Blind source separation, wavelet denoising and discriminant analysis for EEG artefacts and noise cancelling

Rebeca Romo-Vázquez, Hugo Velez-Perez, Radu Ranta, Valérie Louis-Dorr, Didier Maquin, Louis Maillard

Blind source separation, wavelet denoising and discriminant analysis for EEG artefacts and noise cancelling

R. Romo Vázquez[a,b*], H. Vélez-Pérez[a,b], R. Ranta[b], V. Louis Dorr[b], D. Maquin[b], L. Maillard[b,c]

[a]: *Universidad de Guadalajara, CUCEI, Departamento de Electrónica, Av. Revolución 1500, Guadalajara, Jalisco, México*

[b]: *Centre de Recherche en Automatique de Nancy (CRAN), Nancy-University, CNRS, 2, Avenue de la Forêt de Haye F-54516 Vandoeuvre-les-Nancy*

[c]: *Centre Hospitalier Universitaire Nancy (CHU), Neurology Service, 29, Av. du Ml. de Lattre de Tassigny F-54000 Nancy*

Abstract

This paper proposes an automatic method for artefact removal and noise elimination from scalp electroencephalogram recordings (EEG). The method is based on blind source separation (BSS) and supervised classification and proposes a combination of classical and news features and classes to improve artefact elimination (ocular, high frequency muscle and ECG artefacts). The role of a supplementary step of wavelet denoising (WD) is explored and the interactions between BSS, denoising and classification are analyzed. The results are validated on simulated signals by quantitative evaluation criteria and on real EEG by medical expertise. The proposed methodology successfully rejected a good percentage of artefacts and noise, while preserving almost all the cerebral activity. The "denoised artefact-free" EEG presents a very good improvement compared with recorded raw EEG: 96% of the EEGs are easier to interpret.

Key words: EEG, artefact elimination, BSS, supervised classification, wavelet denoising, medical validation

*corresponding author

Email address: rebecca.romo@red.cucei.udg.mx (R. Romo Vázquez[a,b])

1. Introduction

The electrical activity produced by the brain is recorded by the electroencephalogram (EEG) using several electrodes placed on the scalp, generally according to predefined international standards (10-20 for example). Signals characteristics vary from one state to another, such as wakefulness/ sleep or normal/ pathological. Classically, five major brain waves can be distinguished by their frequency ranges: delta (δ) 0.5 – 4Hz, theta (θ) 4 – 8Hz, alpha (α) 8 – 13Hz, beta (β) 13 – 30Hz and gamma (γ) 30 – 128Hz [1].

The informative cortically generated signals are contaminated by extracerebral artefact sources: ocular movements, eye blinks, electrocardiogram (ECG), muscular artefacts. Generally the mixture between brain signals and artefactual signals is present in all sensors, although not necessarily in the same proportions (depending on the spatial distribution). Moreover, the EEG recordings are also affected by other unknown basically random signals (instrumentation noise, other physiological generators, external electromagnetic activity, etc) which can be modeled as additive random noise. These phenomena make difficult the analysis and interpretation of the EEGs, and a first important processing step would be the elimination of the artefacts and noise.

Several methods for artefact elimination were proposed in the literature, a brief review being presented in the next section. Most of them consists in two main steps: artefact extraction from the multichannel recorded signals, generally using some signal separation methods, followed by signal classification (visual or automatic) and "clean EEG" reconstruction. Our goal is to contribute to EEG artefact rejection by proposing an original and more complete automatic methodology consisting in an optimized combination of several signal processing and data analysis techniques.

This paper is organized as follows: the second section presents a brief state of the art of the EEG artefact rejection solutions proposed in the literature, and concludes with the main original contributions of this paper. The third section is dedicated to our proposed approach. This section consists in a first part that briefly introduces the methodological steps included in our approach: BSS, supervised classification and wavelet denoising WD. A second part discusses their mutual interactions and proposes an optimal combination. The fourth section presents the main results and it is also divided

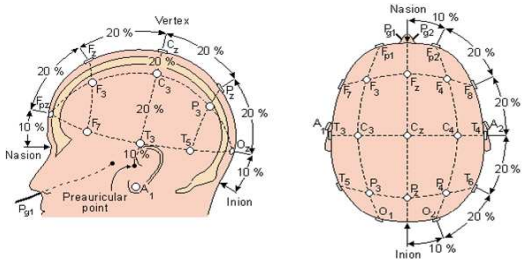


Figure 1: International 10-20 system

in two parts: the first one aims to choose an optimal BSS - WD combination algorithm using a set of semi-simulated realistic EEG signals, while the second one is dedicated to a comparative medical validation on large database of real EEGs. Finally, conclusion and research perspectives are given.

In order to define the framework of this paper, we must specify that we address here classical scalp EEGs, as recorded in clinical routine. More precisely, we deal with signals recorded using the 10-20 international system, with the reference electrode placed at the Fpz position, close to the nasion (see figure 1).

2. State of the art

The most common artefact rejection method is based on linear regression and aims to eliminate the most frequent artefact, due to eye movement or blinking. This method is capable of estimating the influence of the simultaneously measured electrooculogram (EOG) signals in the EEG and to remove them [2]. Still, as the EOG electrodes are frequently placed close to the EEG electrodes, they contain a certain amount of brain activity. Thus, removing these signals might cause distortions in the EEG signals. Besides, this method needs supplementary channels for EOG recording and, of course, does not deal with other artefacts.

Source separation. More complete methods, based on linear decomposition of the multichannel EEG recordings, were proposed in the literature. The most common method is the blind source separation (BSS). The main BSS hypothesis states that the artefact sources are independent from brain sources, either normal or pathologic; the goal is to recover the original sources (brain and artefactual), given only sensor observations. Several algorithms for BSS

were developed in the last years, either using high order statistics (HOS) and thus explicitly addressing the 'independence' (Independent Component Analysis - ICA) or second order statistics (SOS) on time delayed or windowed signals, as well as combined (ICA + SOS) approaches (see for example [3] for a review).

In the EEG artefact identification setups, there is no consensus for a best BSS algorithm. Several ICA based algorithms [4, 5, 6, 7, 8] have been applied. However, some more restrictive (but realistic) assumptions allow the use SOS only: if the sources have different non-random temporal structures, these methods proved efficient. Indeed, as it was pointed out by different authors [9, 10, 11, 12, 13] SOBI [14] is an appropriate algorithm to separate the EEG signals. This SOS algorithm is based on the 'joint diagonalization' of an arbitrary set of time delayed covariance matrices of the measured signals. It allows the separation of Gaussian sources. The SOBI-RO algorithm is adapted for noisy signals by introducing a robust orthogonalization step [15].

Feature extraction and classification. In the artefact elimination framework, BSS is used to separate cerebral and extra-cerebral sources. After separation, the identification of the artefact sources can be done by visual inspection [16, 17], but this approach becomes very impractical when processing many EEGs.

Recently, several methods for automatic identification of artefact sources were proposed. During the feature extraction step, the main source characteristics employed in the literature are:

- **Statistical characteristics** as kurtosis, entropy, trends or extreme values. The kurtosis is used to detect peaky activity distributions characterizing specific artefacts (*i.e.* ocular and cardiac). The entropy also helps to identify the signals concentrated in small temporal intervals, therefore likely to be artefacts [4, 5].

- **Template signals**. If available, artefact measured signals are used as templates (EOG, ECG). A high correlation between a source and the template will indicate an artefact [7, 18, 8, 6, 19, 20].

- **Frequency characteristics**. The sources can be characterized by the energy contained in the different frequency bands [7, 10].

- **Spatial characteristics**. The features based on the mixing characteristics are connected to the scalp topography: the projection strength of

the signals onto the scalp allows the examination of its biological origin [10, 6].

Often, these features are used to determine, with the help of human expertise, specific thresholds for every artefact type. These 'expert systems' are further on used to label specific sources as artefacts. More elaborated methods tend to automatically determine the frontiers between artefact and informative brain sources using different classifiers (Bayes rule based linear or quadratic discriminant analysis [7] or support vector machines [20]).

Similar methods, based on BSS and supervised classification, were equally proposed for magnetoencephalographic signals (MEG) [11, 21]. In all methods, after separation and classification, the artefactual sources are removed and the EEG is reconstructed with the informative brain sources only. As for all classification problems, the accuracy of the solution depends both on the features and on the classifier thresholds.

