, the random initialization for the iteration of the algorithm was used in our study (such a procedure is often used for an adaptive algorithm). The maximum number of iterations was 1000 and the maximum fit of the model is 0.99. If the fit difference between two adjacent iterations was less than 1.0e-8, the iteration stopped. Regarding such an adaptive algorithm and the implemented initialization, the stability of results should be analyzed. In this study, we analyzed the NTF models with J from 1 to 50. For each J, the data tensor was decomposed 100 times using uniformly distributed random initial conditions. The extracted features and factors over 100 runs for each J were then used for evaluation. Indeed, the interest of this study is just of the stability results of NTF according to DIFFIT on the averaged fits.

2.3.2.2. Examination of decomposition of 100 runs according to selected model of NTF
Regarding the selected model of NTF, J components were extracted in each factor. Then, J rank-1 and fourth-order tensors can be formulated. Next, correlation coefficient (CC) between each of such rank-1 tensors and the template tensor defined above is given as

\[ \rho(j, J, r) = \frac{\left[ u^{(f)}_{j,J,r} \right]^T u^{(t)}_{j,J,r} \cdot u^{(i)}_{j,J,r} \cdot u^{(c)}_{j,J,r} \cdot \left[ t^{(f)}_{j,J,r} \right]^T t^{(t)}_{j,J,r}}{\left[ u^{(f)}_{j,J,r} \right]^T u^{(f)}_{j,J,r} \cdot u^{(i)}_{j,J,r} \cdot u^{(c)}_{j,J,r} \cdot \left[ t^{(f)}_{j,J,r} \right]^T t^{(f)}_{j,J,r}} \],
\]

where \( j = 1, 2, \ldots, J, J = 1, 2, \ldots, 50, r = 1, 2, \ldots, 100 \) and each component has been normalized to its standard deviation and its mean has been removed. Then, given a J and a r, J CCs were produced. Subsequently, the maximal CC among the J CCs was chosen as

\[ q(J, r) = \rho(k, J, r) = \max(\rho(1, J, r), \rho(2, J, r), \ldots , \rho(J, J, r)) \],
\]

where \( k \in [1, J] \). Finally, the difference between two groups in the Jth multi-domain feature of model-J of run-r was analyzed.

It should be noted that if the maximal CC mentioned above was close to ‘1’, it would mean that the desired multi-domain feature of MMN as well as its corresponding three factors was also extracted in the run of NTF possessing the CC. If the maximal CCs of many runs of NTF were close to ‘1’, it would
were reported by the previous study, as the peak measurements of MMN in our dataset were larger than that of children with AD in each run.

3. Results

As the peak measurements of MMN in our dataset were reported by the previous study,11 they are not presented here. Nevertheless, for the completeness of this study, the grand averaged difference wave of MMN is shown for the demonstration. Then, the feature of MMN in the time-frequency domain and the multi-domain feature extracted by NTF are reported here.

3.1. Conventional analysis

Figures 2(a)–2(c) display the grand averaged difference wave for each group of children. Topography of MMN peak amplitude and topography of MMN feature in the time-frequency domain, respectively. In Fig. 2(a), the negative peak between 100 ms and 160 ms describes the MMN activity, and the positive peak between 200 ms and 300 ms presents the P3a. Although the visual inspection of this figure reveals that the peak amplitude of MMN of children with RD was larger than that of children with AD in the left hemisphere in the grand averaged difference wave, Huttunen et al.11 failed to note that the difference was statistically significant. Thus, this is the issue we can elaborate on more deeply. In the time-frequency domain, the difference in the attended feature of MMN between the two groups of children was not significant at any of the four locations (F3 (F(1, 40) = 2.39, p = 0.13), F4 (F(1, 40) = 3.07, p = 0.09), C3 (F(1, 40) = 2.35, p = 0.13) and C4 (F(1, 40) = 2.79, p = 0.10)). Hence, based on MMN features in the time and time-frequency domains, we cannot tell that the RD group had a statistically larger magnitude in the feature of interest of MMN in the time-frequency domain than that of the AD group in the left hemisphere.

Through visual inspection of the topographies of MMN in the time and the time-frequency domain, we observe that the RD group had greater power of the MMN brain activity, particularly at C3. Hence, the tendency was that MMN was bigger in the left hemisphere in the RD group compared to the AD group, as concluded by Huttunen et al.11 in the peak amplitude of MMN.

