

## A CONSTRAINED APPROACH FOR EXTRACTION OF PRE-ICTAL DISCHARGES FROM SCALP EEG

*Shahrzad Shapoori, Wenwu Wang and Saeid Sanei*

Faculty of Engineering and Physical Sciences, University of Surrey, Guildford, UK  
{s.shapoori, w.wang & s.sanei}@surrey.ac.uk

### ABSTRACT

A constrained blind source separation (BSS) approach for separation of intracranial spikes from scalp electroencephalogram (EEG) has been proposed in this paper. This method is based on creating a template from intracranial data, which is then used in the form of a constraint in a BSS algorithm. To generate a suitable template, the segments during which the brain discharges are labelled are used to generate the necessary templates. Approximate entropy followed by peak detection and thresholding is used for this purpose. Constrained BSS is then applied to scalp data to extract the desired source and to evaluate its effect on scalp electrodes. The effectiveness of such a constrained approach has been demonstrated by comparing its outcome with that of the unconstrained method.

**Index Terms**— EEG, intracranial recording, approximate entropy, interictal discharges, and constrained BSS.

### 1. INTRODUCTION

Onset of tiny pre-ictal discharges originating from the brain hippocampus can be the start of developing seizure in human. Early detection of such waveforms is indeed useful in clinical assessment and treatment of the disease. Numerous long-lasting neurological disorders are characterized by seizure incidences. Occurrences of seizures have negative effects on the lives of patients such as decreasing awareness, impairing physical abilities or causing abnormal behaviour. In later stages of seizure, the patients need to regularly use anticonvulsants and even go under surgical operation in order to remove the affected parts of their brain [1]. Information regarding the place where seizures happen is necessary before surgery. It has been shown that there are some spikes (discharges) in the EEG long before the start of seizure, which can be used in its prediction [2]. These discharges look much more pronounced in the intracranial EEG than in scalp EEG.

Some research has been undertaken on extraction of intracranial spikes from scalp EEG using different techniques [3]. Extracting spikes from scalp EEG has been done by other

methods such as the morphology-based method by Yadav et al. [4] and the averaging method by Nayak et al. [5].

In the work proposed in [5], averaging has been applied to scalp and intracranial signals recorded instantaneously. These recordings have been synchronised with the discharges recorded using foramen ovale (FO) electrodes from the hippocampus. Signal to noise ratio has been improved due to this averaging. In order to assess the patients before temporal lobe resection at the centre, they go under FO telemetry. Topography, distribution, and amplitude of averaged scalp signals of these patients have been studied and analysed.

In independent component analysis (ICA), the source signals are considered as random variables and the statistical properties of these signals are used to obtain the unmixing matrix. In fact, the joint probability density of sources is considered as the multiplication of marginal probability of sources. In other words, it considers the statistical dependency between sources. Constrained source separation using ICA (cICA) defined in [6], introduces the use of a constraint in the separation algorithm, in order to extract a source, which as well as being statistically independent from other sources, is also the closest to a reference signal. This reference signal should not be necessarily an exact match, but should be close enough to be able to direct the algorithm to the desired output. Therefore, a single independent component (IC) can be extracted based on prior expectations of the required source signal. Classic ICA has been modified by adding prior information, as a constraint, previously in order to apply it to EEG signals for separation and analysis.

In the approach proposed by Corsini et al. [7] ICA has been used to separate the seizure signals for prediction. Their aim was to indirectly apply the traditional nonlinear methods to the scalp EEG. Therefore, only having scalp EEG, intracranial signals have to be separated correctly. It was expected to acquire signals similar to intracranial signal recordings after noise and artefact removal. Then, traditional nonlinear methods would be applied to these signals. In their method, they considered long recordings and analysed the data section by section. For this reason, the continuity of the algorithm failed in some cases where a section of scalp EEG was corrupted or electrical activity was not recorded correctly. The number of

segments that had to be used would need to be more and more for the longer recordings of data. Therefore, a new methodology to maintain the continuity of the estimated sources for particular segments would have to be developed.