Our approach follows the same general idea: EEG signals are processed by BSS, a set of features is extracted from the obtained sources, these features are fed into a classifier and the artefact sources are eliminated from the reconstruction. The originality of our work consists in several aspects:

- we perform an extended evaluation of several BSS algorithms on physiologically plausible EEGs;
- we propose an original feature set for source classification, containing physiologically significant features extracted from the estimated mixing matrix and from time and frequency characteristics of the sources. The proposed method does not need EOG electrodes, as required by most of the methods presented in the literature. The electrocardiogram signal, used as reference signal for the characteristics extraction is routinely recorded simultaneously with the EEG.
- we introduce a wavelet denoising (WD) step and we analyze its interactions with the two other classical elements of the processing chain (BSS and classification), proposing an optimal sequencing of these three processing steps;
- finally, we perform a throughout validation of the processing chain on a large database of real EEGs, using both objective criteria (classification rate) and a systematic medical evaluation.

3. Proposed EEG artefact rejection methodology

As said in the introduction, this section is divided in two parts: the first one briefly recalls the individual methodological steps included in our method (BSS, Bayes classifier, wavelet denoising), while the second one analysis their interactions and proposes an optimal order of these processing steps.

3.1. Methodological steps

3.1.1. Blind Source Separation – BSS

The goal of the BSS is to recover original sources, given only sensor observations. The most simple and widely used assumption in EEG processing is the linear instantaneous mixing: source signals reach the sensors simultaneously [1].

The noisy mixture in this case is written as:

$$\mathbf{X} = \mathbf{AS} + \mathbf{N} \tag{1}$$

where $\mathbf{X}$ is the matrix of the mixed signals ($\boldsymbol{x}_i$ corresponds to a row of $\mathbf{X}$, *i.e.*, a sensor signal), $\mathbf{A}$ is the unknown nonsingular mixing matrix, $\mathbf{S}$ is the matrix of independent sources ($\boldsymbol{s}_i$ corresponds to a source), $\mathbf{N}$ is an additive noise matrix.

The aim of BSS is to find a linear transformation $\mathbf{B}$ of the sensor signals $\mathbf{X}$ that makes the outputs as independent as possible:

$$\mathbf{Y} = \mathbf{BX} = \mathbf{BAS} + \mathbf{BN} \tag{2}$$

where $\mathbf{Y}$ is the estimation of the sources $\mathbf{S}$. We assume here that the number of sources is equal to the number of sensors Q. In this case, $\mathbf{A} \in \mathbb{R}^{Q \times Q}$ and the ideal separation is obtained when $\mathbf{B} = \mathbf{A}^{-1}$ and, consequently, $\mathbf{Y}$ is a (noisy) estimate of $\mathbf{S}$. In practice, BSS algorithms search a $\mathbf{B}$ matrix such as the product $\mathbf{BA}$ is a permuted diagonal and scaled matrix. Therefore, original sources can be recovered except for their order (permutation) and their amplitude (scale): the estimated sources $\mathbf{Y}$ will be permuted and normalized to unitary standard deviation.

BSS evaluation. On simulated signals, one can validate the BSS results by using the separability index *SI* [3, 22, 17]. This index is computed from the transfer matrix $\mathbf{G} = \mathbf{BA}$ between the original sources and the estimated ones. The *SI* is computed starting from the absolute values of the $\mathbf{G}$ elements. The

rows $\mathbf{g_i}$ and the columns $\mathbf{g_j}$ of this matrix are normalized to obtain $\mathbf{g_i}'$ and $\mathbf{g_j}'$ respectively:

$$\mathbf{g_i}' = \frac{|\mathbf{g_i}|}{\max|\mathbf{g_i}|} \qquad \mathbf{g_j}' = \frac{|\mathbf{g_j}|}{\max|\mathbf{g_j}|} \tag{3}$$

The SI_1 and the SI_2 are obtained from the elements of the resulting $Q \times Q$ matrix $\mathbf{G}'$:

$$SI_1 = \frac{\sum_{i=1}^{Q}\left(\sum_{j=1}^{Q}(\mathbf{G}'(i,j)) - 1\right)}{Q(Q-1)} \qquad SI_2 = \frac{\sum_{j=1}^{Q}\left(\sum_{i=1}^{Q}(\mathbf{G}'(i,j)) - 1\right)}{Q(Q-1)}$$

A synthetic separability index SI writes as:

$$SI = \frac{SI_1 + SI_2}{2}, \tag{4}$$

its purpose being to measure the degree to which $\mathbf{G}$ is close to a permutation matrix. For perfect source recovering $SI = 0$.

3.1.2. Wavelet denoising – WD

Nowadays a classical solution for noise removal from non-stationary signals is WD. The decomposition of a noisy signal on a wavelet basis (discrete orthogonal wavelet transform, DWT) have the property to "concentrate" the informative signal in few wavelet coefficients having large absolute values without modifying the noise random distribution. Therefore, denoising can be achieved by thresholding the wavelet coefficients.

Consider the model $x(k) = c(k) + n(k)$, where $x(k)$ is the noisy discrete-time signal (length M), c is the noise-free unknown version of $x(k)$ and $n(k)$ the noise. For the EEG application, the $\boldsymbol{x}$ signal corresponds to a single EEG channel (a measurement electrode). As the DWT is a linear transform, the wavelet coefficient vectors of $\boldsymbol{x}$, $\boldsymbol{c}$ and $\boldsymbol{n}$ are related by:

$$\boldsymbol{w}_x = \boldsymbol{w}_c + \boldsymbol{w}_n \tag{5}$$

Denoising is performed by separating the wavelet coefficients $\boldsymbol{w}_x$ by thresholding. Generally, the wavelet coefficients corresponding to low frequencies (approximation coefficients) remain unchanged. The main problem is to estimate the threshold T between small and large wavelet coefficients (*i.e.*, noise

$\widehat{\boldsymbol{w}}_n$ and informative $\widehat{\boldsymbol{w}}_c$ respectively). Several algorithms have been proposed in the last years, and an extensive review can be found in [23, 24].

The most widely used is the universal threshold $T_U = \sigma_n\sqrt{2\log N}$ proposed by Donoho and Johnstone in their algorithm *VisuShrink* [25] (σ_n is an estimate of the noise standard deviation). *VisuShrink* is used to achieve complete asymptotic elimination of the normal Gaussian noise. Consequently, it might also lead to a less precise reconstruction of the signal of interest.

A different approach is proposed by the *SureShrink* algorithm [26] (*Stein Unbiased Risk Estimator*), that aims to estimate as precisely as possible the "clean" signal by minimizing an estimate of the mean squared error (*MSE*) between the denoised signal and the original one. The obtained threshold T_S (or thresholds, as the method is usually implemented by scale) are lower than the T_U.

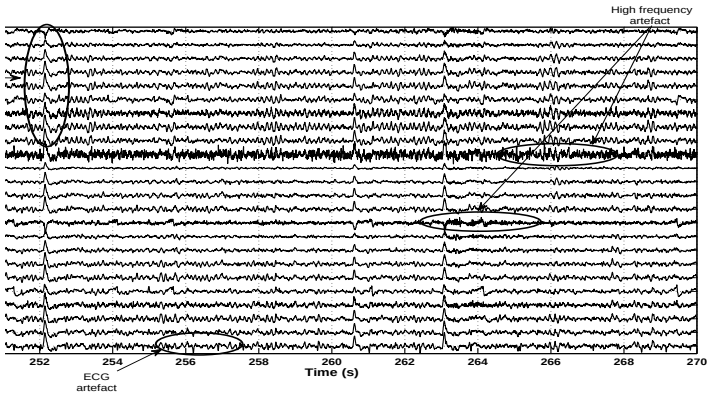
In the EEG case, not loosing information potentially useful to medical diagnosis is of great importance. Moreover, as in EEG the signal to noise ratio is weak, the wavelet coefficients of neuronal signals can have small values, comparable to noise. Therefore, algorithms with a low thresholding as *Sure* are the most appropriate for our application: theoretically, it insures the closest possible reconstruction (in a mean squared error sense) of the informative signal.

