ABSTRACT
Epilepsy surgery outcome strongly depends on the localization of epileptic focus. The analysis of ictal EEG (scalp or intracranial) is a gold standard for definition of localization of epileptic focus. In order to automate visual analysis of large amounts of EEG data, we examine the correlations among electrodes captured by linear, nonlinear and multi-linear data analysis techniques. We study the performance of these statistical tools to understand the complex structure of epilepsy seizure and localize seizure origin. Our analysis results on four patients with temporal lobe epilepsy reveal that multiway (Tucker3 [1]) and nonlinear multiway (Kernelized Tucker3) analysis techniques are capable of capturing epileptic focus precisely when validated with clinical findings whereas linear and nonlinear analysis techniques (SVD, Kernel PCA [2]) fail to localize seizure origin.

KEY WORDS
biomedical computing, data mining, unsupervised learning, multiway analysis, epileptic focus

1 Introduction

Epilepsy is manifested as spontaneous clinical seizures as a result of paroxysmal, abnormally synchronous neuronal activity. The localization of the initial seizure discharge defines the region generating that abnormal activity; conversely, electrical manifestations should uniquely define and elucidate mechanisms of the underlying abnormal neuronal function and structure. Localization and electrical characteristics of spontaneous epileptic seizures in terms of visually identified and quantified parameters (distribution, morphology, energy, frequency, propagation and termination locations and latencies) and quantified measures of signals are the goals of our studies. In refractory epileptic patients, detection and localization of seizures’ origin is important for epilepsy surgery outcome.

1.1 Related Work

The majority of the research devoted to automated detection of epileptic events concentrates around spike detection techniques. These studies essentially differ in how they characterize epileptic spikes and waves. While some focus on features like amplitude, width and slope, others integrate this information with the state of EEG and build expert systems [3, 4, 5]. Some studies, on the other hand, describe epileptic spikes and waves in terms of the information obtained by Fourier Transformation (FT) [6] or Wavelet Transformation [7, 8]. Complex structure of epilepsy has also often attempted to be explained by nonlinear dynamics characterizing brain activity [9, 10].

In order to have a broader understanding of the structure of epileptic events, we will explore how linear, multi-linear and nonlinear statistical tools perform on EEG recordings of epileptic patients. Multi-linear models have been previously applied in the analysis of both EEG and event related potential (ERP) of EEG recordings. PARAFAC (Parallel Factor Analysis), one of the most popular multiway techniques, has been proved to be an efficient tool on EEG analysis [11]. EEG recordings have been transformed into three-way data containing frequency, time and channel information. Similar approach has been employed in an application of variety of PARAFAC algorithms on ERP data, which have been transformed into a multiway dataset by wavelet decomposition [12]. Another application [13], studying the effect of a new drug on brain activity, also uses multi-linear models on EEG. Results demonstrate that significant information has been successfully extracted from a complex drug dataset by using a multi-linear model (Tucker3), rather than two-way models such as Principal Component Analysis (PCA). To our knowledge, multiway models have not been applied for exploratory analysis of epilepsy. Nonlinear analysis using kernel methods have been performed on epileptic EEG data in the context of seizure prediction. These methods employ supervised learning methods [14] and focus on classifica-
tion of epilepsy periods. On the other hand, we study unsupervised learning models, which explore data without any prior training phase.

We are particularly interested in localization of epileptic focus. Although most spike detection techniques are based on single channel data, context knowledge from 16-channel EEG data has been incorporated in building a detection system for epileptic sharp waves in [5]. Sharp wave source localization on multichannel EEG data has also been applied to determine the areas of interest with epileptic activity in [15] and then visual inspection is used to find out clusters. Recently, Independent Component Analysis (ICA), is combined with spatio-temporal clustering methods to identify average location and time series of significant clusters of dipoles on MEG data [16].

1.2 Our Approach and Contributions

The main goal of this work is to understand the characteristics of epilepsy seizures by exploring the nature of the relationship, which arises between electrodes during seizure and use this information to localize seizure origins.

- We demonstrate that rank reduction is possible on EEG data arranged as either two-way matrices or three-way tensors. We choose the factors or combination of factors that account for around 90% of the variance for both two-way and three-way methods. In each factor, the electrodes with large coefficients are observed to be correlated during epilepsy seizure.