3.2. Multi-domain feature of MMN

In this study, DIFFIT suggested that 36 components were appropriate for NTF (this will be illustrated in the next subsection). According to Secs. 2.2.5 and 2.3.2.1, after the feature selection, the feature # 20 was chosen as the desired multi-domain feature of MMN among all 36 features extracted by NTF. Figure 3 shows the desired multi-domain feature as well as the corresponding temporal, spectral, and the spatial components. The feature # 20 indicates that the RD group had statistically significantly larger multi-domain feature of MMN than the AD group had (F(1, 41) = 4.96, p = 0.032). The corresponding temporal component and the spectral component associated with the feature # 20 match the property of MMN in the time and frequency domains, i.e. the peak latency is around 150 ms and the spectrum peaks around 5 Hz. Indeed, it is not surely known that the difference spatial map of the MMN between RD children and AD children is like the spatial component of the feature # 20. This is because either the feature in the time-domain or that in the time-frequency domain did not succeed in significantly discriminate the two groups of children. The spatial component of the feature # 20 indicates that the
difference in MMN between two groups of children may more evidently appear in the central and left hemisphere.

Figure 3 also demonstrates other two extracted features and their temporal, spectral, and spatial components for the corresponding factors (36 features were extracted and we do not think it is necessary to show all of them). For example, regarding the feature # 21, the spectral component is very similar to that of the feature # 20, and meets the expectation of properties of MMN in frequency domain. However, since the temporal component of feature # 21 in contrast to the waveform in Fig. 2a is not sufficiently separated out of mixtures from the view of blind source separation and the spatial component of feature # 21 is different from that of feature # 21.
Figure 3. Multi-domain feature as well as the corresponding temporal, spectral, and the spatial components.

#20, the difference between two groups of children in the feature #21 is not as significant as that in feature #20. Indeed, through visual inspection, we have found that none of the other 35 features in the NTF model concurrently possess the desired temporal component and spectral component as the feature #20 has. It should be noted that the multi-domain feature #20 and the corresponding temporal, spectral, and spatial components are regarded as the templates to formulate the template tensor as Eq. (7) for the following stability analysis in this study.

3.3. Stability analysis of multi-domain feature of MMN

3.3.1. Fit analysis

Figure 4 shows the fits of different NTF models. The solid curve represents the maximal fit among 100 runs for each model, and the dash dot curve denotes the averaged fit over 100 runs for each model. The averaged fit and the maximal fit are surprisingly close to each other. This indicates that the fit varies little among 100 runs of decomposition regarding one NTF model.

Furthermore, DIFFIT was performed on the curve of the averaged fit to estimate the appropriate model of NTF. For DIFFIT, the lower and upper bounds of the number of components should be defined. The upper bound was fixed to be 50 in this study. When the lower bound of number of components ranged from 1 to 10, ‘3’ and ‘11’ were estimated as the proper number of components for NTF. However, after feature selection, no desired feature of
MMN was found. When the lower band of number of components was set to 11, ‘36’ was suggested by DIFFIT as the appropriate number of components for NTF regarding our dataset. Hereinafter, we denote the appropriate model for NTF as model-36.

3.3.2. Decomposition of desired model of NTF

According to the procedure stated in Sec. 2.3.2.2, results of 100 runs decomposition under model-36 are analyzed with the template tensor referenced as the following. Figure 5 shows CC between a rank-1 tensor selected by Eq. (9) for each run of NTF under model-36 and the template tensor. It is evident that the selected rank-1 tensor with the desired multi-domain feature of MMN is very similar to the template tensor in most of runs of NTF. Furthermore, the difference between two groups of children in the desired multi-domain feature of MMN is significant in 82 of 100 runs of NTF. These mean that the extraction of the desired multi-domain feature of MMN and its corresponding temporal, spectral, and spatial components was stable under the model-36 in this study.