3.1.3. Feature extraction

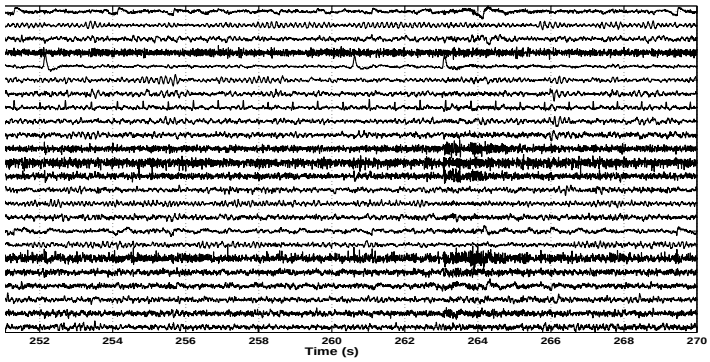
Several sets of features listed in the literature (see section 2) were investigated and a combination of spatial, frequency and template correlation characteristics was selected. The retained features cover the spectral content of the sources, their spatial projections and their correlation to a template routinely recorded signal (ECG).

Frequency characteristics. Spectral analysis was performed for all estimated sources. The power spectral density for the 5 frequency bands of each source y_j was calculated. After normalization using the total energy (frequency range 1 – 128Hz), the 5 features are no longer independent, so only 4 are considered in the sequel: $\overline{E}_{\delta_j}$, $\overline{E}_{\alpha_j}$, $\overline{E}_{\beta_j}$ and $\overline{E}_{\gamma_j}$. For each source, the main frequency f_{m_j} (estimated from the maximum value of the power spectral density) was also calculated. An example for a real EEG recording is depicted in figure 2.

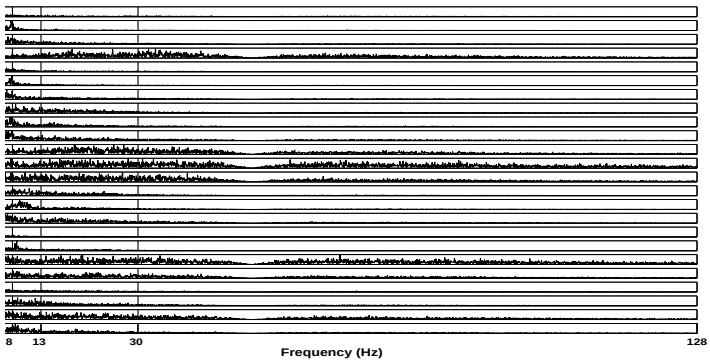
Spatial characteristics. In practice, an EEG signal is the measure of the potential difference between the measure and reference electrodes, which plays



(a) raw artefacted EEG



(b) separated sources (time course)



(c) separated sources (frequency spectra)

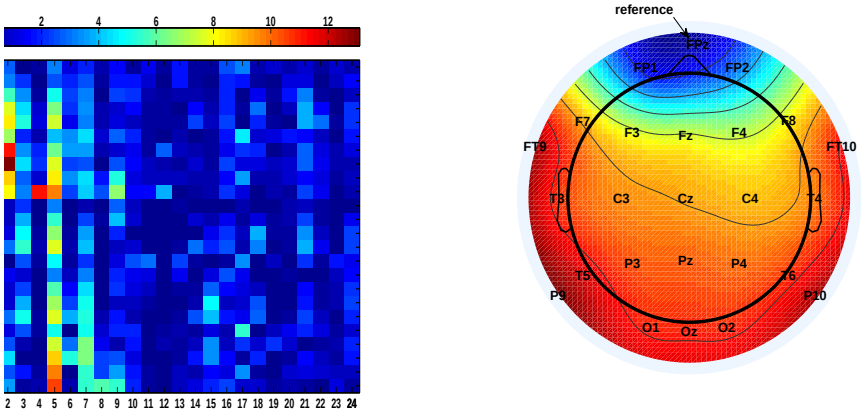
Figure 2: Example of raw EEG and estimated sources by SOBI-RO. Vertical lines in (c) delimit physiological frequency bands

an important role in the mixing matrix coefficients. In this work, the reference is placed near the eyes at Fpz (figures 1 and 3(b)), so it records the eye activity also. Consequently, there is a strong presence of ocular artefacts in all measured signals. After BSS, the sources are normalized and the strong presence of a source in a measured signal is reflected by an important absolute value of the corresponding mixing matrix coefficient. In particular, the elements of the column corresponding to the ocular artefact source should have significantly higher and more constant values than the other columns (figures 3(a) and 3(b)). After BSS estimation of the mixing matrix $\mathbf{A}$, it is possible to evaluate the strength of an source y_j by computing the average $\overline{a}_j$ of the coefficient absolute values of column $\widehat{\mathbf{a}}_j$.

$$\overline{a}_j = \frac{1}{Q}\sum_{i=1}^{Q} |\widehat{a}_{ij}| \qquad j = 1, 2, \ldots, Q \tag{6}$$

The normalized standard deviation σ_j of each column $\widehat{\mathbf{a}}_j$ indicates the spreading:

$$\overline{\sigma_j} = \frac{\sigma_j}{\overline{a}_j} \tag{7}$$



(a) estimated mixing matrix coefficients (absolute values)

(b) scalp projection of the ocular artefact source (source $5 \rightarrow 5^{th}$ mixing matrix column)

Figure 3: Topographic characteristics

As mentioned above, these spatial features are influenced by the reference

electrode: if the reference electrode is different to FPz (and thus the eye artefacts do not highly contaminate it), the EEG must be re-referenced.

Template correlation characteristic. The ECG is usually recorded simultaneously with the EEG, therefore it can be used as a template signal. The correlations between the template and the separated sources y_j write:

$$\rho_{ecg,j} = \frac{\mathtt{cov}(ecg, y_j)}{\sigma_{ecg}\sigma_{y_j}} \tag{8}$$

Finally, each estimated source is characterized by 8 features: 5 spectral features ($\overline{E}_{\delta_j}$, $\overline{E}_{\alpha_j}$, $\overline{E}_{\beta_j}$, $\overline{E}_{\gamma_j}$ and f_{m_j}), 2 spatial features (the normalized standard deviation $\overline{\sigma_j}$ and the average projection on the electrodes $\overline{a}_j$) and one template correlation feature ($\rho_{ecg,j}$).

3.1.4. Classification

The extracted features aim to code and quantify medical expertise. A common solution would be a rule based expert system: according to the feature extraction paragraph, one can expect a rather good identification of the ocular artefact AO and of the ECG artefact $AECG$ using only the spatial and template correlation characteristics respectively. Nevertheless, parameterizing this system is not trivial and not all the source types are addressed (high frequency muscle and noise artefacts HFA, brain B)[1]. We have therefore implemented in this paper two classification methods based on the classical the Bayes classifier, as described next.

After feature extraction, the sources $\mathbf{y}$ are viewed as a vector $\mathbf{sf}$ of dimension d (number of features). These vectors are fed to the classifier. A classical approach is the discriminant analysis, derived from the general Bayes rule. For a vector $\mathbf{sf}$, the probability of belonging to the class C_j is written:

$$\begin{aligned} p(C_j|\mathbf{sf}) &= \frac{p(\mathbf{sf}|C_j)p(C_j)}{p(\mathbf{sf})}, \quad \text{with} \\ p(\mathbf{sf}) &= \sum_{j=1}^{k} p(\mathbf{sf}|C_j)p(C_j), \end{aligned} \tag{9}$$

[1]Such a rule based classifier was implemented in [27], but the results were less convincing than the ones presented in this paper.

where $p(\mathbf{sf}|C_j)$ is the conditional probability of $\mathbf{sf}$ given C_j. Considering Gaussian classes and after some manipulation, one decides if a vector $\mathbf{sf}$ belongs to the class C_j by minimizing:

$$d_B(\mathbf{sf}, C_j) = \frac{1}{2}(\mathbf{sf} - \mu_\mathbf{j})^T {\mathbf{\Sigma_j}}^{-1}(\mathbf{sf} - \mu_\mathbf{j}) + \frac{1}{2}\ln|\mathbf{\Sigma_j}| - \ln p(C_j), \tag{10}$$

The mean $\mu_\mathbf{j}$, the covariance matrix $\mathbf{\Sigma_j}$ and the *a priori* probability $p(C_j)$ are estimated from the training classes defined by the experts. This **training set** was consensually constructed by two neurologists, which identified the 4 types of sources. Although not usually performing source visual analysis, they were able to identify artefact sources by comparing their time courses with the recorded EEGs (bipolar and referential montages). To preserve all information potentially useful to medical diagnosis, only the obvious artefact sources were selected.