In the method proposed by Jing et al. [8], in order to separate epileptic seizure signal sources from all the channels of scalp EEG, a constrained topographic ICA has been used. Application of constraints has been exploited to improve the performance of topographic ICA as an efficient means to group the source signals coming from a particular brain zone. Constraint is generally selected based on some particular physiological or statistical properties of the desired signal sources. In their method, the constraint is based on the spatial-frequency information of the seizure signals. Their method called constrained topographic ICA (cTICA) method has proved to outperform the conventional ICA and TICA algorithms in terms of signal to interference ratio and correlation measurement. Constrained ICA has also been used for artifact rejection in EEG [9].

In order to extract intracranial discharges from scalp EEG, a new constrained ICA (cICA) is developed. Measuring chaotic properties of intra channels, some sections are selected which contain reasonable number of discharges. The discharges are pre-emphasised and templates containing only the discharges are constructed. These templates are used in the constrained cost function. The constrained problem is then turned into an unconstrained problem using a penalty term such as Lagrange multiplier. The experimental results show that the constraint significantly improves detection of pre-ictal discharges from the EEG signals.

It has to be mentioned that although in this work a valuable intracranial dataset jointly/simultaneously recorded with scalp EEG is used, the objective of this research is to derive a useful model, which can be directly applied to the EEG signals for extraction of such spikes. This will be the agenda of our next research step.

## 2. METHODOLOGY

The overall algorithm here follows the block diagram in Fig. 1. Various blocks are explained in the following subsections.

### 2.1. Approximate Entropy

Approximate entropy (ApEn) [10], [11], is a technique for determining the amount of regularity and unpredictability of time series. Several experiments have shown that it can classify complex systems very well in both deterministic and stochastic data. It has two parameters; window length  $m$  and a tolerance  $r$ . ApEn procedure for calculating the complexity of time series  $x$  is as follows:

1. Form a sequence of  $m$ -vectors by following procedure:

$$\mathbf{y}(i) = [x(i), x(i+1), \dots, x(i+m-1)] \quad (1)$$

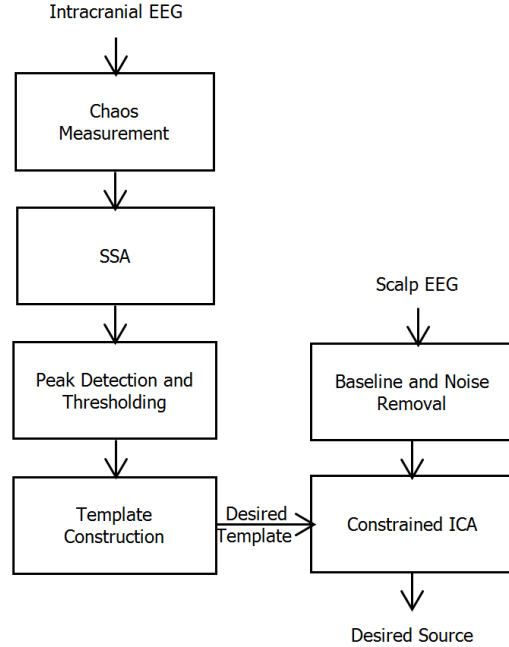


Figure 1. Block diagram of the proposed method.

2. Compute the following value for each constructed vector:

$$C_i^m(r) = \text{num}(d(y(i), y(j)) \leq r) / (N - m + 1) \quad (2)$$

where  $d(\dots)$  indicates the maximum absolute distance between the vector elements.

3. Then, compute:

$$\Phi^m(r) = \sum_{i=1}^{N-m+1} C_i^m(r) / (N - m + 1) \quad (3)$$

4. Finally, the ApEn is determined as follows:

$$\log(\Phi^m(r)) - \log(\Phi^{m+1}(r)) \quad (4)$$

In ApEn,  $m$  as suggested by Pincus [10] can have the value 2, and  $r$  is better to have a value between 0.1 and 0.25 of the standard deviation (SD) of the original data sequence.