3.3.3. Decomposition of multiple models of NTF

In terms of Sec. 2.3.2.3, results of 100 runs decomposition under each model are analyzed with the template tensor referenced as below. Figure 6(a) shows the maximal, averaged and minimal of the selected correlation coefficients (according to Eq. (9)) among 100 runs under each model. We have found that when the number of components is larger than 30, the maximal CC and the averaged CC are very close to each other. Figure 6(b) describes the
maximal, averaged and minimal \( p \)-value of the \( p \)-values for the selected features among 100 runs under each model (\( p \)-value is to reveal the degree of the difference between two groups of children in the selected multi-domain feature according to Eq. (9) in each run of each model). Surprisingly, when the number of components is larger than 30, the minimal \( p \)-value and the averaged \( p \)-value are very close to each other. According to the threshold (\( p = 0.05 \)) in Fig. 6(b), we can tell that when the number of components was larger than 30, the difference between two groups of children in the selected multi-domain feature of MMN was significant in most runs of NTF.

Furthermore, in some models especially when the number of component is smaller than 30, Fig. 6(b) shows that the difference in a multi-domain feature can be significant between two groups as well. While, Fig. 6(a) demonstrates that the corresponding rank-1 tensor is not closely correlated with the template tensor. This means the multi-domain feature is not associated with MMN, but other brain activities. This is out of the range of this study and will be examined in the future.

4. Discussion

To study MMN of children, one challenge is how to correctly, robustly and reliably represent MMN due to the complexity of children’s EEG. Indeed, the generally used peak amplitude is derived from the one-way analysis and is based on the information of the ERP waveform observed in the time-domain. TFR is from the two-way analysis and is in terms of the information of an ERP in the time-frequency domain. In this study, the new multi-domain feature of an ERP is derived from the multi-way analysis using the information of an ERP in the time, frequency and spatial domains simultaneously. In other words, the new multi-domain feature spans the information of MMN in the multi-domain, revealing MMN’s properties in those domains concurrently. Moreover, as shown in Fig. 3, the comparison between groups with the new feature is in terms of the energy of MMN activity. This keeps consistent with the peak amplitude in the time-domain and the feature in the time-frequency domain from the view of comparison of MMN related brain activity between two groups.

Between children with RD and children with AD, as reported by Huttunen et al., the difference of the time-domain peak amplitude of MMN was not considerable; the difference of the time-frequency feature of MMN formulated in this study was not evident either; while the difference of the multi-domain feature of MMN extracted by NTF was significant. Hence, the logic here is that when more information of brain responses in more domains is simultaneously exploited to formulate the feature for representing the MMN activity, the difference between children with RD and children with AD may be more reliably identified. It should be noted that analyzing the peak amplitude in the time-domain or the feature in the time-frequency domain is not limited to data at one electrode; however it may be probably subjective to choose data at several electrodes among all sensors to examine the difference of MMN between the two groups of children, which can bring the uncertainty to the conclusion. The proposed method objectively exploits data collected at all electrodes in the experiment for analysis, implicitly enhancing the reliability of the multi-domain feature to represent the MMN related brain activity.

The EEG datasets may be heterogeneous in the time-domain due to noise and interference. However, the heterogeneous effect may be reduced in the multi-domain since the pattern of the desired brain activity is not as easily contaminated as its peak amplitude. As a result, the multi-domain feature can outperform the feature in one or two domains in discriminating different groups of children.

The applications of tensor factorization to extract the multi-domain feature of EEG can be divided into three categories: (1) for spontaneous EEG, the single-trial EEG in an ERP experiment, and the third one is mainly for the cognitive research. There are quite many publications for the first two types of applications nowadays. However, we do not find many regarding the third one. Actually, the properties of spontaneous EEG, the single-trial EEG in an ERP experiment, and the ERPs, can be different. Hence, it is worth examining the tensor factorization for ERPs to extract the multi-domain features of interest. For this purpose, determining the number of components and the selection of the desired feature are open for
discussion to extract the multi-domain feature of an ERP before our study. Indeed, the tensor factorization has already been used to decompose the third-order tensor of ERPs consisting of the TFR of ERPs of multiple channels.\textsuperscript{27,28} Then, the temporal, spectral, and spatial components extracted by tensor factorization can be examined among different conditions or different participants. However, such study is different from ours because of two points. One is that we decompose the fourth-order tensor of ERPs to extract the multi-domain feature of an ERP which can be used for group-level statistical analysis like a peak amplitude, the other is that analysis of the temporal, spectral, and spatial components is to seek and validate the multi-domain feature of MMN in our application.