- We transform EEG recordings of epilepsy patients into three-way data with modes of time samples, scales and electrodes using wavelet decomposition based on Mexican-hat wavelet.

- We examine if epilepsy data have (i) linear (ii) multilinear (iii) nonlinear structure. We apply data analysis techniques that capture the above three cases: (i) SVD (ii) Tucker3 (iii) Kernel PCA and Kernelized Tucker3.

- We compare the performance of data analysis methods in terms of identifying seizure origin in epilepsy patients. We demonstrate that multiway analysis techniques, i.e., Tucker3 and Kernelized Tucker3, localize seizures more precisely while linear/nonlinear methods on two-way data spread epileptic focus over large regions.

This paper is organized as follows. Section 2 and Section 3 give a brief overview about data analysis techniques used throughout the paper. Section 4 describes how we process EEG recordings and construct matrices and tensors using epileptic EEG data. In Section 5, we discuss and compare the results of analysis methods with clinical findings.

2 Linear/ Multi-linear Analysis

Data are often represented in the form of matrices, \( A \in \mathbb{R}^{d_1 \times d_2} \), constructed from rows and columns. Multiway datasets (tensors), on the other hand, are high-order datasets, which are in the form of \( A \in \mathbb{R}^{d_1 \times d_2 \times \ldots \times d_r} \). \( r \) is the number of modes, which is equal to \( r \geq 3 \).

We consider one of the most common techniques in multiway analysis literature, i.e., Tucker3. This method is the generalization of Singular Value Decomposition (SVD) to high-order datasets. It gives us a substantially smaller set representing the original data. Compared to SVD, it generates component matrices containing singular vectors for each mode. On the other hand, Tucker3 represents the relationships between the modes with a core tensor as opposed to the presence of diagonal core matrix in SVD.

2.1 Tucker3

Tucker3 model decomposes a tensor \( A \in \mathbb{R}^{m \times n \times l} \) into three component matrices and a core tensor \( G \), which shows the relationships between different modes. Tucker3 decomposition of \( A \) is expressed as in equation 1.

\[
A_{ijk} = \sum_{r_1=1}^{R_1} \sum_{r_2=1}^{R_2} \sum_{r_3=1}^{R_3} G_{r_1r_2r_3} X_{ir_1} Y_{jr_2} Z_{kr_3} + E_{ijk} \quad (1)
\]

where \( R_1 \), \( R_2 \) and \( R_3 \) indicate the number of components extracted from mode 1, mode 2 and mode 3, respectively. \( X \in \mathbb{R}^{m \times R_1} \), \( Y \in \mathbb{R}^{n \times R_2} \) and \( Z \in \mathbb{R}^{l \times R_3} \) are the component matrices, \( G \in \mathbb{R}^{R_1 \times R_2 \times R_3} \) is the core tensor and \( E_{ijk} \) is the error term (Figure 1).

Tucker3 is the most flexible model among multiway analysis techniques. Although it suffers from core rotations, which result in non-unique solutions [17], Tucker3 model may be preferred over other multiway analysis methods such as PARAFAC because of its flexibility in extracting different number of components from each mode.

3 Nonlinear Analysis using Kernel Functions

It is not always possible to explore linear patterns in data. Data may contain nonlinear structure, which cannot be discovered by well-studied techniques capable of detecting
linear relations. One way to alter a linear pattern detection algorithm to a nonlinear one is to add attributes, which are nonlinear functions of input data. Any kernel method consists of two steps. First step is the mapping of input data to feature space through some kernel function (Table 1). Equation 2, as an example, illustrates what additional attributes are added when input data, x, are mapped to feature space using a second-degree polynomial kernel function. The idea of mapping is quite useful but explicit mapping to feature space is computationally inefficient. The famous solution to this problem is the “kernel trick” approach [18]. Second step is the application of a classical learning algorithm on the data mapped to the feature space.

\[
x = (x_1, x_2) \rightarrow \phi(x) = (x_1, x_2, x_1^2, x_2^2)
\]

Kernel Principal Component Analysis (Kernel PCA) [2] is an unsupervised learning method capable of finding nonlinear patterns in two-way data. Similar to other kernel methods, Kernel PCA applies two steps. The initial step is the mapping of input data to a higher dimensional feature space using a kernel function and the following step is the application of the well-known PCA model, which is quite similar to SVD in principle.

