Original article

Interictal EEG noise cancellation: GEVD and DSS based approaches versus ICA and DCCA based methods

S. Hajipour Sardouie [a,b,c,*], M.B. Shamsollahi [a], L. Albera [b,c,d], I. Merlet [b,c]

[a] BiSIPL, Sharif University of Technology, Tehran, Iran
[b] Inserm, UMR 1099, Rennes, F-35000, France
[c] LTSI, University of Rennes 1, Rennes, F-35000, France
[d] INRIA, Centre Inria Rennes-Bretagne Atlantique, 35042 Rennes Cedex, France

Abstract

Denoising is an important preprocessing stage in some ElectroEncephaloGraphy (EEG) applications. For this purpose, Blind Source Separation (BSS) methods, such as Independent Component Analysis (ICA) and Decorrelated and Colored Component Analysis (DCCA), are commonly used. Although ICA and DCCA-based methods are powerful tools to extract sources of interest, the procedure of eliminating the effect of sources of non-interest is usually manual. It should be noted that some methods for automatic selection of artifact sources after BSS methods exist, although they imply a training supervised step. On the other hand, in cases where there are some *a priori* information about the subspace of interest, semi-blind source separation methods can be used to denoise EEG signals. Among them the Generalized EigenValue Decomposition (GEVD) and Denoising Source Separation (DSS) are two well-known semi-blind frameworks that can be used with *a priori* information on the subspace of interest. In this paper, we compare the ICA and DCCA-based methods, namely CoM_2 and SOBI, respectively, with GEVD and DSS in the application of extracting the epileptic activity from noisy interictal EEG data. To extract *a priori* information required by GEVD and DSS, we propose a series of preprocessing stages including spike peak detection, extraction of exact time support of spikes and clustering of spikes involved in each source of interest. The comparison of these four methods in terms of performance and numerical complexity shows that CoM_2 give better performance for very low SNR values but require visual inspection to select the sources of interest. For higher SNR values, GEVD and DSS based approaches give similar results but with lower numerical complexity and without requiring a visual selection of the sources of interest.
© 2014 Published by Elsevier Masson SAS.

1. Introduction

In the process of recording brain activity, the signal of interest is usually contaminated by different activities arising from various sources of noise and artifacts. Some of these artifacts are externally generated such as power line noise and instrumentation noise produced by inappropriate electrode connections. In addition to externally generated artifacts, there is noise that is generated by physiological sources, external to the brain, such as eye blinks, eye movements, muscle activity and heart pulse. Moreover, in some applications, part of the brain activity itself such as EEG background activity is considered as a noise that should be removed. Therefore, denoising is an important preprocessing stage in some EEG applications such as source localization or Brain Computer Interface (BCI).

In this paper, we focus on noise cancellation of epileptic interictal EEG data where the main assumption is that the subspace of interest has a spike-like morphology. The subspace of non-interest includes the background EEG signals and the internal noises such as muscle activity. Recently, different methods have been applied to denoise EEG interictal data. Blind Source Separation (BSS) and Independent Component Analysis (ICA) are useful tools in the EEG signal processing applications including EEG denoising [1,2]. As mentioned in [3], only few ICA algorithms such as InfoMax [4] and FastICA [5] are used

nowadays to process biomedical signals. The authors in [3] examined fifteen ICA algorithms in terms of performance and computational complexity in the context of removal of muscle artifacts from interictal epileptiform activity. The results have shown that CoM_2 [6] offers the best compromise between performance and numerical complexity. In [7,8], in addition to ICA algorithms, some other denoising algorithms, such as Canonical Correlation Analysis (CCA) [9], Empirical Mode Decomposition (EMD) [10] and Wavelet-based approach were studied for removal of the muscular artifacts from surface EEG signals recorded in epileptic patients. This study showed that CCA and ICA approaches outperform other methods in denoising of interictal signals with moderate noise levels. Although BSS methods are powerful tools to extract sources of interest, the procedure of eliminating the effect of sources of non-interest is usually manual. It should be noted that some methods for automatic selection of artifact sources after BSS methods exist, although they imply a training supervised step [11,12].

In addition to completely-blind source separation methods, semi-blind approaches can be used when there is *a priori* information about the subspace of interest. The Generalized Eigen-Value Decomposition (GEVD) and Denoising Source Separation (DSS) [13] are two well-known frameworks that can be used as denoising methods. In this paper, we use the GEVD-based and DSS-based methods to denoise interictal EEG data and compare them with ICA and DCCA algorithms. Indeed, we want to test if *a priori* information on the spike-like subspace can be used in the denoising process. For EEG signals, it has been previously shown that different assumptions about sources of interest, such as spatial constraints [14], locations of known sources [15], shape and latency of the signal of interest [16] and time support of spikes [17], can be considered in semi-blind or constrained source separation methods. In this paper, we use the timing information of the epileptic interictal sources (i.e. the time samples corresponding to each epileptic source) and propose a series of preprocessing stages to extract the useful information to be used in the GEVD and DSS frameworks. These preprocessing stages consist of the detection and clustering of the epileptic spikes involved in each source of interest. It should be mentioned that a GEVD-based method was previously proposed in [17] in order to determine epileptic regions from epileptic intracerebral EEG signals. This method has a manual preprocessing stage to extract periods including interictal epileptiform discharges. On the contrary, in a GEVD-based method used in this paper to denoise interictal EEG data, all the preprocessing stages to extract the time samples corresponding to each epileptic source, including spike detection, extraction the exact time support of spikes and spike clustering, are fully automated. It should also be mentioned that both works have been done in parallel and independently. Some parts of our work were previously presented in [18].

Between the ICA and DCCA methods, we choose CoM_2 [6] and SOBI [19], respectively. The reason for selecting these algorithms is that the CoM_2 algorithm offers the best compromise between performance and numerical complexity as shown in [3] and SOBI is one of the most common DCCA algorithms used in biomedical applications.

The rest of the paper is organized as follows. First the problem formulation is presented. After that, the four denoising methods are introduced. Since the CoM_2 and SOBI algorithms have been extensively explained in the other articles, we provide a brief description of them. The proposed preprocessing stages for extracting *a priori* information on the epileptic subspace are explained in details and then the integration of this information into the GEVD and DSS based denoising methods is presented. The numerical complexity of the four denoising methods is also computed. In the next section, we compare the performance and numerical complexity of the four methods for denoising of both simulated and real interictal data. This section is followed by the conclusion.

In the following sections, bold uppercase letters (e.g. A) and bold lowercase letters (e.g. a) are used to denote matrices and vectors, respectively. The superscripts $\sharp$ and T stand for the pseudoinverse and the transpose, respectively.

2. Problem formulation

In the context of recording the epileptic interictal EEG activity, the recorded signal can be considered as a combination of (i) the signal of interest arising from the epileptic brain regions and (ii) the signal of non-interest composed of different types of noise, background activity and artifacts. We model the EEG signal recorded from N electrodes as one realization $\{x[k]\}$ of an N-dimensional random vector process $\{\tilde{x}[k]\}$. We can represent the observation process $\{\tilde{x}[k]\}$ as follows:

$$\tilde{x}[k] = \tilde{x}^{(e)}[k] + \tilde{x}^{(b)}[k] + \tilde{x}^{(m)}[k] + \tilde{v}[k]$$
$$\tilde{x}[k] = A^{(e)}\tilde{s}^{(e)}[k] + A^{(b)}\tilde{s}^{(b)}[k] + A^{(m)}\tilde{s}^{(m)}[k] + \tilde{v}[k] \qquad (1)$$

where $\{\tilde{s}^{(e)}[k]\}$, $\{\tilde{s}^{(b)}[k]\}$, $\{\tilde{s}^{(m)}[k]\}$ and $\{\tilde{v}[k]\}$ are the random vector processes representing the activity of P_e epileptic sources, P_b background sources, P_m muscular sources and the N-dimensional instrument noise, respectively. The mixing matrices $A^{(e)}$, $A^{(b)}$ and $A^{(m)}$ of size $(N \times P_e)$, $(N \times P_b)$ and $(N \times P_m)$, model the transfer function from all possible sources of activity to scalp electrodes. All these mixing matrices can be combined in a global mixing matrix A of size $(N \times P)$ where $P = P_e + P_b + P_m$. Therefore, (1) can be rewritten as follows:

$$\tilde{x}[k] = A\tilde{s}[k] + \tilde{v}[k] \qquad (2)$$

where $\{\tilde{s}[k]\}$ is the random vector process representing all sources.

As far as the statistical properties of vector random processes $\{\tilde{s}^{(e)}[k]\}$, $\{\tilde{s}^{(b)}[k]\}$, $\{\tilde{s}^{(m)}[k]\}$ and $\{\tilde{v}[k]\}$ are concerned, we can assume that they are independent as they correspond to different physiological/physical phenomena. Nevertheless, such an assumption is not valid within each vector random process regarding its components. In particular, some of epileptic activity sources may be statistically mutually dependent. We will then assume that, for every time index, the three vectors $\tilde{s}^{(e)}[k]$, $\tilde{s}^{(b)}[k]$ and $\tilde{s}^{(m)}[k]$ can be factorized as $\tilde{s}^{(e)}[k] = B^{(e)}\tilde{r}^{(e)}[k]$, $\tilde{s}^{(b)}[k] = B^{(b)}\tilde{r}^{(b)}[k]$ and $\tilde{s}^{(m)}[k] = B^{(m)}\tilde{r}^{(m)}[k]$, respectively, where $\tilde{r}[k] = [\tilde{r}^{(e)}[k]^\mathsf{T}, \tilde{r}^{(b)}[k]^\mathsf{T}, \tilde{r}^{(m)}[k]^\mathsf{T}]^\mathsf{T}$ is a P'-dimensional vector of mutually independent random variables such that $P' \leq N$ and where $B = [B^{(e)^\mathsf{T}}, B^{(b)^\mathsf{T}}, B^{(m)^\mathsf{T}}]^\mathsf{T}$

is a ($P \times P'$) matrix such that the product $\boldsymbol{G} = \boldsymbol{A}\,\boldsymbol{B}$ is a full column rank matrix. Eventually, the $\{\tilde{\boldsymbol{v}}[k]\}$ vector random process can be assumed to be Gaussian as most of instrument noises [3].