When the advanced signal processing methods are applied to study EEG, one critical concern is that how to validate the new findings not to be technically produced.\textsuperscript{17} The straightest manner is to interpret the new findings through the theoretical and earlier empirical knowledge of brain responses. This means that the prior knowledge of the brain responses is a must. Otherwise, it becomes difficult to determine the reliability and the rationale of the new feature. Since MMN was identified,\textsuperscript{1} there have been thousands of publications devoted to reporting the study of MMN, providing enough prior knowledge to interpret the new finding in this study.

In this study, the initialization of NTF algorithm was based on the randomly generated factors. Indeed, the singular value decomposition and the fibers from the decomposed tensor can be used for the initialization as well.\textsuperscript{15} We will report the difference in determining the appropriate number of components for NTF using DIFFIT and the difference in the multi-domain feature of MMN among the three initialization methods later.

In the application of NTF to study ERPs, one of the key parameters for decomposition is the number of components/features to be extracted, which is a challenging problem. Indeed, this parameter plays the same role as the frequency band to a digital filter from a systematic view of signal processing. For example, when the digital filter is used to improve the signal to noise ratio for MMN, different studies may use different band-pass filters.\textsuperscript{52} Moreover, since it is hard to precisely know the real frequency range and the true spectral structure of an ERP, the frequency band for a band-pass digital filter is actually empirical. Usually, as long as the output of the filter meets the theoretical expectation of one ERP, the setup of the digital filter should be acceptable for that ERP study. In this study, we executed an empirical study using a range of numbers of components for NTF models and DIFFIT\textsuperscript{17,48} suggested that the model of 36 components was appropriate. We chose this method because this method is easy to implement and the linear transform model of our data collected by the low-density array is highly probably underdetermined and we cannot use the model order selection methods\textsuperscript{58} which are usually for the over-determined model to estimate such a parameter (a single integer) with the available dataset. Recently, in our independent work, we have gained some EEG data collected by high-density array including 128 sensors and we have used a model order selection method called GAP\textsuperscript{59} to estimate the number of sources in the EEG data.\textsuperscript{48} We have found that the number of sources was about 120 in the conventionally averaged MMN trace and was about 14 in the MMN trace which was cleaned by an optimal filter.\textsuperscript{56-59} Indeed, there are other available and robust methods and procedures to estimate the number of components extracted by NTF, including CORCONDIA,\textsuperscript{49} and cross-validation of component models.\textsuperscript{60} We attempt to collect more data, implement those methods and procedures, and document a study for estimating the number of components to be extracted by NTF from EEG recordings collected by the high-density array in a future publication.

Furthermore, as only the evoked activity of MMN is studied here, it is worth investigating the single trials of MMN recordings including the evoked and induced activities to examine whether the induced activity is able to assist in discriminating the children with RD and children with AD. This is really to test whether more types of brain responses can be richer to contribute a more discriminative multi-domain feature of an ERP or not since the results from single trials and from the average over single trials might be different.\textsuperscript{61} Two ways to proceed can be designed to achieve this goal. One is to still use the fourth-order tensor as that in this study, but the TFR is substituted by the event-related spectral perturbation\textsuperscript{51}; the other is first to obtain the TFR of single trials, and then formulate a fifth-order tensor including the dimensions of time...
by frequency by channel by trial by subject. When the latter procedure is adopted, a fast NTF algorithm is desired because the computing load will dramatically increase in case the single-trial EEG data is exploited. Another interesting topic is related to the method to produce the TFR. Except the wavelet transform, Hilbert–Huang transform (HHT) has been a promising method to analyze ERPs in the time-frequency domain. Hence, it would be very interesting to examine the multi-domain feature of MMN extracted by NTF from TFR of ERPs derived from HHT.

Another topic which is worth further investigation is to perform non-negative Tucker decomposition (NTD)\(^1\) on the multi-way representation of MMN data to extract the desired multi-domain feature of MMN instead of NTF. Regarding NTF under the CP model, the numbers of components in different factors should keep invariable, however, for NTD, the numbers of components in different factors can be different. For example, Fig. 3 shows that the spectral components are duplicative, thus, it is not necessary to extract redundant spectral components. We predict that using NTD can reduce the number of components to be extracted. Moreover, both CP and Tucker model with non-negativity constraint conform to the linear transform model of data without any strong assumptions on the probability density functions of sources in the model. This is different from some other two-way analysis methods, namely, ICA which requires the sources to be independent from each other. Hence, it would be interesting to compare the two-way feature of an ERP extracted by ICA and the multi-way feature of the ERP extracted by NTF under CP and Tucker models. We will report such studies in the near future.