### 2.2. Singular Spectrum Analysis

Singular spectrum analysis (SSA) is a powerful method for analysis of real valued time series [12]. It is based on the singular value decomposition (SVD) of time series. Moreover, it is functioning without any parametric model. In addition, SSA does not require the stationarity condition. Therefore, it is not model based and has a wide range of applications. SSA consists of two main stages: decomposition and reconstruction. In the first stage, the time series is decomposed into a

number of components. In order to achieve this, the embedding procedure, followed by SVD, is applied to the signal. The embedding procedure maps a  $\mathbf{1D}$  vector  $\mathbf{x}$  of length  $n$  to an  $l \times l_1$  matrix  $\mathbf{Y}$ :

$$\mathbf{Y} = [\mathbf{s}_1, \mathbf{s}_2, \dots, \mathbf{s}_l], \quad (5)$$

$$= \begin{pmatrix} x_1 & x_2 & \cdots & x_l \\ x_2 & x_3 & \cdots & x_{l+1} \\ \vdots & \vdots & \ddots & \vdots \\ x_{l_1} & x_{l_1+1} & \cdots & x_{l+l_1-1} \end{pmatrix} \quad (6)$$

where vectors  $\mathbf{s}_k = [x_k, x_{k+1}, \dots, x_{k+l_1-1}]^T \in R^l$ ,  $l_1 = n - l + 1$ ,  $l$  is the window length and  $(\cdot)^T$  stands for transpose. To keep the information regarding variation of data, window length should be large enough. Next, the SVD of trajectory matrix  $\mathbf{Y}$ , is computed and rewritten as:

$$\mathbf{Y} = \sum_{i=1}^d \mathbf{Y}_i = \sum_{i=1}^d \sqrt{\lambda_i} \mathbf{u}_i \mathbf{v}_i^T \quad (7)$$

where  $\lambda_i$  is the  $i$ th eigenvalue of  $\mathbf{Y}\mathbf{Y}^T$ 's covariance matrix,  $\mathbf{u}_i$  shows the eigenvector corresponding to  $\lambda_i$ ,  $d$  is the number of eigenvalues, and  $\mathbf{v}_i = \mathbf{Y}^T \mathbf{u}_i / \sqrt{\lambda_i}$ .

At last, in order to reconstruct the time series, the time series is rebuilt using groups of desired eigenvalues. Therefore, the matrices obtained in the previous stage are grouped into several sub matrices:

$$\mathbf{Y} = \sum_{q=q_1}^Q \hat{\mathbf{Y}}_q \quad (8)$$

where index  $q$  shows the  $q$ th subgroup of eigenvalues,  $\hat{\mathbf{Y}}_q$  the sum of  $\mathbf{Y}_i$  within group  $q$ , and  $Q$  refers to the total number of groups. The calculated matrix is further transformed into a Hankel matrix, which in turn is converted to a time series. The hankelization operator,  $\mathcal{H}$  for an  $i \times j$  matrix  $\mathbf{X}$  is defined as [11]:

$$\mathcal{H}\mathbf{X} = \begin{pmatrix} \tilde{x}_1 & \tilde{x}_2 & \cdots & \tilde{x}_l \\ \tilde{x}_2 & \tilde{x}_3 & \cdots & \tilde{x}_{l+1} \\ \vdots & \vdots & \ddots & \vdots \\ \tilde{x}_{l_1} & \tilde{x}_{l_1+1} & \cdots & \tilde{x}_{l+l_1-1} \end{pmatrix} \quad (9)$$

$$\tilde{x} = \sum_{i,j \in D_k} /num(D_k) \quad (10)$$

$$D_k = (i, j) : 1 \leq \hat{i} \leq i, 1 \leq \hat{j} \leq j, \hat{i} + \hat{j} = k + 1 \quad (11)$$

### 2.3. Proposed Method

The procedure of the proposed method is illustrated as a block diagram in Fig. 1. Approximate entropy is applied to the pre-seizure part of intracranial data, which contains discharges,

as a measure of chaosity. Seizure starts when the chaosity of signal decreases to its global minimum and then increases. Discharges happen when the chaosity of the signal reaches its local minimums and then increases. The approximate entropy algorithm applied here segments the data with an overlap. Segments that follow exactly where the local minimums happen are chosen as data segments, which include discharges.

SSA has been applied to these segments to remove the signal baseline and noise, therefore keep mostly the discharges. The result is zero meaned and normalized.

In the next stage, a modified peak detection algorithm is applied to these processed segments, considering the time distance of actual discharges and their amplitude. In addition, the threshold of peak detection is adjusted. The result of the final peak detection stage is a generated template based on the location and amplitude of actual discharges in the intra signal segments. The scalp data is also processed in order to remove the noise and baseline.

