



**HAL**  
open science

# A Penalized Semialgebraic Deflation ICA Algorithm for the Efficient Extraction of Interictal Epileptic Signals

Hanna Becker, Laurent Albera, Pierre Comon, Amar Kachenoura, Isabelle Merlet

► **To cite this version:**

Hanna Becker, Laurent Albera, Pierre Comon, Amar Kachenoura, Isabelle Merlet. A Penalized Semi-algebraic Deflation ICA Algorithm for the Efficient Extraction of Interictal Epileptic Signals. *IEEE Journal of Biomedical and Health Informatics*, 2017, 21 (1), pp.94-104. 10.1109/JBHI.2015.2504126 . hal-01246030

**HAL Id: hal-01246030**

**<https://hal.science/hal-01246030>**

Submitted on 17 Dec 2015

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

# A penalized semi-algebraic deflation ICA algorithm for the efficient extraction of interictal epileptic signals

Hanna Becker, Laurent Albera, *Senior Member, IEEE*, Pierre Comon, *Fellow, IEEE*, Amar Kachenoura, and Isabelle Merlet

**Abstract**—As a non-invasive technique, Electroencephalography (EEG) is commonly used to monitor the brain signals of patients with epilepsy such as the interictal epileptic spikes. However, the recorded data are often corrupted by artifacts originating, for example, from muscle activities, which may have much higher amplitudes than the interictal epileptic signals of interest. To remove these artifacts, a number of Independent Component Analysis (ICA) techniques were successfully applied. In this paper, we propose a new deflation ICA algorithm, called Penalized Semi-Algebraic Unitary Deflation (P-SAUD) algorithm, that improves upon classical ICA methods by leading to a considerably reduced computational complexity at equivalent performance. This is achieved by employing a penalized semi-algebraic extraction scheme, which permits us to identify the epileptic components of interest (interictal spikes) first and obviates the need of extracting subsequent components. The proposed method is evaluated on physiologically plausible simulated EEG data and actual measurements of three patients. The results are compared to those of several popular ICA algorithms as well as second order blind source separation methods, demonstrating that P-SAUD extracts the epileptic spikes with the same accuracy as the best ICA methods, but reduces the computational complexity by a factor of 10 for 32-channel recordings. This superior computational efficiency is of particular interest considering the increasing use of high-resolution EEG recordings, whose analysis requires algorithms with low computational cost.

**Index Terms**—deflation ICA, EEG, denoising, artifact removal, epilepsy.

## I. INTRODUCTION

Electroencephalography (EEG) is a non-invasive and cheap technique that allows for the monitoring of brain signals with a high temporal resolution, making it a routinely used method for the diagnosis and management of epilepsy. Epilepsy is one of the most common neuronal diseases and leads to temporary dysfunctions of the electrical brain activity, the epileptic seizures, as a result of sudden abnormal electric discharges, called paroxysmal discharges. Two different types

of epileptic paroxysms can be recorded. During seizures, ictal discharges last several seconds to a few minutes and are characterized by a rhythmic activity. Between seizures, brief paroxysms, called interictal spikes, occur in irregular intervals and are considered as interictal signatures of the underlying pathology. Interictal spikes can be frequently observed on the scalp with a high Signal-to-Noise-Ratio (SNR), whereas the ictal discharges occur more randomly. Therefore, the analysis of EEG recordings of epileptic patients includes the study of interictal epileptic activity. However, the epileptic activity recorded by surface electrodes is usually corrupted by eye and muscle artifacts that hinder its interpretation. Thus, before the interictal epileptic data can be properly analyzed, the first step consists in removing these artifacts. As little information on the underlying sources is available a priori, this is a typical application for Blind Source Separation (BSS) methods [1], [2].

Due to their different physiological origins, it is reasonable to assume that the artifacts are statistically independent of the interictal epileptic activity. This motivates the application of BSS methods that are based on Second Order (SO) statistics of the measurement data only, exploiting uncorrelatedness of the source signals, and of Independent Component Analysis (ICA), which also involves higher order statistics. In fact, SO methods such as SOBI [3], CCA [4], and Generalized Eigenvalue Decomposition (GEVD)-based techniques [5], [6], and a large number of ICA algorithms have been successfully applied to separate the artifacts from neuronal activities (see, e.g., [7], [8], [4], [9]). Note that the method proposed in [6] is dedicated to ictal epileptic signals while the one presented in [5] requires a higher SNR compared with ICA techniques. The ICA algorithms can be distinguished into two types of approaches: joint methods, that separate all sources simultaneously, and deflation methods, which extract the sources sequentially. Among the joint methods, one can find popular algorithms such as Infomax [10], JADE [11], and CoM<sub>2</sub> [12], whereas deflation methods include, for example, the well-known FastICA algorithm<sup>1</sup> [13] and RobustICA [14].

Recent studies [15], [16] have compared the performance of a number of popular ICA algorithms for EEG denoising. The authors of [15] concluded that CoM<sub>2</sub> [12] is the method that yields the best compromise between performance and

H. Becker is with Univ. Nice Sophia Antipolis, CNRS, I3S, UMR 7271, Sophia Antipolis, France and GIPSA-Lab, CNRS UMR5216, Grenoble Campus, St Martin d’Heres, France.

L. Albera is with INSERM, U1099, Rennes, F-35000, France, Université de Rennes 1, LTSI, Rennes, France, and INRIA, Centre Inria Rennes - Bretagne Atlantique, France.

P. Comon is with GIPSA-Lab, CNRS UMR5216, Grenoble Campus, St Martin d’Heres, France.

A. Kachenoura and I. Merlet are with INSERM, U1099, Rennes, F-35000, France and Université de Rennes 1, LTSI, Rennes, France.

The work of H. Becker was supported by Conseil Régional PACA and by CNRS. P. Comon is supported by the FP7 European Research Council Programme, DECODA project, under grant ERC-AdG-2013-320594.

<sup>1</sup>Please note that a version of the FastICA algorithm which jointly extracts all sources also exists.

computational complexity. However, this method extracts all sources simultaneously and considering a sufficiently large number of sources, this can easily lead to the separation of a hundred signals, resulting in a high computational burden of ICA in the context of high-resolution EEG. Thus, the denoising of large amounts of EEG recordings remains time-consuming even with modern high-performance computers. Considering that the use of high-resolution EEG systems can be expected to increase over the next few years, it is desirable to develop more computationally efficient algorithms to limit the requirements on computing resources. Since we are only interested in the interictal epileptic activity (interictal spikes), this could be achieved by limiting the algorithm to the extraction of the interictal spikes.

To extract a reduced number of ICA components, deflation methods can be employed. However, a remaining difficulty in deflationary approaches consists in ensuring that the signals of interest are extracted first such that the algorithm can be stopped after the extraction of a small number of components from the mixture. This requires the exploitation of prior knowledge about the signals of interest. In [17], the constrained ICA (cICA) framework has been developed to this end and ICA methods that work with a reference signal, generally referred to as ICA-R, have been put forward [17], [18], [19] to extract the signals with the highest resemblance to the references. These methods are based on a Newton-like learning scheme to solve the constrained optimization problem.

In this paper, we propose a new deflation ICA algorithm that improves upon the deflation method presented in [20] (in the following referred to as DeIL<sup>R</sup>) by resorting to a semi-algebraic optimization scheme based on a two-source contrast and alternating Jacobi iterations [12]. Furthermore, in the proposed ICA algorithm, a penalization term is added in the contrast function to ensure that the epileptic sources are extracted first. Contrary to the ICA-R approach, that relies on a reference signal, which is not always available in practice, to extract the sources of interest, we make use of the temporal structure of the source time signals. More particularly, we exploit the high autocorrelation of the epileptic spikes compared to muscle artifacts, leading to a penalization term that is related to the CCA approach. The proposed method, called Penalized Semi-Algebraic Unitary Deflation (P-SAUD) algorithm, is compared to DeIL<sup>R</sup> as well as CoM<sub>2</sub>, FastICA, CCA, and SOBI in terms of performance and computational complexity.

This paper is organized as follows: in Section II, we describe the data model and present the proposed method. We also describe the simulation setup and the real data that is analyzed in this paper and provide some background information on the patients in which the data were recorded. Section III contains the results obtained on the simulated and the real EEG data. Conclusions are drawn in Section IV.

## II. METHODS

### A. Problem formulation

In this paper, we assume that the measurements  $\tilde{\mathbf{x}}[t] \in \mathbb{R}^N$ ,  $t = 1, \dots, T$ , of the electric potential recorded by  $N$  sensors placed on the scalp for  $T$  time samples constitute a

mixture of epileptic activity  $\tilde{\mathbf{s}}_e[t]$ , muscular activity  $\tilde{\mathbf{s}}_m[t]$ , and background activity  $\tilde{\mathbf{s}}_b[t]$  of the brain. Fig. 1 displays examples of these physiological signals, showing in particular the characteristic form of an interictal epileptic spike, which has a high autocorrelation, and the appearance of muscle artifacts, which resemble uncorrelated Gaussian noise. The propagation of the source signals to the surface is characterized by the matrix  $\mathbf{A}^{(e)} \in \mathbb{R}^{N \times P_e}$  for the epileptic sources, the matrix  $\mathbf{A}^{(m)} \in \mathbb{R}^{N \times P_m}$  for the muscle artifacts, and the matrix  $\mathbf{A}^{(b)} \in \mathbb{R}^{N \times P_b}$  for the background activity. Furthermore, we consider the instrumentation noise  $\tilde{\mathbf{n}}[t]$  that is added during the measurement process. This leads to the following model for the EEG data:

$$\tilde{\mathbf{x}}[t] = \mathbf{A}^{(e)}\tilde{\mathbf{s}}_e[t] + \mathbf{A}^{(m)}\tilde{\mathbf{s}}_m[t] + \mathbf{A}^{(b)}\tilde{\mathbf{s}}_b[t] + \tilde{\mathbf{n}}[t]. \quad (1)$$

As the number of sources  $P_e + P_m + P_b$  generally exceeds the number of sensors  $N$ , the mixture (1) is underdetermined. In this paper, our objective is to separate the epileptic signals from the muscle and background activity.