It should be noted that the conventionally defined EEG actually includes two categories which are spontaneous EEG and ERPs. Since EEG was found in 1920s,\(^6\) the spontaneous EEG-based diagnosis of neurological disorders has been one of the fundamental research topics in brain study.\(^7\) However, ERPs are not studied for such a purpose as extensively as spontaneous EEG. This is because the experimental design and data collection of spontaneous EEG are much more applicable than those of ERPs and is because the use of advanced machine learning methods to study spontaneous EEG is consequently more straightforward in the field of computer science and artificial intelligence. For example, the parametric features,\(^8\) nonlinear and wavelet-based features,\(^9\) higher order cumulant features,\(^10\) recurrence quantification analysis,\(^11\) intrahemispheric, interhemispheric and distal EEG coherence,\(^12\) functional community analysis,\(^13\) and probabilistic neural networks,\(^14\) etc., have been used for the analysis of spontaneous EEG. However, they are seldom applied for the study of ERPs. From this perspective, the application of NTF for MMN research in this study provides one framework for the application of other advanced machine learning methods as mentioned above for the research of ERPs. The framework is that the cognitive knowledge of ERPs should be merged into the method and should be revealed by the results obtained from the advanced methods besides the results of prediction, detection, diagnosis and classification.

5. Conclusion

In this presentation, through NTF conforming to the CP model, we extract a multi-domain feature of MMN to discriminate children with RD and children with AD by exploiting the temporal, spectral, TFR, and spatial information of MMN simultaneously. The novel feature of MMN is more discriminative than the peak amplitude of MMN\(^15\) and the feature derived from the TFR of MMN. Moreover, the temporal, spectral and spatial components parallel with the new feature of MMN extracted by NTF do match the temporal, spectral and spatial properties of the MMN in our study. This indicates that the derived feature fulfills expectations of MMN related brain activity in this study, which is the fundamental for the extension of our proposed methodology for the research of other MMN and other ERPs.

In summary, as long as the brain signals including EEG, magnetoencephalography (MEG), functional magnetic resonance imaging (fMRI), and so on, can be modeled by the linear transform of latent variables, we suggest using non-negative tensor factorization and other multi-way analysis techniques to extract the multi-domain feature for more reliable representation of the underlying source of brain activity, particularly in the study of children.

The demo including MATLAB codes and ERP data to extract multi-domain feature of an ERP is available via http://users.jyu.fi/~fcong/
Appendix A. Outer Product of Vectors

Given three vectors \( \mathbf{a}, \mathbf{b}, \mathbf{c} \in \mathbb{R}^d \), their outer product yields a third-order rank-one tensor:
\[
\mathbf{z} = \mathbf{a} \otimes \mathbf{b} \otimes \mathbf{c} \in \mathbb{R}^{d \times d \times d},
\]
where, \( z_{ijk} = a_i b_j c_k \).

Appendix B. Mode-\( n \) Tensor Matrix Product

The mode-\( n \) product \( \mathbf{Y} = \mathbf{G} \times_n \mathbf{A} \) of a tensor \( \mathbf{G} \in \mathbb{R}^{n \times J_1 \times \cdots \times J_N} \) and a matrix \( \mathbf{A} \in \mathbb{R}^{J_n \times d} \) is a tensor \( \mathbf{Y} \in \mathbb{R}^{n \times J_1 \times \cdots \times J_{n-1} \times d} \), with elements
\[
y_{ijk_1 \cdots k_{n-1} \cdot j} = \sum_{k_n=1}^{J_n} a_{kj} g_{ik_1 \cdots k_{n-1}k_n},
\]
where \( [a] = [a_1, a_2, \ldots, a_J] \).

