Multiview classification and dimensionality reduction of EEG data through tensor factorisation

Loukianos Spyrou · Samaneh Kouchaki · Saeid Sanei

Abstract Electroencephalography (EEG) signals arise as a mixture of various neural processes that occur in different spatial, frequency and temporal locations. In classification paradigms, algorithms are developed that can distinguish between these processes. In this work, we apply tensor factorisation to a set of EEG data from a group of epileptic patients and factorise the data into three modes; space, time and frequency with each mode containing a number of components or signatures. We train separate classifiers on various feature sets corresponding to complementary combinations of those modes and components and test the classification accuracy of each set. The relative influence on the classification accuracy of the respective spatial, temporal or frequency signatures can then be analysed and useful interpretations can be made. Additionally, we show that through tensor factorisation we can perform dimensionality reduction by evaluating the classification performance with regards to the number mode components and by rejecting components with insignificant contribution to the classification accuracy.

Keywords tensor factorisation · multiview · EEG · epilepsy

L. Spyrou
Department of Computer Science
University of Surrey
E-mail: l.spyrou@surrey.ac.uk

S. Kouchaki
Department of Computational and Evolutionary Biology
University of Manchester

S. Sanei
Department of Computer Science
University of Surrey

1 Introduction

The application of tensor factorisation to signal processing has has found recent applications to several areas such as antenna array processing, biomedical signal processing, feature extraction, and classification [12, 23, 11]. A tensor is a multi-way representation of data or a multidimensional array. Each direction in the tensor is called mode or way. Using tensor factorisation, the true underlying structure of that data can be preserved. Tensor factorisation methods has been shown to be powerful tools to describe signals which in general change in time, frequency, and space. Tensor analysis can provide a good way to discover the main features of the data and extract the hidden underlying information especially in the case of having big data size.

Several tensor based methods have been suggested for decomposition and multi-way representation of data. Tucker decomposition [35, 24] is one of the common tensor factorisation which is a generalisation of singular value decomposition (SVD) to higher order tensors. Using the Tucker model, data are decomposed into a number of factors and a core tensor of lower dimensions than the original data. Therefore, as suggested in [10], it can be employed to compress the high dimensional data to extract significant features.

The application of tensor decomposition can be significant for biomedical signals, such as EEG, where many transient events and movement related sources and artifacts are involved and most sources are inherently nonstationary. Moreover, the related brain neural process exhibit specific space-time-frequency locations. EEG signals in particular, consist of multichannel recordings with good temporal resolution which subsequently offers good time-frequency resolution. The application of tensor analysis is then natural and the

data can be factorised into space, time and frequency modes [45, 13, 20, 1]. Unlike those studies, we perform tensor factorisation within a machine learning context and the utility of the method is enhanced when considering multi-subject data where data can be factorised in the group level, identifying common components [34].

Typically, in EEG classification tasks, data are collected for each subject in an offline manner and then a pattern classification algorithm is trained on user specific data. This has the benefit of adapting the algorithm to the exact patterns of that user. However, such a procedure is time consuming and delays the actual use of the system. To overcome this problem, subject-independent (SI) algorithms can be useful. Such methods have the advantage of being used immediately without having to know any prior knowledge or subject specific data information, albeit with reduced performance. To alleviate this problem a new direction of SI identification for EEG, using tensor factorisation has been suggested here.

Interictal epileptic discharges (IEDs) are sharp transients of electrical activity that occur in the period between seizures of epileptic patients and provide clinical information regarding the type and location of seizures. IED analysis can provide useful presurgical information. IED detection methods can be categorised in several groups. They are typically based on methods such as template matching [27, 44, 38, 18], classification [46, 26], through independent component analysis (ICA) [22, 14], signal modelling [19] and many other methods [36] in common with the well established field of spike detection [42, 41]. The common characteristic of all these methods is that a description of an IED/spike signal is obtained either through modelling or through the use of real data and some similarity-based algorithm developed in order to detect those signals. This is often facilitated by obtaining useful representations of the signal that can better exploit its structure. The feature representations that are useful for IED detection has been a rather unresolved issue [36]. They can be grouped into the following categories: a) Mimetic, where the extracted features are designed to mimic human evaluation and methodology for the presence of a spike [17, 2], b) Time-Frequency represensations such as Wavelet, Fourier, or Hilbert transform have been used in IED detection and with much success in EEG signal processing in general [6, 4, 3, 37], c) Non-Linear features such as Hjorth parameters, fractal dimension, mean energy [8] and through d) decomposition methods such ICA [39, 29] and kernel type methods[43].

An epileptic scalp and intracranial EEG dataset of 18 patients is utilised in this work which consists of multiple interictal epileptic discharge (IED) events per subject. There are two primary goals in this study. Firstly, the use of tensor factorisation in order to enable the detection of the specific IED data signatures that exist in the group level. That way, a clinician can choose which specific type or property of the signal they require to be detected. Secondly, the estimation of a subspace that reduces the dimensionality of the classification task. Both of these methodologies are evaluated in a SI framework demonstrating the generalisation performance of our proposed algorithms. In Section 2 we describe the scalp-intracranial EEG dataset while Section 3 describes the theory behing tensor analysis and factorisation. Section 4 describes the preprocessing steps used to facilitate tensor decomposition and classification and the algorithm that detects the obtained modes and components. Section 5 shows results of multiview classification and dimensionality reduction on both scalp and intracranial EEG. Section 6 discusses the results while Section 7 concludes the paper.

2 Dataset

2.1 Recording

The study included 18 patients studied with scalp EEG and simultaneous intracranial multicontact foramen ovale (FO) recordings. For the intracranial data, two flexible bundles of 6 electrodes each were inserted through the left and right FO under general anaesthesia and fluoroscopic control, following the technique recommended by [40]. Each individual electrode consisted of a 0.1 mm fully insulated stainless steel wire. Recording contacts of the 3 deepest electrodes were 3 mm long and those of the most superficial electrodes were 5 mm long. Distance between contiguous electrodes was 10 mm except for the two most superficial electrodes whose interelectrode distance was 15 mm. For each electrode bundle, the two deepest electrodes lay next to medial temporal structures. The positions of the FO bundles were confirmed with post-insertion radiography.

