

# An Adaptive Filtering Approach Using Supervised SSA for Identification of Sleep Stages From EEG

Samaneh Kouchaki <sup>1,\*</sup>, Konstantinos Eftaxias <sup>1</sup>, Saïed Sanei <sup>1</sup>

1. Faculty of Engineering and Physical Sciences, University of Surrey, Guildford, UK

Received: July 28 2014

Accepted: October 23 2014

## Keywords:

Adaptive Filters,  
Empirical Mode  
Decomposition,  
Recursive Least Squares,  
Singular Spectrum Analysis,  
Sleep EEG.

## ABSTRACT

**Purpose:** Sleep is a complex physiological state and an indicator of the changes in the brain function similar to those occurring in many psychiatric and neurological conditions. Since visual sleep scoring consuming process, automatic sleep staging methods, also called scoring, hold promise in diagnosing alterations in the sleep process and the sleep EEG more effectively.

**Method:** In this paper, a supervised approach for sleep scoring from single channel EEG signals is proposed. First, a supervised singular spectrum analysis (SSA) which is a subspace based method is used to extract the desired signal for each stage of sleep. Then, two recursive least squares (RLS) adaptive filters are trained and used to identify first and deep sleep stages.

**Result:** The proposed system which can be considered as a filter bank for separating multiple signal subbands is tested using real EEG where the results verify the accuracy of the proposed method.

**Conclusion:** The overall result show the effectiveness of algorithm for detection of sleep stages from EEG signals often characterised by a sharp increase in delta and a rapid decrease in alpha as sleep deepens.

## 1. Introduction

The signals recorded in nature are generally the mixtures of a number of source signals and interferences. Hence, our task is to unmix them in order to retrieve the underlying sources. In multichannel data, specially when the number of sources is equal or less than the number of electrodes, this problem is efficiently handled by employing blind source separation (BSS) techniques [1, 2, 3]. However, most of such techniques fail for single channel or generally underdetermined (where the number of sources is larger than the number of mixtures) recordings.

In several applications, just one channel is used in the separation process. Extracting various sleep stages with different prominent frequency bands can be considered

as a good example, where each frequency band prevails in a special recorded channel [4, 5]. Sleep is a complex state which is characterized by reduced responsiveness to external stimuli, a reduction in body movements, and the changes in metabolic rate [6]. It is an unconsciousness state from which a person can be aroused. Interactions between activating and inhibiting systems within the brain is used to identify brain states during wakefulness and sleep. There are two distinct states with different levels of neuronal activity: nonrapid eye movement (NREM) and rapid eye movement (REM) sleep. Most prominently, the alpha rhythm (8-13 Hz) is attenuated and delta (up to 4 Hz) waves evolve as the NREM sleep deepens. Delta rhythm is often called slow wave (SW). Since visual sleep scoring is a time consuming process, automatic sleep staging methods, also called scoring, hold promise in diagnosing alterations in the sleep process and the sleep EEG more effectively.

### \* Corresponding Author:

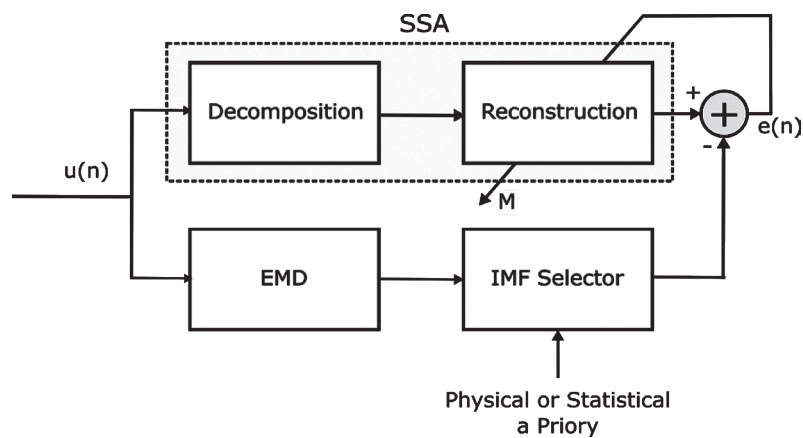
Samaneh Kouchaki, PhD  
Faculty of Engineering and Physical Sciences, University of Surrey, Guildford, UK.  
Tel: +447450217402  
E-mail: s.kouchaki@surrey.ac.uk