### 3.1 Kernelized Tucker3

Although multiway analysis methods are powerful enough to capture multi-linear structure of data, which cannot be discovered by linear detection models, their power is limited to a subset of nonlinear structures (multi-linearity addresses only some specific sort of nonlinearity). We want to point out that multiway datasets may possess more complex structures than multi-linearity. This structure cannot be captured by classical nonlinear analysis techniques because current nonlinear methods are limited to two-way data.

We introduce a novel technique, which combines kernel methods with Tucker3 model on three-way tensors. The idea originates from indirect fitting approach, which refers to algorithms applied on cross-product or covariance matrices [19].

Let \( A \in \mathbb{R}^{mxnxl} \) be a three-way tensor. First, second and third modes represent variables (e.g. electrodes), first-class features (e.g. frequency) and second class features (e.g. time samples) respectively. Our goal is to capture the structure among variables given first and second class features. We form two different kernel tensors \( K^1 \) and \( K^2 \) to find the structure in the first mode of tensor \( A \):

Let \( A_i \) be the \( i \)th slice, then \( i \)th slice of \( K, K_i \), is computed as \( K_i = \phi(A_i) \ast \phi(A_i)^T \), where \( \phi \) is the kernel function. Kernel matrices are all centered. We then compute \( K^1 \) and \( K^2 \) by slicing tensor \( A \) along second and third mode, respectively as shown in Figure 2. After the computation of \( K^1 \) and \( K^2 \), next step is to fit Tucker3 model to these kernel tensors for exploratory analysis.

### 4 Data Construction and Analysis

Complex structure of epilepsy has not been characterized with linear/nonlinear relations yet. We explore the performance of linear, multi-linear and nonlinear analysis techniques in extracting information from epilepsy data. We are specifically interested in the success of these techniques in localization of seizure origin. We focus on the portion of the data, which corresponds to seizure period of each patient since it has been shown that strong correlation arises between electrodes when a seizure develops [20].

#### 4.1 Two-way Data Analysis

We apply 2-way linear (SVD) and 2-way nonlinear (Kernel PCA) methods on two types of matrices. First type of dataset contains information from time domain of each electrode. We compute instantaneous signal power for each time sample from each electrode. We then construct a matrix, \( T \), where each entry \( T_{ij} \) represents the instantaneous power of \( j \)th time sample for the \( i \)th electrode. Second type of data contains signal power at different frequencies. We use Fast Fourier Transformation (FFT) to find frequency content of the signal and then form matrix \( F \), where each entry \( F_{ij} \) represent the signal power at \( j \)th frequency component at \( i \)th electrode (Figure 3).

We perform SVD and Kernel PCA on the two matrices given in Figure 3. In Kernel PCA, we apply Gaussian RBF (Radial Basis Function) kernel with a parameter deter-
deviation of the data recorded on each electrode.

To construct our dataset as three-way data, $A$, where each entry $A_{ijk}$ represents the power of wavelet coefficient at $i^{th}$ electrode for $j^{th}$ scale (contains frequency information [12]) at $k^{th}$ time sample.

The data are transformed using a Mexican-hat wavelet in the frequency band of 1-35Hz. Other studies, which have previously applied wavelet decomposition techniques on EEG data to form multiway datasets, have used complex Morlet wavelet [12]. However, we make use of Mexican-hat as a mother wavelet since we aim to capture epileptic events and Mexican-hat wavelet has been particularly shown to fit epileptic events well [7].

Our datasets contain single seizure period for each patient. After computing wavelet coefficients, we apply downsampling using a certain downsampling factor due to memory problems encountered while fitting Tucker3 model to large datasets. Another approach would be indirect fitting of Tucker3. Indirect-fitting would be computationally more efficient than direct fitting of Tucker3 model if we choose the electrodes with large ($\geq 0.45$) correlation values with that component.

## 4.2 Three-way Data Analysis

Even though FT is a widely used technique for frequency spectrum analysis, it is not sufficient to represent information content of non-stationary signals, e.g., EEG. FT assumes that all frequencies identified in frequency spectrum are available during the whole time duration. However, it is not the case for non-stationary signals. We use wavelet decomposition that is capable of identifying which frequency component is available during which time periods. Similar to [11, 12], we construct our dataset as three-way data, $A$, where each entry $A_{ijk}$ represents the power of wavelet coefficient at $i^{th}$ electrode for $j^{th}$ scale (contains frequency information [12]) at $k^{th}$ time sample.