### B. ICA approach

Let us assume that the data  $\{\tilde{\mathbf{x}}[t]\}$  constitute one finite-length realization of a random vector process  $\{\mathbf{x}[t]\}$ . Since epileptic, muscle, and background activity have different physiological origins, their signals can be assumed to be represented by random vector processes that are statistically independent. This property can be exploited for the separation of the three types of activity such that the underlying random vector processes are maximally statistically independent. This is the objective of ICA, which computes the decomposition:

$$\mathbf{x}[t] = \mathbf{A}\mathbf{s}[t] \quad (2)$$

where  $\mathbf{A} \in \mathbb{R}^{N \times P}$  is the mixing matrix. The  $P \leq N$  components of the random vector process  $\{\mathbf{s}[t]\}$  are mutually independent and can be divided into three subgroups that form bases for the signal subspaces of the epileptic, muscle, and background activity. Although the different sources of the same type of activity might not be independent, ICA still permits us to separate the subspaces of the different types of activity if the following three conditions hold: i) the three subspaces are statistically independent, ii) the sum of their dimensions does not exceed the number of sensors, and iii) the instrumentation noise is Gaussian. Note that the independent components can only be extracted up to a scale and permutation indeterminacy.

The identification of independent components by linear transformation is accomplished by maximizing a contrast function. In this paper, we concentrate on the following two-source contrast, introduced in [12], that is used to achieve statistical independence of a pair of random variables  $y$  and  $z$ :

$$\psi(y, z) = C_{4,y}^2 + C_{4,z}^2. \quad (3)$$

Here  $C_{4,y}$  denotes the Fourth Order (FO) cumulant of the random variable  $y$ . To facilitate the separation of the sources, we work on the prewhitened data, that are obtained as:

$$\mathbf{z}[t] = \mathbf{F}^+ \mathbf{x}[t]. \quad (4)$$

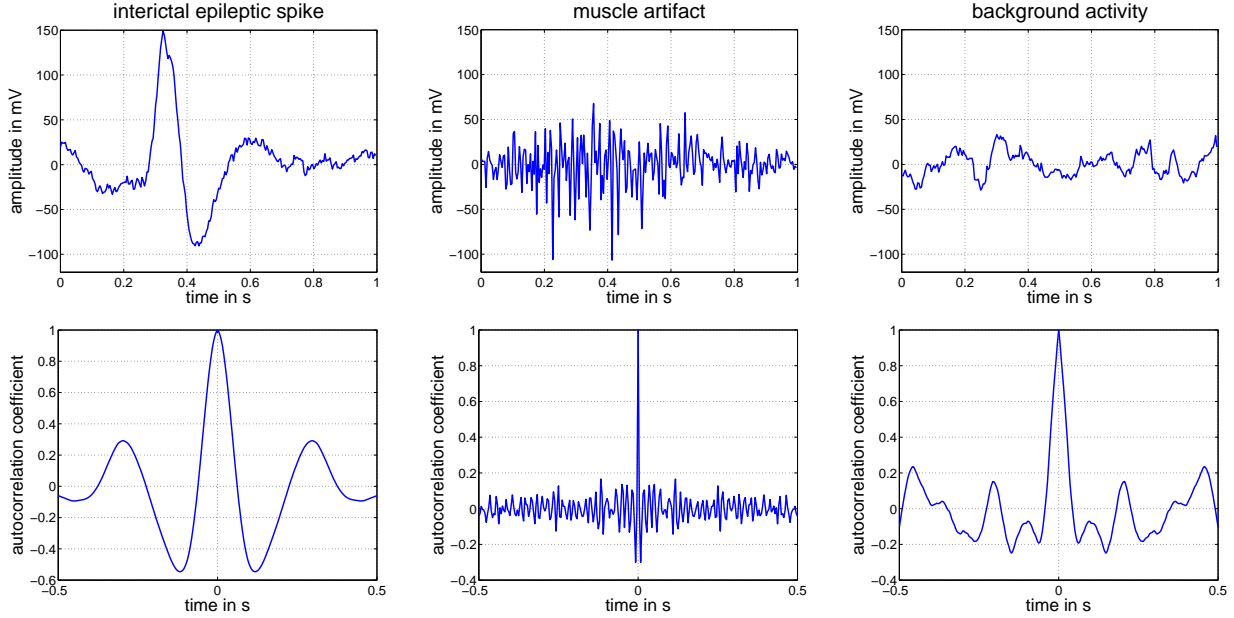


Fig. 1. Example of an interictal epileptic spike (left), a muscle artifact (middle), and background activity of the brain (right) and their autocorrelation functions.

Here,  $\mathbf{F}^+$  denotes the Moore-Penrose pseudo inverse of the matrix  $\mathbf{F} \in \mathbb{R}^{N \times P}$ , which corresponds to a square root of the (possibly denoised) data covariance matrix  $\mathbf{C}_x$  of  $\{\mathbf{x}[t]\}$ . This has the advantage that the sought mixing matrix becomes an orthonormal matrix, subsequently called  $\mathbf{H} \in \mathbb{R}^{P \times P}$ . Once an estimate  $\hat{\mathbf{H}}$  of this matrix has been determined, the original mixing matrix can be obtained as  $\hat{\mathbf{A}} = \mathbf{F}\hat{\mathbf{H}}$ . Furthermore, we employ a parameterization of the mixing vectors, which is described in appendix A. This reduces the identification of the  $p$ -th mixing vector to the estimation of a certain number of Givens rotation angles  $\phi_{p,k}$ .

### C. The P-SAUD algorithm

1) *Presentation of the algorithm:* The idea of the P-SAUD algorithm consists in extracting the sources consecutively by determining the optimal rotation angles for all signal pairs. More particularly, to extract the  $p$ -th source, the algorithm updates the rotation angles  $\phi_{p,k}$  for  $k = 1, \dots, P - p$  alternatingly over several sweeps. After each deflation step, the dimension of the ICA problem is decreased by resorting to the orthogonal projection approach used by DeLL<sup>R</sup> as explained below. In order to ensure that the epileptic activity is extracted first, we employ the following penalized contrast function, which exploits the fact that the autocorrelation of the interictal spikes is higher than that of the muscle artifacts:

$$\psi_c(s_p, s_k) = C_{4,s_p}^2 + C_{4,s_k}^2 + \lambda \text{cum}(s_p[t], s_p[t + \tau])^2. \quad (5)$$

Here,  $\text{cum}(a, b)$  stands for the SO cumulant (covariance) of random variables  $a$  and  $b$ . Furthermore,  $\lambda$  is a penalization parameter that determines the influence of the covariance penalty term and  $\tau$  denotes a fixed signal delay. In practice, the value of the penalization parameter  $\lambda$  needs to be adjusted depending on the kurtosis  $K_1$  and the autocorrelation  $A_1$  of the source to extract. As the magnitudes of these factors are

generally unknown, we propose to estimate them based on the signal  $\{s_p[t]\}$  retrieved at the previous iteration. To adjust the balance between original contrast and penalization, we introduce the relative penalization parameter  $\alpha$ : for  $\alpha < 1$ , the original contrast dominates, for  $\alpha > 1$  the autocorrelation preponderates, and for  $\alpha = 1$  both terms are equally important. The penalization parameter  $\lambda$  is then given by  $\lambda = \alpha \frac{K_1}{A_1}$ . In order to ensure that the epileptic activity is extracted first, we use a high value of  $\alpha$  during the first iterations ( $\alpha_{max} > 1$ ) and reduce the penalization parameter with increasing number of iterations until it reaches a final value that manages a balance between the two-source contrast (3) and the penalization term (generally  $0 \leq \alpha_{min} \leq 1$ ). The signal delay  $\tau$  has to be adjusted depending on the autocorrelation profile of the epileptic spikes and may range from 1 to several samples (such that the autocorrelation between the epileptic signal and its delayed version is still high). For the simulations (cf. Section III-A), we chose  $\tau = 1$  as is typically used for CCA, whereas we observed that higher delays (e.g.,  $\tau = 5$ ) lead to more robust results on the real data examples (see Section III-B).

The signals  $\{s_p[t]\}$  and  $\{s_k[t]\}$  depend on the parameter  $\phi_{p,k}$  that characterizes the Givens rotation. Setting  $\theta_{p,k} = \tan(\phi_{p,k})$ , the signals are obtained from the corresponding elements of the data vector as:

$$\begin{bmatrix} s_k[t] \\ s_p[t] \end{bmatrix} = \sqrt{\frac{1}{1 + \theta_{p,k}^2}} \begin{bmatrix} 1 & -\theta_{p,k} \\ \theta_{p,k} & 1 \end{bmatrix} \begin{bmatrix} z_k[t] \\ z_{P-p+1}[t] \end{bmatrix}. \quad (6)$$

The determination of the parameter  $\theta_{p,k}$  that maximizes (5) is detailed in Appendix B.