Many methods have been proposed for automatic sleep scoring using EEG data. As it is hard to score the EEG sleep data automatically using time or frequency information separately, most techniques try to use both time and frequency information to improve the scoring results. One of the common techniques is wavelet transform (WT) [6-9]. Using WT the extracted components vary significantly with changes in the mother wavelet. This has a negative impact on the robustness and accuracy of the outcome. Empirical mode decomposition (EMD) [10] is another alternative which automatically decomposes the signal to its constituent components intrinsically matched with the signal structure. Moreover, EMD does not use any fixed or predefined function in the decomposition stage. Therefore, it may help have a better result although it has its inherent mode change problem. This may not necessarily link each mode function to an event in the brain.

SSA is another powerful subspace based method for analysing real valued time series [11]. It combines the multi-variate statistics, classical time series analysis, dynamical systems, and signal processing. SSA is becoming an effective method in various areas such as economics [11] and biomedical signal processing [12, 13]. Basically, SSA decomposes a data into a number of interpretable el-

ements with different subspaces, such as noise and trend, and can be used for any time series with complex structure [14]. For instance, for decades, SSA has been used for time series analysis for both trend detection and prediction in financial data [11, 14]. Basic SSA has problem in selecting the desired subspace automatically. Hence, a constrained SSA should be employed to generate an estimate of the required trend, which later can be used as a reference for the adaptive filters described in section 2.

Recently, we proposed a technique able to produce the desired subspace for narrowband signals [15] which can extract the desired component of each stage automatically using EMD. This method called SSA-EMD uses prior knowledge about the desired data to select the corresponding subspace. A block diagram of this method can be seen in Fig. 1. Hence, SSA-EMD can be considered as a powerful time-frequency method and, therefore, it can be used for sleep scoring. Using SSA-EMD, a learning system can be developed and trained based on the EEG from different sleep stages. Since each stage is characterized by a distinct set of physiological and neurological related features and a dominant frequency band, it can be considered as one good application of this method.



**Figure 1.** Block diagram of the single channel source separation system using an adaptive procedure for selecting the desired subspace. This is carried out by tuning a set of weights governed by the EMD process.

Adaptive filtering (AF) has several applications in many areas such as communications, sound, image, and signal processing [16, 17]. They have the ability of adjusting their properties according to the desired features of the input data. Therefore, they can filter the unwanted components of input data specially in time varying signals. Least mean squares (LMS) is the common algorithm for AFs [17]. However, they have slow convergence due to the eigenvalue spread of input data

correlation matrix. On the other hand, recursive least squares (RLS) filters have a decoupled route to convergence. RLS filters can work well only if their output error is white and Gaussian.

A hybrid model is proposed here which uses RLS and SSA-EMD and applied to EEG for sleep scoring. The block diagram of the proposed method can be seen in Figure 2. At each stage, the signal segment is passed

to both RLS and SSA-EMD. SSA-EMD produces the desired, also called target, signal and the RLS updates the filter weights accordingly.

The remainder of the paper is structured as follows: First, RLS, SSA, and other related techniques are explained and the proposed method is introduced. Then, the results of applying the method to sleep scoring is shown. Finally, the last section concludes the paper.

## 2. Materials and Methods

In this section, first, RLS and SSA-EMD are explained, and then, the proposed method is introduced.