A simplified version of (10) is the Mahalanobis distance:

$$d(\mathbf{sf}, C_j) = \frac{1}{2}(\mathbf{sf} - \mu_\mathbf{j})^T {\mathbf{\Sigma_j}}^{-1}(\mathbf{sf} - \mu_\mathbf{j}) \tag{11}$$

Source classification evaluation. Classification methods are classically evaluated using the number of the true positive (TP), true negative (TN), false negative (FN) and false positive (FP) detections. The values are used to compute the sensitivity (Sen), the specificity (Spe) and the false positive rate (FPR) of the classifier (eq. 12). Alternatively, one can compute the positive predictive value (PPV) and the negative predictive value (NPV), who give a complementary view on the results (eq. 13).

$$Sen = \frac{TP}{TP + FN} \quad Spe = \frac{TN}{TN + FP} \quad FPR = 1 - Spe \tag{12}$$

$$PPV = \frac{TP}{TP + FP} \quad NPV = \frac{TN}{TN + FN} \tag{13}$$

3.2. Interactions and optimal ordering

This subsection analyses the interaction between the three processing steps previously described (BSS, classification and WD) and proposes an optimized implementation, both in terms of chosen algorithms and in terms of ordering.

3.2.1. BSS and WD

Clearly, the first processing step must be chosen between BSS (to identify artefact sources) and denoising (to eliminate additive noise). Naturally, source separation algorithms perform better in no noise situations (even if specific algorithms as SOBI-RO were designed to deal with the noise), therefore a first approach would be to denoise the raw EEG signals before BSS. On the other hand, the WD step might lead to a loss of information by eliminating also a part of the informative signal, possibly useful for the BSS and/or for the classification steps.

Our first aim is thus to study the interaction between BSS and WD. Rather few references combine them, either for ECG [28] or for EEG preprocessing [29, 30, 31, 32]. However, the only detailed study on the interaction is presented in [33].

First, observe that both techniques are linear, and thus the mixing equation can be combined to the wavelet transform to obtain the mixture of wavelet coefficients:

$$\boldsymbol{W}_x = \boldsymbol{A}\boldsymbol{W}_s + \boldsymbol{W}_n, \tag{14}$$

with $\boldsymbol{W}_x$, $\boldsymbol{W}_s$ and $\boldsymbol{W}_n$ being the wavelet coefficients matrices of the measured signals, sources and noise respectively.

In particular, the vector of the wavelet coefficients of the ith measured signal $\boldsymbol{w}_{x,i}$ (row i of $\boldsymbol{w}_x$) can be written as a noisy linear combination of the wavelet coefficients of the sources $\boldsymbol{w}_{s,j}$:

$$\boldsymbol{w}_{x,i} \;=\; [w_{x,i}(1), w_{x,i}(2), \ldots w_{x,i}(M)] = \boldsymbol{a}_i \begin{bmatrix} \boldsymbol{w}_{s,1} \\ \boldsymbol{w}_{s,2} \\ \vdots \\ \boldsymbol{w}_{s,N} \end{bmatrix} + \boldsymbol{w}_{n,i} \tag{15}$$

with $\boldsymbol{a}_i$ the corresponding row of the mixing matrix $\boldsymbol{A}$.

Denoising this signal implies thresholding of the wavelet coefficients vector $\boldsymbol{w}_{x,i}$ using a threshold T to obtain the wavelets coefficients of the denoised

estimate $\widehat{\boldsymbol{w}}_{c,i}$. In a matrix form, thresholding can be written as:

$$\widehat{\boldsymbol{w}}_{c,i} = \boldsymbol{w}_{x,i}\boldsymbol{F}_{T,x_i} = \boldsymbol{w}_{x,i}\begin{bmatrix} f(w_{x,i}(1)) & 0 & 0 & \dots & 0 \\ 0 & f(w_{x,i}(2)) & 0 & \dots & 0 \\ 0 & 0 & f(w_{x,i}(3)) & \dots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \dots & f(w_{x,i}(M)) \end{bmatrix} \tag{16}$$

with $f(w_x)$ being the thresholding (shrinkage) function implementing the chosen threshold (see for example [23] for details).

Considering the mixing equation (14), one obtains:

$$\widehat{\boldsymbol{w}}_{c,i} = \boldsymbol{w}_{x,i}\boldsymbol{F}_{T,x_i} = \boldsymbol{a}_i \begin{bmatrix} \boldsymbol{w}_{s,1} \\ \boldsymbol{w}_{s,2} \\ \vdots \\ \boldsymbol{w}_{s,P} \end{bmatrix} \boldsymbol{F}_{T,x_i} + \boldsymbol{w}_{n,i}\boldsymbol{F}_{T,x_i} \tag{17}$$

In other words, thresholding the wavelet coefficients of $\boldsymbol{x}_i$ implies applying the same threshold to all the sources contributing to it. Which means that the individual sources risk to be distorted by the threshold, adapted to the mixture but not to the sources characteristics.

For example, a source that has a small contribution to the measured signal (*i.e.*, a small mixing coefficient) risks to be lost or highly distorted after thresholding. According to this analysis, if there is a distortion risk (depending on the value of the threshold), it seems better to insert the WD step after the BSS. In order to confirm the presented analysis, we performed several simulations, the results being presented in the next section.

3.2.2. Classifier position

A last issue is the position of the classifier in the processing chain. In our opinion, there is no general answer: as for the BSS case, denoising might distort the signals and mislead both the experts and the automatic classification. Indeed, to avoid possible interpretation errors[2], clinical experts need all the available information to do a correct source identification. Consequently, we propose to place the classification step immediately after the

[2]According to our tests on real EEGs, this problem mainly occurs for muscle artefacts containing high frequency components, often distorted by denoising.

BSS. It is noteworthy that, in order to preserve the coherence of the method, the training set was also constructed from non denoised sources.

4. Results

This section consists on two parts: the first one presents simulated signals and leads to the final design of the processing chain. The second one focuses on real EEGs and on medical validation.

4.1. Simulation

The goal of this simulation step is to validate the analysis presented in section 3.2.1: which is the optimal BSS–WD combination?

Nine (Q=9) semi-simulated sources were used (figure 4(a)):

- 6 brain sources (S1, S2, S3, S4, S5 and S6) extracted from an intracerebral (depth) encephalogram (SEEG) recording. As the SEEG signals are recorded by electrodes placed inside the brain, we can safely assume that these signals are artefacts free. In order to insure maximum independence, these 6 sources were extracted from different brain locations and different time windows.

- 3 artefacts sources (one ocular artefact source (S7), an ECG artefact source (S8) and a high frequency artefact source (S9)) were also used as semi-simulated sources. These sources were extracted from scalp EEG recordings (different patients) by source separation. They were identified by the clinicians as artefacts.

The semi-simulated sources (figure 4(a)) were mixed using a random matrix (uniform distribution in [-1, 1]) (see example in figure 4(b)). Several BSS algorithms (ICAlab and EEGlab toolbox [34, 35]) were tested on ideal and on noisy simulated mixtures, with 4 SNR values (5dB, 10dB, 15dB, 20dB). Both white and colored noises were used, with similar results. The noises were independently generated for the 9 measures (spatially white noise). The results for the best 14 algorithms are presented in table 1 (average values over 1000 simulations).

Analyzing this table, we can see that the combined algorithm (ICA + SOS) COMBI is the most performant for the noise free signals ($SI = 0.0428$). For low noise level (SNR=20dB), it is the EEGLab's RUNICA (Extended Infomax) algorithm that presents the best results ($SI = 0.1065$), however,

this result is close to the result of the SOBI-RO algorithm ($SI = 0.1118$). On the other hand, SOBI-RO has the best performance for higher noise levels (15dB $\rightarrow$ SI= 0.1247, 10dB $\rightarrow$ SI=0.1452 and 5dB $\rightarrow$ SI=0.1757). From these results we can conclude that the COMBI algorithm is the best adapted for the noise free signals, while SOBI-RO is the most performant for noisy EEG signals.