The aim of the proposed statistical approaches, is to extract the signal of interest, $\{\boldsymbol{x}^{(e)}[k]\}$, which represents one realization of the epileptic vector random process $\{\tilde{\boldsymbol{x}}^{(e)}[k]\}$, from $\{\boldsymbol{x}[k]\}$. Indeed, by computing from $\{\boldsymbol{x}[k]\}$ the corresponding realization, $\{\boldsymbol{r}^{(e)}[k]\}$, of $\{\tilde{\boldsymbol{r}}^{(e)}[k]\}$ and the corresponding mixing matrix $\boldsymbol{G}^{(e)} = \boldsymbol{A}^{(e)}\,\boldsymbol{B}^{(e)}$, the N-dimensional signal $\{\boldsymbol{x}^{(e)}[k]\}$ given by $\boldsymbol{x}^{(e)}[k] = \boldsymbol{G}^{(e)}\boldsymbol{r}^{(e)}[k]$ will represent the denoised EEG data.

3. Methodology

3.1. The ICA and DCCA-based algorithms

The ICA and DCCA problems consist in retrieving unobserved sources $\tilde{\boldsymbol{r}}[k] = [\tilde{r}_1[k], \ldots, \tilde{r}_P[k]]^{\mathsf{T}}$ from observed mixtures $\tilde{\boldsymbol{x}}[k] = [\tilde{x}_1[k], \ldots, \tilde{x}_N[k]]^{\mathsf{T}}$ which can linearly be modeled as given by $\tilde{\boldsymbol{x}}[k] = \boldsymbol{G}\tilde{\boldsymbol{r}}[k]$. The fundamental assumption of ICA is that the unknown sources are statistically independent. The ICA problem can also be stated as finding the left inverse of the mixing matrix $\boldsymbol{G}$. On the other hand, the DCCA methods use the assumption of spatial decorrelation and time color of the sources to extract them. There exist a lot of different ICA and DCCA algorithms used for different applications such as biomedical signal processing, communications, speech processing, etc. In this paper, we use a well-known ICA method, namely CoM$_2$. CoM$_2$ relies on the maximization of a contrast function derived from Fourth-Order (FO) cumulants of the observations [6]. On the other hand, SOBI is a DCCA method that uses Second-Order (SO) cumulants to extract the sources [19]. By applying the ICA or DCCA-based algorithms on the observation signals, the estimate of the mixing matrix and P' sources including both sources of interest and non-interest are obtained. P'_e components are visually selected as epileptic sources. Then, by using the estimate of the mixing matrix, the epileptic subspace corresponding to each epileptic source can be obtained. Note that for the ICA and DCCA-based methods as well as the GEVD and DSS ones the number of epileptic sources has to be known and it is assumed that it has been already estimated. Estimating the number of sources of interest is not in the scope of this paper. It is by itself a signal processing problem which was addressed in other researches [20,21].

3.2. The GEVD/DSS-based methods

In order to denoise interictal EEG signals, we show how the GEVD/DSS approach can be used through two main stages: (1) a preprocessing stage and (2) the subspace decomposition stage. In the preprocessing stage, some *a priori* knowledge about the occurrence times of spikes (spike time samples involved in each epileptic source) are extracted. This information is used in the GEVD or DSS method to separate the epileptic subspace from the noise subspace. Finally, the estimated denoised signal is achieved by reconstructing the observation signals only from estimated epileptic sources. The preprocessing stage has three parts. We first detect the spike peak times.

Then the exact time support of spikes is identified. After that, we cluster the extracted spikes into P'_e groups corresponding to P'_e epileptic interictal sources. In the rest of this section, these three stages are explained in detail and the GEVD and DSS-based methods to denoise interictal data are then studied.

3.2.1. Spike detection

In the context of EEG signal processing, various methods have been proposed for spike detection [22–28]. These methods may use the information obtained from a single channel [22] or from several (multi-) channels [28,25]. They may also be categorized by the features they used: morphological features [28] or time–frequency ones [22]. Most of the spike detection methods have an enhancement stage that generates an output signal in which the distinction between the spikes and the noise is increased by some filtering methods such as Wavelet Transform [29,27], matched filters [30] or Kalman filter [31]. At this stage, the output signal is used in a decision procedure in order to extract the spike peak times. This decision procedure may be implemented by a simple thresholding algorithm [26] or may have a more complex structure [22].

In this paper, we use a three-stage multi-channel spike detection method that is a generalized version of the previously proposed single-channel spike detection method [22,23]. This single-channel wavelet-based spike detection method consists of two stages (S_1 and S_2) that detect the candidates for spike peak times of each channel of observation (Fig. 1). In [23], it was shown that this wavelet-based algorithm has a better performance than other spike detection algorithms. To generalize this method to multi-channel detection, we add a third stage (S_3) in which a clustering algorithm combines the extracted information from all channels to detect the spike peak times. Fig. 1 shows the three stages of the algorithm.

The detailed procedure of this three-stage multi-channel spike detection method can be described as follows. First, note that the signal of interest, the spikes in our application, are contaminated by piece-wise stationary signals, such as background EEG and measurement noise as well as by transient signals such as muscular artifacts. Therefore, this algorithm eliminates the signals of non-interest in two successive stages: the first stage (S_1) separates transient signals (including the spikes) from the piece-wise stationary signals and the second stage (S_2) tries to separate the spikes from muscular artifacts (rejection of false alarms). To this end, two measures $T_k^{(n)}$ and $G_k^{(n)}$ are calculated for each channel of observation $\{x_n[k]\}$ and for each time sample k. Then by using the appropriate thresholding processes, the candidates for spike peak times are obtained for each channel. The detailed procedure of these two stages is presented in [22, 23] and is also given in Appendix A.

Up to this point, we estimate the candidates for the spike peak times for each observation channel. Since there may be false detections and since the peak time might be misestimated by a few samples, a decision algorithm has to be performed at a stage S_3 by using the information of all channels. We use a hierarchical clustering method [32] to group all possible candidates for spike peaks of all channels. The hierarchical grouping starts by defining a number N_s of clusters equal to the total number
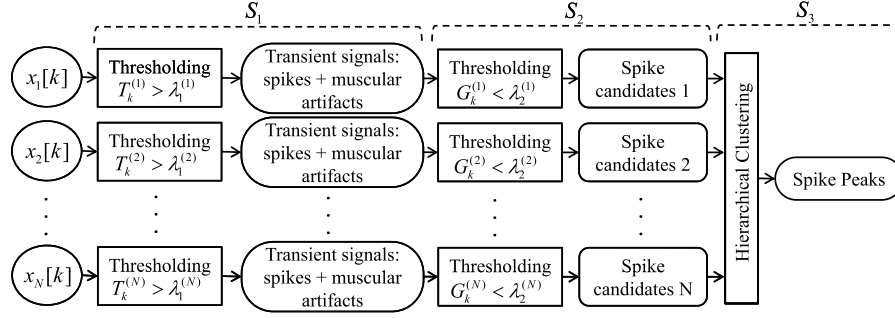
Fig. 1. The spike peak detection algorithm.

of candidates. In the first grouping step, the number of clusters is reduced to $N_s - 1$ by combining the two nearest spike peak times. In successive stages the newly formed groups are linked to other spikes or spike groups with the lowest distance (i.e. the distance between the centers of the clusters). In each step we save the center and the maximum intra-distance, i.e. the distance between the individuals in one group, for each cluster. This procedure is repeated until the maximum intra-distance of all clusters is less than a given threshold Th_{spike} chosen as the length of a spike. After extracting all clusters, we select only the clusters for which the population is more than N_{eff} (i.e. the minimum number of channels which are affected by one spike) and accordingly the centers of these clusters are the estimated peak times.

3.2.2. Extraction of the exact time support of the spikes

Only a few studies have worked on extracting the time support of a spike. In most applications, such as extracting the morphological features, only the spike peak times and the distance between adjacent peaks are calculated. To extract the important morphological points, the curvature features based on the second derivative are usually used. In this paper, we use a curvature measure proposed in [33] to extract the time support of the spikes. First, in order to eliminate the ineffective channels, i.e. channels with little spike characteristics, we calculate a measure of spikiness based on the fourth standardized moment of each observation channel as follows:

$$Spikiness\big(\{\tilde{x}_n[k]\}\big) = \frac{\langle E[(\tilde{x}_n[k] - E[\tilde{x}_n[k]])^4]\rangle}{\langle E[(\tilde{x}_n[k] - E[\tilde{x}_n[k]])^2]\rangle^2} \qquad (3)$$

where $E[.]$ and $\langle.\rangle$ denote the mathematical expectation and temporal mean operators, respectively. We then eliminate the first N_{ue} channels with minimum spikiness. The remaining channels are called "epileptic channels" in the following. Then we consider a symmetric window around each peak time (extracted in the previous step) and calculate the average value of the epileptic channels for all samples in the defined window. Then each signal is low pass filtered with a moving average filter of length $L_{ma,1}$. Up to this point, for each spike we have a smooth signal around the spike peak time. To find the start and end samples of each spike, we define a curvature measure as follows:

$$Curv_x[k] = \sum_{i=1}^{5} \cos\big(\theta_i^{(k)}\big) \qquad (4)$$
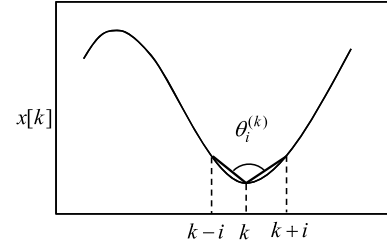
Fig. 2. The angle θ_i used in the curvature definition.

where $\theta_i^{(k)}$ is the angle between two line segments that connect the points $(k - i, x[k - i])$ and $(k + i, x[k + i])$ to $(k, x[k])$ as shown in Fig. 2. Then we use Th_θ as a threshold to define the start and end points of the spike. Since the data are noisy, we must ensure that the curvature of at least N_θ consecutive points are greater than Th_θ. Thus, the start point of a spike is chosen as a first sample where its curvature and the curvature of the $N_\theta - 1$ following samples are greater than Th_θ. Similarly, the end point of a spike is chosen as a last sample where its curvature and the curvature of the $N_\theta - 1$ previous samples are greater than Th_θ.