**Recursive Least Squares:** The RLS is a recursive method that adjusts the filtering weights by minimizing a weighted linear least squares cost function. If we have a desired signal  $\mathbf{d}$  which is transmitted over a noisy environment and record a noisy mixture  $\mathbf{u}$  and consider the following linear equation between  $\mathbf{d}$  and  $\mathbf{u}$  at each time instant  $t$ :

$$u(t) = \sum_{i=0}^p a_i(i)d(t-i) + v(t) \tag{1}$$

where  $v(t)$  represents noise. Then, RLS tries to recover the desired signal using a  $p+1$  tap filter  $\mathbf{w}$ :

$$\hat{d}(n) = \mathbf{w}_n^T \mathbf{u}_n \tag{2}$$

where  $\mathbf{u}_n$  is a vector including  $p$  recent samples of  $u(n)$ . RLS has the aim of estimating the filter weights at each time sample but without performing the least square algorithm for each  $n$ . Filter coefficients,  $\mathbf{w}$ , for both filters are estimated using the following recursive operations [16]:

$$\begin{aligned} a^{(k)}(n) &= d^{(k)}(n) - \mathbf{u}_n^T \mathbf{w}_{n-1}^{(k)} \\ \mathbf{w}_n^{(k)} &= \mathbf{w}_{n-1}^{(k)} + a^{(k)}(n) \mathbf{g}_n^{(k)} \end{aligned} \tag{3}$$

where

$$\begin{aligned} \mathbf{g}_n^{(k)} &= \mathbf{P}_{n-1}^{(k)} \mathbf{u}_n (\rho + \mathbf{u}_n^T \mathbf{P}_{n-1}^{(k)} \mathbf{u}_n) \\ \mathbf{P}_n^{(k)} &= \rho^{-1} \mathbf{P}_{n-1}^{(k)} - \mathbf{g}_n^{(k)} \mathbf{u}_n^T \rho^{-1} \mathbf{P}_{n-1}^{(k)} \end{aligned} \tag{4}$$

and  $\rho$  is the forgetting factor and  $k = 1, 2$ .

**Singular Spectrum Analysis:** SSA is a subspace decomposition algorithm with two stages: decomposition and reconstruction. First step involves embedding fol-

lowed by singular value decomposition (SVD). In the embedding stage a vector  $\mathbf{u}$  of length  $n$  is mapped to a matrix  $\mathbf{Y}$  using a sliding window:

$$\begin{aligned} \mathbf{Y} &= [\mathbf{s}_1, \mathbf{s}_2, \dots, \mathbf{s}_l] \\ &= \begin{pmatrix} u_1 & u_2 & \dots & u_l \\ u_2 & u_3 & \dots & u_{l+1} \\ \vdots & \vdots & \ddots & \vdots \\ u_l & u_{l+1} & \dots & u_{l+l-1} \end{pmatrix} \end{aligned} \tag{5}$$

where  $\mathbf{s}_k = [u_k, u_{k+1}, \dots, u_{k+l-1}]^T R^l$ ,  $l$  is the window length, and  $(.)^T$  stands for transpose.  $l$  should be large enough in order to contain the information about the data variation.

Next, the SVD of this matrix is computed as:

$$\mathbf{Y} = \sum_{i=1}^d \mathbf{Y}_i = \sum_{i=1}^d \sqrt{\lambda_i} \mathbf{r}_i \mathbf{v}_i^T \tag{6}$$

where  $\lambda_i$  is the  $i$ th eigenvalue of covariance matrix  $\mathbf{Y}\mathbf{Y}^T$ ,  $\mathbf{r}$  is the corresponding eigenvector,  $d$  is the total number of eigenvalues, and  $\mathbf{v}_i = \mathbf{Y}^T \mathbf{r}_i / \sqrt{\lambda_i}$