For the extraction of the first source, i.e.,  $p = 1$ , we initialize the matrix  $\mathbf{G}^{(1)} = \mathbf{I}_P$  and the vector  $\mathbf{y}^{(1)} = \mathbf{z}$  that correspond to a temporary unmixing matrix estimate and a temporary data vector, respectively. Then, the Givens rotation matrix  $\mathbf{G}_g^{(1,1)}(\theta_{1,1})$  for the first pair of sources,  $p = 1$ ,  $k = 1$ , is computed based on the value  $\theta_{1,1}$  that maximizes the contrast (5).

The temporary matrix  $\mathbf{G}^{(1)}$  and vector  $\mathbf{y}^{(1)}$  are then updated by replacing them by the matrix  $\mathbf{G}_g^{(1,1)}\mathbf{G}^{(1)}$  and the vector  $\mathbf{G}_g^{(1,1)}\mathbf{y}^{(1)}$ , respectively. Thereafter, the algorithm proceeds with the estimation of the optimal Givens rotation matrix for the next pair of sources, i.e., for  $k = 2$ , which is based on the new temporary matrix and vector. The procedure is repeated for all pairs of sources with  $p = 1$  and  $k = 1, \dots, P - p$ , and over  $I$  sweeps. At the end, the mixing vector for the first source and the corresponding signal vector are extracted as:

$$\hat{\mathbf{h}}_1 = \mathbf{g}^{(1)} \quad (7)$$

$$\hat{\mathbf{s}}_1[t] = \hat{\mathbf{h}}_1^T \mathbf{z}[t] \quad (8)$$

where  $\mathbf{g}^{(1)T}$  denotes the last row of the matrix  $\mathbf{G}^{(1)}$ .

As the mixing matrix is orthonormal, the mixing vector of the  $p$ -th source is sought within the subspace that is orthogonal to all previously extracted mixing vectors. To this end, to extract the  $p$ -th component,  $p = 2, \dots, P$ , we first find the matrix  $\tilde{\mathbf{G}}^{(p-1)}$  and the vector  $\tilde{\mathbf{y}}^{(p-1)}$  such that:

$$\mathbf{G}^{(p-1)} = \begin{bmatrix} \tilde{\mathbf{G}}^{(p-1)T} \\ \mathbf{g}^{(p-1)T} \end{bmatrix} \quad \text{and} \quad \mathbf{y}^{(p-1)} = \begin{bmatrix} \tilde{\mathbf{y}}^{(p-1)} \\ y_1^{(p-1)} \end{bmatrix}$$

and initialize the temporary mixing matrix and data vector for the  $p$ -th source by  $\mathbf{G}^{(p)} = \tilde{\mathbf{G}}^{(p-1)T}$  and  $\mathbf{y}^{(p)} = \tilde{\mathbf{y}}^{(p-1)}$ . The estimation then follows the same lines as those described above for the first source. The mixing and signal vectors of the  $p$ -th source are extracted as:

$$\hat{\mathbf{h}}_p = \mathbf{g}^{(p)} \quad (9)$$

$$\hat{\mathbf{s}}_p[t] = \hat{\mathbf{h}}_p^T \mathbf{z}[t]. \quad (10)$$

This procedure is inspired by the deflation algorithm described in [20]. The algorithm is stopped after  $M$  components have been extracted (including the sources of interest). To automatically determine the number of components to extract, one could employ a spike detection method (see, e.g., [21]) that is run on the extracted components and stop the algorithm once several components without epileptic activity have been identified. Alternatively, the algorithm could be stopped as soon as the autocorrelation of the extracted signals falls below a given threshold. The P-SAUD algorithm is summarized in Fig. 2.

2) *Analysis of the computational complexity:* To determine the computational complexity of P-SAUD, we analyze the number of floating point operations (FLOPs) that are required for its completion. As the number of additions is usually of the same order as the number of multiplications, subsequently, we consider only the number of multiplications.

The first step of P-SAUD consists in prewhitening the data, which can be accomplished either by an Eigenvalue Decomposition (EVD) or by a Singular Value Decomposition (SVD), leading to a computational complexity of  $\min(TN^2/2 + 4N^3/3 + PMT, 2TN^2)$  FLOPs [15].

The highest computational cost of P-SAUD can be associated with the estimation of the SO and FO cumulants that are exploited in the penalized contrast. These can be obtained in two different ways: i) by estimating the cumulants for each pair of analyzed sources from the temporary data vectors at each iteration, or ii) by estimating first the complete cumulant

```

Prewhitening:  $\mathbf{C}_x = \mathbf{U}\Sigma^2\mathbf{U}^T$ ,  $\mathbf{z}[t] = (\mathbf{U}\Sigma)^+ \mathbf{x}[t]$ 
initialize  $\mathbf{y}^{(1)}[t] = \mathbf{z}[t]$ ,  $\mathbf{G}^{(1)} = \mathbf{I}_P$ ,  $\alpha = \alpha_{max}$ 
for  $p = 1$  to  $M$  do
  for  $i = 1$  to  $I$  do
     $\alpha \leftarrow \alpha - \frac{1}{I}(\alpha_{max} - \alpha_{min})$ 
    for  $k = 1$  to  $P - p$  do
      estimate kurtosis and penalty of the current
      source estimate to adjust the penalization pa-
      rameter:
         $K_1 = \text{kurt}(y_1^{(p)}[t])^2$ 
         $A_1 = \text{cum}(y_1^{(p)}[t], y_1^{(p)}[t + \tau])^2$ 
         $\lambda = \alpha \frac{K_1}{A_1}$ 
      maximize (5) with respect to  $\theta_{p,k}$  and build
      Givens rotation matrix  $\mathbf{G}_g^{(p,k)}$ 
       $\mathbf{G}^{(p)} \leftarrow \mathbf{G}_g^{(p,k)} \mathbf{G}^{(p)}$ 
       $\mathbf{Y}^{(p)} \leftarrow \mathbf{G}_g^{(p,k)} \mathbf{Y}^{(p)}$ 
    end for
  end for
  extract  $\hat{\mathbf{h}}_p = \mathbf{g}^{(p)}$ ,  $\hat{\mathbf{s}}_p[t] = \hat{\mathbf{h}}_p^T \mathbf{z}[t]$ 
   $\mathbf{G}^{(p+1)} = \tilde{\mathbf{G}}^{(p)T}$ ,  $\mathbf{y}^{(p+1)}[t] = \tilde{\mathbf{y}}^{(p)}[t]$ 
end for

```

Fig. 2. Description of the P-SAUD algorithm.

matrices and by deriving the statistics required for the analysis of a certain pair of sources using the orthogonal transformation matrix  $\mathbf{G}^{(n)}$ . For both ways of estimating the statistics, the computational cost for all  $\sum_{n=1}^M (P - n)$  considered pairs has to be summed up over all iterations  $I$ . Considering the first method for estimating the statistics, we note that the estimation of the cumulants for one pair of components that is associated with two rows of the matrix  $\{\mathbf{y}^{(n)}[t]\}$  requires  $O(8T)$  FLOPs if the Leonov-Shiryaev formula for zero-mean data with unit-variance is used. Furthermore, the estimation of the covariance requires  $O(4T)$  FLOPs for the penalization term. On the whole, considering that the number of epileptic components is small compared to the number of sensors and that these components are extracted first, this leads to  $O(8TIMN)$  and  $O(4TIMN)^2$  FLOPs for cumulant estimations for the extraction of the first  $M$  sources. If the second way of estimating the statistics is used,  $O(N^4/24)$  FO cumulants have to be estimated, corresponding to  $O(N^4/24)$  different elements of the FO cumulant matrix for real-valued data. For each FO cumulant, this requires  $O(3T)$  FLOPs. Exploiting the multilinearity property of cumulants, the FO cumulants of the data after one Givens rotation can be derived from the quadricovariance matrix with 16 FLOPs per cumulant, i.e.,  $O(2N^4/3)$  FLOPs in total. On the whole, this corresponds to  $O(TN^4/8 + 2N^5MI/3)$  FLOPs. Similar considerations reveal a computational cost of  $O(N^2T + 4MN^2I)$  for the estimation and transformation of the time-delayed covariance matrix.

Furthermore, at each iteration and for each pair of components, the maximization of (5) requires the rooting of an 8-th degree polynomial (cf. appendix) which can be accomplished with  $IMNQ_8$  FLOPs, where  $Q_n$  denotes the computational complexity associated with the rooting of a polynomial of

TABLE I  
COMPUTATIONAL COMPLEXITY IN TERMS OF REAL-VALUED  
MULTIPLICATIONS FOR P-SAUD, CCA, SOBI, FASTICA, CoM<sub>2</sub>, AND  
DeIL<sup>R</sup>. THE RESULTS FOR SOBI, FASTICA, AND CoM<sub>2</sub> ARE  
REPRODUCED FROM [15].