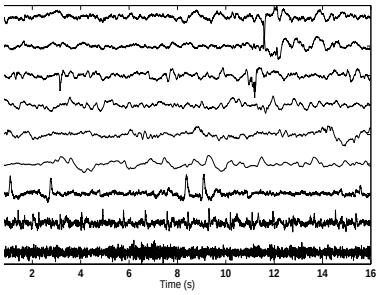
Table 1: Comparison of BSS algorithms on semi-simulated EEG (no noise and noisy)

Algorithme SAS	**no noise**	**20dB**	**15dB**	**10dB**	**5dB**
AMUSE [36]	0.0565	0.2410	0.2726	0.2946	0.3098
SOBI-RO [15]	0.0941	0.1118	**0.1247**	**0.1452**	**0.1757**
EVD [37]	0.0600	0.2219	0.2553	0.2803	0.2968
SOBI [38, 14]	0.0882	0.1407	0.1673	0.1989	0.2322
SOBI-BPF[39]	0.0753	0.1124	0.1346	0.1663	0.2085
WASOBI [40]	0.0654	0.1874	0.2119	0.2348	0.2507
EWASOBI [41]	0.0614	0.1563	0.1881	0.2246	0.2555
SEONS2[42]	0.0962	0.1398	0.1502	0.1652	0.1877
RUNICA[43]	0.0744	**0.1065**	0.1280	0.1634	0.2191
JADE[44]	0.1533	0.1798	0.1947	0.2131	0.2370
JADE-TD [45]	0.1857	0.2160	0.2300	0.2465	0.2664
EFICA[46]	0.0815	0.1162	0.1351	0.1622	0.2032
COMBI[41]	**0.0428**	0.1266	0.1509	0.1777	0.2029
MULTICOMBI[47]	0.1219	0.1553	0.1817	0.2091	0.2300

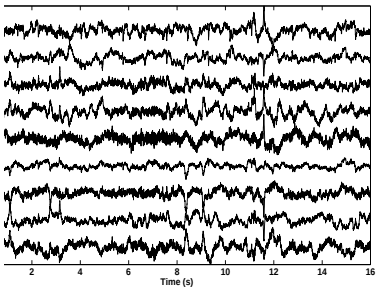
The BSS algorithms were next tested in combination with the WD. Denoising was implemented before and after BSS. As for the BSS only, the combination BSS-WD was evaluated using the SI (eq. 4) averaged over 1000 simulations (mixing matrices). Mean SI values are presented in table 2. For conciseness, only the two best algorithms are presented, the best for the no noise situation (**COMBI**) and the best for the noisy signals (**SOBI-RO**) (see example figure 4). Also, only the results for the *Sure* algorithm [26], known to minimally distort the informative signals (see section 3.1.2) are presented here. More detailed results can be found in [48, 27].

From these results we can conclude that:

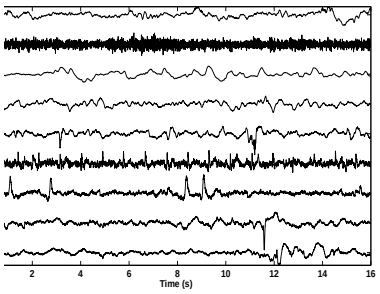
- For the COMBI algorithm, the denoising insertion distorts the informative sources (even with a low threshold algorithm), so the SI increases.



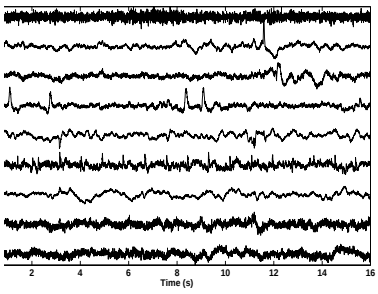
(a) semi - simulated sources



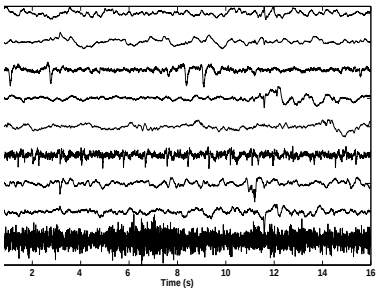
(b) simulated EEG



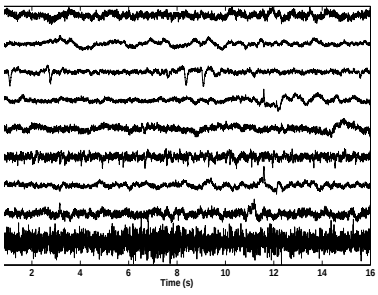
(c) COMBI without noise



(d) noisy COMBI



(e) SOBI-RO without noise



(f) noisy SOBI-RO

Figure 4: Simulation example

- Inserting the denoising before BSS damages the *SI* obtained without denoising by SOBI-RO also, and for all noise levels, confirming the previous analysis (probable source distortion).
- Without denoising, the results obtained with SOBI-RO are higher than those obtained with COMBI, both for denoised and noisy signals.

Table 2: Comparison of BSS algorithms on simulated EEG, before and after denoising

	COMBI		**SOBI-RO**	
no noise	**0.0428**		0.0941	
SNR	Noisy	Denoised	Noisy	Denoised
20dB	0.1266	0.1861	**0.1118**	0.1309
15dB	0.1509	0.2173	**0.1247**	0.1424
10dB	0.1777	0.2616	**0.1452**	0.1608
5dB	0.2029	0.3126	**0.1757**	0.1867

Therefore, our proposed method is based on the SOBI-RO (for this algorithm 100 time delays were considered) separation and the *Sure* denoising, with BSS applied on the raw signals (before denoising).

Considering the discussion section 3.2.2, the final order of the three processing steps is **BSS** → **Classification** → **Denoising**.

4.2. Real EEG validation

This subsection is dedicated to the evaluation of the presented method on a database of real signals. As said previously, two types of validation are possible: the first one aims to quantitatively evaluate the source classification performances, while the second is clinically oriented (blind medical validation of the "clean" reconstructed EEGs).

Data base. Thirty-eight scalp EEGs from 19 epileptic adults (aged 16 to 51), were recorded using a Micromed system at CHU-Nancy (2 by patient, 24 electrodes, 10-20 system, common reference at Fpz, notch filter at 50 Hz). Sampling frequency was 256 Hz. The data-base contains one sleep recording. Since the very low frequency artefacts (baseline shifts, slow ocular movements) perturb the source classification, a high-pass filter with a cut-off frequency of 1 Hz was applied.

The proposed method was applied on 20 seconds EEG sequences, in order to have a sufficient number of samples and thus a reliable BSS result. All recordings were seizure and inter-ictal spikes free.

By clinical inspection, the experts identified and labeled artefact and brain informative sources for the 38 electroencephalograms. Half of them were included in the **training set**, the other half being used to evaluate the performances of the classification.

4.2.1. Quantitative evaluation: source classification performances

The classifiers were evaluated using the sensitivity (*Sen*) and the false positive rate (*FPR*) (eq. 12), both on training and on testing subsets. Tables 3(a) and 3(b) show the global performances, all classes included, and the detailed performances by class ((a) Malahanobis distance, (b) Bayes classifier).

Table 3: Classifier performances (*Sen/FPR*) (a) Distance de Mahalanobis classifier (b) Bayes classifier

(a)

	Training		**Testing**	
	Sen	**FPR**	**Sen**	**FPR**
global	85.3%	4.9%	82.2%	5.9%
AO	95.4%	2.5%	85.2%	4.4%
HFA	77.6%	10%	68.1%	7.7%
AECG	100%	0%	100%	0%
B	87.4%	14.9%	89.0%	21.8%

(b)

	Training		**Testing**	
	Sen	**FPR**	**Sen**	**FPR**
global	80.2%	6.6%	78.3%	7.6%
AO	100%	6.4%	88.9%	7.9%
HFA	80.8%	13.6%	71.5%	11.9%
AECG	100%	0.2%	100%	0%
B	78.1%	10.4%	80.5%	16.1%

As it can be seen in tables 3(a) and 3(b), the two methods are quite performant (and thus implicitly validate our features choice). The global performances slightly better for the classifier based on Mahalanobis distance, who shows both the best *Sen* and the best *FPR*. This observation is valid both for testing and for training data, although the performances are slightly different.