For scalp EEG, 20 standard chlorided silver cup electrodes were applied according to the Maudsley electrode placement system [28, 16, 21, 31]. The advantage of the Maudsley System with respect to the standard $10-20$ System is that it provides more extensive coverage of the lower part of the cerebral convexity and adapts itself to cranial asymmetries [9], increasing the sensitivity for recording from basal sub-temporal structures. This system is essentially similar to 10-20 system, naming of electrodes is identical, but mid-temporal, posterior-temporal and occipital electrodes in the Maudsley System are approximately 20 mm lower than in the $10-20$ System.

Cable telemetry of 32 channels was used for data acquisition (Telefactor Beekeeper system). Data were digitised at 200 Hz and bandpass filtered (high pass cut-off frequency at 0.3 Hz and low pass cut-off frequency at 70 Hz). The system input range was 2 mV and data were digitised with a 12 bit analog-to-digital converter (an amplitude resolution of 0.488 mV). From each patient, a period of 20 min of intracranial EEG recordings were transcribed onto a digital file.

3 Methods

3.1 Tucker model

The Tucker factorisation was introduced in [35] and later refined in [25]. It is a generalisation of SVD to higher dimensions. As an example, Tucker [35] models each element of a four-way tensor $\underline{\mathbf{X}}$ with I_1, I_2, and I_3 factors in first, second, and third modes respectively similar to the SVD as:

$$x_{ijkl} \approx \sum_{i_1=1}^{I_1} \sum_{i_2=1}^{I_2} \sum_{i_3=1}^{I_3} u_{(1)ii_1} u_{(2)ji_2} u_{(3)ki_3} g_{i_1 i_2 i_3 l} \tag{1}$$

where i, j, k, l are the indices of the four modes. Equation (1) can be rewritten in tensor form as:

$$\underline{\mathbf{X}} \approx \underline{\mathbf{G}} \times_1 \mathbf{U}_1 \times_2 \mathbf{U}_2 \times_3 \mathbf{U}_3 \tag{2}$$

As can be seen, the Tucker model decomposes a tensor into a core tensor ($\underline{\mathbf{G}}$) multiplied by a matrix ($\mathbf{U}_1$, $\mathbf{U}_2$, or $\mathbf{U}_3$) along each mode.

Using the unfolded version of tensor $\underline{\mathbf{X}}$, Tucker model can be shown as

$$\begin{aligned} \mathbf{X}_{(1)} &= \mathbf{U}_1 \mathbf{G}_{(1)} (\mathbf{U}_3 \odot \mathbf{U}_2)^T + E_{(1)}, \\ \mathbf{X}_{(2)} &= \mathbf{U}_2 \mathbf{G}_{(2)} (\mathbf{U}_3 \odot \mathbf{U}_1)^T + E_{(2)}, \\ \mathbf{X}_{(3)} &= \mathbf{U}_3 \mathbf{G}_{(3)} (\mathbf{U}_2 \odot \mathbf{U}_1)^T + E_{(3)} \end{aligned} \tag{3}$$

The Tucker model parameters can be optimised using alternating least squares (ALS) by considering all except the parameter at each optimisation step [30] as:

$$\begin{aligned} \mathbf{U}_1 &= \mathbf{X}_{(1)} (\mathbf{G}_{(1)} (\mathbf{U}_3 \odot \mathbf{U}_2)^T)^\dagger, \\ \mathbf{U}_2 &= \mathbf{X}_{(2)} (\mathbf{G}_{(2)} (\mathbf{U}_3 \odot \mathbf{U}_1)^T)^\dagger, \\ \mathbf{U}_3 &= \mathbf{X}_{(3)} (\mathbf{G}_{(3)} (\mathbf{U}_2 \odot \mathbf{U}_1)^T)^\dagger, \\ \underline{\mathbf{G}} &= \underline{\mathbf{X}} \times_1 \mathbf{U}_1^\dagger \times_2 \mathbf{U}_2^\dagger \times_3 \mathbf{U}_3^\dagger \end{aligned} \tag{4}$$
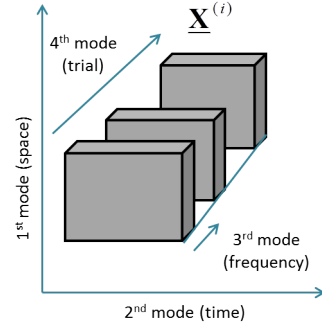

Fig. 1 Representation of 4-way tensor built for each subject.

By imposing the orthogonality constraint on all the modes, SVD can be used to update Tucker3 parameters as:

$$\begin{aligned} \mathbf{U}_1 &= \mathrm{LSing}(\mathbf{X}_{(1)} (\mathbf{U}_3 \odot \mathbf{U}_2)^T), \\ \mathbf{U}_2 &= \mathrm{LSing}(\mathbf{X}_{(2)} (\mathbf{U}_3 \odot \mathbf{U}_1)^T), \\ \mathbf{U}_3 &= \mathrm{LSing}(\mathbf{X}_{(3)} (\mathbf{U}_2 \odot \mathbf{U}_1)^T), \\ \underline{\mathbf{G}} &= \underline{\mathbf{X}} \times_1 \mathbf{U}_1^T \times_2 \mathbf{U}_2^T \times_3 \mathbf{U}_3^T, \end{aligned} \tag{5}$$

where $\mathrm{LSing}(\cdot)$ stands for left singular vectors of the input matrix. Equation (5) is called higher order orthogonal iteration (HOOI) and may converge to a local optimum solution.

3.2 EEG tensor factorisation

In this work, the EEG data of each subject are converted to a 4-way tensor $\underline{\mathbf{X}}^{(i)} \in \mathbb{R}^{N_s \times N_t \times N_f \times N_i}$ where i is the subject index and N_s, N_t, N_f, N_i are the spatial, time, frequency, and trial dimensions as illustrated in Fig. 1. The trial dimension contains both IED and non-IED segments (c.f. Section 4.1). In order to extract the common information for a group of K subjects, the individual tensors are concatenated (as suggested in [10]) along the forth dimension (Fig. 2) resulting in a group level tensor $\underline{\mathbf{X}}^g \in \mathbb{R}^{N_s \times N_t \times N_f \times N_{1:K}}$ where $N_{1:K}$ is the total number of trials of the K subjects.