	Number of FLOPs
P-SAUD	$\min(TN^2/2 + 4N^3/3 + PMT, 2TN^2) + IP^2Q_8/2 + 4P^2IM + \min(4TIMP, P^2T + 2IP^3) + PTM + \min(2IP^5M/3 + P^4T/8, 6ITPM)$
CCA	$T(3N^2 + 7P) + 32P^3/3 + NP^2$
SOBI	$\min(TN^2/2 + 4N^3/3 + PMT, 2TN^2) + 4N^3/3 + IP(P-1)(17(N_\tau-1) + 75 + 4P + 4P(N_\tau-1))/2 + (N_\tau-1)N^3/2$
FastICA	$\min(TN^2/2 + 4N^3/3 + PMT, 2TN^2) + (2(P-1)(P+T) + 5TP(P+1))/2I$
CoM <sub>2</sub>	$\min(TN^2/2 + 4N^3/3 + PMT, 2TN^2) + IP^2Q_4/2 + \min(IP^6/6 + 2IP^3 + P^4T/8 + TP^2, 6ITP^2)$
DeIL <sup>R</sup>	$\min(TN^2/2 + 4N^3/3 + PMT, 2TN^2) + 3ITN^2$

degree  $n$ . The update of the temporary unmixing matrix and data matrix necessitates  $4T + 4N$  FLOPs for each pair of components, i.e.,  $(4T + 4N)IMN$  FLOPs for the extraction of  $M$  components. Finally, the computation of the signal and mixing vectors of the extracted components adds up to  $4N^2IM + NTM$  FLOPs.

Table I summarizes the computational complexity of P-SAUD in comparison to the computational cost of CCA, SOBI, FastICA, CoM<sub>2</sub>, and DeIL<sup>R</sup>. For a fair comparison, we used a deterministic version of DeIL<sup>R</sup>, which takes into account the whole data length, instead of the adaptive scheme described in [20].

#### D. Simulations

To compare the performance of P-SAUD to other methods, we conduct a simulation study where we analyze how accurately the epileptic activity is extracted from the noise. To this end, we generate 32 s of realistic EEG data for 32 electrodes with a sampling rate of 256 Hz originating from two source regions with independent interictal epileptic spike-like activity. To do this, we employ a realistic head model with three compartments representing the brain, the skull, and the scalp, whose surfaces are obtained from a normal MRI. To characterize the source regions, we consider a source space that is defined by the triangularized inner cortical surface (grey matter / white matter interface), where a grid dipole is placed at the centroid of each of the triangles. The grid consists of 19626 triangles (9698 for the left hemisphere and 9928 for the right hemisphere) and on average, each triangle describes 5 mm<sup>2</sup> of the cortical surface. The lead field matrix, which characterizes the attenuations between the source dipoles and the sensors, is then computed numerically for all grid dipoles using a boundary element method (ASA, ANT, Enschede, Netherlands). We consider two source regions located in the superior frontal gyrus and the superior occipital gyrus. Each source region, also called patch, is composed of 100 adjacent grid dipoles corresponding to a cortical area of approximately 5 cm<sup>2</sup>. This patch size was chosen because it is known that about 5 cm<sup>2</sup> of cortex need to be active in order to observe

a signal at the surface. Two independent sets of simulated epileptic signals for the two patches are obtained using a model of coupled neuronal populations as previously described [22] that generates highly-correlated epileptiform spike-like signals for the dipoles within each patch. The epileptic data are then obtained by multiplying the lead field vectors of all patch dipoles with the associated spike-like signals and by summing up the results. Finally, a mixture of artifacts recorded during an EEG session (containing mostly muscle activity but also background activity and other artifacts as well as instrumentation noise) is added to the epileptic data according to a given Signal to Noise Ratio (SNR). The SNR was varied from -30 dB to -5 dB because in practice, the muscle artifacts are often of much higher amplitude than the neuronal signals.

We apply the P-SAUD algorithm to 50 realizations of simulated data with different epileptic spikes and different artifacts and compare its results to those of SAUD (P-SAUD without penalization, i.e.,  $\lambda = 0$  for all iterations), SOBI, CoM<sub>2</sub>, CCA, FastICA, and DeIL<sup>R</sup>. The relative penalization parameter of P-SAUD is varied from  $\alpha_{max} = 4$  to  $\alpha_{min} = 0$  over  $I = 20$  sweeps, which we deemed to be sufficient for the algorithm to converge. The maximal delay for the autocorrelation matrix considered in SOBI is fixed to 15 time samples. For CCA and P-SAUD, we use a delay of  $\tau = 1$ . In order to analyze the performance of the various methods, we compute the correlation coefficients of the original and estimated signals and of the original and estimated mixing vectors of the epileptic sources averaged over the two patches and the 50 realizations.

#### E. Real data

The performance of P-SAUD was also evaluated on real data obtained in three patients, 2 females and 1 male, aged 42-63 years, suffering from drug-resistant temporal lobe epilepsy. Their clinical features are described in Table II. During the presurgical evaluation of these patients, EEG was continuously recorded during daytime for one week using a 32-channel video-EEG monitoring system (BrainQuick, Micromed). Signal was sampled at 256 Hz and high-pass filtered at 0.03 Hz. The antiepileptic drug polytherapy (see Table II for details) was progressively withdrawn during the EEG monitoring to stimulate the occurrence of spikes and seizures. The three patients illustrated in this paper were specifically chosen because of the presence of interictal spikes on their EEG both in a noise-free (reference) and in a noisy (muscular activity) condition. Video-EEG is always carried out as part of normal clinical care of patients who give informed consent to the protocol approved by the local ethical board. Patients are informed that their data may be used for research purposes.

### III. RESULTS

#### A. Simulation results

Fig. 3 (left and middle) shows the spatial mixing vector correlation coefficients and the signal correlation coefficients of P-SAUD in comparison to several ICA methods as a function of the SNR. In addition, we determine how many sources have to be extracted to identify the epileptic activity

TABLE II  
CLINICAL FEATURES OF THE THREE PATIENTS WHOSE EEG RECORDINGS ARE ANALYZED IN THIS PAPER.

	Patient 1	Patient 2	Patient 3
Gender	M	F	F
Age	45	63	42
Precipitating factors	No	cranial trauma (16 months old)	No
Age at first seizure	24	18	11
Seizure frequency	2-4 / Month	< 5 / Month	1-5 / Month
Antiepileptic Drug Therapy (AED) (dose, mg/day)	Sodium Valproate (1000)	Levetiracetam (250), Clonazepam (0.5)	Lamotrigin (400), Oxcarbazepin (600) Clobazam (10)
Interictal spikes on EEG	Right temporal	Right temporal	Right+left temporal
Electrode recording spike of max. amplitude	FT10	FT10	FT9, FT10
MRI	-	Right hippocampal sclerosis (T2 and IR signal abnormality)	Right temporal-parietal dysplasia
Ictal SPECT	Increased CBF in right insular, temporal and orbitofrontal regions	Increased CBF in right temporal regions (lateral neocortex + pole)	-
Intracerebral recordings	-	-	<ul style="list-style-type: none"> <li>• Left anterior middle temporal gyrus involving sometimes the posterior part of the left superior and inferior temporal gyri</li> <li>• Right temporal-occipital-parietal junction + right anterior middle and posterior inferior temporal gyri</li> </ul>

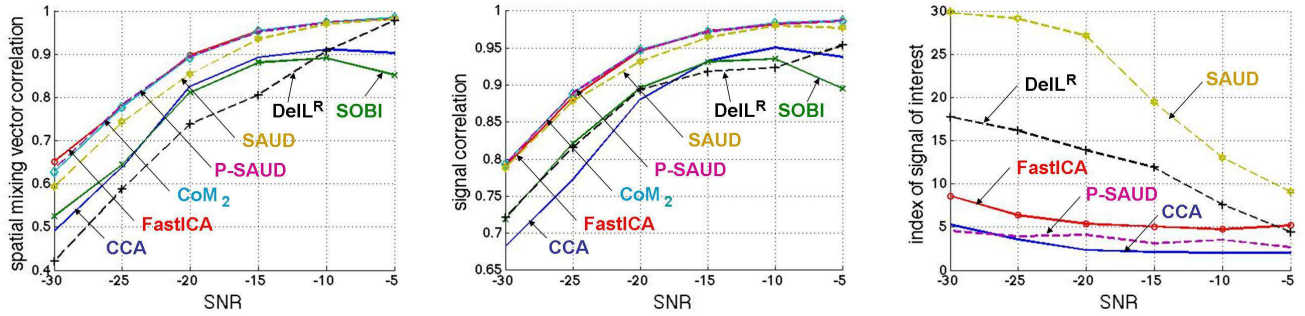


Fig. 3. (Left and middle) Correlation of recovered mixing vectors and signal vectors and (right) index of recovered sources.<sup>1</sup>

using the P-SAUD algorithm compared to SAUD and the other considered deflation algorithms, i.e., FastICA and DelL<sup>R</sup>. To this end, we identify the indices of the extracted components whose signals show the highest correlation with the original, noiseless epileptic signals. The maximal index, averaged over all realizations, is plotted in Fig. 3 (right) as a function of the SNR. Even though CCA extracts the sources jointly, we show the corresponding indices after ordering the extracted components according to their autocorrelation to demonstrate the interest in using the autocorrelation as a penalization term for P-SAUD.