For a more precise analysis of the two automatic methods, a detailed analysis by class is presented in tables 3(a) and 3(b). This analysis shows for the *AO* that the two classification methods make a good identification of the artefactual sources however, the Bayes classifier method increases the

good artefactual source identification but the false detections also (highest *Sen* and highest *FPR*). These observations are valid both for testing and training data, with of course better performances on the training subset.

The two automatic methods were also analyzed for the high frequency artefact source identification. However, in this case, the 2 methods have very similar sensitivity results. The Mahalanobis distance method presents the smallest *FPR* (less false detections), whereas the Bayes method presents the higher *Sen* (better artefactual source identification). Some of the sources identified by the experts as *HFA* are classified in other classes. In particular, there is an important *FPR* for the brain class: this result indicates that the two classes (*HFA* and *B*) are close in the feature space. This drawback should be nuanced, as the main objective is to have a *minimal loss* of brain information (although some of the EEGs reconstructed from the sources classified as brain might still be noisy).

For the ECG artefact identification the two methods in both subsets, testing and training groups, have an excellent performance with a good *Sen* rate. The *FPR* is equal to 0 for all the cases for the ECG artefact source identification. It must be noted here that very few *ECG* sources were identified by the experts, so this "perfect" classification should be validated on a larger data-base.

Finally, for the brain sources identification, the Bayes method has the smallest *FPR* (as well as the smallest *Sen*), that is this method identifies just a few false brain sources. On the opposite, the Mahalanobis distance method identifies correctly more brain sources than the Bayes method (highest *Sen*) but has a *FPR* slightly higher than the Bayes classifier. Interestingly, for the (only) EEG recording without artefacts (sleep), the classifier identified all sources as brain (no loss of information) [3].

As the classes are rather unbalanced (very few ECG artefact sources, mid-sized ocular and high-frequency artefact classes and a big brain class), one might use the *PPV* and the *NPV* (eq. 13) values to have a complementary view.

The *PPV* and *NPV* values presented in tables 4(a) and 4(b) confirm the previous *Sen/FPR* analysis and offer a complement of information. The most interesting case is the ocular artefact class (AO). According to the *PPV*

[3]The sources for this EEG (represented by the 8 dimensional feature vector $\mathbf{sf}$) are in the middle of the B class.

Table 4: Classifier performances (PPV/NPV) (a) Distance de Mahalanobis classifier (b) Bayes classifier

(a)

	Training		**Testing**	
	PPV	**NPV**	**PPV**	**NPV**
global	85.3%	95.1%	82.2%	94.1%
AO	65.6%	99.8%	54.8%	99.0%
HFA	74.6%	91.4%	80.3%	86.2%
$AECG$	100%	100%	100%	100%
B	92.0%	77.5%	86.8%	81.4%

(b)

	Training		**Testing**	
	PPV	**NPV**	**PPV**	**NPV**
global	80.2%	93.4%	78.3%	92.8%
AO	44.0%	100%	41.4%	99.2%
HFA	69.2%	92.3%	73.6%	81.0%
$AECG$	87.5%	100%	100%	100%
B	93.6%	67.6%	89.0%	72.6%

values, the number of false positives (FP) in the AO class is bigger for the Bayes classifier than for the Mahalanobis distance (as said previously). But one can notice that, quantitatively, there are more FP (false positives) than TP (true positives) in this class when using Bayes (this is not the case for the Mahalanobis classifier)! Considering the values for the other classes, one can conclude that a certain amount of brain labeled signals were classified as AO. The problem might be important, as it might lead to some information loss (*i.e.* brain sources elimination). Still, in practice, this phenomenon has very little influence on the final results (see also next subsection): first of all, as classes are unbalanced, rather few brain labeled sources are misclassified (high Sen for the brain class); second, this analysis deals with sources *labeled* by the neurologists and, as explained before, they avoided to label as artefacts sources that were not obviously artefacts. In other words, both automatic classifiers classed as AO sources having rather ambiguous characteristics, some of them maybe real ocular artefacts not selected by the neurologists.

Regardless of the classifier and of the employed criteria, global performances reflect mostly the brain source classification. This is quite expected, as the brain sources largely outnumber the artefact sources, so their good or bad classification rate highly influence the global classification.

A detailed comparison of the the two discriminant analysis classifiers (Bayes and Mahalanobis) reveals that for all 3 artefact classes, the Bayes classifier has a higher Sen but a highest FPR also (more artefacts are detected, but more brain sources are also labeled as artefacts). For brain sources, the situation is inverted (as expected). Generally, we can conclude that Bayes classifier should be used when 'brain only' sources are needed in the reconstruction (*i.e.* all artefacts must be eliminated). When this procedure is too rough (*i.e.* when all information should be preserved), the Mahalanobis based classifier is the best choice. Therefore, for the medical validation step (final validation of the processing chain), we used the reconstructed EEG after the Mahalanobis distance classifier.

4.2.2. Qualitative evaluation: medical validation of the reconstructed EEG

This last subsection presents the medical validation results in terms of visual interpretation of the reconstructed denoised artefact-free EEGs. A last issue is then the reconstruction method. The EEG reconstruction was made using the estimated matrix $\mathbf{A}$ ($\hat{\mathbf{A}}$) and the estimates sources $\mathbf{Y}$ that

is:

$$EEGrec = \sum_{i \in R} \hat{a}_i \hat{y}_i \tag{18}$$

where R represent the indices of the sources classified as brain sources. For consciousness, denoising was implemented either on the selected sources $\hat{\boldsymbol{y}}$ or on the reconstructed EEG signals. Both versions were visually analyzed by the neurologists, who found them completely equivalent (we remind that the this validation was made consensually by two neurologists).

Three versions of cleaned EEGs were reconstructed and compared to the raw EEG: after BSS and visual identification of the artefact sources by the experts (**clinical inspection**), after BSS and Mahalanobis classifier (**BSSC**) and after BSS, Mahalanobis classifier and wavelet denoising (**BSSCD**). All processing steps were developed using MATLAB ® 7.9.0. The processing time for each EEG (20 seconds long) was 2 seconds using an Intel CORE^{TM} 2 duo processor.

The clinical validation was made by blindly comparing the improvement between the raw EEG and each cleaned EEG (clinicians had no information on the noise and artefact removal method). The improvement was measured in two ways: the artefact removal quality and the physiological interpretation improvement. Clearly, they have a strong relation: an artefact free EEG allows a better interpretation.

As expected, the artefact sources having a good sensitivity (AO et ECG) were more easily eliminated. Still, the clinically validated results offer a complement of information, as they globally classify 24 channels reconstructed EEGs and not individual sources. For example, if all misclassified artefact sources belong to a single EEG, it would indicate that this particular EEG had a recording or separation problem. On the contrary, if the misclassified artefact sources are equally distributed over the whole EEG data-set, all reconstructions would be affected and none would be easier to interpret.

Consequently, clinicians noted the reconstructed EEGs, from an artefact elimination point of view, with: **1** no elimination, **2** partial elimination, **3** complete elimination, and, from an interpretation point of view, with: **1** much harder to interpret (loss of information), **2** harder (partially removed artefacts confusing the interpretation), **3** equal, **4** easier and **5** much easier. The raw EEG was always noted 3.

The percentage of EEGs by category is presented in the following table (5). The $AECG$ artefact elimination validation is not presented, as all these

sources were completely eliminated (Sen=1, tables 3(a) and 3(b), to compare with the results from [20], where the ECG was reduced in 98.4% of the EEGs).