In the next stage, initially, the elementary matrices from the previous stage are grouped into a number sub-matrices:

$$\mathbf{Y} = \sum_{q=1}^Q \hat{\mathbf{Y}}_q \tag{7}$$

where  $Q$  determines the total number of groups, index  $q$  refers to  $q$ th subgroup of eigenvalues, and  $\hat{\mathbf{Y}}_q$  indicates the sum of  $\mathbf{Y}_i$  within group  $q$ . Finally, the resulting matrix is transformed into the form of a Hankel matrix that can be converted to a time series. The hankelization operator  $\hat{h}$  for an  $i \times j$  matrix  $\mathbf{Y}$  is obtained as:

$$\begin{aligned} \hat{h}\mathbf{Y} &= \begin{pmatrix} \tilde{y}_1 & \tilde{y}_2 & \dots & \tilde{y}_i \\ \tilde{y}_2 & \tilde{y}_3 & \dots & \tilde{y}_{i+1} \\ \vdots & \vdots & \ddots & \vdots \\ \tilde{y}_j & \tilde{y}_{j+1} & \dots & \tilde{y}_{j+i-1} \end{pmatrix} \\ \tilde{y}_k &= \sum_{i,j \in D_k} y_{i,j} / \text{num}D_k \\ D_k &= \{(\hat{i}, \hat{j}) : 1 \leq \hat{i} \leq i, 1 \leq \hat{j} \leq j, \hat{i} + \hat{j} = k + 1\} \end{aligned} \tag{8}$$

where num abbreviates number.

Basic SSA has problems with selecting the desired subspace that has a significant impact on the final results. We use our method proposed for automatically

selecting the desired subspaces using EMD [15]. In this method, the first stage of SSA is performed. At the same time, EMD is applied to the data and a number of intrinsic mode functions (IMFs) falling within particular frequency band are selected according to some pre-set criteria. For instance, for the sleep data the prominent frequency band is known for each sleep stage. Therefore, the IMF with maximum power in the desired frequency band can be selected and used in the SSA reconstruction stage. Then, the resulted component can help to mask the unwanted subspaces in the SSA reconstruction stage:

$$J(\mathbf{M}) = \|\mathbf{H} - \mathbf{R}\mathbf{M}\mathbf{\Lambda}^{1/2}\mathbf{V}^T\| \tag{9}$$

Where  $\|\cdot\|$  denotes Frobenius norm,  $\mathbf{R}$ ,  $\mathbf{\Lambda}$ , and  $\mathbf{V}$  are the SVD factors obtained in the first stage of SSA,  $\mathbf{H}$  is the Hankel matrix of component obtained from EMD, and  $\mathbf{M}$  is a diagonal matrix of adaptive weights:

$$\mathbf{M}_{k+1} = \mathbf{M}_k - \mu \text{diag}(\nabla_{\mathbf{M}_k} J) \tag{10}$$

where  $\text{diag}(\cdot)$  changes a vector into a gradient matrix. Using a simple gradient approach leads to the following recursive equation:

$$\mathbf{M}_{k+1} = \mathbf{M}_k - \mu \mathbf{R}\mathbf{M}_k\mathbf{\Lambda}^{1/2}\mathbf{V}^T(\mathbf{H} - \mathbf{R}\mathbf{M}_k\mathbf{\Lambda}^{1/2}\mathbf{V}^T)^T \tag{11}$$

where  $\mu$  is the step size which is set empirically.

**Proposed Sleep Scoring Method:** The proposed adaptive system can be seen in Figure 2. The aim is to train two RLS filters using SSA-EMD as a tool for estimating the desired signal at each time instant. In other words, by first using SSA-EMD, the desired signal is obtained and used for training the filters at each stage. After training, we have a set of weight vectors which can be used to classify the sleep EEG data into three stages. Therefore, the original data is separated to two parts. One part is used in the training stage and the second part for testing the system. After training, the whole set of weight vectors can be used to classify data in test stage. After passing test data to the resulted filters and obtaining the results from both filters, the signal's power is calculated and used for determining the segment stage.