We use the Tucker model for tensor factorisation into 3 modes (space, time, frequency) and a core tensor. Each mode is set to consist of a number (I_s, I_t, I_f) of components. Therefore, for a tensor $\underline{\mathbf{X}}$ ($\underline{\mathbf{X}}^g$ or $\underline{\mathbf{X}}^{(i)}$), the decomposition can be described as:

$$\underline{\mathbf{X}} = \underline{\mathbf{G}} \times_1 \mathbf{U}_s \times_2 \mathbf{U}_t \times_3 \mathbf{U}_f \tag{6}$$

where $\underline{\mathbf{G}} \in \mathbb{R}^{I_s \times I_t \times I_f \times N}$ shows the core tensor, $\mathbf{U}_s \in \mathbb{R}^{N_s \times I_s}$, $\mathbf{U}_t \in \mathbb{R}^{N_t \times I_t}$, and $\mathbf{U}_f \in \mathbb{R}^{N_f \times I_f}$ are the spatial, temporal and frequency modes respectively. When
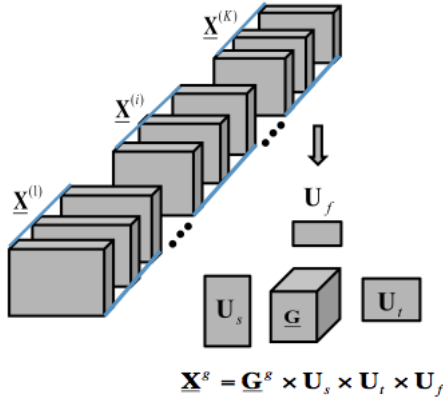
Fig. 2 Concatenating the data along the forth dimension and factorising the resulting group level tensor using Tucker decomposition.

performing tensor factorisation on $\underline{\mathbf{X}}^g$, the core sub-tensors of individual subjects can be selected by choosing the trials that correspond to that subject: $\underline{\mathbf{G}}^{(i)} \in \mathbb{R}^{I_s \times I_t \times I_f \times \{N_i\}}$. These core tensors describe the contribution of the obtained group level modes to the data of individual subjects and also are used as features for classification (c.f. Section 4.1).

Additionaly, for any predefined modes $\mathbf{U}_s$, $\mathbf{U}_t$, and $\mathbf{U}_f$ the core tensor of a subject -j- can be obtained by:

$$\underline{\mathbf{G}}^{(j)} = \underline{\mathbf{X}}^{(j)} \times_1 \mathbf{U}_s^T \times_2 \mathbf{U}_t^T \times_3 \mathbf{U}_f^T \qquad (7)$$

where $\underline{\mathbf{G}}^{(j)}$ shows the core tensor obtained by projecting the data $\underline{\mathbf{X}}^{(j)}$ of subject j onto the space spanned by $\mathbf{U}_s$, $\mathbf{U}_t$ and $\mathbf{U}_f$. These modes can be obtained from the factorisation procedure of the K training subjects. That way, the data of a test subject can be evaluated in terms of the common charecteristics of a training set of K subjects.

Tensor factorisation enables the following two important methods for EEG data:

1. By choosing specific components ($L_s \subseteq 1 : I_s, L_t \subseteq 1 : I_t, L_f \subseteq 1 : I_f$) of each mode the classification scheme can be targeted to detect only those components (see Figure 4). Therefore, classification performance can be evaluated for different aspects of the data (Section 4.2).
2. Since the dimension of the core tensor is $\underline{\mathbf{G}}^{(i)} \in \mathbb{R}^{I_s \times I_t \times I_f \times N_i}$, we can perform dimensionality reduction by selecting the number of components of each mode i.e I_s, I_t, I_f to be smaller than the dimensions of the original representation N_s, N_t, N_f. Similarly, by evaluating the classification accuracy for different components separately, we can discard
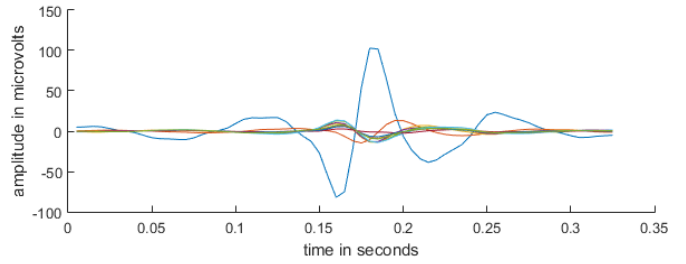

Fig. 3 Example IED discharge for one subject. Channels are superimposed on the same plot and the signal is an average over all IED segments.

those which do not reach a significant classification accuracy (Section 4.3).

4 Classification schemes

4.1 Pre-processing and classifier

An expert epileptologist inspected the intracranial data and marked the timing information where the intracranial data exhibited visually noticable epileptic discharges. The raw scalp and intracranial EEG data were band-pass filtered in the [1 70]Hz range and notch filtered at 50 Hz. Then, the data were sliced in a $\pm 162.5ms$ window centered on the intracranial timing scores and baselined on the preceding $162.5ms$ with the resulting signal finally being linearly detrended to remove undesired drifts. An example IED can be seen in Figure 3. Non-IED segments were also obtained from time segments where there are no scored IEDs. For each subject, the number of sliced IED and non-IED segments was chosen to be the same. The spectrogram method was used to convert the time-domain signal into time frequency features with a hanning-tapered window length of $80ms$ and an overlap of 50% between windows. These time frequency features are obtained for each IED and non-IED segment. Each IED event then consists of 756 (12 FO channels × 7 temporal × 9 frequency) features for intracranial and 1260 (20 scalp channels × 7 temporal × 9 frequency) features for scalp data.