The final template is incorporated into the constraint part of ICA cost function to be able to extract the most similar source from the scalp mixtures.

Consider the following linear mixing model for EEG data:

$$\mathbf{X} = \mathbf{A}\mathbf{S} + \mathbf{V} \quad (12)$$

where  $\mathbf{X}$  is an  $n \times T$  matrix of mixture signals,  $\mathbf{S}$  is an  $m \times T$  matrix of source signals,  $\mathbf{A}$  is the mixing matrix of  $n \times m$ , and  $\mathbf{V}$  is noise.  $n$  is the number of scalp electrodes,  $T$  is the number of samples and  $m$  is the number of sources. Then, the objective is to find the desired sources using the unmixing vector  $\mathbf{w}$ :

$$\mathbf{y} = \mathbf{w}\mathbf{X} \quad (13)$$

$\mathbf{w}$  is a  $1 \times n$  vector and  $\mathbf{y}$  is a  $1 \times T$  vector which indicates the estimated source.

Here, the Hyvarinen's fixed-point algorithm namely, fastICA, in deflation mode [13] is used to obtain the ICA cost function,  $J(\mathbf{w})$ , since it has been shown to be effective for EEG source separation.

The main problem with most blind source separation techniques such as ICA is that they do not produce unique outputs without using some prior knowledge. Therefore, in order to extract a precise source, which helps in seizure prediction, constraints can be added to original cost function. Lagrange multipliers are used to incorporate the constraint function,  $J_c(\mathbf{w})$ , into the original cost function, which in turn changes the problem into an unconstrained one:

$$\min_{\mathbf{w}} J(\mathbf{w}) \text{ subject to } \max_{\mathbf{w}} J_c(\mathbf{w}) \quad (14)$$

In this paper, the inner product between the estimated signal ( $\mathbf{y}$ ) and the reference template is used. This cost function indicates that the objective is to find the source, which has the maximum correlation to this template. Hence, the constraint cost function used in this method is:

$$J_c(\mathbf{w}) = \mathbf{w}^T \mathbf{X} \mathbf{y}_t \quad (15)$$

where  $\mathbf{y}_r$  is the reference template. Therefore, using iterative methods, the following update function can be used:

$$\begin{aligned}\mathbf{w}_{k+1} &= \mathbf{w}_k - \mu(\nabla_{\mathbf{w}}J - \lambda\nabla_{\mathbf{w}}J_c) \\ \mathbf{w}_{k+1} &= \mathbf{w}_k - \mu(\nabla_{\mathbf{w}}J - \lambda\mathbf{X}\mathbf{y}_r)\end{aligned}\quad (16)$$

where  $\mu$  indicates the step size and  $\lambda$  is a fixed Lagrange multiplier.

In order to test the extracted component, one source is selected and the effect of its corresponding component on the scalp electrodes is depicted using topoplot and is analyzed. The effect of deep intracranial discharges should be emphasized. Therefore, the topoplot should contain the highest energy near the temporal area. However, the effect of this component is not very local as it is related to a deep source.

### 3. EXPERIMENTAL RESULTS

#### 3.1. Data

Data is recorded using both scalp and FO electrodes to measure scalp mixtures and intracranial sources respectively. The recording has been done simultaneously with patients under anesthesia. The sampling rate is 200 samples per second [5].

#### 3.2. Results

Approximate entropy is applied to the pre-seizure part of intracranial data (3000 samples, sampling-rate: 200 samples/second), which contains discharges, as a measure of chaosity. The algorithm applied here segments the data into 200 sample segments with an overlap of 100 samples. The results can be seen in Fig. 2.

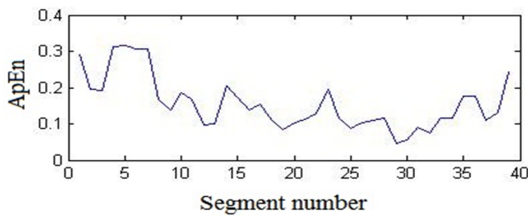


Figure 2. The result of applying the approximate entropy to a relevant channel of intracranial data including 39 segments.

As illustrated in Fig. 2, one possible solution for the chosen segment is the data part corresponding to the segment after segment 3.