3.2.3. Spike clustering

In this stage, the spikes extracted in the previous stages are clustered into P'_e clusters corresponding to P'_e epileptic sources. The morphological features of spikes (including the peak location, the zero-crossing sample, slope and ...) originating from different epileptic sources are slightly different. These features may be used to cluster the spikes according to their corresponding epileptic sources. In this paper, we use the ratio between the spike amplitudes on all electrodes at two main peak times. Since the location of epileptic sources in the brain has a direct impact on the spike amplitudes of different channels, this vector of features can effectively cluster the spikes. Since for the simulated data of this paper we use biphasic spikes, we consider two peaks for each spike in this section. To extract this vector of features for each spike, we do the following procedure on each spike truncated at its extracted time support. These truncated spikes are represented by $\{x^{(\ell)}[k_\ell]\}$ with $1 \le \ell \le N_s$ where N_s is the number of spikes. So, we should first detect the occurrence time of two main peaks as shown in Fig. 3. Note that the spike peak times extracted in the first preprocessing stage may match one of these peaks, but in this stage we determine all of the peak times. To this end, for each spike, we calculate
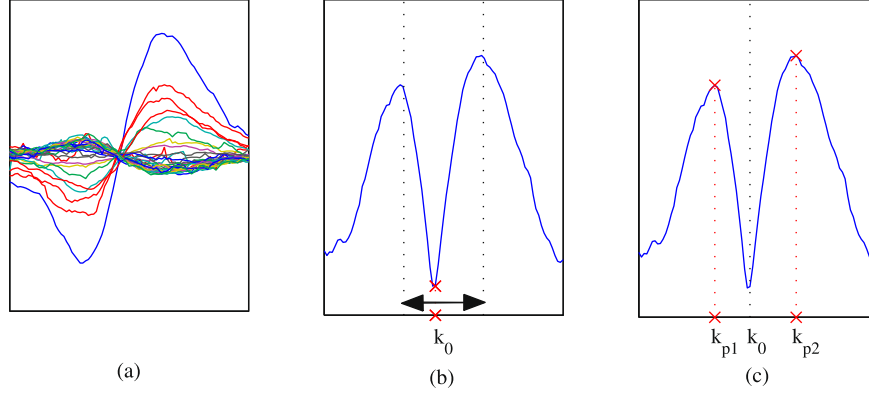
Fig. 3. The procedure of feature extraction from one spike: (a) the low-pass filtered signals of all channels, (b) extracting the zero-crossing sample of $\overline{|x^{(\ell)}|}[k_\ell]$, (c) extracting the occurrence time of two negative and positive peaks.

the moving average low pass filtered signal (with the filter of length $L_{ma,2}$) of each channel as shown in Fig. 3(a). Then, we calculate the average of absolute values of amplitudes of all channels in each time sample, represented by $\{\overline{|x^{(\ell)}|}[k_\ell]\}$ as shown in Fig. 3(b). Then, the time of zero-crossing in approximately middle of the signal is determined. To find this point, the sample with minimum value in the middle third of the signal is extracted (k_0) as shown in Fig. 3(b). After that, for $\{\overline{|x^{(\ell)}|}[k_\ell]\}$, we consider two segments: one from the start point to k_0 and the other one from k_0 to end, and for each segment we find the point with maximum value as shown in Fig. 3(c). These two arguments of maximum, $k_{p1}^{(\ell)}$ and $k_{p2}^{(\ell)}$, correspond to the occurrence time of the negative and positive peaks of the spike. Then we use the values of observations of different channels in the time samples $k_{p1}^{(\ell)}$ and $k_{p2}^{(\ell)}$ to generate a normalized vector of features for each spike as follows:

$$\forall 1 \le \ell \le N_s,$$
$$f[\ell] = \left[\frac{x[k_{p1}^{(\ell)}]^\mathsf{T}}{\sqrt{\sum_{n=1}^{N} x_n[k_{p1}^{(\ell)}]}}, \frac{x[k_{p2}^{(\ell)}]^\mathsf{T}}{\sqrt{\sum_{n=1}^{N} x_n[k_{p2}^{(\ell)}]}} \right]^\mathsf{T} \quad (5)$$

Then by using the feature vectors of all spikes, $f[\ell]$ for $\ell \in \{1, \ldots, N_s\}$, and the Fuzzy C-Means (FCM) clustering algorithm [34], the spikes are classified into P'_e groups corresponding to P'_e epileptic sources. We represent the cluster containing all time samples corresponding to the p-th epileptic source by $cluster(p)$ and the sets $T_1^{(p)}$ and $T_2^{(p)}$ are defined as follows:

$$T_1^{(p)} = \{k : k \in cluster(p)\} \quad (6)$$
$$T_2^{(p)} = \{k : k \notin cluster(p)\} \quad (7)$$

3.2.4. The GEVD-based denoising algorithm

Let us assume that the vector random process $\{\tilde{x}[k]\}$ can be approximately decomposed as a linear combination of decorrelated random processes, say as $\tilde{x}[k] = G\tilde{r}[k]$ where P'_e components of $\{\tilde{r}[k]\}$ correspond to the components of $\{\tilde{r}^{(e)}[k]\}$, i.e. they span the epileptic source subspace, and the other $P' - P'_e$ components span the noise subspace. To extract these P'_e epileptic components as well as their corresponding columns in the
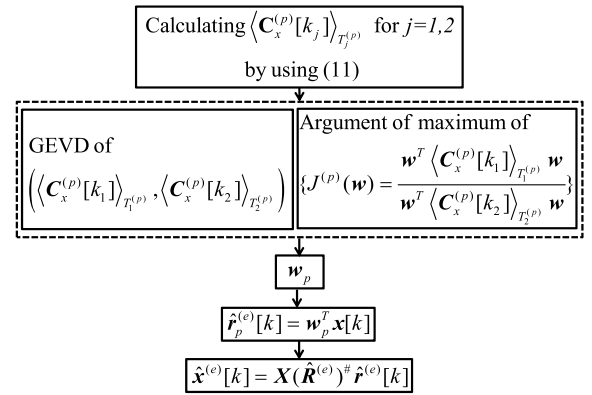

Fig. 4. The GEVD framework: the procedure of extracting the epileptic subspace.

mixing matrix G, we use a GEVD-based subspace decomposition algorithm. In this algorithm, we extract each epileptic component separately and then reconstruct the denoised observation signal corresponding to epileptic sources. The whole procedure to extract the epileptic subspace is shown in Fig. 4. For the epileptic source $p \in \{1, \ldots, P'_e\}$ and for each time sample, we define two types of covariance matrices of the random vector $\{\tilde{x}[k]\}$ as follows:

$$\forall k_1 \in T_1^{(p)},$$
$$C_x^{(p)}[k_1] = \mathrm{E}\left[\left(\tilde{x}[k_1] - \mathrm{E}[\tilde{x}[k_1]]\right)\left(\tilde{x}[k_1] - \mathrm{E}[\tilde{x}[k_1]]\right)^\mathsf{T}\right]$$
$$= G C_r^{(p)}[k_1] G^\mathsf{T} \quad (8)$$
$$\forall k_2 \in T_2^{(p)},$$
$$C_x^{(p)}[k_2] = \mathrm{E}\left[\left(\tilde{x}[k_2] - \mathrm{E}[\tilde{x}[k_2]]\right)\left(\tilde{x}[k_2] - \mathrm{E}[\tilde{x}[k_2]]\right)^\mathsf{T}\right]$$
$$= G C_r^{(p)}[k_2] G^\mathsf{T} \quad (9)$$

where the sets $T_1^{(p)}$ and $T_2^{(p)}$ are defined in (6) and (7) and the matrices $C_r^{(p)}[k_1]$ and $C_r^{(p)}[k_2]$ are the diagonal covariance matrices of the source vector $\{\tilde{r}[k]\}$ at two time samples $k_1 \in T_1^{(p)}$ and $k_2 \in T_2^{(p)}$, respectively.

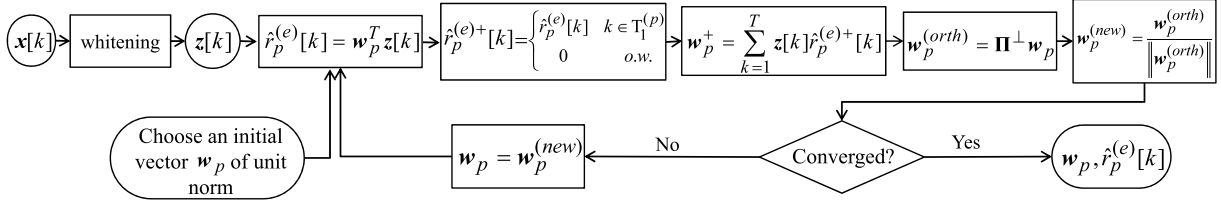To identify the p-th epileptic component $\{r_p^{(e)}[k]\}$, we maximize the Rayleigh quotient defined as follows:

Fig. 5. The DSS framework: this framework used to extract one epileptic source and its corresponding epileptic subspace.

$$J^{(p)}(\boldsymbol{w}) = \frac{\boldsymbol{w}^{\mathsf{T}} \langle \boldsymbol{C}_x^{(p)}[k_1] \rangle_{T_1^{(p)}} \boldsymbol{w}}{\boldsymbol{w}^{\mathsf{T}} \langle \boldsymbol{C}_x^{(p)}[k_2] \rangle_{T_2^{(p)}} \boldsymbol{w}} \tag{10}$$

where $\langle \boldsymbol{C}_x^{(p)}[k_j] \rangle_{T_j^{(p)}}$ is the time average of the covariance matrices $\boldsymbol{C}_x^{(p)}[k_j]$ on the interval $T_j^{(p)}$ and $\boldsymbol{w}$ is the estimated separator. In practice and with ergodicity hypothesis, this average covariance matrix can be estimated from a realization $\{\boldsymbol{x}[k]\}_{k \in T_j^{(p)}}$ of the vector random process $\{\tilde{\boldsymbol{x}}[k]\}_{k \in T_j^{(p)}}$ as follows:

$$\forall j \in \{1, 2\},$$

$$\langle \boldsymbol{C}_x^{(p)}[k_j] \rangle_{T_j^{(p)}} \approx \frac{1}{n(T_j^{(p)})} \sum_{k \in T_j^{(p)}} \left(\{\boldsymbol{x}[k]\} - \langle \{\boldsymbol{x}[k]\} \rangle_{T_j^{(p)}} \right)$$

$$\times \left(\{\boldsymbol{x}[k]\} - \langle \{\boldsymbol{x}[k]\} \rangle_{T_j^{(p)}} \right)^{\mathsf{T}} \tag{11}$$