Classification of a subject's EEG data is performed by using as features the obtained core tensor for that subject: $\underline{\mathbf{G}}^{(i)}$ (as well as using the original representation $\underline{\mathbf{X}}^{(i)}$ for comparison purposes only). We employ regularised linear logistic regression in order to train classifiers that can distinguish between IED and non-IED events both for the core tensor features and for the original representations. Logistic regression has been widely used in EEG data obtaining state of the art performance [15]. One benefit of logistic regression is that it provides a natural way to express the class member-

ship probability from the observed data. The estimation of the classifier is performed by minimising a quadratically regularised logistic regression loss function where there is a trade-off in classification accuracy against over-fitting and generalisation performance. The regularisation parameter, which controls the trade-off is estimated with cross-validation [33]. Cross validation (CV) is a method for model selection that works by partitioning the training data in a number of complementary subsets (folds). In k-fold cross-validation the data are split into k subsets. Each subset is used once to validate the model, while the rest $k-1$ subsets are used to train the classifier. For each fold we test the following regularisation parameters $[125, 25, 5, 1, 0.2, 0.04, 0.008]$ scaled by 10% of the variance of the data and select the one that achieves the highest performance. Significance testing of classification accuracies is performed with the binomial proportion confidence interval [5].

In order to facilitate SI classification of the IEDs and evaluate the generalisation performance of our algorithm we split our dataset of 18 patients into training and test sets. Each training set consists of the data from all subjects except the data of a single test subject. We perform this leave-subject-out procedure for each subject resulting in 18 different training-test combinations.

4.2 Multiview Classification

In this work, multiview classification entails the evaluation of classification performance for specific components of the obtained modes through tensor factorisation. In group level analysis, the multiview paradigm consists of choosing a subset of the components from the group level tensor factorisation. For each training set as described in Section 4.1, we perform the factorisation of Equation (6) by using the group level tensor $\underline{\mathbf{X}}^g$ and extract a core tensor for each training subject and common spatial, time and frequency modes. We further select the desired subsets of the core tensor corresponding to various selections (views) of factor components $(L_s \subseteq 1 : I_s, L_t \subseteq 1 : I_t, L_f \subseteq 1 : I_f)$. For example, the influence of the 1st spatial component on the classification accuracy can be examined by selecting $L_s = 1$, $L_t = 1 : I_t$ and $L_f = 1 : I_f$. The core tensors $\underline{\mathbf{G}}^{(i)}(L_s, L_t, L_f)$ that correspond to our selection of L_s, L_t and L_f are used to train classifiers $C^{(i)}(L_s, L_t, L_f)$ that correspond only to them. Then, the trained classifiers are used to classify the test data for the same L_s, L_t, L_f by obtaining $\underline{\mathbf{G}}^{(test)}(L_s, L_t, L_f)$ through Equation 7. By training separate classifiers for different selections of L_s, L_t, L_f different aspects of the data can be evaluated separately. An important feature of our multiview classification scheme is that apart
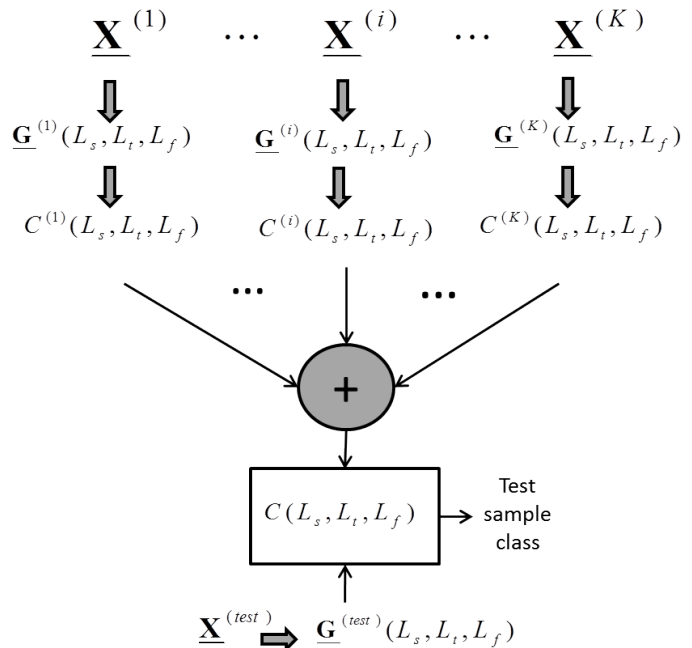

Fig. 4 The procedure of subject-independent detection of specific components through the selection of spatial, temporal and frequency components: L_s, L_t, and L_f respectively. Each training subject's $(i = 1 : K)$ are converted to their core tensors through group level tensor factorisation. Separate classifiers C^i are trained for each subject and then combined in an ensemble with the average sum rule. Unseen data from a test subject are similarly converted to features corresponding to the same selection of components and subsequently classified with the ensemble classifier. This procedure is repeated for all test subjects.

from the classification accuracy the algorithm outputs the probability that a segment is an IED. As a result, each view can be assigned a certainty value regarding its discriminability for IEDs. A summary of the proposed procedure can be seen in Fig. 4. These subject-specific classifiers are combined according to the average-sum rule [32] in order to classify the test data. We also train classifiers on unfactorised data in order to be able to contrast the performance of the various feature sets with the original representation.