### 3. Results

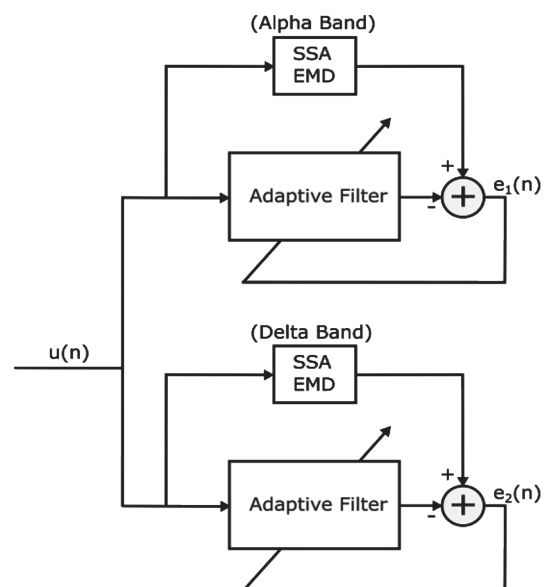
Thirty-six healthy men and women participated in two laboratory sessions, one involving a sleep extension protocol and the other a sleep restriction protocol. During each session polysomnography (PSG) measures were recorded at a sampling rate of 256Hz for a baseline (BL)

night (8 hours), seven condition nights (sleep extension (ES), 10 hours; sleep restriction (SR), 6 hours) and a recovery night (12 hours) following a period of total sleep deprivation. In this paper, the proposed method and also one commonly used methods for EEG sleep processing which uses Morlet wavelet were applied to sleep data to extract different frequency bands (alpha, theta and delta).

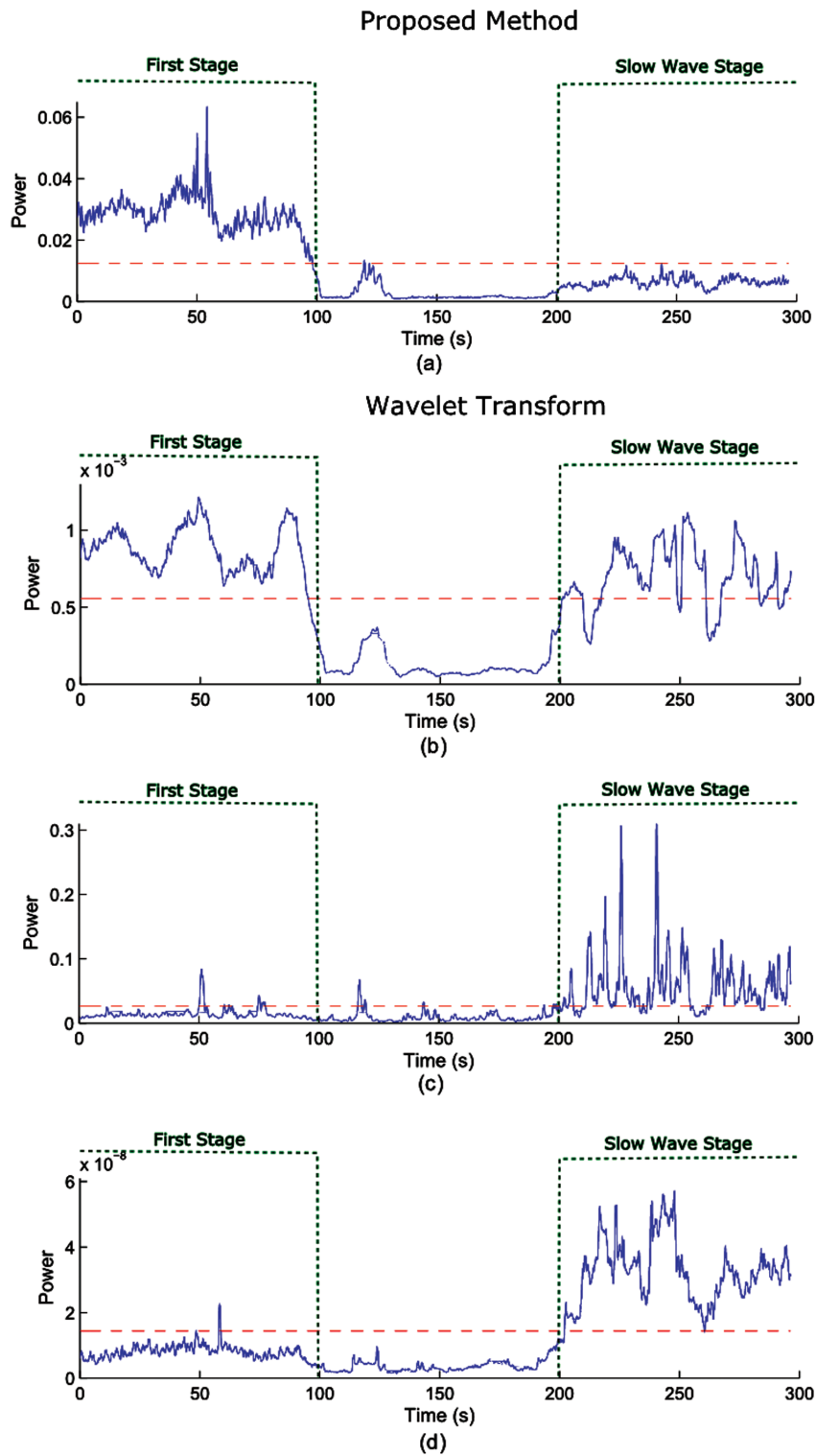
Delta rhythm is a slow brain wave which tends to have its highest amplitude during deep sleep in adults and is usually prominent frontally. Alpha can be seen in the posterior regions of the head on both sides and emerges with closing of the eyes and relaxation. Therefore, only one channel for which more variations in alpha and another one with more variations in delta were chosen.