Appendix C. Mode-\( n \) Tensor-Vector Product

The mode-\( n \) product \( \mathbf{Y} = \mathbf{G} \hat{\times}_n (\mathbf{a}) \) of a tensor \( \mathbf{G} \in \mathbb{R}^{n \times J_1 \times \cdots \times J_N} \) and a set of \( N \) column vectors is a tensor
\[
\mathbf{Y} = \mathbf{G} \hat{\times}_n (\mathbf{a}) = \mathbf{G} \hat{\times}_n \mathbf{a} = (a_1, a_2, \ldots, a_J),
\]
where \( [a] = [a_1, a_2, \ldots, a_J] \).

Appendix D. Tensor Factorization Algorithms [Refs. 39 and 40 (Eq. (4))]

To estimate a component \( \mathbf{u}^{(n)} \), we assume all other components are fixed. The approximation tensor \( \mathbf{Y}^{(n)} \) is split into two parts as follows: A rank-one tensor \( \mathbf{Y}^{(n)} \) is built up from components \( \mathbf{u}^{(n)} \) to be estimated
\[
\mathbf{Y}^{(n)} = \mathbf{u}^{(n)} \otimes \mathbf{u}^{(n)} \otimes \ldots \otimes \mathbf{u}^{(N)}.
\]

A rank \((J-1)\) tensor \( \hat{\mathbf{Y}}^{(n)} \) is composed of \( N \) factors
\[
\hat{\mathbf{Y}}^{(n)} = [\mathbf{u}^{(n)} \mathbf{u}^{(n)} \mathbf{u}^{(n)} \ldots \mathbf{u}^{(n)}],
\]
that is
\[
\hat{\mathbf{Y}}^{(n)} = \sum_{k=1}^{J_n} \mathbf{u}^{(n)}(k) \otimes \mathbf{u}^{(n)}(k) \otimes \ldots \otimes \mathbf{u}^{(n)}(k).
\]

Appendix D. Tensor Factorization Algorithms [Refs. 39 and 40 (Eq. (4))]

To estimate a component \( \mathbf{u}^{(n)} \), we assume all other components are fixed. The approximation tensor \( \mathbf{Y}^{(n)} \) is split into two parts as follows: A rank-one tensor \( \mathbf{Y}^{(n)} \) is built up from components \( \mathbf{u}^{(n)} \) to be estimated
\[
\mathbf{Y}^{(n)} = \mathbf{u}^{(n)} \otimes \mathbf{u}^{(n)} \otimes \ldots \otimes \mathbf{u}^{(n)}.
\]

A rank \((J-1)\) tensor \( \hat{\mathbf{Y}}^{(n)} \) is composed of \( N \) factors
\[
\hat{\mathbf{Y}}^{(n)} = [\mathbf{u}^{(n)} \mathbf{u}^{(n)} \mathbf{u}^{(n)} \ldots \mathbf{u}^{(n)}],
\]
that is
\[
\hat{\mathbf{Y}}^{(n)} = \sum_{k=1}^{J_n} \mathbf{u}^{(n)}(k) \otimes \mathbf{u}^{(n)}(k) \otimes \ldots \otimes \mathbf{u}^{(n)}(k).
\]

Appendix D. Tensor Factorization Algorithms [Refs. 39 and 40 (Eq. (4))]

To estimate a component \( \mathbf{u}^{(n)} \), we assume all other components are fixed. The approximation tensor \( \mathbf{Y}^{(n)} \) is split into two parts as follows: A rank-one tensor \( \mathbf{Y}^{(n)} \) is built up from components \( \mathbf{u}^{(n)} \) to be estimated
\[
\mathbf{Y}^{(n)} = \mathbf{u}^{(n)} \otimes \mathbf{u}^{(n)} \otimes \ldots \otimes \mathbf{u}^{(n)}.
\]