4.3 Dimensionality reduction

We demonstrate two ways that dimensionality reduction can be achieved through tensor factorisation. Firstly, we estimate the classification accuracy for an increasing number of components of each mode I_s, I_t, I_f. The classification accuracy for each case is then evaluated and a certain point can be chosen that achieves a smaller number of features without reducing the accuracy. Secondly, we show that rejecting components that do not achieve a significant classification accuracy in the train-
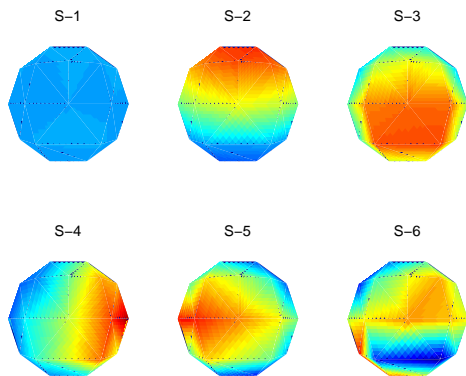
Fig. 5 Group level spatial components for the scalp data obtained through tensor factorisation. Each topography shows a common spatial characteristic between the patients' EEG data. It can be seen that expected spatial signatures were identified by the tensor factorisation procedure, such as left and right lateralisation of S-4 and S-5.
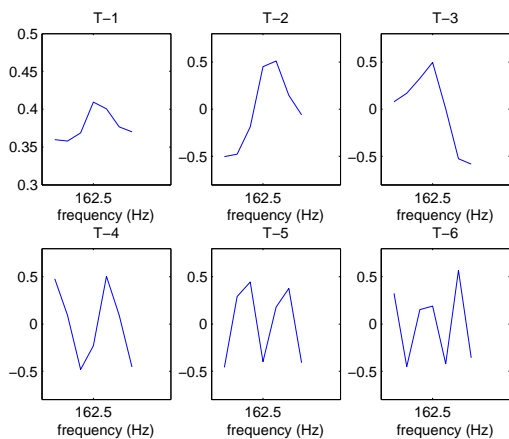

Fig. 6 Group level temporal components obtained through tensor factorisation. Each plot shows a common temporal characteristic between the patients' EEG data. Different activities peaking at around the middle of the segment were identified in accordance with the various expected IED morphologies.

ing set results in a small but significant increase in classification accuracy in the training set and equivalent performance on the test set.

5 Results

5.1 Factor analysis

In this section we show the obtained components of 6 spatial, time and frequency modes of the group level tensor in Figures 5, 6 and 7 respectively.
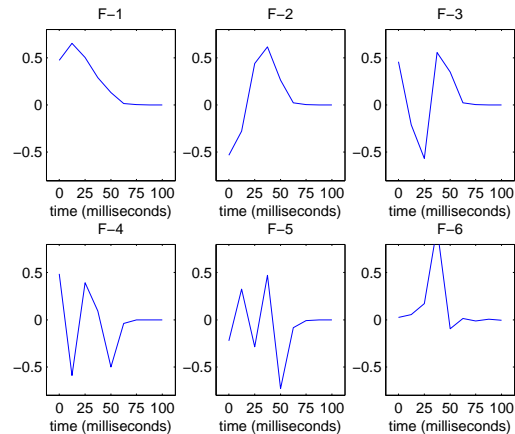

Fig. 7 Group level temporal components obtained through tensor factorisation. Each plot shows a common frequency characteristic between the patients' EEG data.

5.2 Multiview classification

We demonstrate results for both scalp and intracranial data highlighting the utility for both datasets. For scalp data, the classification accuracy for the components of the frequency and temporal modes can be seen in Tables 1 and 2. The column 'benchmark' denotes the accuracy of the original representation (i.e. without tensor factorisation) for both tables. The columns F-1 to -6 and T-1 to -6 the classification accuracy when only that component is used of that mode. For example the column F-1 denotes the accuracy when $L_f = 1, L_s = 1 : 6, L_t = 1 : 6$. It is observed that different components of different modes have different classification accuracy for different subjects. For intracranial data, we show the group average ratio of IEDs that were detected according to their brain lateralisation (Figure 8). It is a known characteristic that IEDs of the same subject can originate from different sides of the brain for different trials. The identification of which side the IED belongs to was performed in a post-hoc fashion by picking the side of the brain that the IED exhibits the highest amplitude (resulting in right- and left-sided segments). This information was not used in the training of the classifiers. SP-2 exhibited the highest detection rate for the right-sided segments, 88%, while the detection rate was 10% for left-sided segments.

5.3 Dimensionality reduction

In Figure 9 we show the group average classification accuracy with respect to the number of components (I_s, I_t, I_f).

In Table 3 the within subject classification accuracy for the 6 spatial segments of the scalp data in columns
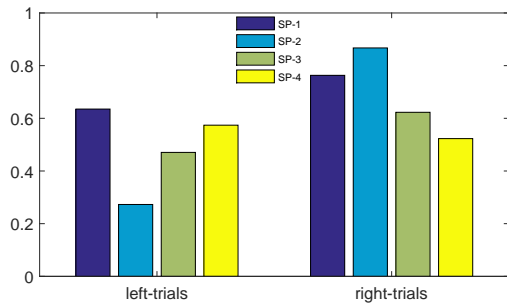
Fig. 8 Proportion of intracranial IEDs that were detected for each of the subject-independent classifiers trained separately on the 4 spatial components. The IEDs were split into left- and right-sided segments according to the side of the brain they had the highest amplitude. SP-2 provides the highest discriminability for distinguishing left from right segments.
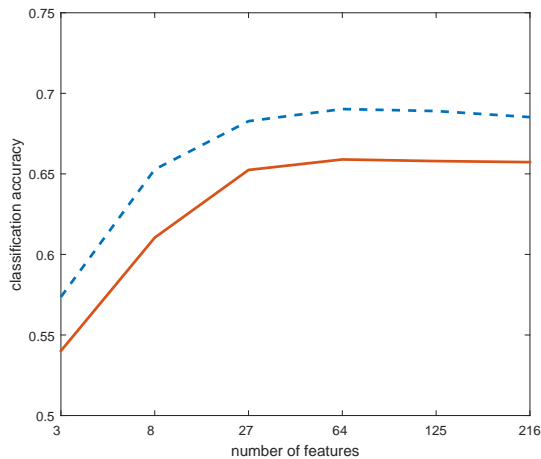

Fig. 9 Classification accuracy by progressively increasing the number of components (I_s, I_t, I_f) from 3 to 216. We show the group average accuracy of the test set (thick) and training set (dashed). A ceiling level can be observed at 64 features where classification accuracy does not increase any further.

(S-1 to -6) is shown. In the column denoted as 'reduced' we show the accuracy when we perform the dimensionality reduction method of Section 4.3. The spatial components that did not reach a significance classification accuracy at a 99% confidence were rejected and shown in bold. When performing that method for components from all three modes the resulting subject independent group accuracy was 65.9% while on the full original representation it was 66%.