The data are transformed using a Mexican-hat wavelet in the frequency band of 1-35Hz. Other studies, which have previously applied wavelet decomposition techniques on EEG data to form multiway datasets, have used complex Morlet wavelet [12]. However, we make use of Mexican-hat as a mother wavelet since we aim to capture epileptic events and Mexican-hat wavelet has been particularly shown to fit epileptic events well [7].

Our datasets contain single seizure period for each patient. After computing wavelet coefficients, we apply downsampling using a certain downsampling factor due to memory problems encountered while fitting Tucker3 model to large datasets. Another approach would be indirect fitting of Tucker3. Indirect-fitting would be computationally more efficient than direct fitting of Tucker3 model if we choose the electrodes with large ($\geq 0.45$) correlation values with that component.

## 5 Results and Interpretation

We studied 4 patients with temporal lobe epilepsy with different pathology substrates. Patient 1 had right temporal lobe epilepsy with a cavernoma on the right temporal pole. Patient 2 and 3 had right temporal lobe epilepsy with temporal tumor. Patient 4 had right temporal epilepsy with mesial temporal sclerosis. All patients were investigated for intractable partial epilepsy at Marmara and Yeditepe University Epilepsy Centers. Ictal EEG recordings were done with long term video EEG monitoring with scalp electrodes. The data analyzed with referentially or bipolar recorded surface EEG obtained with a nominal band pass of 1-35 Hz, a time scale of 30 mm per second and an optimal display of gain will be used for visual analyses of seizure characteristics. The recording of EEG with referential electrode Cz was used for computational analyses.

For each patient, we evaluate the performance of:

- **Linear Techniques on 2-way data**: SVD on matrix $T$ and SVD on matrix $F$ (Figure 3),
- **Nonlinear Techniques on 2-way data**: Kernel PCA on matrix $T$ and Kernel PCA on matrix $F$,
- **Multiway Techniques on 3-way data**: Tucker3 on tensor $A$, Kernelized Tucker3 on tensors $K_1$ and $K_2$ (Figure 2, 3).

We determine a set of electrodes (Figure 4), which are highly-correlated with significant components extracted using computational analysis techniques (Table 2). These electrodes are compared with clinical findings given in Table 3. Results demonstrate that Tucker3 and Kernelized Tucker3 both define the location of seizure origin precisely and close to clinical findings for all 4 patients. On the other hand, linear and nonlinear techniques on 2-way data, distinguish electrodes spread over a larger region around epileptic focus and therefore, provide less precise results.

Both Kernelized Tucker3 and Tucker3 capture epileptic focus precisely. However, Kernelized Tucker3 is computationally more expensive than Tucker3 because in addition to Tucker3 modeling, it also involves kernel function calculations. Therefore, our results suggest that Tucker3 is the model, which captures epileptic focus both precisely and efficiently. In future studies, we will investigate...
decade with the ready availability of high-speed computer 
implementation of short-term interventions to abort a seizure. 

do not improve the localization of the epileptic tissue which is structurally and 

whether Kernelized Tucker3 outperforms Tucker3 when 
kernels other than Gaussian RBF are used.

| PATIENT | SVD on T | SVD on F | Tucker3 on A | KPCA on T | KPCA on F | K. Tucker3 on $K^1$ | K. Tucker3 on $K^2$
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>$F_7, T_5, T_{15}, F_8, T_4$</td>
<td>$T_5, T_5, O_1, F_8, T_4$</td>
<td>$F_7, T_3, F_8, T_4$</td>
<td>$F_7, O_1, F_8, T_4, T_5$</td>
<td>$F_7, T_5, F_8, T_4, T_5, O_1$</td>
<td>$F_7, T_3, F_8, T_4$</td>
<td>$F_7, T_3, F_8, T_4$</td>
</tr>
<tr>
<td>2</td>
<td>$F_{P2}, T_4$</td>
<td>$F_{P2}, T_4, C_4, T_5$</td>
<td>$F_{P2}, F_4, F_8, C_4, T_4, T_6$</td>
<td>$F_{P1}, F_{P2}, T_3, C_4, T_5, T_6$</td>
<td>$F_{P1}, F_{P2}, T_3, O_2, P_1, P_4, T_6$</td>
<td>$F_{P2}, F_6, C_4, T_4, T_6$</td>
<td>$F_{P2}, F_6, C_4, T_4, T_6$</td>
</tr>
</tbody>
</table>
| 3       | $F_{P1}, F_{P2}, T_4, T_5, P_6, F_4$ | $F_{P1}, F_{P2}, T_5, T_6, F_4, P_6$ | $F_{P1}, F_{P2}, F_6, T_6, T_6, P_6$ | $F_{P1}, F_{P2}, F_6, T_6, T_6, P_6$ | $F_{P1}, F_{P2}, F_6, T_6, T_6, P_6$ | $F_{P1}, F_{P2}, F_6, T_6, T_6, P_6$ | $F_{P1}, F_{P2}, F_6, T_6, T_6, P_6$
| 4       | $T_6$ | $T_6$ | $F_7, F_8, T_6$ | $F_7, F_8, T_6$ | $F_7, F_8, T_6$ | $F_7, F_8, T_6$ |