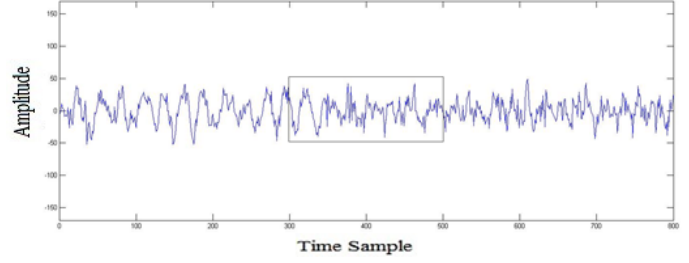


Figure 3. A channel of EEG with a segment including pre-ictal discharges

This segment is shown in Fig. 3. After choosing the desired segment, SSA is applied for noise removal and the locations of spikes are determined by the procedure explained in the previous section, which is peak detection and thresholding. The results are shown in Fig. 4.

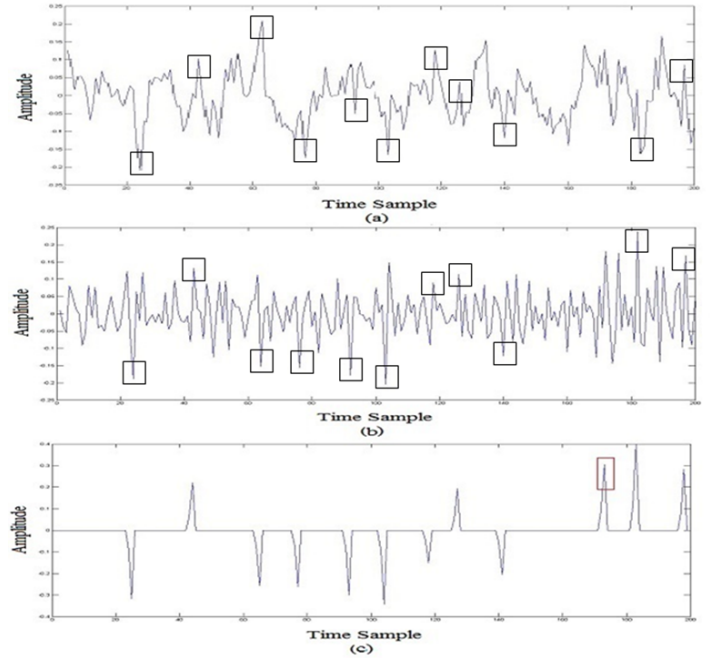


Figure 4. The location of discharges in an intracranial data segment, which is shown by squared boxes; (a) intracranial data (b) output of SSA, and (c) location of spikes.

As depicted, using this method all of the discharges within the chosen segment are located in addition to a false peak as indicated in Fig. 4(c). Finally, the constrained ICA is applied to the data. The convergence graph is shown in Fig. 5. The effect of the extracted source with conventional ICA and constrained ICA on the scalp electrodes is shown in Fig. 6. It is illustrated in Fig. 6(a) that the maximum power is in the location of interest, as shown with orange colour. In Fig. 6(b), the power is scattered on the brain. Therefore, it can be concluded that this procedure can be used to emphasize the intracranial discharges in the scalp signals.

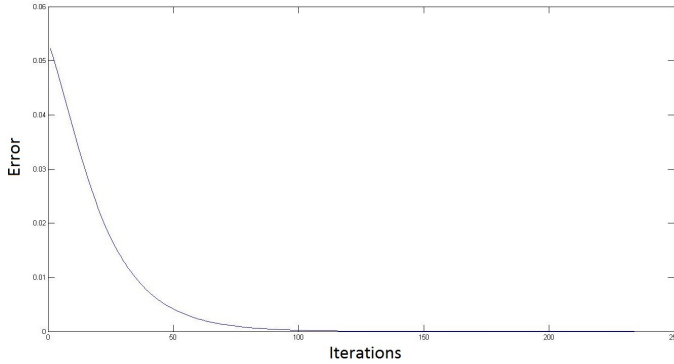


Figure 5. Convergence graph of cICA. Error is the difference between  $\mathbf{w}_k$  and  $\mathbf{w}_{k+1}$ .