Table 5: Artefact elimination and facility of interpretation performances

	Identification methods		
category	**Clinical inspection**	**BSSC**	**BSSCD**
	***AO* elimination**		
3	42% (16)	50% (19)	50% (19)
2	50% (19)	48% (18)	48% (18)
1	8% (3)	2% (1)	2% (1)
	***HFA* elimination**		
3	13% (5)	21% (8)	63% (24)
2	66% (25)	63% (24)	35% (13)
1	21% (8)	16% (6)	2% (1)
	Facility of interpretation		
5	32% (12)	24% (9)	48% (18)
4	39% (15)	63% (24)	48% (18)
3	18% (7)	8% (3)	2% (1)
2	11% (4)	5% (2)	2% (1)
1	0%	0%	0%

A first conclusion appears from table 5: the denoising doesn't contribute to the *AO* elimination (98% completely or partially eliminated both by **BSSC** and **BSSCD**). Similar results were obtained by [20] (96.8%), but with the help of an electrooculogram. On the contrary, the *HFA* was the most difficult to identify (as it was for the experts also). The results obtained by clinical inspection and **BSSC** are very similar (79% and 84% respectively), with a slightly better performance for the latter. On the contrary, the denoising step improves considerably the *HFA* elimination. As it can be seen, the percentage of EEG showing a total elimination of the *HFA* (noted 3) increases from 21% to 63% and the percentage of the overall improvement (categories 2 and 3) increases from 84% to 98%.

Concerning the facility of interpretation, **BSSC** seems to offer better performances than the clinical inspection method (87% compared to 71%). This rather surprising result is coherent with the artefact elimination results from the same table 5: some of the artefact sources were not visually identified by the experts during the labelling step, but the automatic method correctly

classified them. This can be explained, as said previously, by the fact that clinical experts have only chosen the evident artefact sources to avoid potential loss of information (see also quantitative results of *Sen*, *FPR*, *PPV* and *NPV* in the previous subsection). As for *HFA* elimination, further denoising (**BSSCD**) increases the percentage of EEGs noted 5 from 24% to 48% and the overall improvement percentage (categories 4 and 5) increases from 87% to 96%.

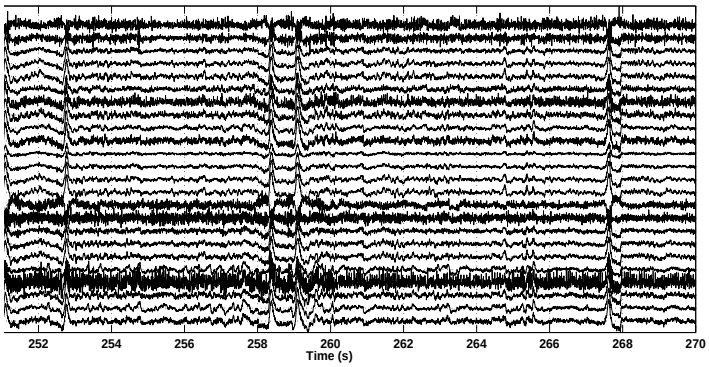
An example of EEG processing is shown in figure 5.

5. Conclusion and future research

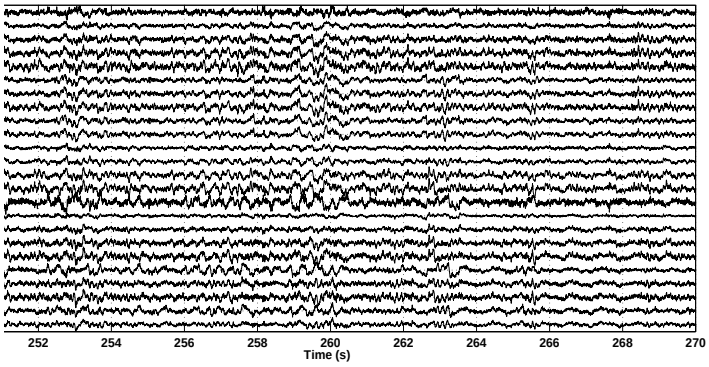
In this paper, we propose a method for eliminating several types of artefacts and noise based on blind source separation (SOBI-RO), wavelet denoising (*SureShrink*) and supervised classification (Mahalanobis). The analysis of the interaction between the three methods yielded an optimal preprocessing chain, validated on simulated and real signals. In particular, we show that BSS must be performed on the noisy raw signals, *i.e.* before wavelet denoising. Concerning the choice of the BSS algorithm, our simulations (on highly realistic signals) led to the conclusion that in normal clinical setups (noisy environment) second order statistics algorithms perform better than HOS algorithms. This conclusion is supported by extensive simulation results (see also recent studies by [13, 22]).

The obtained preprocessing chain was applied on inter-ictal EEG and was compared with a semi-automatic approach based on BSS and visual classification. The results show a significant improvement of both artefact elimination and EEG interpretation. According to the clinicians, high frequency artefact sources (muscles and noise) are difficult to identify (visually, thus also by supervised classification). Moreover, it is also possible that BSS fails to separate these sources from the brain activity. The WD step introduction addresses this problem, with very good results from an interpretation point of view (with only 2% of the EEGs showing a loss of information, compared to 11% for clinical inspection).

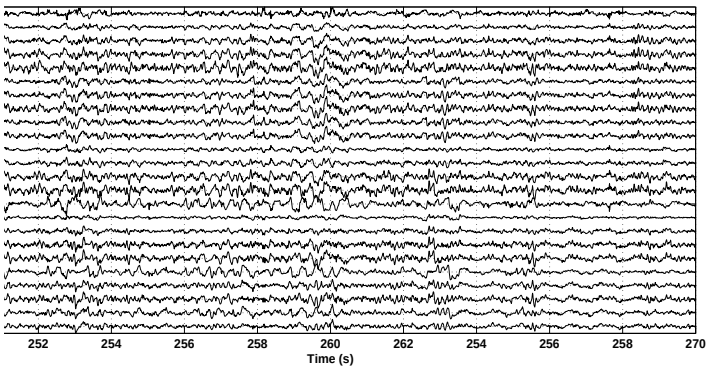
The proposed method can be applied on any EEG window, seizure included. First results of "clean" ictal EEG, obtained using the same feature space and training set as for the inter-ictal case, are very encouraging. Moreover, the method can be adapted further, in order to take into account the specificities of the seizure signals (*i.e.*, new classes or new features dedicated to seizure sources).



(a) raw EEG



(b) **BSSC**, noted 4



(c) **BSSCD**, noted 5

Figure 5: Raw and clean reconstructed EEGs

References

[1] S. Sanei, J. Chambers, EEG Signal Processing, John Wiley & Sons, 2007.

[2] R. Croft, R. Barry, Removal of ocular artifact from the EEG: a review, Neurophysiologie Clinique/Clinical Neurophysiology 30 (1) (2000) 5–19.

[3] A. Cichocki, S. Amari, Adaptive Blind Signal and Image Processing Learning Algorithms and Applications, John Wiley & Sons, New York, USA, 2002.

[4] A. Delorme, T. Sejnowski, S. Makeig, Enhanced detection of artifacts in EEG data using higher-order statistics and independent component analysis, NeuroImage 34 (2007) 1443–1449.

[5] A. Greco, N. Mammone, F. Morabito, M. Versaci, Kurtosis, Renyi's entropy and independent component scalp maps for the automatic artifact rejection from EEG data, International Journal of Signal Processing 2 (4) (2006) 240–244.

[6] Y. Li, Z. Ma, W. Lu, Y. Li, Automatic removal of the eye blink artifact from EEG using an ICA-based template matching approach, Physiological Measurement 27 (4) (2006) 425.

[7] P. LeVan, E. Urrestarazu, J. Gotman, A system for automatic artifact removal in ictal scalp EEG based on independent component analysis and Bayesian classification, Clinical Neurophysiology 117 (4) (2006) 912–927.

[8] C. James, O. Gibson, Temporally constrained ICA: an application to artifact rejection in electromagnetic brain signal analysis, IEEE Transactions on Biomedical Engineering 50 (9) (2003) 1108–1116.

[9] J. Kierkels, G. van Boxtel, L. Vogten, A model-based objective evaluation of eye movement correction in EEG recordings, IEEE Transactions on Biomedical Engineering 53 (2) (2006) 246–253.

[10] S. Romero, M. Mañanas, M. Barbanoj, A comparative study of automatic techniques for ocular artifact reduction in spontaneous EEG signals based on clinical target variables: A simulation case, Computers in Biology and Medicine 38 (3) (2008) 348–360.

[11] J. Escudero, R. Hornero, D. Abasolo, A. Fernandez, M. Lopez-Coronado, Artifact removal in magnetoencephalogram background activity with independent component analysis, IEEE Transactions on Biomedical Engineering 54 (11) (2007) 1965–1973.