The signals were segmented into two groups for training and testing the system. The resulting power spectrum obtained from one segment of the test data can be seen in Figure 3. The result for other available signals resembles the same structure.

The proposed method here attempts primarily to enhance and highlight the sleep-related indicators and features using SSA-EMD rather than feature extraction. In other words, we extracted a new signal using both wavelet and SSA-EMD. Then, we compared the results using adaptive filtering for classification. This is in line with the wavelet-based or other similar methods for decomposition of the signals to facilitate extraction of the relevant frequency bands or features.



**Figure 2.** Training stage of the proposed method. The desired signal is extracted from SSA-EMD for two specific frequency bands.



**Figure 3.** (a) and (b) show alpha variations using proposed and wavelet-based methods respectively and (c) and (d) show delta variations using proposed and wavelet-based methods respectively.

The main difference here is to use the processed time sequence followed by an adaptive filter, rather than some statistical features such as mean, std, and power in different frequency bands followed by non-adaptive

classifiers (support vector machine (SVM), etc.). Obviously, such statistical features are not suitable to be used by our proposed adaptive filter.



**Table 1.** Start of second and fourth (SW) stages for the examined subject obtained by manual scoring.

Start Stage	First	Second	SW
Time Segment	1	101	201

**Table 2.** Comparison between precision and recall of the proposed method and WT for real sleep data

Method	Performance	Precision		Recall	
		$\alpha$	$\delta$	$\alpha$	$\delta$
Proposed		0.99	0.91	0.99	0.99
WT		0.55	0.90	0.97	0.99
SVM		0.98	0.88	0.86	0.98

For the sake of overall comparison, we used SVM which is commonly used to classify sleep EEG stages. Here, the result of SSA-EMD was segmented and passed to SVM classifier. Using SVM, only the class of each segment can be determined and there is no adaptation opposed to the RLS method.