where $n(T_j^{(p)})$ represents the number of samples in the set $T_j^{(p)}$ and $\langle \boldsymbol{x}[k] \rangle_{T_j^{(p)}} = \frac{1}{n(T_j^{(p)})} \sum_{k \in T_j^{(p)}} \boldsymbol{x}[k]$. Finding the argument of the maximum of (10) is equivalent to jointly diagonalize the matrices $\langle \boldsymbol{C}_x^{(p)}[k_1] \rangle_{T_1^{(p)}}$ and $\langle \boldsymbol{C}_x^{(p)}[k_2] \rangle_{T_2^{(p)}}$. Consequently, the separator vector can be found by solving a problem of joint diagonalization by congruence. To this end, the GEVD of the pair of matrices $\langle \boldsymbol{C}_x^{(p)}[k_1] \rangle_{T_1^{(p)}}$ and $\langle \boldsymbol{C}_x^{(p)}[k_2] \rangle_{T_2^{(p)}}$ is solved in order to maximize (10). More particularly, the vector $\boldsymbol{w}_p$ maximizing $J^{(p)}$ (10) is computing as the eigenvector associated with the largest eigenvalue of matrix $\langle \boldsymbol{C}_x^{(p)}[k_2] \rangle_{T_2^{(p)}}^{-1} \langle \boldsymbol{C}_x^{(p)}[k_1] \rangle_{T_1^{(p)}}$. An estimate of $\{r_p^{(e)}[k]\}$ is then given by $\{\hat{r}_p^{(e)}[k]\}$ where $\hat{r}_p^{(e)}[k] = \boldsymbol{w}_p^{\mathsf{T}} \boldsymbol{x}[k]$ for any time index k. Once the estimate $\{\hat{r}_p^{(e)}[k]\}$ of the P_e' epileptic components $\{r_p^{(e)}[k]\}$ has been computed, the estimated denoised EEG signal $\{\hat{\boldsymbol{x}}^{(e)}[k]\}$ is achieved by reconstructing the observation signals only from the estimated epileptic subspace as follows:

$$\hat{\boldsymbol{x}}^{(e)}[k] = \hat{\boldsymbol{G}}^{(e)} \hat{\boldsymbol{r}}^{(e)}[k] \tag{12}$$

where $\hat{\boldsymbol{r}}^{(e)}[k] = [\hat{r}_1^{(e)}[k], ..., \hat{r}_{P_e'}^{(e)}[k]]^{\mathsf{T}}$, $\hat{\boldsymbol{G}}^{(e)} = \boldsymbol{X}(\hat{\boldsymbol{R}}^{(e)})^{\sharp}$ and where $\boldsymbol{X}$ and $\hat{\boldsymbol{R}}^{(e)}$ are the $(N \times T)$ and $(P_e' \times T)$ matrices standing for T samples of the N-dimensional signal $\{\boldsymbol{x}[k]\}$ and T samples of the P_e'-dimensional reconstructed signal $\{\hat{\boldsymbol{r}}^{(e)}[k]\}$, respectively, with the assumption that $T \geq N$. Therefore, $\{\hat{\boldsymbol{x}}^{(e)}[k]\}$ contains essentially the contribution of the epileptic activity on the scalp electrodes.

3.2.5. The DSS-based denoising algorithm

The DSS [13] method is an almost recent framework which can be used to design new source separation algorithms. This framework can be optimized to generate a wide range of source separation algorithms, from completely to partially blind methods to solve specific problems. In this framework, the source separation algorithms are generated around the denoising methods, such that various kinds of prior knowledge are formulated in terms of denoising [13]. The DSS framework is based on the Expectation–Maximization (EM) algorithm [35].

In this paper, as mentioned before, we use the information extracted from the preprocessing stages of the algorithm to design an appropriate denoising method in the DSS framework. The model $\tilde{\boldsymbol{x}}[k] = \boldsymbol{G}\tilde{\boldsymbol{r}}[k] + \tilde{\boldsymbol{v}}[k]$ for the observation is considered. The flowchart of the proposed DSS algorithm to extract one epileptic source is shown in Fig. 5. In this flowchart, $\{\hat{r}_p[k]\}$ and the vector $\boldsymbol{w}_p$ represent the p-th extracted source and its corresponding separator, respectively.

As shown in Fig. 5, the first step of the algorithm is a whitening of the observations $\{\boldsymbol{x}[k]\}$. The whitened signal $\{\boldsymbol{z}[k]\}$ is given by $\boldsymbol{z}[k] = \boldsymbol{\Theta}^{\sharp} \boldsymbol{x}[k]$ where $\boldsymbol{\Theta}$ is a square root of the covariance matrix of the noiseless observation vector process $\{\tilde{\boldsymbol{x}}[k] - \tilde{\boldsymbol{v}}[k]\}$. By means of this whitening procedure, the matrix $\boldsymbol{G}$ is transformed into an orthogonal matrix $\boldsymbol{W}$ of size $(P' \times P')$. The five consequent steps are the main steps of the DSS method which are repeated in an iterative procedure. First, by using a random initial value of separator vector $\boldsymbol{w}_p$ of unit norm of size $(P' \times 1)$, a noisy estimate of the p-th source is calculated as follows:

$$\forall k, \quad \hat{r}_p^{(e)}[k] = \boldsymbol{w}_p^{\mathsf{T}} \boldsymbol{z}[k] \tag{13}$$

The next step, called the denoising step, is the main stage of the algorithm. In this stage, by using the pre-obtained information, the current estimated noisy source is modified or denoised. The general form of this stage is given by $\hat{r}_p^{(e)+}[k] = f(\hat{r}_p^{(e)}[k])$. In the DSS algorithm used in this paper, the denoised source $\{\hat{r}_p^{(e)+}[k]\}$ is produced from the noisy source $\{\hat{r}_p^{(e)}[k]\}$ as follows:

$$\hat{r}_p^{(e)+}[k] = \begin{cases} \hat{r}_p^{(e)}[k] & \text{if } k \in T_1^{(p)} \\ 0 & \text{otherwise} \end{cases} \tag{14}$$

where $T_1^{(p)}$ is the set containing the time samples corresponding to the p-th source as defined in (6).

In the third, fourth and fifth steps, by using the denoised source $\{\hat{r}_p^{(e)+}[k]\}$, a new estimation of the separator vector is calculated and normalized as follows:

Table 1
The numerical complexity of each step of the GEVD and DSS methods.

	Multiplications	Comparisons
Spike detection	$NT(MF_w + 12) + 2(N-1)N_s$	$2T\log(T) + NN_s\log(NN_s) + \frac{1}{2}(N-1)N_s(N(N_s+1)-3) + N_s$
Extraction of the exact time support	$N(2T+2) + 44N_sL_w$	N_sL_w
Spike clustering	$N(L_{ma,2}T_s + 2N_sP_e'^2)$	$\frac{4}{3}T_s$
The GEVD algorithm	$\frac{91}{3}N^3P_e' + N^2P_e'(T+T_s+1)$	
The DSS algorithm	$B + P_e'(N^2T + NT + N) + \sum_{p=1}^{P_e'} It_p(2NT+N)$	

Table 2
The numerical complexity of the denoising algorithms.

	Numerical complexity
GEVD ($P_e'=1$)	$91N^3/3 + N^2(T+T_S+1) + 44N_sL_w + N(14T + MF_wT + 2 + 2N_s) - 2N_s$
GEVD ($P_e'\neq 1$)	$91N^3P_e'/3 + N^2P_e'(T+T_S+1) + 44N_sL_w + N(14T + MF_wT + L_{ma,2}T_s + 2 + 2N_s(1+P_e'^2)) - 2N_s$
DSS ($P_e'=1$)	$B + It_1(2NT+N) + N(16T + MF_wT + 4 + 2N_s) + 44N_sL_w - 2N_s$
DSS ($P_e'\neq 1$)	$B + N^2TP_e' + \sum_{p=1}^{P_e'} It_p(2NT+N) + 44N_sL_w - 2N_s + N(15T + MF_wT + P_e'(T+1) + L_{ma,2}T_s + 3 + 2N_s(1+P_e'^2))$
SOBI	$B + KTN^2/2 + 4N^3/3 + (K-1)N^3/2 + It_s P'^2[4P'(K-1) + 17(K-1) + 4P' + 75]/2$
CoM$_2$	$B + \min(12It_c f_4(P')P'^2 + 2It_cP'^3 + 3Tf_4(P') + TP'^2, 13It_cTP'^2/2) + It_cP'^2Q/2$

$$\boldsymbol{w}_p^+ = \sum_{k=1}^{T} \boldsymbol{z}[k]\hat{r}_p^{(e)+}[k] \qquad (15)$$

$$\boldsymbol{w}_p^{(orth)} = \boldsymbol{\Pi}^\perp \boldsymbol{w}_p^+ \qquad (16)$$

$$\boldsymbol{w}_p^{(new)} = \frac{\boldsymbol{w}_p^{(orth)}}{\|\boldsymbol{w}_p^{(orth)}\|} \qquad (17)$$

where $\boldsymbol{\Pi}^\perp$ is defined by:

$$\boldsymbol{\Pi}^\perp = \boldsymbol{I} - \boldsymbol{V}(\boldsymbol{V}^\mathsf{T}\boldsymbol{V})^{-1}\boldsymbol{V}^\mathsf{T} = \boldsymbol{I} - \boldsymbol{V}\boldsymbol{V}^\mathsf{T}$$
$$\text{with } \boldsymbol{V} = [\boldsymbol{w}_1, \dots, \boldsymbol{w}_{p-1}] \qquad (18)$$

It is clear that the matrix $\boldsymbol{\Pi}^\perp$ represents the orthogonal projector onto the subspace orthogonal to the previously extracted sources, so by multiplying this matrix by $\boldsymbol{w}_p^+$, the effect of previously extracted sources is removed. By using this deflation step in the DSS framework, the convergence to previously extracted sources is prevented [13]. Then, the present estimated separator vector $\boldsymbol{w}_p$ is passed to the first block and this procedure continues until convergence. The outputs of the algorithm are $\boldsymbol{w}_p$ and $\hat{r}_p^{(e)}[k] = \boldsymbol{w}_p^\mathsf{T}\boldsymbol{z}[k]$. By using the same procedure, all epileptic sources are extracted. The estimated denoised EEG signal can then be achieved as follows:

$$\forall k, \quad \hat{\boldsymbol{x}}^{(e)}[k] = \hat{\boldsymbol{G}}^{(e)}\hat{\boldsymbol{r}}^{(e)}[k] \qquad (19)$$

where $\hat{\boldsymbol{G}}^{(e)} = \boldsymbol{\Theta}\hat{\boldsymbol{W}}^{(e)}$ with $\hat{\boldsymbol{W}}^{(e)} = [\boldsymbol{w}_1, ..., \boldsymbol{w}_{P_e'}]$.