6 Discussion

Tensor factorisation enables the visualisation of multiway EEG data and it can be considered the multiway extension of principal component analysis. As seen from Figures 5, 6 and 7 the obtained components of the spa-

tial, temporal and frequency modes identified the main characteristics of the data. For example, the S-4 describes activity originating from the right side of the brain. An important characteristic of our methodology is that the data a test subject can be projected to the modes obtained from tensor factorisation from a different subject or subjects. This way, the test subject's data can be evaluated with regards to that subject or group.

In a machine learning context, tensor factorisation allows the training of classifiers that operate on the projection of a test subject's data to a subset of the modes. It is clear from Tables 1 and 2 that different components of the frequency and temporal modes obtain different classification accuracies. A typical application would be to evaluate the discriminatory characteristics of a subject's data with respect to the modes extracted from a training group. In our epileptic scalp EEG dataset the frequency component with the highest and most consistent accuracy was F-1, a very typical IED frequency profile. By extracting the corresponding values from the core tensor we observed that the values for IED segments were much higher than for non-IED segments. Similarly, in the temporal domain (Table 2) the components T-1 to -3 obtained the highest performance. From Figure 6 we can see that these components exhibit spike like characteristics. Since each subject has their highest accuracy in one or two of those three, we can subsequently group the patients according to their temporal IED profile. Finally, for each view, our algorithm outputs the posterior probability that a segment is an IED. As seen from Tables 1 and 2 the two highest average probability views (shown in bold) coincide with the two highest accuracy views in most cases. That enables the user of such a system, an epilepsy clinician in this case, to select the views with the highest probability in an unsupervised fashion. Although this procedure does not increase the classification accuracy it is useful to estimate the important characteristics of the dataset that is analysed.

Furthermore, for the intracranial dataset, the proposed procedure was able to identify a spatial component that could determine the side of the brain that the target signal occured (see Figure 8). That was done without using any such information in the algorithm development and was a result of the tensor factorisation method. This has implications for epileptic data detection where the IEDs display variable locations and more generally for brain data when they exhibit multimodal characteristics.

In terms of dimensionality reduction, tensor factorisation was able to reduce the features required on the scalp EEG dataset from 1260 to 64 without any loss in

average classification accuracy. Similarly, we proposed an automatic method which rejects components that do not achieve a significant classification accuracy on the training set. That way the algorithm achieved similar accuracy (65.9%) on the test as the original representation (66%).

7 Conclusions

We developed a combined tensor factorisation and machine learning framework that allows enhanced analysis of EEG data. Firstly, a classification methodology was developed that can facilitate separate processing of the main factors underlying multiway data. Tensor factorisation using the Tucker decomposition model was used here to extract meaningful features from our multiway epileptic data. Within that framework it was possible to compare the contribution of those factors to the final classification accuracy. As was observed in our epileptic dataset, patients exhibited variable performance for different underlying factors. Such information is hidden when multiway data are not separated into factors. This can be generalised in the EEG field in cases that the detection of particular neural processes of a particular task is desired. Furthermore we proposed two ways that the dimensionality of the data can be reduced an issue which is becoming increasingly important for real world applications. Both of these methodologies can be extended to the EEG field for a variety of different datasets and tasks such as brain computer interfaces. It can also be extended to Big Data applications as well and can reduce the distributed load in large scale systems [7].

References

1. Acar, E., Aykut-Bingol, C., Bingol, H., Bro, R., Yener, B.: Multiway analysis of epilepsy tensors. Bioinformatics **23**(13), i10–i18 (2007)
2. Acir, N., Oztura, I., Kuntalp, M., Baklan, B., Güzeli, C.: Automatic detection of epileptiform events in EEG by a three-stage procedure based on artificial neural networks. IEEE Transactions on Biomedical Engineering **52**(1), 30–40 (2005)
3. Adeli, H., Ghosh-Dastidar, S., Dadmehr, N.: A wavelet-chaos methodology for analysis of EEGs and EEG sub-bands to detect seizure and epilepsy. IEEE Transactions on Biomedical Engineering **54**(2), 205–211 (2007)
4. Adeli, H., Zhou, Z., Dadmehr, N.: Analysis of EEG records in an epileptic patient using wavelet transform. Journal of Neuroscience Methods **123**(1), 69 – 87 (2003)
5. Alan Agresti, B.A.C.: Approximate is better than "exact" for interval estimation of binomial proportions. The American Statistician **52**(2), 119–126 (1998)
6. Argoud, F., De Azevedo, F.M., Neto, J., Grillo, E.: SADE3 : an effective system for automated detection of epileptiform events in long-term EEG based on context information. Medical Biological Engineering Computing **44**, 459–470 (2006)
7. Baker, B., Silva, R., Calhoun, V.D., Sarwate, A.D., Plis, S.M.: Large scale collaboration with autonomy: Decentralized data ica. Machine Learning for Signal Processing (2015)
8. Bao, F.S., Gao, J.m., Hu, J., Lie, D.Y.C., Zhang, Y., Oommen, K.J.: Automated epilepsy diagnosis using interictal scalp EEG. Conference Proceedings IEEE Engineering in Medicine Biology Society pp. 6603–6607 (2009)
9. Binnie, C., Dekker, E., Smit, A., Van der Linden, G.: Practical considerations in the positioning of EEG electrodes. Electroencephalogry Clinical Neurophysiology **53**, 453–458 (1982)
10. Cichocki, A.: Tensor decompositions: a new concept in brain data analysis? arXiv preprint arXiv:1305.0395 (2013)
11. Cichocki, A., Zdunek, R., Phan, A.H., Amari, S.: Nonnegative matrix and tensor factorizations: applications to exploratory multi-way data analysis and blind source separation. John Wiley & Sons (2009)
12. Comon, P.: Tensor decompositions, state of the art and applications. arXiv preprint arXiv:0905.0454 (2009)
13. Cong, F., Lin, Q.H., Kuang, L.D., Gong, X.F., Astikainen, P., Ristaniemi, T.: Tensor decomposition of eeg signals: A brief review. Journal of Neuroscience Methods **248**, 59 – 69 (2015)
14. De Lucia, M., Fritschy, J., Dayan, P., Holder, D.S.: A novel method for automated classification of epileptiform activity in the human electroencephalogram-based on independent component analysis. Medical & Biological Engineering & Computing **46**, 263–272 (2008)
15. Farquhar, J., Hill, J.: Interactions between pre-processing and classification methods for event-related-potential classification : best-practice guidelines for Brain-Computer Interfacing. Neuroinformatics **11**(December), 175–192 (2013)
16. Fernandez Torre, J., Alarcon, G., Binnie, C., Polkey, C.: Comparison of sphenoidal, foramen ovale and anterior temporal placements for detecting interictal epileptiform discharges in presurgical assessment for temporal lobe epilepsy. Clinical Neurophysiology **110**, 895–904 (1999)
17. Gotman, A., Gloor, P.: Automatic recognition and quantification of interictal epileptic activity in the human scalp EEG. Electroencephalography Clinical Neurophysiology pp. 513–529 (1976)
18. Grouiller, F., Thornton, R., Groening, K., Spinelli, L., Duncan, J., Schaller, K., Siniatchkin, M., Lemieux, L., Seeck, M., Michel, C., Vulliemoz, S.: With or without spikes: localization of focal epileptic activity by simultaneous electroencephalography and functional magnetic resonance imaging. Brain **134**, 2867–2886 (2011)
19. Janca, R., Jezdik, P., Cmejla, R., Tomasek, M., Worrell, G.a., Stead, M., Wagenaar, J., Jefferys, J.G.R., Krsek, P., Komarek, V., Jiruska, P., Marusic, P.: Detection of interictal epileptiform discharges using signal envelope distribution modelling: application to epileptic and non-epileptic intracranial eecordings. Brain Topography **28**(1), 172–183 (2014)
20. Karahan, E., Rojas-Lopez, P.A., Bringas-Vega, M.L., Valdes-Hernandez, P.A., Valdes-Sosa, P.A.: Tensor analysis and fusion of multimodal brain images. arXiv:1506.06040 (2015)
21. Kissani, N., Alarcon, G., Dad, M., Binnie, C., Polkey, C.: Sensitivity of recordings at sphenoidal electrode site for detecting seizure onset: evidence from scalp, superficial