Table 1 shows the time points where we expect to see the stage changes. Moreover, the average precision (positive productive rate) and recall (sensitivity) rate using the proposed method, WT, and SVM for different data segments from different subjects can be seen in Table 2 where  $precision = tp / (tp + fp)$  and  $recall = tp / (tp + fn)$  and  $tp$ ,  $fp$ , and  $fn$  indicate true positive, false positive, and false negative respectively.

#### 4. Discussion

Sleep EEG data usually include several psychophysiological states each characterized by a different frequency band and the corresponding signals are corrupted by noise. SSA is a powerful subspace based technique but has problem in extracting the desired subspace automatically. At the same time, EMD automatically decomposes the signal into components sorted in terms of frequency and does not need any predefined parameter selection. SSA can perform well in a noisy environment. However, the original EMD cannot work the same in such situations. Ensemble EMD has been reported to have a better result for noisy data [18]. However, adjusting its parameter (noise ratio to be added to signal for cancelling the noise effect) is not easy. Additionally, its resulted components do not necessarily contain the desired components. A combination of these two however uses both time and frequency information of the data which can be more effectively used for sleep data scoring compared to the case where only SSA or EMD is used. On the other hand, AF techniques adjust their properties according to the desired features of the in-

put data. Therefore, they can be trained based on the data features and then used to classify different sleep stages. Simple filtering cannot produce the same results as the filter parameters need to be adaptive to the signal pattern, rather than frequency, variabilities. Moreover, RLS is used for tracking the input changes and train the system accordingly. Therefore, the convergence of the algorithm does mean that the changes in the input file have been tracked accurately. We empirically adjusted the RLS parameters to have good results.

The complexity of the proposed method is the maximum complexity of SSA and RLS stages. The complexity of both methods are  $O(n^2)$  where  $n$  is the signal length. The complexity of CWT is  $O(n)$ . However, for the classification stage again we need other classifiers such as RLS or SVM which have the same or higher complexity.

Moreover, In most of sleep data analysis applications such as ours, the processing is offline. Therefore, the small changes in complexity are well compromised by an increase in accuracy.

As indicated in Figure 3, the proposed method can determine the transitions between the stages of sleep by more accurately evaluating the alpha and delta (SW) brain activity variations. This achievement complies with an accurate manual scoring the sleep data by clinical experts. For alpha rhythm, the proposed method is able to better highlight the alpha before and in the first stage of the sleep and show alpha depression during deep sleep. The wavelet transform misclassifies sleep spindles which are waves within the 12-14 Hz band. This results in some dominant peaks over the slow wave duration. SSA can isolate the spindles and we can achieve better precision score compared to wavelet transform. In other words, using wavelet transform, we have false alpha detection during the slow wave stage.

## 5. Conclusions

Combination of EMD and SSA is an elegant approach to automatically identify the subspace of relatively narrowband signals in an eigen-based approach often used for single channel data decomposition and time series analysis, i.e. SSA. For classification purposes, such a structure can produce the necessary target signal for an adaptive filter which can be trained and later used as classifier. The overall result is indeed suitable for detection of sleep stages and also detecting other narrowband information from EEG signals often characterised by a sharp increase in delta and a rapid decrease in alpha as sleep deepens. Comparing the results with both manual scoring and those achieved using wavelet transform the effectiveness of the proposed method is verified. Understanding the sleep process in both normal and abnormal cases and various sleep scenarios such as sleep enforcement and deprivation however, requires further work. Accurate detection of spindles, k-complexes and the correlation between the sleep state and the changes in other physiological and biological human metrics are of significant importance