3.3. Numerical complexity

In this section, we analyze the numerical complexity of the four algorithms which will be expressed in flops. A flop corresponds to a multiplication followed by an addition, but in practice the number of multiplications is computed because they are more expensive than additions [36]. In this section, the total numerical complexity of the GEVD-based and DSS-based

methods is computed and compared with the numerical complexity of CoM$_2$ [6] and SOBI [19]. Note that the results of this section will be used later in the section "Results" in order to compare the denoising algorithms.

The numerical complexity of each step of the GEVD-based and DSS-based methods is shown in Table 1. Table 2 shows the numerical complexity of the four denoising algorithms expressed in flops. In these tables, N, T and P_e' are the number of channels, time samples and epileptic sources, respectively. M and F_w are the number of wavelet decompositions and the length of discrete mother wavelet in spike detection algorithm. N_s is the total number of spikes in all epileptic sources and T_s is the total number of samples in the spike durations. L_w is the length of window used in extracting the exact time support of the spikes. The length of the moving average filters used in Sections 3.2.2 and 3.2.3 is denoted by $L_{ma,1}$ and $L_{ma,2}$, respectively. It_p is the number of iterations in the DSS algorithm to extract the p-th epileptic source. $B = \min\{TN^2/2 + 4N^3/3 + P'NT, 2TN^2\}$ is the number of flops required to perform spatial whitening. For CoM$_2$, $f_4(P') = P'(P'+1)(P'+2)(P'+3)/24$ is equal to the number of free entries in a fourth order cumulant tensor of dimension P' enjoying all symmetries. It_s and It_c are the number of sweeps required by a joint diagonalization process in SOBI and by contrast function optimization algorithms (CoM$_2$), respectively. K is the number of delay lags used in SOBI. Q is the complexity required to compute the roots of a real 4th degree polynomial by Ferrari's technique (we may take $Q \approx 30$ flops) [36,3].

Since the comparison is much simpler than multiplications, we consider only the number of multiplications for calculating the required flops. It should be noted that for the GEVD-based and DSS-based methods, we separately show the numerical complexities of two cases (1) $P_e' = 1$ and (2) $P_e' \neq 1$, because for the case of $P_e' = 1$ the spike clustering stage is discarded.

4. Results

4.1. Simulated data generation

The simulated EEG data are generated using a realistic model developed in our team [37,38]. We build a mesh of the cortical surface from a 3D MRI T1 image of a subject (Brain-Visa, SHFJ, Orsay, France). This mesh is composed of 40 500 triangles of mean surface 5 mm^2. A current dipole is placed at the barycenter of each triangle and oriented orthogonally to the triangle surface, leading to a field of current dipoles. From this mesh, P distributed sources, called "patches", generating interictal spikes, are defined. Each patch is composed of 100 dipole sources to which we have assigned hyper-synchronous spike-like activities generated from a model of neuronal populations [37]. From this setup, and considering 32 electrodes in the cortical region, the forward problem is then solved using a realistic head model made of three nested homogeneous volumes shaping the brain, the skull and the scalp (Boundary Element Method, ASA, ANT, Enschede, Netherlands). The epileptic activity on the electrodes, namely the signal of interest, is then calculated by using the forward model and the epileptic patches. On the other hand the signal of non-interest, composed of muscle activity, background EEG and instrument noise, is extracted from a real EEG data. Finally these real EEG data are added to the generated epileptic activity with a specified Signal-to-Noise Ratio (SNR) defined as follows:

$$\text{SNR} = 10\log_{10}\left(\frac{\sum_{k=1}^{T} ||\boldsymbol{x}^{(e)}[k]||^2}{\sum_{k=1}^{T} ||\boldsymbol{x}^{(b)}[k] + \boldsymbol{x}^{(m)}[k] + \boldsymbol{v}[k]||^2}\right) \quad (20)$$

We considered three scenarios that can account for some of the interesting source configurations encountered in partial epilepsy. So, we ran three different tests: one with $P_e = 1$ and two with $P_e = 2$. For each test, we change the values of SNR in the range -15 dB to 5 dB. For each SNR, 50 trials of the simulated EEG are generated. In the first scenario, we consider one single patch localized in the superior temporal gyrus in the left hemisphere as shown in Fig. 6(a). In the second scenario, two patches are considered, the first patch is located in the superior temporal gyrus and the second patch is in the inferior frontal gyrus as shown in Fig. 6(b). In this scenario, these two patches have uncorrelated dynamics. In the third scenario, we consider the same locations (Fig. 6(b)) for two epileptic patches but with correlated dynamics. For correlated dynamics, we include a short time delay between patch one and two to mimic propagation.

4.2. The denoising procedure and setting the parameters

The denoising procedure for the four methods has three steps: (1) extracting P' sources by using the observations, (2) choosing P'_e sources of interest, say epileptic interictal sources, and eliminating sources of non-interest and (3) reconstructing the observation signal by using only the sources of interest. It is clear that for the GEVD-based and DSS-based methods, the epileptic interictal sources are automatically extracted by these algorithms and there is no need for extra stages
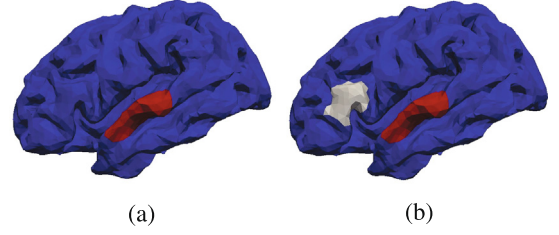

Fig. 6. The location of the epileptic sources in different scenarios: (a) scenario 1 and (b) scenarios 2 and 3.

to choose the epileptic sources. But for CoM$_2$ and SOBI, the epileptic sources should be selected from all extracted ones. To this end, we select visually P'_e components as epileptic sources. Moreover, for CoM$_2$ and SOBI, we first set the number of sources equal to the number of observation channels ($P' = N$), but the simulations showed that for these algorithms for higher SNRs, considering fewer number of sources generates better results. As a result, for these algorithms, in each trial of each scenario, we generated the results by choosing the value of P' yielding the best result.

In order to optimize the performance of the GEVD and DSS methods, we set the value of the parameters Th_{spike}, N_{ue}, L_w, N_θ, $L_{ma,1}$ and $L_{ma,2}$ equal to 80, 17, 81, 7, 9 and 19, respectively. The value of N_{eff} was set to be 4 and 6 for the first scenario and the other scenarios, respectively. Th_θ was set to be $0.4\overline{curv}_x$ where $\overline{curv}_x$ is the average of the curvature measure in the examined interval. Note that these values were estimated from a few trials of the simulated data and then we used these constant values dealing with all data.

4.3. Evaluation criteria

To compare the results of the denoising algorithms, we use four measures. The first two measures evaluate the performance of the denoised signals and the two other measures calculate their required flops and CPU time.

The first measure, called $Error_1$, shows the error in separating the signal of interest, the epileptic interictal activity, from the signal of non-interest. The measure $Error_1$ is given by:

$$Error_1 = \frac{\sum_{n=1}^{N}\sum_{k=1}^{T}(x_n^{(e)}[k] - \hat{x}_n^{(e)}[k])^2}{\sum_{n=1}^{N}\sum_{k=1}^{T}(x_n^{(e)}[k])^2} \quad (21)$$

where $\{x_n^{(e)}[k]\}$ and $\{\hat{x}_n^{(e)}[k]\}$ are the actual and estimated denoised EEG signals on the n-th channel, respectively.

The second measure, called $Error_2$, shows the ability of the denoising algorithms to identify the epileptic subspaces corresponding to different patches. $Error_2$ is defined as follows:

$$Error_2 = \frac{1}{P'_e}\sum_{p=1}^{P'_e}\left(\frac{\sum_{n=1}^{N}\sum_{k=1}^{T}(x_n^{(e,p)}[k] - \hat{x}_n^{(e,p)}[k])^2}{\sum_{n=1}^{N}\sum_{k=1}^{T}(x_n^{(e,p)}[k])^2}\right) \quad (22)$$