and deep foramen ovale recordings. Clinical Neurophysiology **112**, 232–240 (2001)

22. Kobayashi, K., James, C., Nakahori, T., Akiyama, T., Gotman, J.: Isolation of epileptiform discharges from unaveraged eeg by independent component analysis. Clinical neurophysiology **110**(10), 1755–1763 (1999)

23. Kolda, T.G., Bader, B.W.: Tensor decompositions and applications. SIAM review **51**(3), 455–500 (2009)

24. Kroonenberg, P.M., De L., J.: Principal component analysis of three-mode data by means of alternating least squares algorithms. Psychometrika **45**(1), 69–97 (1980)

25. Levin, J.: Three-mode factor analysis. Psychological Bulletin **64**(6), 442 (1965)

26. Liu, Y.C., Lin, C.C.K., Tsai, J.J., Sun, Y.N.: Model-based spike detection of epileptic EEG data. Sensors **13**(9), 12,536–12,547 (2013)

27. Lodder, S.S., van Putten, M.J.A.M.: A self-adapting system for the automated detection of inter-ictal epileptiform discharges. PLoS ONE **9**(1), e85,180 (2014)

28. Margerison, J., Binnie, C., McCaul, I.: Electroencephalographic signs employed in the location of ruptured intracranial arterial aneurysms. Electroencephalogry Clinical Neurophysiology **28**, 296–306 (1970)

29. Martis, R., Acharya, U., Yan, J., Petznick, A., Chua, C., Ng, E.: Application of Intrinsic Time-scale Decomposition (ITD) to EEG Signals for Automated Seizure Prediction. International Journal of Neural Systems **23**(5) (2013). 1350023

30. Mørup, M.: Applications of tensor (multiway array) factorizations and decompositions in data mining. Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery **1**(1), 24–40 (2011)

31. Nayak, D., Valentín, A., Alarcón, G., García Seoane, J.J., Brunnhuber, F., Juler, J., Polkey, C.E., Binnie, C.D.: Characteristics of scalp electrical fields associated with deep medial temporal epileptiform discharges. Clinical Neurophysiology **115**(6), 1423–1435 (2004)

32. Rokach, L.: Ensemble-based classifiers. Artificial Intelligence Review **33**(1-2), 1–39 (2010)

33. Schaffer, C.: Selecting a classification method by cross-validation. Machine Learning **13**(1), 135–143 (1993)

34. Spyrou, L., Kouchaki, S., Sanei, S.: Multiview classification of brain data through tensor factorisation. In: Machine Learning for Signal Processing (MLSP), 2015 IEEE 25th International Workshop on, pp. 1–6. IEEE (2015)

35. Tucker, L.R.: Implications of factor analysis of three-way matrices for measurement of change. Problems in measuring change pp. 122–137 (1963)

36. Tzallas, A.T., Tsipouras, M.G., Tsalikakis, D.G., Karvounis, E.C., Astrakas, L., Konitsiotis, S., Tzaphlidou, M.: Automated epileptic seizure detection methods : A review study. In: D. Stevanovic (ed.) Epilepsy - Histological, Electroencephalographic and Psychological Aspects (2009). Available from: http://www.intechopen.com/books/epilepsy-histological-electroencephalographic-and-psychological-aspects/automated-epileptic-seizure-detection-methods-a-review-study

37. Vahabi, Z., Amifrattah, R., Ghassmi, F., Shayegh, D.: Online epileptic seizure prediction using wavelet-based bi-phase correlation of electrical signal tomography. International Journal of Neural Systems **25**(6) (2015). 1550028

38. Vijayalakshmi, K., Abhishek, A.M.: Spike Detection in Epileptic Patients EEG Data using Template Matching Technique. International Journal of Computer Applications **2**(6), 5–8 (2010)