Fig. 3 (left and middle) shows that P-SAUD clearly outperforms DelL<sup>R</sup> for SNR less than -5 dB and achieves a performance that is comparable to that of CoM<sub>2</sub> or FastICA. Considering the spatial mixing vector correlation, for small SNR, SAUD is slightly worse than P-SAUD. As shown in Fig. 3 (right), for SAUD, the index of epileptic signals increases

with diminishing SNR and is very high for SNR below -20 dB. This means that SAUD extracts the signals of interest rather late, which favors the accumulation of errors. On the contrary, P-SAUD ensures that the epileptic activity is extracted first as confirmed by low indices between 2 and 3 for SNR > -20 dB and an only slight increase for smaller SNR. This is due to the selection of the extracted components by the autocorrelation term as can be concluded from the small CCA source indices and explains the better performance of P-SAUD compared to SAUD. For FastICA and DelL<sup>R</sup>, in this simulation, the average numbers of sources that need to be extracted do not exceed 10 and 20, respectively. However, in practice, these two algorithms cannot be stopped after the extraction of a reduced number of components because these methods extract the sources in an arbitrary order and there is no guarantee that all epileptic sources have been extracted after a given number of deflation steps. The SO methods CCA and SOBI do not extract the epileptic components as accurately as higher order algorithms such as P-SAUD and FastICA, in particular

<sup>1</sup>Please see the pdf version of the paper available online for colored versions of Fig. 3, 4, 5, 6, 7, and 8.

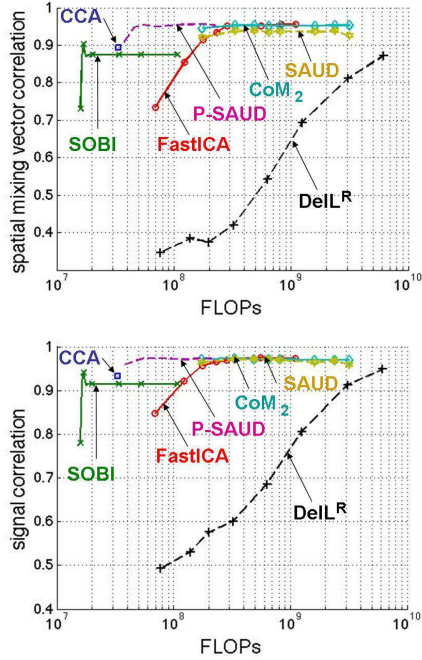


Fig. 4. Performance as a function of computational complexity.<sup>1</sup>

concerning the spatial mixing vectors.

Finally, in Fig. 4, the performance of the tested algorithms is plotted as a function of the number of FLOPs computed according to Table I for an SNR of -15 dB. Except for CCA, which requires a fixed number of FLOPs, the computational complexity is varied by changing the number of iterations performed by the algorithms. We assume that P-SAUD extracted the epileptic activity after  $M = 2$  deflation steps. It can be seen that the computational complexity of the SO methods CCA and SOBI (at the point of convergence) is smaller than that of conventional FO methods such as FastICA and CoM<sub>2</sub> by a factor of approximately 100. However, FastICA, CoM<sub>2</sub>, and SAUD extract the spatial mixing vectors with a higher accuracy than the SO methods. At an order of FLOPs that is comparable to that of the other FO methods, DeIL<sup>R</sup> exhibits a very bad performance. Due to the deflation scheme, P-SAUD extracts the epileptic signals at a significantly reduced computational complexity (by a factor of about 10) compared to the other FO methods while attaining the same accuracy as FastICA and CoM<sub>2</sub>. We have observed that this gain in computational complexity is also reflected in a reduction of the CPU time (by a factor of about 4).

### B. Real data results

For Patient 1, we considered an interval of about 20 s of interictal epileptic spikes corrupted by muscle artifacts and noise (cf. Fig. 5 (left)); due to limited space, we illustrate only a segment of 10 s length of the considered data). To remove the artifacts and noise, we applied both P-SAUD (with  $\tau = 5$ ) and CoM<sub>2</sub> to the data, extracting 32 independent signal components, shown in Fig. 5 (right). Contrary to CoM<sub>2</sub>, which extracts the signals in an arbitrary order, in the case of P-SAUD, the order of the extracted signals depends on

TABLE III  
CORRELATION COEFFICIENT BETWEEN THE P-SAUD USED TO RECONSTRUCT THE DATA AND THE CORRESPONDING CoM<sub>2</sub> COMPONENTS FOR REAL DATA EXAMPLES

	P-SAUD comp.	CoM <sub>2</sub> comp.	correlation coefficient
Patient 1	1	2	0.997
	2	5	0.989
Patient 2	2	1	0.641
	2	2	0.746
	11	32	0.989
	24	14	0.997
Patient 3	2	8	0.988
	3	32	0.996
	4	13	0.989
	6	17	0.930
	9	5	0.921

their autocorrelation. The components that characterize the epileptic spikes are thus extracted first whereas muscle activity, which has a low autocorrelation, is extracted last. The data were reconstructed using the first two P-SAUD components, selected by an EEG expert (see Fig. 5 (center left)). Note that reconstructing the EEG from components 2 and 5 extracted by CoM<sub>2</sub> yields similar results. Indeed, these components are very similar to the first two P-SAUD components as confirmed by their high correlation coefficients (see Table III). A comparison with the original data shows that the muscle activity, which corrupted in particular the recordings of electrodes FC6 and T4, has been removed in the reconstructed data and the noise has been reduced.

For Patient 2, we illustrate the behavior of P-SAUD on about 40 s of EEG recordings of epileptic spikes that are not only corrupted by muscle activity, but also include eye blink artifacts. A 10 s long segment of the noisy data and the signal components extracted by P-SAUD are shown in Fig. 6. As the eye blink artifacts also have a high autocorrelation, the associated ICA components are extracted during the first deflation steps of the P-SAUD algorithm, corresponding to the first and third P-SAUD component in this example, whereas the epileptic spikes are mostly contained in the second component (Note that this component has been split into two components by CoM<sub>2</sub> as indicated in Table III). This shows that the proposed method does not avoid the extraction of artifacts such as eye blinks whose autocorrelation is similar to or higher than that of the epileptic spikes. However, as there is only a small number of ICA components associated with these artifacts (only two components in our example), this is not very annoying because the epileptic spikes can still be extracted within a small number of deflation steps.

We analyzed 8.5 s of EEG recordings of Patient 3 as an example of data containing independent interictal spikes from two sources (see for example clean spike recordings on electrodes FT9 and FT10 in Fig. 7). The clean spikes, noisy data, reconstructed data, and P-SAUD and CoM<sub>2</sub> components for this patient are shown in Figure 7. The data are corrupted by muscle artifacts and eye blinks. As for Patient 2, the eye blinks are contained in the first ICA component extracted



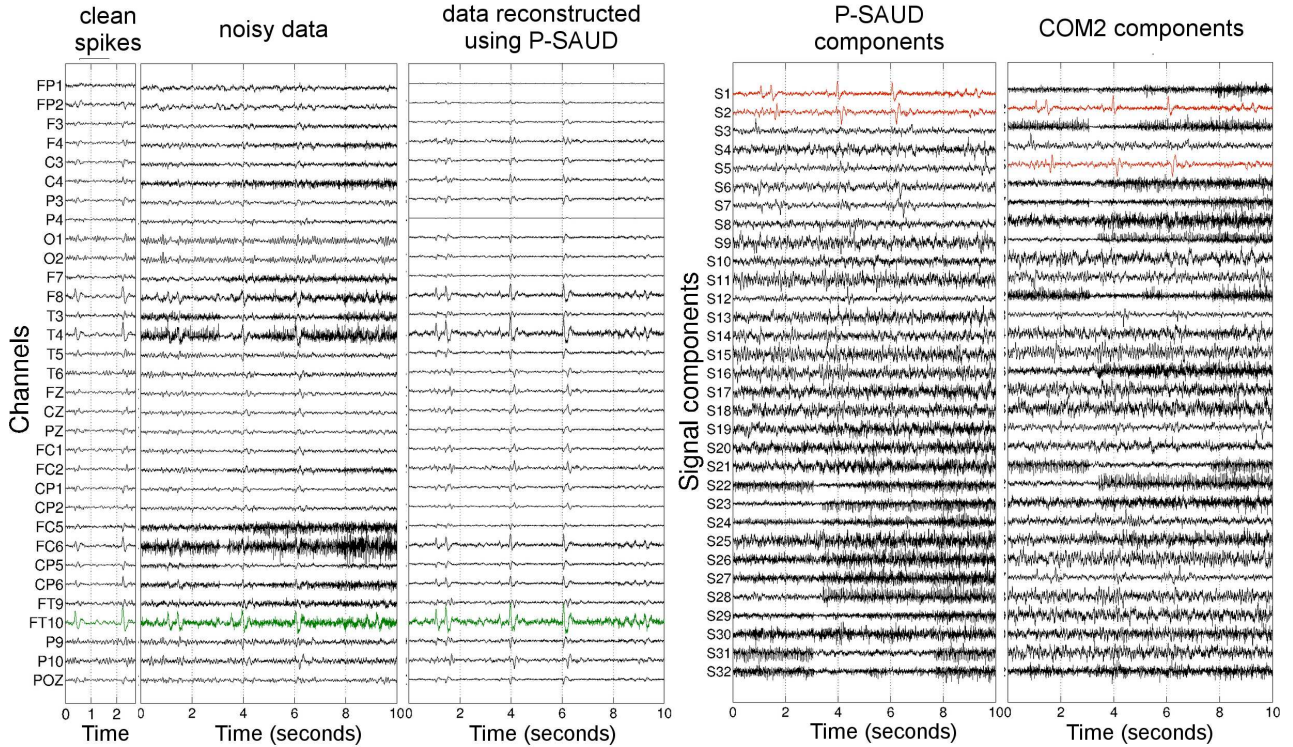


Fig. 5. Real EEG recordings of patient 1: 2.5 s of clean spikes and 10 s segment of noisy data, data reconstructed using P-SAUD, and signal components extracted with P-SAUD and CoM<sub>2</sub>. The signals of the channel FT10, at which the maximum spike amplitude is observed, are marked in green. The P-SAUD components (1 and 2) used to reconstruct the data as well as the corresponding CoM<sub>2</sub> components (2 and 5, cf. Table III), which lead to similar reconstruction results, are marked in red.<sup>1</sup>

by P-SAUD, whereas the epileptic spikes are spread over components 2, 3, 4, 6, and 9.