where $\{x_n^{(e,p)}[k]\}$ and $\{\hat{x}_n^{(e,p)}[k]\}$ are the actual and estimated denoised EEG signals reconstructed from the p-th epileptic component on the n-th channel, respectively. It is clear that, in
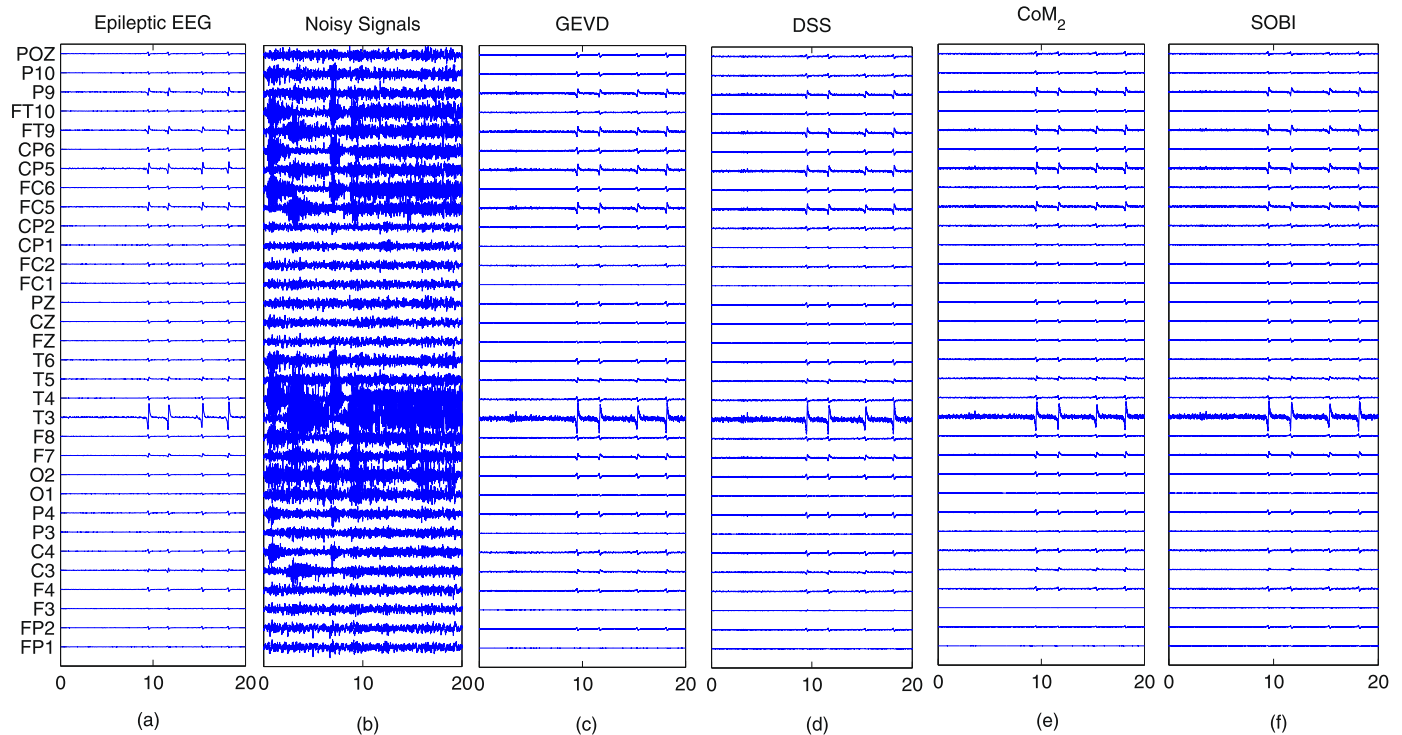
Fig. 7. An example of denoising procedure in the case of simulated data of scenario 1 generated from the activation of a single patch located in the superior temporal gyrus: (a) noise free simulated EEG with interictal spike-like activity, (b) noisy EEG after adding real muscle activity and background EEG for SNR = −15 dB, (c–f) EEG denoised by the GEVD-based, DSS-based, CoM$_2$ and SOBI algorithms, respectively.
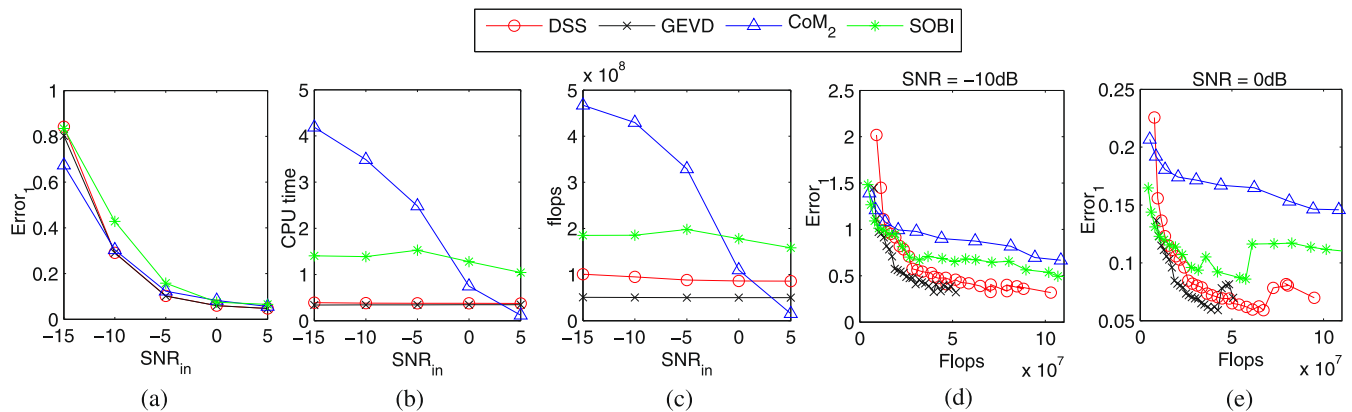

Fig. 8. Denoising results in the case of scenario 1: (a) $Error_1$ as a function of SNR, (b) CPU time as a function of SNR, (c) flops as a function of SNR, (d) $Error_1$ as a function of number of flops for SNR = −10 dB and (e) $Error_1$ as a function of number of flops for SNR = 0 dB.

scenarios with only one patch, $Error_2$ is equal to $Error_1$. Since in scenarios with correlated activities, the activities of different patches cannot be separated, $Error_2$ cannot be computed.

4.4. Results of the denoising methods for simulated data

Fig. 7 shows an example of simulated, noisy and denoised data of the first scenario with SNR = −15 dB. Electrode T3 shows clear spike-like activity on original simulated data but this activity is completely hidden in the simulated noisy data. A visual analysis of denoised data shows that for this example the spike activity at T3 was well retrieved with all four algo-

rithms. The other channels were acceptably denoised by using all four algorithms.

Fig. 8 shows the results obtained with the first scenario by using the four denoising algorithms for different SNRs. These algorithms were compared in terms of $Error_1$ as well as the CPU time and the required flops as functions of SNR. As illustrated in Figs. 8(a–c), the GEVD and DSS methods generated similar results in terms of $Error_1$. Fig. 8(a) shows that all four algorithms generated almost similar results. More particularly, the GEVD and DSS methods generated lower $Error_1$ than CoM$_2$ and SOBI in most cases except for SNR = −15 dB, for which CoM$_2$ gave better results. The main advantage of the GEVD and DSS methods is their low
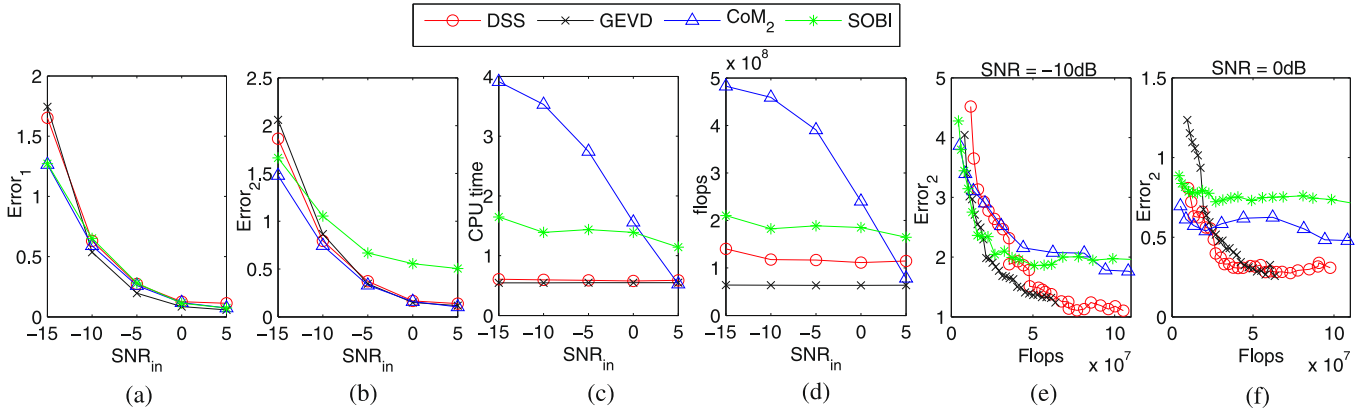
Fig. 9. Denoising results in the case of scenario 2: (a) $Error_1$ as a function of SNR, (b) $Error_2$ as a function of SNR, (c) CPU time as a function of SNR, (d) flops as a function of SNR, (e) $Error_2$ as a function of number of flops for SNR = −10 dB and (f) $Error_2$ as a function of number of flops for SNR = 0 dB.
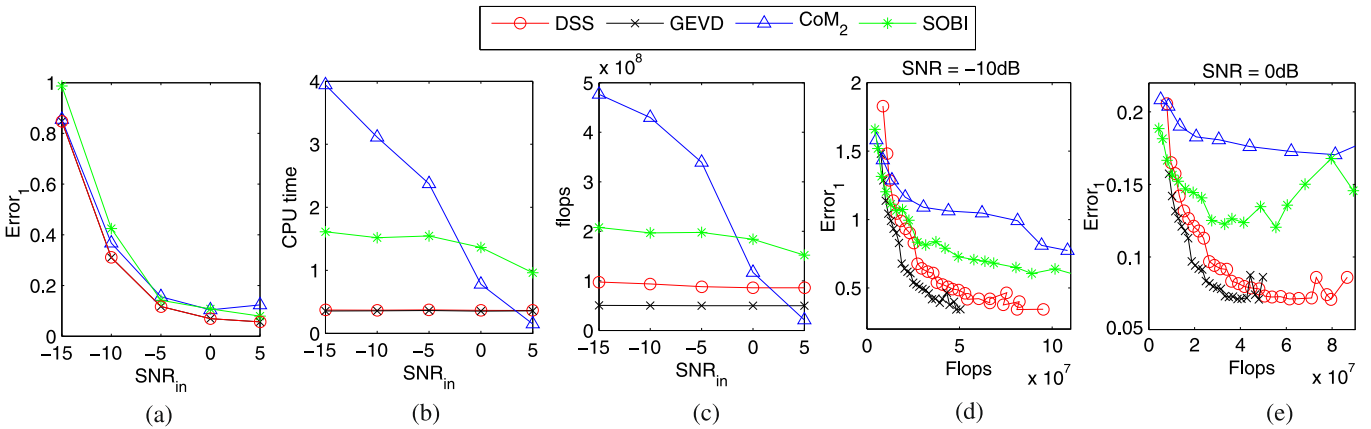

Fig. 10. Denoising results in the case of scenario 3: (a) $Error_1$ as a function of SNR, (b) CPU time as a function of SNR, (c) flops as a function of SNR, (d) $Error_1$ as a function of number of flops for SNR = −10 dB and (e) $Error_1$ as a function of number of flops for SNR = 0 dB.

computational cost in comparison with the CoM_2 and SOBI algorithms; this statement can be easily confirmed by the results shown in Figs. 8(b–c). Note that, for CoM_2, the required flops and consequently the CPU time greatly decreased in higher input SNRs. As mentioned in Section 4.2, this is caused by using fewer number of sources to optimize the performance of the CoM_2 algorithm. To compare the algorithms in terms of performance and speed, we varied the number of channels for each algorithm to have almost similar numerical complexities. Then we compared the performance of the algorithms by using the measure $Error_1$. Figs. 8(d–e) show the average $Error_1$ of the four denoising algorithms as a function of number of flops for two SNR values −10 dB and 0 dB. As shown in these figures, the GEVD and DSS methods have smaller $Error_1$ than CoM_2 and SOBI, especially for high numerical complexities.

Fig. 9 shows the results obtained by using four denoising algorithms in the case of the scenario 2. As shown in Figs. 9(a–b), for SNRs higher than −15 dB, all four algorithms generated similar $Error_1$ but for SNR = −15 dB, CoM_2 and SOBI outperformed GEVD and DSS. In terms of $Error_2$, for SNRs higher than −15 dB, GEVD, DSS and CoM_2 generated approximately similar results, but SOBI could not compete with these algorithms. These results show that although SOBI can separate the epileptic subspace from noise subspace, it cannot

effectively separate the epileptic subspaces corresponding to different epileptic sources from each other. Similar to $Error_1$, for SNR = −15 dB, SOBI and CoM_2 had lower $Error_2$ than GEVD and DSS. As previously mentioned, the GEVD and DSS methods had lower computational cost in comparison with the CoM_2 and SOBI algorithms as shown in Figs. 9(c–d). Similar to scenario 1, for scenario 2 we varied the number of channels for each algorithm to get an equivalent numerical complexity and we compared the performance of the algorithms by using the measure $Error_2$. Figs. 9(e–f) show the average $Error_2$ of four denoising algorithms as a function of number of flops for two SNR values −10 dB and 0 dB. In this scenario, similar to scenario 1, for high numerical complexities the GEVD and DSS algorithms had smaller $Error_2$ than CoM_2 and SOBI.