[12] K. Ting, P. Fung, C. Chang, F. Chan, Automatic correction of artifact from single-trial event-related potentials by blind source separation using second order statistics only, Medical Engineering and Physics 28 (8) (2006) 780–794.

[13] M. Klemm, J. Haueisen, G. Ivanova, Independent component analysis: comparison of algorithms for the investigation of surface electrical brain activity, Med. Biol. Eng. Comput. 47 (2009) 413–423.

[14] A. Belouchrani, K. Abed-Meraim, J. Cardoso, E. Moulines, A blind source separation technique using second-order statistics, IEEE Transactions on signal processing 45 (2) (1997) 434–444.

[15] A. Belouchrani, A. Cichocki, Robust whitening procedure in blind source separation context, Electronics Letters 36 (24) (2000) 2050–2053.

[16] R. Vigario, Extraction of ocular artefacts from EEG using independent component analysis, Electroencephalography and Clinical Neurophysiology 103 (3) (1997) 395–404.

[17] C. Melissant, A. Ypma, E. Frietman, C. Stam, A method for detection of Alzheimer's disease using ICA-enhanced EEG measurements, Artificial Intelligence in Medicine 33 (3) (2005) 209–222.

[18] G. Wallstrom, R. Kass, A. Miller, J. Cohn, N. Fox, Automatic correction of ocular artifacts in the EEG: a comparison of regression-based and component-based methods, International journal of psychophysiology 53 (2) (2004) 105–119.

[19] N. Nicolaou, S. Nasuto, Automatic artefact removal from event-related potentials via clustering, Journal of VLSI Signal Processing 48 (1) (2007) 173–183.

[20] S. Shao, K. Shen, C. Ong, E. Wilder-Smith, X. Li, Automatic EEG Artifact Removal: A Weighted Support-Vector-Machine Approach With

Error Correction, IEEE Transactions on Biomedical Engineering 56 (2) (2009) 336 – 344.

[21] J. Dammers, M. Schiek, F. Boers, C. Silex, M. Zvyagintsev, U. Pietrzyk, K. Mathiak, Integration of amplitude and phase statistics for complete artifact removal in independent components of neuromagnetic recordings, IEEE Transactions on Biomedical Engineering 55 (10) (2008) 2353–2362.

[22] J. Escudero, R. Hornero, D. Abasolo, Consistency of the blind source separation computed with five common algorithms for magnetoencephalogram background activity, Medical Engineering & Physics 32 (10) (2010) 1137–1144.

[23] A. Antoniadis, J. Bigot, T. Sapatinas, Wavelet estimators in nonparametric regression: a comparative simulation study, Journal of Statistical Software 6 (6) (2001) 1–83, `http://www-lmc.imag.fr/lmc-sms/Anestis.Antoniadis/HTTP/publis-anto.html`.

[24] A. Antoniadis, Wavelet methods in statistics: Some recent developements and their applications, Statistics Surveys 1 (2007) 16–55.

[25] D. Donoho, I. Johnstone, Ideal spatial adaptation via wavelet shrinkage, Biometrika 81 (1994) 425–455.

[26] D. Donoho, I. Johnstone, Adapting to unknown smoothness via wavelet shrinkage, Journal of the American Statistical Association 90 (1995) 1200–1224.

[27] R. Romo-Vàzquez, Contribution à la détection et à l'analyse des signaux eeg épileptiques : débruitage et séparation de sources, Ph.D. thesis, Institut National Polytechnique de Lorraine, Nancy - France (2010).

[28] V. Vigneron, A. Paraschiv-Ionescu, A. Azancot, O. Sibony, C. Jutten, Fetal electrocardiogram extraction based on non-stationary ICA and wavelet denoising, in: 7th IEEE International Symposium on Signal Processing and its Applications, Paris, France, 2003.

[29] M. Berryman, S. Messer, A. Allison, D. Abbott, Techniques for noise removal from EEG, EOG, and airflow signals in sleep patients, in: Proceedings-SPIE, Vol. 5467, 2004, pp. 89–97.

[30] Y. Rong-Yi, C. Zhong, Blind source separation of multichannel electroencephalogram based on wavelet transform and ICA, Chinese Physics 14 (11) (2005) 2176–2180.

[31] N. Castellanos, V. Makarov, Recovering EEG brain signals: artifact suppression with wavelet enhanced independent component analysis, Journal of neuroscience methods 158 (2) (2006) 300–312.

[32] H. Ghandeharion, A. Erfanian, A fully automatic ocular artifact suppression from eeg data using higher order statistics: Improved performance by wavelet analysis, Medical Engineering & Physics 32 (7) (2010) 720 – 729.

[33] B. Rivet, V. Vigneron, A. Paraschiv-Ionescu, C. Jutten, Wavelet Denoising for Blind Source Separation in Noisy Mixtures, in: Independent component analysis and blind signal separation: fifth international conference, ICA 2004, Granada, Spain, 2004.

[34] A. Cichocki, S. Amari, K. Siwek, T. Tanaka, A. H. Phan, ICAlab toolboxes, `http://www.bsp.brain.riken.jp/ICALAB` (2009).

[35] A. Delorme, S. Makeig, EEGLAB Wikitorial, `http://sccn.ucsd.edu/wiki/EEGLAB_TUTORIAL_OUTLINE` (2009).

[36] L. Tong, V. Soon, Y. F. Huang, R. Liu, Indeterminacy and identifiability of blind identification, IEEE Trans. CAS 38 (1991) 499–509.

[37] P. Georgiev, A. Cichocki, Blind source separation via symmetric eigenvalue decomposition, in: in Proceedings of Sixth International Symposium on Signal Processing and its Applications, Kuala Lumpur, Malaysia, 2001, pp. 17–20.

[38] A. Belouchrani, K. Abed-Meraim, J. Cardoso, E. Moulines, Second-order blind separation of temporally correlated sources, in: Proc. Int. Conf. Digital Signal Processing, 1993, pp. 346–351.

[39] A. Cichocki, A. Belouchrani, Sources separation of temporally correlated sources from noisy data using a bank of band-pass filters, in: Proc. of Third International Conference on Independent Component Analysis and Signal Separation (ICA-2001), San Diego, USA, 2001, pp. 173–178.

[40] A. Yeredor, Blind separation of gaussian sources via second-order statistics with asymptotically optimal weigthting, IEEE Signal Processing Letters 7 (2000) 197–200.

[41] P. Tichavský, Z. Koldovský, E. Doron, A. Yeredor, G. Gomez-Herrero, Blind signal separation by combining two ica algorithms: Hos-based efica and time structure-based wasobi, in: Proceedings of The 2006 European Signal Processing Conference (EUSIPCO'2006), 2006.

[42] S. Choi, A. Cichocki, A. Belouchrani, Second order nonstationary source separation, Journal of VLSI Signal Processing 32 (1-2) (2002) 93–104.

[43] S. Makeig, A. Bell, T.P.-Jung, T. J. Sejnowski, Independent component analysis of electroencephalographic data, in: D. Touretzky, M. Mozer, M. Hasselmo (Eds.), Advances in Neural Information Processing Systems, Vol. 8, MIT Press, Cambridge, MA, 1996, pp. 145–151.

[44] J. Cardoso, A. Souloumiac, Blind beamforming for non Gaussian signals, IEE Proceedings-F 40 (6) (1993) 362–370.

[45] K. Müller, P. Philips, A. Ziehe, Jadetd: Combining higher-order statistics and temporal information for blind source separation with noise, in: Proc. Int. Workshop on Independent Component Analysis and Blind Separation of Signals (ICA '99), Aussois, 1999.

[46] Z. Koldovský, P. Tichavský, E. Oja, Efficient variant of algorithm fastica for independent component analysis attaining the cramér-rao lower bound, IEEE Trans. on Neural Networks 17 (5) (2006) 1265–1277.

[47] P. Tichavský, Z. Koldovský, A. Yeredor, G. G. Herrero, E. Doron, A hybrid technique for blind non-gaussian and time-correlated sources using a multicomponent approach, Neural Networks, IEEE Transactions 19 (3) (2008) 421–430.

[48] R. Romo-Vázquez, R. Ranta, V. Louis-Dorr, D. Maquin, EEG ocular artefacts and noise removal, in: 29th Annual International Conference of the IEEE Engineering in Medicine and Biology Society, EMBC'07, 2007.