As these examples show, to reconstruct the epileptic spike data using P-SAUD, it would have been sufficient to extract only the first few components, which leads to a reduction of the computational complexity compared to CoM<sub>2</sub> and FastICA.

Finally, to demonstrate the good performance of the proposed denoising method and to illustrate the practical importance of this step in the analysis of surface EEG recordings, we employed distributed source localization to identify the regions involved in the interictal activity (irritative zone). To this end, we applied the Sparse Variation-Based Sparse Cortical Current Distribution (SVB-SCCD) [23] distributed source localization algorithm to the data of the three patients. This algorithm permits to identify spatially extended sources and was applied to the maxima of spikes of the noisy data, of the P-SAUD denoised data, as well as of clean data. Figure 8 shows representative examples of the source localization results obtained for the three patients. It can be seen that the brain regions localized on the denoised spikes correspond very well to the brain regions localized on the clean spikes. The localization of these regions is also in agreement with regions of increased cerebral blood flow during ictal SPECT (patients 1 and 2) or in accordance with findings from intracranial SEEG recordings (in patient 3). On the contrary, the distributed source localization results obtained on the noisy spikes always included activations of remote brain regions, which were not concordant with the functional imaging (patients 1 and 2) or intracerebral (patient 3) data.

#### IV. DISCUSSION AND CONCLUSION

We have presented a new deflation algorithm that efficiently extracts the epileptic source activity from EEG data corrupted by noise and artifacts. This preprocessing step is crucial to ensure good quality EEG data and therefore to provide a reliable interpretation of these data. Accordingly, results obtained as an example in three patients show that the spurious activation of brain regions suggested from the localization of noisy spikes would have impaired the correct interpretation of the source localization results and the proposed denoising procedure improved the result of distributed source localization.

The proposed method is based on the deflation scheme used by DeL<sup>R</sup>, but achieves a clearly improved performance by adopting a semi-algebraic optimization strategy. The quality of the extracted signals and spatial mixing vectors is indeed comparable to that achieved by techniques such as CoM<sub>2</sub> and FastICA as demonstrated by simulations (see Fig. 3) and also observed in the analysis of clinical data (see Table III). However, this performance is achieved at a considerably reduced computational cost. While the computational complexity is reduced only by a factor of 10 for the 32-channel EEG system considered in this paper (see Fig. 4), it can easily be reduced by a factor of 100 for high-resolution EEG recordings. This is due to the fact that the proposed algorithm resorts to a contrast function that is penalized by autocorrelation terms. As shown by simulations and demonstrated on real data examples, this ensures that the epileptic activity is extracted during the first few deflation steps. In particular, the epileptic spikes are

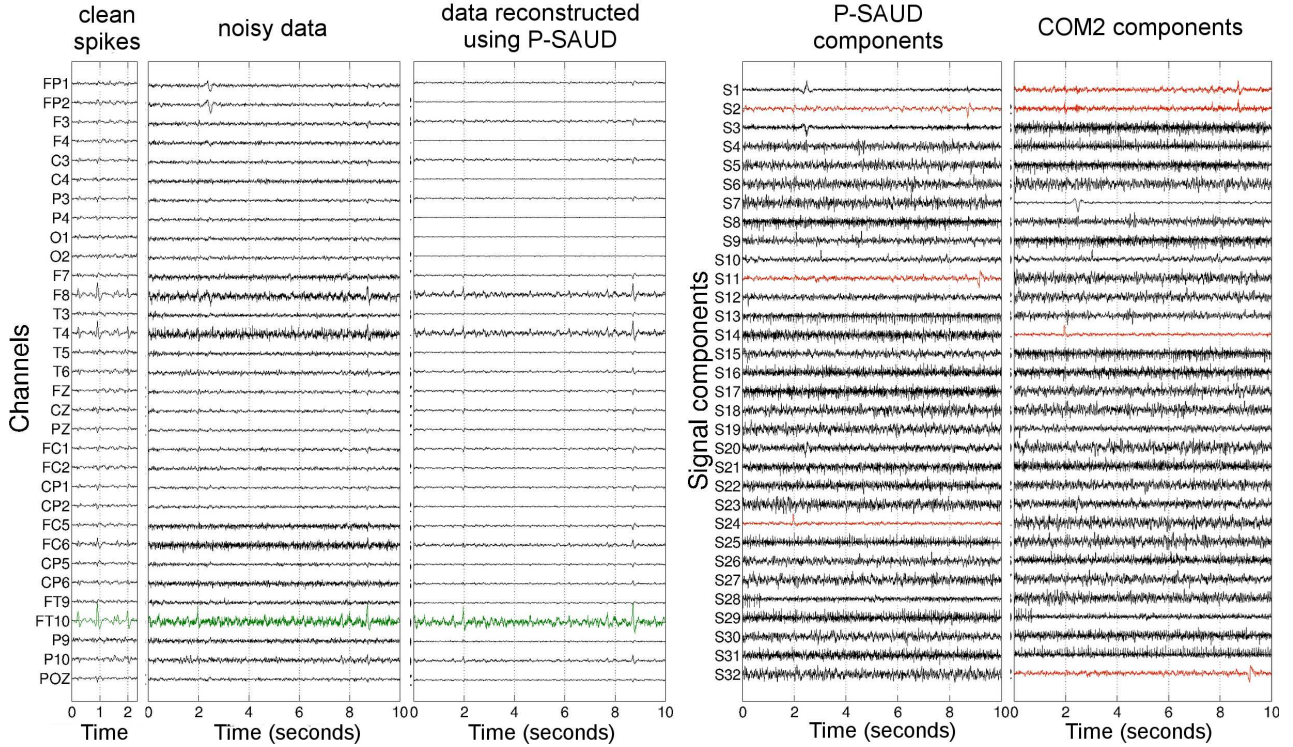


Fig. 6. Real EEG recordings of patient 2: 2.5 s of clean spikes and 10 s segment of noisy data, data reconstructed using P-SAUD, and signal components extracted with P-SAUD and CoM<sub>2</sub>. The signals of the channel FT10, at which the maximum spike amplitude is observed, are marked in green. The P-SAUD components (2, 11, and 24) used to reconstruct the data as well as the corresponding CoM<sub>2</sub> components (1, 2, 14, and 32; cf. Table III), which lead to similar reconstruction results, are marked in red.<sup>1</sup>

extracted before the numerous components associated with muscle artifacts, which constitute the main nuisance in the EEG denoising problem. Even though P-SAUD may also extract other artifacts with high autocorrelation, such as eye blinks or cardiac activity, among the first few ICA components, these artifacts are confined to a small number of components. Therefore, it is acceptable to identify these additional components because the algorithm can still be limited to the extraction of only a small number of components including the activity of interest.

A remaining difficulty consists in choosing an appropriate value for the parameter  $\tau$ , that corresponds to the delay considered in the autocorrelation term. Here, we have chosen  $\tau = 5$  for the real data examples, but this parameter depends on the sampling rate and may vary from patient to patient. In practice, we suggest to test different values of  $\tau$  on a small set of data, extracting all P-SAUD components and selecting the parameter which requires the smallest number of P-SAUD components to be extracted in order to accurately reconstruct the epileptic spike signals. The whole set of data can then be analyzed efficiently using the selected parameter.

In conclusion, the proposed P-SAUD algorithm succeeds in denoising the EEG recordings of epileptic sources with the same performance as CoM<sub>2</sub> and FastICA, but at a considerably reduced computational cost. This makes P-SAUD a promising algorithm to deal with the up-coming challenge of treating high-resolution EEG data.

While the clinical data examples considered in this paper confirm that the good performance of P-SAUD, which has

been observed in the simulations, also holds true for real data, they are not sufficient to definitely validate the proposed method. Future work will therefore consist in further consolidating the performance analysis of the P-SAUD algorithm by applying it to larger sets of clinical EEG recordings and, in particular, to high-resolution EEG data. Moreover, we will explore the usefulness of P-SAUD for denoising ictal EEG signals.