A rank \((J-1)\) tensor \( \hat{\mathbf{Y}}^{(n)} \) is composed of \( N \) factors
\[
\hat{\mathbf{Y}}^{(n)} = [\mathbf{u}^{(n)} \mathbf{u}^{(n)} \mathbf{u}^{(n)} \ldots \mathbf{u}^{(n)}],
\]
that is
\[
\hat{\mathbf{Y}}^{(n)} = \sum_{k=1}^{J_n} \mathbf{u}^{(n)}(k) \otimes \mathbf{u}^{(n)}(k) \otimes \ldots \otimes \mathbf{u}^{(n)}(k).
\]
neuropsychiatric and neurological diseases and in ageing, Brain 134 (2011) 3432-3450.
10. N. Natãınen, The perception of speech sounds by the human brain as reflected by the mismatch negativity (MMN) and its magnetic equivalent (AMN), Psychophysiology 38 (2001) 1-21.
48. M. E. Timmerman and H. A. Kiers, Three-mode
46. F. Cong, Z. He, J. Hamalainen, A. Cichocki and T.
47. M. Morup and L. K. Hansen, Automatic relevance
45. C. F. Beckmann and S. M. Smith, Probabilistic
44. M. I. Garrido, J. M. Kilner, K. E. Stephan and K. J.
43. C. Tallon-Baudry, O. Bertrand, C. Delpuech and
42. C. Tallon-Baudry, O. Bertrand, C. Delpuech and
41. A. Delorme and S. Makeig, EEGLAB: An open
40. A. H. Phan and A. Cichocki, Multi-way nonneg-
39. A. Cichocki and A. H. Phan, Fast local algorithms
38. R. Bro, Multi-way analysis in the food industry —
37. N. Gillis and F. Glineur, Accelerated multiplicative
36. S. Pakarinen, R. Takegata, T. Rinne, M. Huotila-
35. R. Naatanen, S. Pakarinen, T. Rinne
34. R. Naatanen, S. Pakarinen, T. Rinne and R. Takaga,
33. S. Pakarinen, R. Takaga, T. Rinne, M. Huotila-
32. S. Pakarinen, R. Takaga, T. Rinne, M. Huotila-
31. S. Pakarinen, R. Takaga, T. Rinne, M. Huotila-
30. S. Pakarinen, R. Takaga, T. Rinne, M. Huotila-
29. N. Gillis and F. Glineur, Accelerated multiplicative
28. R. Naatanen, S. Pakarinen, T. Rinne
27. M. E. Timmerman and H. A. Kiers, Three-mode
26. F. Cong, Z. He, J. Hamalainen, A. Cichocki and T.
25. M. E. Timmerman and H. A. Kiers, Three-mode
24. F. Cong, Z. He, J. Hamalainen, A. Cichocki and T.
23. M. E. Timmerman and H. A. Kiers, Three-mode
22. R. Bro, Multi-way analysis in the food industry —
21. F. Cong, P. H. Leppanen, P. Astikainen, J. Hamalainen, J. K. Hietanen and T. Ristaniemi, Dimension reduction: Additional benefit of an optim-
20. F. Cong, I. Kalyakin, H. Li, T. Huttunen-Scott, Y. X.
19. R. Bro, Multi-way analysis in the food industry —
18. F. Cong, P. H. Leppanen, P. Astikainen, J. Hamalainen, J. K. Hietanen and T. Ristaniemi, Dimension reduction: Additional benefit of an optim-
17. F. Cong, I. Kalyakin, H. Li, T. Huttunen-Scott, Y. X.
16. F. Cong, P. H. Leppanen, P. Astikainen, J. Hamalainen, J. K. Hietanen and T. Ristaniemi, Dimension reduction: Additional benefit of an optim-
15. F. Cong, I. Kalyakin, H. Li, T. Huttunen-Scott, Y. X.
13. F. Cong, P. H. Leppanen, P. Astikainen, J. Hamalainen, J. K. Hietanen and T. Ristaniemi, Dimension reduction: Additional benefit of an optim-
12. F. Cong, I. Kalyakin, H. Li, T. Huttunen-Scott, Y. X.
11. A. Delorme and S. Makeig, EEGLAB: An open
10. A. H. Phan and A. Cichocki, Multi-way nonneg-
9. A. Cichocki and A. H. Phan, Fast local algorithms
8. R. Bro, Multi-way analysis in the food industry —
7. N. Gillis and F. Glineur, Accelerated multiplicative
6. F. Cong, P. H. Leppanen, P. Astikainen, J. Hamalainen, J. K. Hietanen and T. Ristaniemi, Dimension reduction: Additional benefit of an optim-
5. F. Cong, I. Kalyakin, H. Li, T. Huttunen-Scott, Y. X.
4. M. I. Garrido, J. M. Kilner, K. E. Stephan and K. J.
3. C. Tallon-Baudry, O. Bertrand, C. Delpuech and
2. C. Tallon-Baudry, O. Bertrand, C. Delpuech and
1. A. Delorme and S. Makeig, EEGLAB: An open