Fig. 10 shows the denoising results of the four algorithms in the case of the scenario 3 in which we considered two epileptic patches with correlated activity. In this case, since the sources were highly correlated, we extracted only one epileptic source in each trial. As shown in Fig. 10(a), all four algorithms generated almost similar $Error_1$. More particularly, the results obtained by the GEVD and DSS methods were slightly better than the CoM_2 and SOBI algorithms, especially for the SNRs higher than −15 dB. As the previously studied scenarios, the CPU time and flops required by the GEVD and DSS methods were
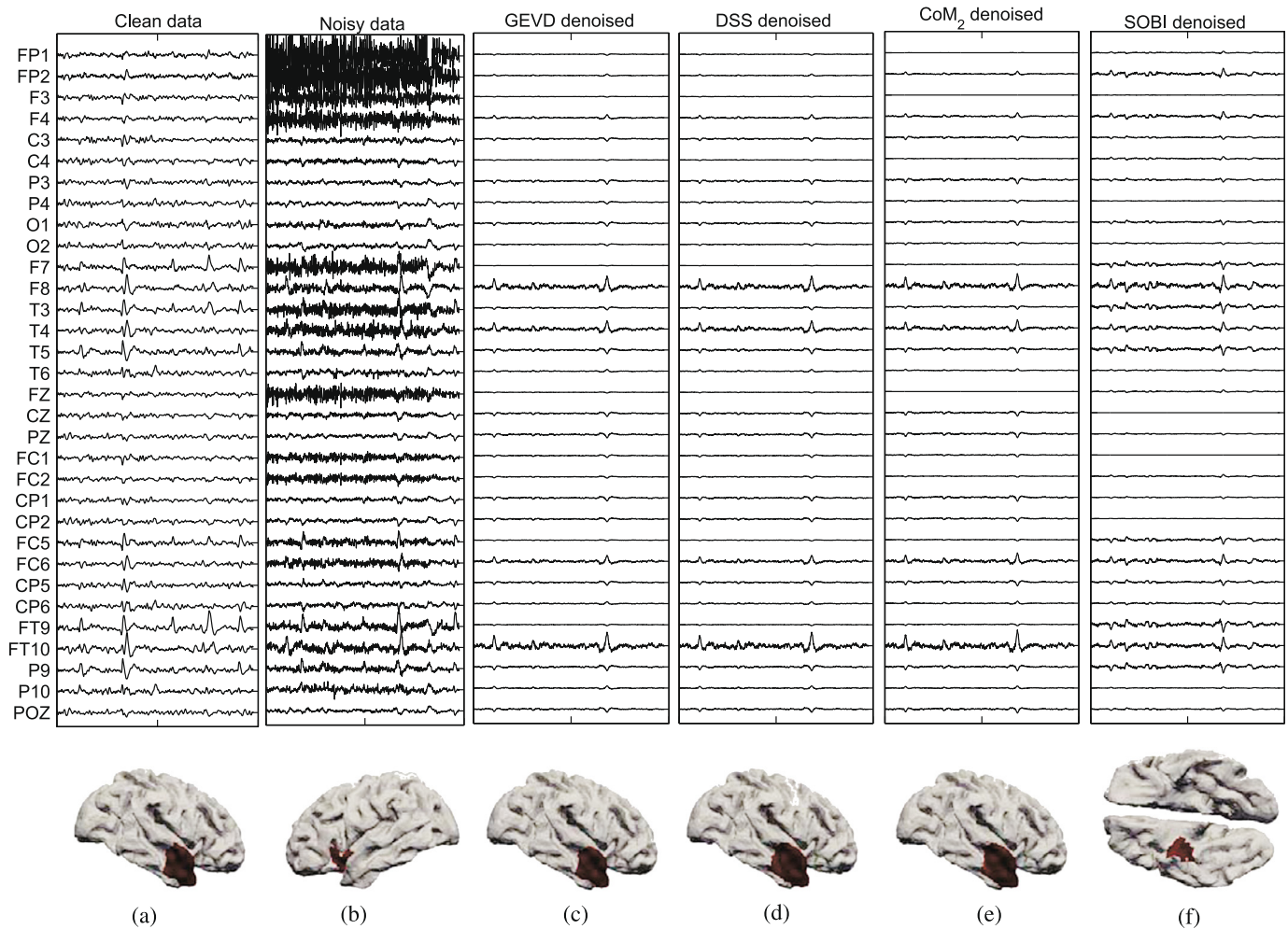
Fig. 11. Denoising of real interictal spikes data: (a) an epoch containing noise-free interictal spikes, (b) an epoch including spikes hidden in muscle activity and EEG denoised by (c) the GEVD-based method, (d) DSS-based method, (e) CoM$_2$ and (f) SOBI. The source localization results obtained from 4-ExSo-MUSIC are depicted at the bottom of each column.

much lower than those of the CoM$_2$ and SOBI algorithms as shown in Figs. 10(b–c). Similar to previous scenarios, for scenario 3, we varied the number of channels for each algorithm to get an equivalent numerical complexity and we compared the performance of the algorithms by using the measure $Error_1$. Figs. 10(d–e) show the average $Error_1$ of four denoising algorithms as a function of number of flops for two SNR values −10 dB and 0 dB. Similar to scenarios 1 and 2, the GEVD and DSS methods generated smaller $Error_1$ than CoM$_2$ and SOBI, especially for high numerical complexity.

In addition to above mentioned results, it is noteworthy that for all the scenarios, the GEVD and DSS methods are entirely automatic, while for CoM$_2$ and SOBI, the appropriate number of sources (P') were defined by changing the number of sources and calculating error for each case. Furthermore for CoM$_2$ and SOBI, the sources of interest were determined visually.

4.5. Application to real data

In this section we evaluate the denoising methods in the case of real data. The GEVD, DSS, CoM$_2$ and SOBI algo-

rithms were applied to denoise interictal spikes obtained from a patient suffering from drug-resistant partial epilepsy. During video-EEG monitoring, scalp-EEG data were acquired from 32 electrodes at a sampling frequency of 256 Hz. These data were reviewed in order to isolate two epochs: i) one epoch containing clean spikes (Fig. 11(a)), and ii) one epoch including spikes partially hidden in muscle activity with very high level of noise (Fig. 11(b)). The same procedure as for simulated data was applied to the noisy real EEG epoch to reconstruct the denoised EEG signals by using the four denoising algorithms as shown in Figs. 11(c–f). Since we do not know the ground truth to evaluate qualitatively the performance of the four methods, a source localization process was performed on the original clean signal (considered as a reference), on the noisy data, as well as on the data denoised by GEVD, DSS, CoM$_2$ and SOBI. As shown in Figs. 11(c–f), the interictal spikes are visible at electrodes F8, T4 and FT10 in the denoised data whereas they are partially hidden in the noisy data. The denoising results of three algorithm GEVD, DSS and CoM$_2$ are almost similar and they show interictal spikes at electrodes F8, T4 and FT10, but SOBI generates also some signals on the remaining electrodes. The

recent 4-ExSo-MUSIC algorithm [39] was used to achieve the source localization. Regarding the source localization results (bottom of each column of Fig. 11), the spike sources are localized in the right anterior temporal region for the clean epoch. For the noisy epoch, the spike source is incorrectly localized in the left temporal region. The interictal data denoised by GEVD, DSS and CoM_2 are localized in the right temporal neocortex in agreement with the source localization obtained from the clean epoch. The localization results of data denoised by SOBI are discordant with the localization of clean spikes.

5. Conclusion

In this paper, we compared two denoising methods based on GEVD and DSS with CoM_2 and DSS, for denoising interictal epileptic EEG data. In order to extract the useful information to be used in the GEVD-based and DSS-based methods, a series of preprocessing stages were proposed, including spike peak detection, extracting the exact time support of spikes and clustering of spikes involved in each epileptic source.

The four denoising algorithms were compared in terms of performance and numerical complexity by using simulated epileptic EEGs in three different scenarios. Results showed that for the scenario 1 with one epileptic source, the GEVD and DSS algorithms denoised the epileptic interictal data as well as CoM_2 and SOBI for the SNR values higher than -15 dB. But for SNR $= -15$ dB, the CoM_2 algorithm outperformed the GEVD and DSS methods. The results also showed that, for the second scenario with two uncorrelated sources, the GEVD, DSS and CoM_2 methods surpassed SOBI in separating the epileptic subspaces corresponding to different epileptic sources for SNR values higher than -15 dB. But for SNR $= -15$ dB, the GEVD and DSS methods could not compete the CoM_2 and SOBI algorithms. In addition, for the scenario 3 with two correlated epileptic sources, the GEVD and DSS methods had better performance than CoM_2 and SOBI. In all scenarios, the GEVD and DSS algorithms had much lower computational cost in comparison with CoM_2 and SOBI in terms of required flops and CPU time. We also examined the feasibility of these denoising algorithms dealing with real interictal data and showed that the GEVD, DSS and CoM_2 algorithms denoised real data as well as simulated ones.

Finally in terms of implementation, an advantage of the GEVD and DSS methods over CoM_2 and SOBI is the fact that an extra stage is not needed to visually select the sources of interest. Further work should consider more complicated scenarios with higher number of epileptic patches with different locations and different correlation states.

Appendix A. The wavelet-based spike peak detection algorithm

In this section, the wavelet-based single-channel spike peak detection algorithm [22,23] is given in detail. According to the definition of the Continuous Wavelet Transform (CWT), the coefficients of the wavelet transform of the observed signal from n-th channel, denoted by $\{x_n[k]\}$, with the mother wavelet $\psi(t)$ are defined as follows:

$$D_{a,b}^{(n)} = \frac{1}{f_s} \sum_{k=1}^{T} x_n[k] \psi_{a,b}^*[k] \tag{A.1}$$

$$\psi_{a,b}[k] \in L^2(\mathbb{R}),$$

$$\psi_{a,b}[k] = \frac{1}{\sqrt{|a|}} \psi\left(\frac{k-b}{a}\right), \quad a \neq 0, b \in \mathbb{R} \tag{A.2}$$

where $*$ denotes the complex conjugate and f_s is the sampling frequency of $\{x_n[k]\}$. In this algorithm, complex Morlet wavelet is used as follows:

$$\psi(t) = C(1 + \cos 2\pi f_0 t)e^{2i\pi \ell f_0 t}, \tag{A.3}$$

$$|t| \leq \frac{1}{2f_0}, \ell \notin \{-1, 0, 1\}$$

where ℓ sets up the number of oscillations of the complex part, f_0 is the normalized frequency and C is a normalization coefficient ($\|\psi\| = 1$) [22].