## APPENDIX A

### PARAMETERIZATION OF THE MIXING VECTORS

To facilitate the estimation of the mixing matrix, we introduce a parameterization of the mixing vectors based on Givens rotations. As has been originally introduced in [24] and used in [20], any unit-norm vector of dimension  $K$  whose last element is non-negative can be parameterized by  $K - 1$  Givens rotation angles, such that the vector corresponds to the last row of the orthonormal matrix:

$$\mathbf{G}_g^{(p)}(\phi) = \mathbf{G}_g^{(p, K-1)}(\phi_{p, K-1}) \cdots \mathbf{G}_g^{(p, 1)}(\phi_{p, 1})$$

which is composed of the Givens rotation matrices

$$\mathbf{G}_g^{(p, k)}(\phi_{p, k}) = \begin{bmatrix} \mathbf{I}_{k-1} & \mathbf{0}_{k-1, 1} & \mathbf{0}_{k-1, q} & \mathbf{0}_{k-1, 1} \\ \mathbf{0}_{1, k-1} & \cos(\phi_{p, k}) & \mathbf{0}_{q, q} & -\sin(\phi_{p, k}) \\ \mathbf{0}_{q, k-1} & \mathbf{0}_{q, 1} & \mathbf{I}_q & \mathbf{0}_{q, 1} \\ \mathbf{0}_{1, k-1} & \sin(\phi_{p, k}) & \mathbf{0}_{1, q} & \cos(\phi_{p, k}) \end{bmatrix}.$$

Here,  $q = K - k - 1$ ,  $\mathbf{I}_K$  denotes the identity matrix of size  $K \times K$ ,  $\mathbf{0}_{k, q}$  is a  $k \times q$  matrix of zeros, and  $p$  is the source index. After prewhitening, each vector of the orthonormal mixing

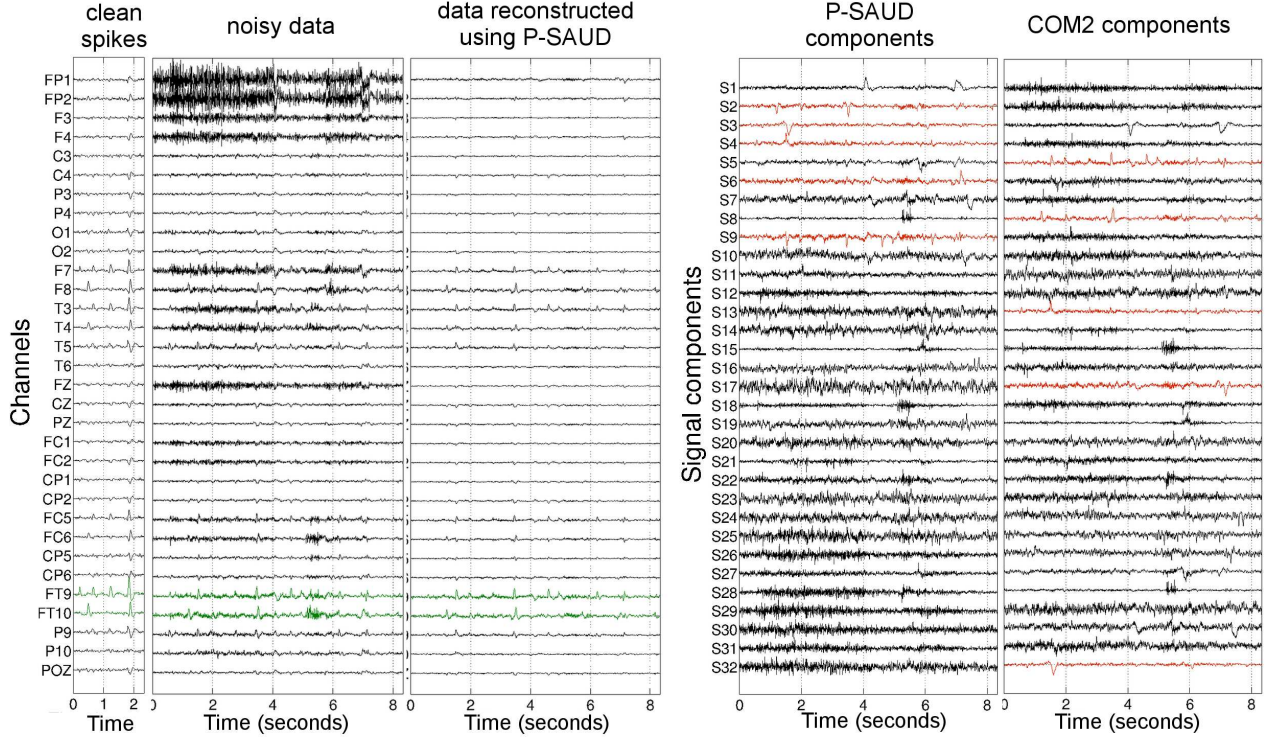


Fig. 7. Real EEG recordings of patient 3: 2.5 s of clean spikes and 8.5 s of noisy data, data reconstructed using P-SAUD, and signal components extracted with P-SAUD and CoM<sub>2</sub>. The signals of the channels FT9 and FT10, at which the maximum spike amplitudes are observed, are marked in green. The P-SAUD components (2, 3, 4, 6, and 9) used to reconstruct the data as well as the corresponding CoM<sub>2</sub> components (5, 8, 13, 17, and 32; cf. Table III), which lead to similar reconstruction results, are marked in red.<sup>1</sup>

matrix  $\mathbf{H}$  can thus be characterized by a sequence of Givens rotations. To identify the mixing vectors, it is sufficient to search for the parameters  $\phi_{p,k}$  of these Givens rotations that maximize the statistical independence of the sources. This is the basis for the P-SAUD algorithm, described in the Section II-C.

## APPENDIX B

### MAXIMIZATION OF THE P-SAUD CONTRAST FUNCTION

Based on equation (6) and the multilinearity property of cumulants, the FO cumulants of the signals  $\{s_p[t]\}$  and  $\{s_k[t]\}$  can be computed from the FO cumulants of the data  $\{z_{P-p+1}[t]\}$  and  $\{z_k[t]\}$  as follows:

$$C_{4,s_p}(1 + \theta^2)^2 = \alpha_4\theta^4 + \alpha_3\theta^3 + \alpha_2\theta^2 + \alpha_1\theta + \alpha_0 \quad (11)$$

$$C_{4,s_k}(1 + \theta^2)^2 = \alpha_0\theta^4 - \alpha_1\theta^3 + \alpha_2\theta^2 - \alpha_3\theta + \alpha_4 \quad (12)$$

where we write  $\theta$  for  $\theta_{p,k}$  to simplify the notation and where:

$$\alpha_0 = \text{cum}(z_{P-p+1}, z_{P-p+1}, z_{P-p+1}, z_{P-p+1}) = C_{4,z_{P-p+1}}$$

$$\alpha_1 = 4 \text{cum}(z_{P-p+1}, z_{P-p+1}, z_{P-p+1}, z_k)$$

$$\alpha_2 = 6 \text{cum}(z_{P-p+1}, z_{P-p+1}, z_k, z_k)$$

$$\alpha_3 = 4 \text{cum}(z_{P-p+1}, z_k, z_k, z_k)$$

$$\alpha_4 = \text{cum}(z_k, z_k, z_k, z_k) = C_{4,z_k}$$

with  $\text{cum}(a, b, c, d)$  the FO cross-cumulant of the four random variables  $a, b, c,$  and  $d$ . The cumulants can be estimated from an estimate of the moments according to the Leonov-Shiryayev formula [25]. The covariance is given by:

$$\text{cum}(s_p(t), s_p(t+1)) \cdot (1 + \theta^2) = \beta_2\theta^2 + \beta_1\theta + \beta_0 \quad (13)$$

with:

$$\beta_0 = \sqrt{\lambda} \cdot \text{cum}(z_{P-p+1}(t), z_{P-p+1}(t+1))$$

$$\beta_1 = \sqrt{\lambda} \cdot \text{cum}(z_{P-p+1}(t), z_k(t+1))$$

$$+ \sqrt{\lambda} \cdot \text{cum}(z_k(t), z_{P-p+1}(t+1))$$

$$\beta_2 = \sqrt{\lambda} \cdot \text{cum}(z_k(t), z_k(t+1))$$

Inserting (11), (12), and (13) into the contrast (5), one obtains a rational function. The optima of this function are obtained at the zeros of the 8-th order polynomial with coefficients:

$$\gamma_0 = 2(\alpha_0\alpha_1 - \alpha_3\alpha_4 + \beta_0\beta_1)$$

$$\gamma_1 = 2(\alpha_2^2 + \alpha_1^2 + 2\alpha_2\alpha_4 + 2\alpha_0\alpha_2 - 2\beta_0^2 + \beta_1^2 + 2\beta_0\beta_2) - 8(\alpha_4^2 + \alpha_0^2)$$

$$\gamma_2 = 6(\alpha_0\alpha_3 + \alpha_1\alpha_2 - \alpha_1\alpha_4 - \alpha_2\alpha_3) - 14(\alpha_0\alpha_1 - \alpha_3\alpha_4) - 2\beta_0\beta_1 + 6\beta_1\beta_2$$

$$\gamma_3 = -6(\alpha_1^2 + \alpha_3^2) + 16(\alpha_0\alpha_4 + \alpha_1\alpha_3) - 12(\alpha_0\alpha_2 + \alpha_2\alpha_4) + 8\alpha_2^2 + 4\beta_0\beta_2 - 8\beta_0^2 + 2\beta_1^2 + 4\beta_2^2$$

$$\gamma_4 = 20(\alpha_1\alpha_4 + \alpha_2\alpha_3 - \alpha_0\alpha_3 - \alpha_1\alpha_2) + 10(\beta_1\beta_2 - \beta_0\beta_1)$$

$$\gamma_5 = 6(\alpha_1^2 + \alpha_3^2) - 16(\alpha_0\alpha_4 + \alpha_1\alpha_3) + 12(\alpha_0\alpha_2 + \alpha_2\alpha_4) - 8\alpha_2^2 - 4\beta_0\beta_2 + 8\beta_2^2 - 2\beta_1^2 - 4\beta_0^2$$

$$\gamma_6 = 6(\alpha_0\alpha_3 + \alpha_1\alpha_2 - \alpha_1\alpha_4 - \alpha_2\alpha_3) - 14(\alpha_0\alpha_1 - \alpha_3\alpha_4) + 2\beta_1\beta_2 - 6\beta_0\beta_1$$