Table 2. Comparison of Results. Electrodes which are found to be highly-correlated with significant factors explaining at least 90% of the variance in data.

Table 3. Clinically Determined Seizure Origins

<table>
<thead>
<tr>
<th>PATIENT</th>
<th>Epileptogenic Focus</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>$F_8, T_4$</td>
</tr>
<tr>
<td>2</td>
<td>$T_3, C_4, T_6, P_6$</td>
</tr>
<tr>
<td>3</td>
<td>$T_3, C_4, P_4, P_4$</td>
</tr>
<tr>
<td>4</td>
<td>$F_8, T_4, T_6, P_4$</td>
</tr>
</tbody>
</table>

Figure 4. 10-20 System of Electrode Placement

Understanding that low dimensional chaos per se is not likely to be manifest in most EEG datasets [23]. This understanding has led to a considerably more sophisticated view of nonlinearities in EEG and to the development of methods of detecting such effects. In particular, work over the past 5 years or so by several groups [24, 25, 26, 27, 28, 29] has focused on a variety of nonlinear measures, mostly based on correlation integrals [30]. Unlike previous work, these more recent investigations do not claim to detect chaos in EEG. Rather they take an empirical approach seeking to correlate values of nonlinear measures with disease states either in space or in time. Our analysis results strongly support this approach by showing that multi-linear analysis techniques (Tucker3) and nonlinear multiway methods (Kernelized Tucker3), which both use space, time and frequency information of EEG recordings, reveal the localization of epileptic focus most precisely.

In our study, potential criticism is the number and homogeneity of the patients. In future studies, we will investigate more patients with the same etiological pathologies. From computational point of view, we will examine how well Tucker3, high-order generalization of SVD, performs compared to Multi-linear ICA, which is the generalization of ICA to multiway data.

6 Conclusion and Future Work

A successful outcome of epilepsy surgery depends on the localization of the epileptic tissue which is structurally and functionally abnormal. The availability of reliable methods of seizure prediction and localization could enhance the quality and safety of patients with epilepsy; facilitate implementation of short-term interventions to abort a seizure. Efforts to predict have gathered momentum over the past decade with the ready availability of high-speed computer processing and the application of sophisticated mathematical techniques to biological processes [22].

In studies of EEG, there is now a generally accepted understanding that high dimensional chaos per se is not likely to be manifest in most EEG datasets [23]. This understanding has led to a considerably more sophisticated view of nonlinearities in EEG and to the development of methods of detecting such effects. In particular, work over the past 5 years or so by several groups [24, 25, 26, 27, 28, 29] has focused on a variety of nonlinear measures, mostly based on correlation integrals [30]. Unlike previous work, these more recent investigations do not claim to detect chaos in EEG. Rather they take an empirical approach seeking to correlate values of nonlinear measures with disease states either in space or in time. Our analysis results strongly support this approach by showing that multi-linear analysis techniques (Tucker3) and nonlinear multiway methods (Kernelized Tucker3), which both use space, time and frequency information of EEG recordings, reveal the localization of epileptic focus most precisely.

In our study, potential criticism is the number and homogeneity of the patients. In future studies, we will investigate more patients with the same etiological pathologies. From computational point of view, we will examine how well Tucker3, high-order generalization of SVD, performs compared to Multi-linear ICA, which is the generalization of ICA to multiway data.

References