48. M. E. Timmerman and H. A. Kiers, Three-mode
47. M. Morup and L. K. Hansen, Automatic relevance
46. F. Cong, Z. He, J. Hamalainen, A. Cichocki and T.
45. C. F. Beckmann and S. M. Smith, Probabilistic
44. M. I. Garrido, J. M. Kilner, K. E. Stephan and K. J.
43. C. Tallon-Baudry, O. Bertrand, C. Delpuech and
42. C. Tallon-Baudry, O. Bertrand, C. Delpuech and
41. A. Delorme and S. Makeig, EEGLAB: An open
40. A. H. Phan and A. Cichocki, Multi-way nonneg-
39. A. Cichocki and A. H. Phan, Fast local algorithms
38. R. Bro, Multi-way analysis in the food industry —
37. N. Gillis and F. Glineur, Accelerated multiplicative
36. S. Pakarinen, R. Takegata, T. Rinne, M. Huotila-
35. R. Naatanen, S. Pakarinen, T. Rinne
34. R. Naatanen, S. Pakarinen, T. Rinne and R. Takaga,
33. S. Pakarinen, R. Takaga, T. Rinne, M. Huotila-
32. S. Pakarinen, R. Takaga, T. Rinne, M. Huotila-
31. S. Pakarinen, R. Takaga, T. Rinne, M. Huotila-
30. S. Pakarinen, R. Takaga, T. Rinne, M. Huotila-
29. N. Gillis and F. Glineur, Accelerated multiplicative
28. R. Naatanen, S. Pakarinen, T. Rinne
27. M. E. Timmerman and H. A. Kiers, Three-mode
26. F. Cong, Z. He, J. Hamalainen, A. Cichocki and T.
25. M. E. Timmerman and H. A. Kiers, Three-mode
24. F. Cong, Z. He, J. Hamalainen, A. Cichocki and T.
23. M. E. Timmerman and H. A. Kiers, Three-mode
22. R. Bro, Multi-way analysis in the food industry —
21. F. Cong, P. H. Leppanen, P. Astikainen, J. Hamalainen, J. K. Hietanen and T. Ristaniemi, Dimension reduction: Additional benefit of an optim-
20. F. Cong, I. Kalyakin, H. Li, T. Huttunen-Scott, Y. X.
19. R. Bro, Multi-way analysis in the food industry —
18. F. Cong, P. H. Leppanen, P. Astikainen, J. Hamalainen, J. K. Hietanen and T. Ristaniemi, Dimension reduction: Additional benefit of an optim-
17. F. Cong, I. Kalyakin, H. Li, T. Huttunen-Scott, Y. X.
16. F. Cong, P. H. Leppanen, P. Astikainen, J. Hamalainen, J. K. Hietanen and T. Ristaniemi, Dimension reduction: Additional benefit of an optim-
15. F. Cong, I. Kalyakin, H. Li, T. Huttunen-Scott, Y. X.
13. F. Cong, P. H. Leppanen, P. Astikainen, J. Hamalainen, J. K. Hietanen and T. Ristaniemi, Dimension reduction: Additional benefit of an optim-
12. F. Cong, I. Kalyakin, H. Li, T. Huttunen-Scott, Y. X.
11. A. Delorme and S. Makeig, EEGLAB: An open
10. A. H. Phan and A. Cichocki, Multi-way nonneg-
9. A. Cichocki and A. H. Phan, Fast local algorithms
8. R. Bro, Multi-way analysis in the food industry —
7. N. Gillis and F. Glineur, Accelerated multiplicative
6. F. Cong, P. H. Leppanen, P. Astikainen, J. Hamalainen, J. K. Hietanen and T. Ristaniemi, Dimension reduction: Additional benefit of an optim-
5. F. Cong, I. Kalyakin, H. Li, T. Huttunen-Scott, Y. X.
4. M. I. Garrido, J. M. Kilner, K. E. Stephan and K. J.
3. C. Tallon-Baudry, O. Bertrand, C. Delpuech and
2. C. Tallon-Baudry, O. Bertrand, C. Delpuech and
1. A. Delorme and S. Makeig, EEGLAB: An open

FA 1
November 15, 2012 13:3 1250025
1250025-18