$$\gamma_7 = -2(\alpha_2^2 + \alpha_1^2 + 2\alpha_2\alpha_4 + 2\alpha_0\alpha_2 - 2\beta_2^2 + \beta_1^2 + 2\beta_0\beta_2) + 8(\alpha_4^2 + \alpha_0^2)$$

$$\gamma_8 = 2(\alpha_0\alpha_1 - \alpha_3\alpha_4 - \beta_1\beta_2)$$

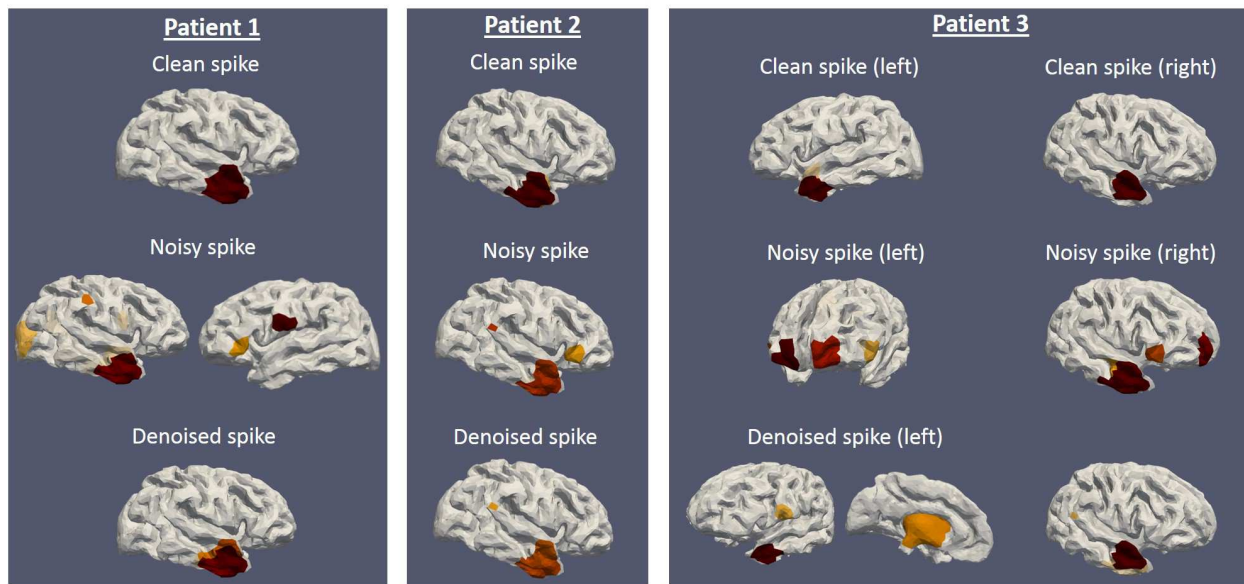


Fig. 8. Source localization results obtained for the three patients on clean spikes, noisy spikes and spikes denoised by P-SAUD. For all patients, source localization of denoised spikes yields results consistent with source localization of spikes uncorrupted by muscular activity as well as with functional imaging and intracerebral data. On the contrary, sources localized from noisy data always involved more brain regions, located in areas that were not consistent with other conclusions of the presurgical evaluation. These spurious regions were, for example, the right occipital pole and inferior frontal and central regions bilaterally (patient 1), the right inferior frontal gyrus (patient 2) or the frontal poles (patient 3).<sup>1</sup>

The optimal rotation angle  $\theta$  corresponds to the real-valued zero for which the rational function attains its highest maximum.

#### REFERENCES

- [1] R. Vigarío and E. Oja, "BSS and ICA in neuroinformatics: from current practices to open challenges," *IEEE Reviews in Biomedical Engineering*, vol. 1, pp. 50–61, 2008.
- [2] P. Comon and C. Jutten, Eds., *Handbook of Blind Source Separation*. New York: Academic Press, 2010, ch. 18.
- [3] A. Belouchrani, K. Abed-Meraim, J.-F. Cardoso, and E. Moulines, "A blind source separation technique using second-order statistics," *IEEE Trans. Signal Processing*, vol. 45, no. 2, pp. 434–444, February 1997.
- [4] W. De Clercq, A. Vergult, B. Vanrumste, W. Van Paesschen, and S. Van Huffel, "Canonical correlation analysis applied to remove muscle artefacts from the electroencephalogram," *IEEE Trans. Biomed. Eng.*, pp. 2583–2587, 2006.
- [5] S. Hajipour Sardouie, M. Bagher Shamsollahi, L. Albera, and I. Merlet, "Interictal EEG noise cancellation: GEVD and DSS based approaches versus ICA and DCCA based methods," to appear in *Elsevier IRBM*, 2014.
- [6] —, "Denoising of ictal EEG data using semi-blind source separation methods based on time-frequency priors," to appear in *IEEE Journal of Biomedical and Health Informatics*, 2014.
- [7] T.-P. Jung, S. Makeig, C. Humphries, T. W. Lee, M. J. McKeown, and V. Iragui, "Removing electroencephalographic artifacts by blind source separation," *Psychophysiology*, vol. 37, no. 2, pp. 163–178, 2000.
- [8] E. Urrestarazu, J. Iriarte, M. Alegre, M. Valencia, C. Viteri, and J. Artieda, "Independent component analysis removing artifacts in ictal recordings," *Epilepsia*, vol. 45, no. 9, pp. 1071–1078, 2004.
- [9] B. W. McMenamin, A. J. Shackman, J. S. Maxwell, D. R. W. Bachhuber, A. M. Koppenhaver, L. L. Greischar, and R. J. Davidson, "Validation of ICA-based myogenic artifact correction for scalp and source-localized EEG," *NeuroImage*, vol. 49, pp. 2416–2432, 2010.
- [10] T. W. Lee, M. Girolami, and T. J. Sejnowski, "Independent component analysis using an extended Infomax algorithm for mixed sub-Gaussian and super-Gaussian sources," *Neural Computation*, vol. 11, no. 2, pp. 417–441, 1999.
- [11] J. F. Cardoso and A. Souloumiac, "Blind beamforming for non-Gaussian signals," *IEE Proceedings F*, vol. 140, no. 6, pp. 362–370, 1993.
- [12] P. Comon, "Independent component analysis - a new concept?" *Signal Processing*, vol. 36, pp. 287–314, 1994.
- [13] A. Hyvärinen, "Fast and robust fixed-point algorithms for independent component analysis," *IEEE Trans. Neural Networks*, vol. 10, no. 3, pp. 626–634, May 1999.
- [14] V. Zarzoso and P. Comon, "Robust independent component analysis by iterative maximization of the kurtosis contrast with algebraic optimal step size," *IEEE Trans. Neural Networks*, vol. 21, no. 2, pp. 248–261, 2010.
- [15] L. Albera, A. Kachenoura, P. Comon, A. Karfoul, F. Wendling, L. Senhadji, and I. Merlet, "ICA-based EEG denoising: a comparative analysis of fifteen methods," *Special Issue of the Bulletin of the Polish Academy of Sciences - Technical sciences*, vol. 60, no. 3, pp. 407–418, 2012.
- [16] A. Delorme, J. Palmer, J. Onton, T. Oostenveld, and S. Makeig, "Independent EEG sources are dipolar," *PLOS ONE*, vol. 7, no. 2, February 2012.
- [17] W. Lu and J. C. Rajapakse, "Approach and applications of constrained ICA," *IEEE Trans. Neural Networks*, vol. 16, no. 1, pp. 203–212, January 2005.
- [18] D. S. Huang and J.-X. Mi, "A new constrained independent component analysis method," *IEEE Trans. Neural Networks*, vol. 18, no. 5, pp. 1532–1535, 2007.
- [19] J.-X. Mi, "A novel algorithm for independent component analysis with reference and methods for its applications," *PLOS ONE*, vol. 9, no. 5, 2014.
- [20] N. Delfosse and P. Loubaton, "Adaptive blind separation of independent sources: a deflation approach," *Signal Processing*, vol. 45, pp. 59–83, 1995.
- [21] K. P. Indiradevi, E. Elias, P. S. Sathidevi, S. Dinesh Nayak, and K. Radhakrishnan, "A multi-level wavelet approach for automatic detection of epileptic spikes in the electroencephalogram," *Comput. Biol. Med.*, vol. 38, no. 7, pp. 805–816, 2008.
- [22] D. Cosandier-Riméle, J.-M. Badier, P. Chauvel, and F. Wendling, "A physiologically plausible spatio-temporal model for EEG signals recorded with intracerebral electrodes in human partial epilepsy," *IEEE Trans. Biomed. Eng.*, vol. 54, no. 3, pp. 380–388, March 2007.
- [23] H. Becker, L. Albera, P. Comon, R. Gribonval, and I. Merlet, "Fast, variation-based methods for the analysis of extended brain sources," in *Proc. of EUSIPCO*, Lisbon, Portugal, 2014.
- [24] P. A. Regalia, "An adaptive unit norm filter with applications to signal analysis and Karhunen-Loève transformations," *IEEE Trans. Circuits Syst.*, vol. 37, no. 5, 1990.
- [25] P. Mc Cullagh, *Tensor methods in statistics*. London, U. K.: Chapman and Hall, 1987, ch. Monographs on statistics and applied probability.