39. Wang, C., Zou, J., Zhang, J., Wang, M., Wang, R.: Feature extraction and recognition of epileptiform activity in eeg by combining pca with apen. Cognitive Neurodynamics **4**(3), 233–240 (2010)

40. Wieser, H., Elger, C., Stodieck, S.: The foramen ovale electrode: a new recording method for the preoperative evaluation of patients suffering from mesio-basal temporal lobe epilepsy. Electroencephalography Clinical Neurophysiology **61**, 314–322 (1985)

41. Wilson, S.B., Emerson, R.: Spike detection: a review and comparison of algorithms. Clinical Neurophysiology **113**(12), 1873–1881 (2002)

42. Wilson, S.B., Turner, C.A., Emerson, R.G., Scheuer, M.L.: Spike detection II: automatic, perception-based detection and clustering. Clinical Neurophysiology **110**(3), 404–411 (1999)

43. Yuan, Q., Zhou, W., Yuan, S., Li, X., Wang, J., Jia, G.: Epileptic EEG classification based on kernel sparse representation. International Journal of Neural Systems **24**(4) (2014). 1450015

44. Zhanfeng, J., Sugi, T., Goto, S., Wang, X., Nakamura, M.: Multichannel template extraction for automatic EEG spike detection. 2011 IEEE/ICME International Conference on Complex Medical Engineering (CME) pp. 179–184 (2011)

45. Zhao, Q., Caiafa, C., Cichocki, A., Zhang, L., Phan, A.: Slice oriented tensor decomposition of eeg data for feature extraction in space, frequency and time domains. In: C. Leung, M. Lee, J. Chan (eds.) Neural Information Processing, *Lecture Notes in Computer Science*, vol. 5863, pp. 221–228. Springer Berlin Heidelberg (2009)

46. Zhou, J., Schalkoff, R.J., Dean, B.C., Halford, J.J.: Morphology-based wavelet features and multiple mother wavelet strategy for spike classification in EEG signals. Annual International Conference of the IEEE Engineering in Medicine and Biology Society **2012**, 3959–3962 (2012)

Table 1 Classification accuracy for scalp data (%) when no tensor factorisation is performed denoted as 'benchmark', and when each of the 6 components of the frequency mode are evaluated separately (F-1 to -6). The classification accuracy is evaluated with the subject-independent methodology of Figure 4. In bold, we show the two components that resulted in IED detections with the highest average posterior probability.

Subject	Benchmark	F-1	F-2	F-3	F-4	F-5	F-6
1	70	**68**	38	60	56	63	**73**
2	79	**77**	**70**	65	55	62	69
3	60	**62**	54	61	50	53	**62**
4	70	**66**	**62**	54	54	49	57
5	61	**56**	47	50	50	50	**59**
6	69	**68**	55	62	59	58	**65**
7	64	**59**	53	57	51	53	**60**
8	59	**58**	**52**	52	50	49	56
9	67	**64**	52	**60**	53	52	56
10	79	**73**	60	**70**	61	63	75
11	61	**58**	**57**	56	52	52	53
12	72	**69**	**67**	56	53	51	57
13	66	**65**	55	56	56	53	**60**
14	59	**58**	**59**	57	52	51	54
15	45	49	47	50	**52**	**55**	45
16	73	**73**	**62**	54	54	53	57
17	70	**73**	**62**	55	58	53	57
18	64	**61**	52	51	58	52	**56**

Table 2 Classification accuracy for scalp data (%) when no tensor factorisation is performed denoted as 'benchmark', and when each of the 6 components of the temporal mode are evaluated separately (T-1 to -6). The classification accuracy is evaluated with the subject-independent methodology of Figure 4. In bold, we show the two components that resulted in IED detections with the highest average posterior probability.

Subject	Benchmark	T-1	T-2	T-3	T-4	T-5	T-6
1	70	43	**70**	**67**	55	59	53
2	79	66	**76**	**76**	59	63	56
3	60	48	**59**	**55**	57	53	52
4	70	**65**	56	**60**	45	54	49
5	61	**56**	52	**57**	50	51	52
6	69	51	**67**	**67**	49	57	54
7	64	**58**	55	**61**	51	56	48
8	59	**60**	**52**	54	47	51	46
9	67	56	**62**	**62**	52	57	55
10	79	51	**75**	**77**	57	69	57
11	61	**60**	**56**	56	50	51	52
12	72	**67**	**68**	53	58	48	56
13	66	50	**59**	**64**	45	58	48
14	59	**56**	**56**	54	50	54	50
15	45	50	47	49	**51**	48	**53**
16	73	58	**65**	59	51	**56**	50
17	70	62	**64**	59	48	58	54
18	64	**59**	52	**57**	50	55	55

Table 3 Within subject classification accuracy of the six components of the spatial mode (S-1 to -6) for scalp data. Rejected components for each subject are shown in bold according to a 99% confidence level criterion. The significance level is different for each subject since the number of trials is different.

Subject	Benchmark	Reduced	S-1	S-2	S-3	S-4	S-5	S-6
1	80	83	81	65	**55**	64	62	58
2	83	86	78	**57**	**60**	68	64	**59**
3	59	62	62	**57**	**48**	**59**	**51**	**50**
4	73	73	73	**56**	**50**	58	58	**53**
5	55	63	63	**52**	**55**	59	**50**	**53**
6	74	76	74	61	57	55	**53**	57
7	60	63	60	56	**54**	**53**	**53**	56
8	65	64	65	57	**53**	**54**	**50**	**51**
9	70	68	66	56	60	**55**	58	58
10	82	82	84	61	69	64	59	**53**
11	63	63	62	**52**	60	**51**	**52**	55
12	74	74	73	61	57	55	61	61
13	72	72	69	56	55	64	59	56
14	59	59	59	**53**	49	54	54	**50**
15	48	56	**50**	**51**	**54**	**51**	**50**	56
16	77	77	77	58	55	**53**	**52**	61
17	75	74	75	**53**	**54**	61	**55**	**56**
18	62	65	64	**56**	**56**	59	60	**56**
mean	68.4	70	68.6	56.6	55.6	57.6	55.6	55.5