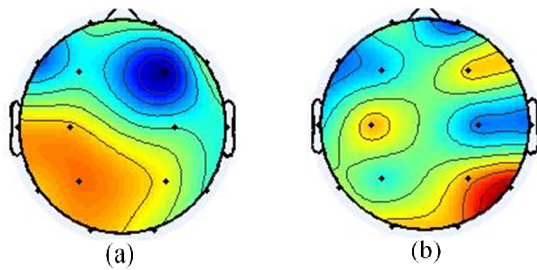


Figure 6. Topoplot of extracted source using (a) constrained ICA and (b) conventional ICA.

#### 4. CONCLUSION AND DISCUSSION

In this paper, a constrained approach for extraction of pre-ictal discharges from scalp EEG has been proposed. In order to achieve this, first a template based on intracranial recordings, is generated to be used in the constraint cost function, which is incorporated into the ICA formula. Constrained ICA is then applied to the scalp EEG to extract the corresponding desired source.

In conventional ICA, independency of sources is considered in order to separate them, which is not necessarily useful when a particular source is required. By adding extra information as a constraint, the closeness of the extracted source to the reference signal (template) is also considered as well as independency of the sources.

The results show that the proposed method works well with real data and outperforms unconstrained ICA.

#### 5. REFERENCES

- [1] G. Alarcon, Diagnosis and classification of the epilepsies: pathophysiological and psychiatric aspects, Published in Psychiatric Comorbidity and Epilepsy. Basic Mechanisms, Diagnosis, and Treatment. Edited by H. W. McConnell and P. J. Snyder. American Psychiatric Press, Inc. pp. 37-83, 1998.
- [2] Z. Rogowski, I. Gath, and E. Bental, "On the prediction of epileptic seizures," *Biological Cybernetics*, vol. 42, pp. 9-15, 1981.
- [3] S. Sanei, *Adaptive Processing of Brain Signals*, John Wiley & Sons, 2013.
- [4] R. Yadav, R. Agarwal, and M. N. S. Swamy, "A novel morphology-based classifier for automatic Detection of Epileptic Seizures," *Proceedings of the 32nd Annual International Conference of the IEEE EMBS Buenos Aires, Argentina*, pp. 5545-5548, 2010.
- [5] D. Nayaka, A. Valentn, G. Alarcon, J. J. G. Seoaneb, F. Brunnhubera, J. Julera, and C. E. Polkeya, "Characteristics of scalp electrical fields associated with deep medial temporal epileptiform discharges," *Clinical Neurophysiology*, vol. 115, pp. 1423-1435, 2004.
- [6] W. Lu and J. C. Rajapakse, "ICA with reference," *Proceedings of the 3rd Int. Conf. Independent Component Analysis and Blind Signal Separation: ICA2001*, pp. 120-125.
- [7] J. Corsini, L. Shoker, S. Sanei, and G. Alarcon, "Epileptic Seizure Predictability from Scalp EEG Incorporating Constrained Blind Source Separation," *IEEE Trans. on Biomedical Engineering*, vol. 53, pp. 790-799, 2006.
- [8] M. Jing and S. Sanei, "A Novel Constrained Topographic Independent Component Analysis for Separation of Epileptic Seizure Signals," *Hindawi Journal of Computational Intelligence and Neuroscience*, vol. 2007, 2007.
- [9] C. J. James, and O. J. Gibson, "Temporally Constrained ICA: An Application to Artifact Rejection in Electromagnetic Brain Signal Analysis," *IEEE Trans. on Biomedical Engineering*, vol. 50, pp. 1108-1116, 2003.
- [10] S. M. Pincus, "Approximate entropy as a measure of system complexity," *Proceedings of the National Academy of Science United States of America*, vol. 88, pp. 2297-2301, 1991.
- [11] D. Abasolo, R. Hornero, P. Espino, J. Poza, C. I. Sanchez, and R. de la Rosa, "Analysis of regularity in the EEG background activity of Alzheimers disease patients with Approximate Entropy," *Clinical Neurophysiology*, vol. 116, pp. 1826-1834, 2005.
- [12] N. Golyandina, V. Nekrutkin, and A. Zhigljavsky, *Analysis of Time Series Structure: SSA and related techniques*, Chapman & Hall/CRC, 2001.
- [13] A. Hyvarinen, J. Karhunen, and E. Oja, *Independent Component Analysis*, John Wiley & Sons, 2004.