In the first stage, S_1, by using the wavelet coefficients $D_{a,b}^{(n)}$ in the n-th channel, we calculate the measure $T_k^{(n)}$ for each time sample k and compare it with the decision threshold $\lambda_1^{(n)}$ as follows:

$$T_k^{(n)} = \sum_{i=1}^{M} |D_{a_i,k}^{(n)}|^2 > \lambda_1^{(n)} \tag{A.4}$$

where parameters M and a_is are chosen such that the coefficients $D_{a_i,k}^{(n)}$ approximately cover the bandwidth of the spikes. Consequently, at the output of the stage S_1, the transient signals are enhanced compared to the background activity without any distinction between interictal spikes and muscular activity [22]. On the other hand, the experiments show that for a muscular artifact, compared with a spike, the squared modulus increases for high resolution scales [22]. This remark gives the idea to build a decision parameter $G_k^{(n)}$ which is calculated as the mean gravity center of the abscissa $\frac{1}{a_i}$'s weighted by $|D_{a_i,k}^{(n)}|^2$ on the detection interval around each sample point. Therefore, in the stage S_2, the measure $G_k^{(n)}$ is calculated in the n-th channel, in the window of length $2w + 1$, as follows [22]:

$$G_k^{(n)} = \frac{\sum_{j=k-w}^{k+w} g_j^{(n)}}{2w + 1} \quad \text{where } g_j^{(n)} = \frac{\sum_{i=1}^{M}(\frac{1}{a_i})|D_{a_i,j}^{(n)}|^2}{\sum_{i=1}^{M} |D_{a_i,j}^{(n)}|^2} \tag{A.5}$$

to separate the interictal spikes from muscular activity. Experimentally, this quantity takes distinct values in the presence of an artifact or a spike. By comparing $G_k^{(n)}$ with a threshold $\lambda_2^{(n)}$, the stage S_2 separates the useful signals from artifacts [22]. In [22], a practical method to choose the adaptive thresholds $\lambda_1^{(n)}$ and $\lambda_2^{(n)}$ was proposed.

References

[1] Vigario R, Oja E. BSS and ICA in neuroinformatics: from current practices to open challenges. IEEE Rev Biomed Eng 2008;1:50–61.
[2] Delorme A, Palmer J, Onton J, Oostenveld R, Makeig S. Independent EEG sources are dipolar. PLoS ONE 2012;7(2):e30135.

[3] Albera L, Kachenoura A, Comon P, Karfoul A, Wendling F, Senhadji L, et al. ICA-based EEG denoising: a comparative analysis of fifteen methods. Bull Pol Acad Sci 2012:407–18 [special issue].

[4] Lee TW, Girolami M, Sejnowski TJ. Independent component analysis using an extended infomax algorithm for mixed sub-Gaussian and super-Gaussian sources. Neural Comput 1999;2:417–41.

[5] Comon P, Jutten C. Handbook of blind source separation: independent component analysis and applications. Academic Press, Elsevier; 2010.

[6] Comon P. Independent Component Analysis, a new concept? Signal Process 1994;36(3):287–314.

[7] Safieddine D, Kachenoura A, Albera L, Birot G, Wendling F, Senhadji L, et al. ICA versus CCA pour le débruitage de signaux épileptiques intercritiques: une étude comparative de performances basée sur la localisation de la zone épileptogène. IRBM 2011;32:298–301.

[8] Safieddine D, Kachenoura A, Albera L, Birot G, Karfoul A, Pasnicu A, et al. Removal of muscle artifact from EEG data: comparison between stochastic (ICA and CCA) and deterministic (EMD and wavelet-based) approaches. EURASIP J Adv Signal Process 2012(1):1–15.

[9] De Clercq W, Vergult A, Vanrumste B, Paesschen WV, Huffel SV. Canonical correlation analysis applied to remove muscle artifacts from the electroencephalogram. IEEE Trans Biomed Eng 2006;53(12):2583–7.

[10] Huang NE, Shen Z, Long SR, Wu MC, Shih HH, Zheng Z, et al. The empirical mode decomposition and the Hilbert spectrum for nonlinear and non-stationary time series analysis. Proc R Soc Lond 1998;454:903–95.

[11] LeVan P, Urrestarazu E, Gotman J. A system for automatic artifact removal in ictal scalp EEG based on independent component analysis and Bayesian classification. Clin Neurophysiol 2006;117(4):912–27.

[12] Romo Vázquez R, Vélez-Pérez H, Ranta R, Louis Dorr V, Maquin D, Maillard L. Blind source separation, wavelet denoising and discriminant analysis for EEG artefacts and noise cancelling. Biomed Signal Process Control 2012;7(4):389–400.

[13] Sarela J, Valpola H. Denoising source separation. J Mach Learn Res 2005;6:233–72.

[14] Hesse CW, James CJ. On semi-blind source separation using spatial constraints with applications in EEG analysis. IEEE Trans Biomed Eng 2006;53(12):2525–34.

[15] Latif MA, Sanei S, Chambers J, Shoker L. Localization of abnormal EEG sources using blind source separation partially constrained by the locations of known sources. IEEE Signal Process Lett 2006;13(3):117–20.

[16] Ahmadian P, Sanei S, Ascari L, Gonzalez Villanueva L, Umilta MA. Constrained blind source extraction of readiness potentials from EEG. IEEE Trans Neural Syst Rehabil Eng 2013;21(4):567–75.

[17] Samadi S, Amini L, Cosandier-Rimélé D, Soltanian-Zadeh H, Jutten C. Reference-based source separation method for identification of brain regions involved in a reference state from intracerebral EEG. IEEE Trans Biomed Eng 2013;60(7):1983–92.

[18] Hajipour S, Shamsollahi MB, Albera L, Merlet I. Noise cancelation of epileptic interictal EEG data based on generalized eigenvalue decomposition. In: TSP 12. 2012 IEEE thirty-fifth international conference on telecommunications and signal processing, Prague, Czech Republic. 2012. p. 591–5.

[19] Belouchrani A, Abed-meriam K, Cardoso JF, Moulines E. A blind source separation technique using second-order statistics. IEEE Trans Signal Process 1997;45(2):434–44.

[20] Josse J, Husson F. Selecting the number of components in principal component analysis using cross-validation approximations. Comput Stat Data Anal 2012;56(6):1869–79.

[21] Velicer WF. Determining the number of components from the matrix of partial correlations. Psychometrika 1976;41(3):321–7.

[22] Senhadji L, Dillenseger JL, Wendling F, Rocha C, Kinie A. Wavelet analysis of EEG for three-dimensional mapping of epileptic events. Ann Biomed Eng 1995;23:543–52.

[23] Senhadji L, Bellanger J-J, Carrault G, et al. EEG spike detectors based on different decompositions: a comparative study. In: Time frequency and wavelets in biomedical signal processing. 1997. p. 407–21.

[24] Wilson SB, Emerson R. Spike detection: a review and comparison of algorithms. Clin Neurophysiol 2002;113:1873–81.

[25] Acir N, Oztura I, Kuntalp M, Baklan B, Guzelis C. Automatic detection of epileptiform events in EEG by a three-stage procedure based on artificial neural networks. IEEE Trans Biomed Eng 2005;52:30–40.

[26] Bourien J, Ruel J, Senhadji L, Puel JL. Comparison of three spike detectors dedicated to single unit action potentials of the auditory nerve. In: Proceedings of the 29th annual international conference of the IEEE EMBS. Lyon, France. 2007. p. 1430–3.

[27] Indiradevia K, Eliasa E, Sathidevia P, Nayakb SD, Radhakrishnanb K. A multi-level wavelet approach for automatic detection of epileptic spikes in the electroencephalogram. Comput Biol Med 2008;38:805–16.

[28] Ji Z, Sugi T, Goto S, Wang X, Ikeda A, Nagamine T, et al. An automatic spike detection system based on elimination of false positives using the large-area context in the scalp EEG. IEEE Trans Biomed Eng 2011;58:2478–88.

[29] Senhadji L, Wendling F. Epileptic transient detection: wavelets and time–frequency approaches. Clin Neurophysiol 2002;32:175–92.

[30] Pfurtscheller G, Fischer G. A new approach to spike detection using a combination of inverse and matched filter techniques. Electroencephalogr Clin Neurophysiol 1978;44:243–7.

[31] Oikonomou V, Tzallas A, Fotiadis D. A Kalman filter based methodology for EEG spike enhancement. Comput Methods Programs Biomed 2007;8:101–8.

[32] Ward J. Hierarchical grouping to optimize an objective function. J Am Stat Assoc 1963;58(301):236–44.

[33] Rosenfeld A, Johnston E. Angle detection on digital curves. IEEE Trans Comput 1973;2:875–8.

[34] Bezdek JC, Ehrlich R, Full W. FCM: the fuzzy c-means clustering algorithm. Comput Geosci 1984;10(2):191–203.

[35] Dempster A, Laird N, Rubin D. Maximum likelihood from incomplete data via the EM algorithm. J R Stat Soc B 1977;39(1):1–38.

[36] Kachenoura A, Albera L, Senhadji L, Comon P. ICA: a potential tool for BCI systems. IEEE Signal Process Mag 2008;25(1):57–68 [special issue on brain–computer interfaces].

[37] Wendling F, Bellanger JJ, Bartholomei F, Chauvel P. Relevance of nonlinear lumped-parameter models in the analysis of depth-EEG epileptic signals. Biol Cybern 2000;83:367–78.

[38] Cosandier-Rimélé D, Merlet I, Badier JM, Chauvel P, Wendling F. The neuronal sources of EEG: modeling of simultaneous scalp and intracerebral recordings in epilepsy. NeuroImage 2008;42(1):135–46.

[39] Birot G, Albera L, Wendling F, Merlet I. Localization of extended brain sources from EEG/MEG: the ExSo-MUSIC approach. NeuroImage 2011;56(1):102–13.