## References

- [1] T. P. Jung, S. Makeig, C. Humphries, T. W. Lee, M. J. Mckeown, V. Iragui, and T. J. Sejnowski, "Removing electroencephalographic artifacts by blind source separation," *Psychophysiology*, vol. 37, no. 2, pp. 163–178, 2000.
- [2] V. D. Calhoun, J. Liu, and T. Adali, "A review of group ICA for fMRI data and ICA for joint inference of imaging, genetic, and ERP data," *Neuroimage*, vol. 45, no. 1, pp. S163–S172, 2009.
- [3] M. De Vos, A. Vergult, L. De Lathauwer, W. De Clercq, S. Van Huffel, P. Dupont, A. Palmieri, and W. Van Paesschen, "Canonical decomposition of ictal scalp EEG reliably detects the seizure onset zone," *NeuroImage*, vol. 37, no. 3, pp. 844–854, 2007.
- [4] E. Niedermeyer, "Alpha rhythms as physiological and abnormal phenomena," *International Journal of Psychophysiology*, vol. 26, no. 1–3, pp. 31–49, 1997.
- [5] R. Mistlberger, B. Bergmann, and A. Rechtschaffen, "Relationships among wake episode lengths, contiguous sleep episode lengths, and electroencephalographic delta waves in rats with suprachiasmatic nuclei lesions," *Sleep*, vol. 10, no. 1, pp. 12–24, 1987.
- [6] S. Sanei, *Adaptive processing of brain signals*. John Wiley, 2013.
- [7] S. G. Mallat, "A theory for multiresolution signal decomposition: the wavelet representation," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 11, no. 7, pp. 674–693, 1989.
- [8] E. Sitnikova, E. H. Alexander, A. K. Alexey, and V. L. Gilles, "Sleep spindles and spikewave discharges in EEG: their generic features, similarities and distinctions disclosed with Fourier transform and continuous wavelet analysis," *Journal of neuroscience methods*, vol. 180, no. 2, pp. 304–316, 2009.
- [9] F. Ebrahimi, M. Mikaeili, E. Estrada, and H. Nazeran, "Automatic sleep stage classification based on EEG signals by using neural networks and wavelet packet coefficients," *EMBC2008*, 2008.
- [10] N. E. Huang, Z. Shen, S. R. Long, M. C. Wu, H. H. Shih, Q. Zheng, N. C. Yen, C. C. Tung, and H. H. Liu, "The empirical mode decomposition and the hilbert spectrum for nonlinear and non-stationary time series analysis," *Proceedings of the Royal Society of London. Series A: Mathematical, Physical and Engineering Sciences*, vol. 454, no. 1971, pp. 903–995, 1998.
- [11] N. Golyandina, V. Nekrutkin, and A. A. Zhigljavsky, *Analysis of time series structure: SSA and related techniques*. CRC Press, 2010.
- [12] S. Sanei, T. K.-M. Lee, and V. Abolghasemi, "A new adaptive line enhancer based on singular spectrum analysis," *IEEE Transactions on Biomedical Engineering*, vol. 59, no. 2, pp. 428–434, 2012.
- [13] F. Ghaderi, H. R. Mohseni, and S. Sanei, "Localizing heart sounds in respiratory signals using singular spectrum analysis," *IEEE Transactions on Biomedical Engineering*, vol. 58, no. 12, pp. 3360–3367, 2011.
- [14] T. Alexandrov and N. Golyandina, "The automatic extraction of time series trend and periodical components with the help of the caterpillar-SSA approach," *Exponenta Pro* 3, vol. 4, pp. 54–61, 2004.
- [15] S. Kouchaki and S. Sanei, "Supervised single channel source separation of electroencephalography signals," *MLSP 2013*, 2013.
- [16] S. S. Haykin, *Adaptive Filter Theory*, 4th edition. Pearson Education India, 2005.
- [17] B. Widrow and S. D. Stearns, *Adaptive signal processing*. IET, 1985.
- [18] Z. Wu and N. E. Huang, "Ensemble empirical mode decomposition: a noise-assisted data analysis method," *Advances in Adaptive Data Analysis*, vol. 1, no. 01, pp. 1–41, 2009.