An Efficient Jacobi-Like Deflationary ICA Algorithm: Application to EEG Denoising

Sepideh Hajipour Sardouie, *Student Member, IEEE*, Laurent Albera, *Senior Member, IEEE*, Mohammad Bagher Shamsollahi, *Senior Member, IEEE*, and Isabelle Merlet

Abstract—In this paper, we propose a Jacobi-like Deflationary ICA algorithm, named JDICA. More particularly, while a projection-based deflation scheme inspired by Delfosse and Loubaton's ICA technique (DelL$^{\mathbb{R}}$) is used, a Jacobi-like optimization strategy is proposed in order to maximize a fourth order cumulant-based contrast built from whitened observations. Experimental results obtained from simulated epileptic EEG data mixed with a real muscular activity and from the comparison in terms of performance and numerical complexity with the FastICA, RobustICA and DelL$^{\mathbb{R}}$ algorithms, show that the proposed algorithm offers the best trade-off between performance and numerical complexity when a low number (~ 12) of electrodes is available.

Index Terms—Deflation, denoising, ElectroEncephaloGraphy, higher order statistics, independent component analysis, interictal epileptic data, Jacobi-like optimization.

I. INTRODUCTION

INDEPENDENT COMPONENT ANALYSIS (ICA) [8], [9] is a very useful tool in signal processing especially to process biomedical signals such as ElectroEncephaloGraphic (EEG) data [1]–[5]. The ICA problem consists of retrieving unobserved realizations of a P-dimensional random vector $\boldsymbol{s} = [s_1, \ldots, s_P]^\top$ from observed realizations of an N-dimensional random vector $\boldsymbol{x} = [x_1, \ldots, x_N]^\top$ that can linearly be modeled as follows:

$$\boldsymbol{x} = \sum_{p=1}^{P} \boldsymbol{a}_p s_p + \boldsymbol{\nu} = \boldsymbol{A}\boldsymbol{s} + \boldsymbol{\nu} \qquad (1)$$

where $\boldsymbol{\nu}$ represents an N-dimensional noise independent of $\boldsymbol{s}$. The fundamental assumption of ICA is that the P unknown random variables s_p (called sources) are statistically independent, i.e. their joint Probability Density Function (PDF) can be factorized as the product of their marginal PDFs.

ICA algorithms can be divided into two groups: i) "joint" or "symmetric" approaches jointly extract the independent components ii) "deflationary" techniques estimate sources one by one. Joint algorithms seem to converge to the expected solution in practice, but no theoretical result is available. On the other hand, the convergence of most of deflationary algorithms have been proved analytically [4], [6], [10]. In addition, in deflationary algorithms, a penalty term can be added to the contrast function [7] to force the algorithm to extract the sources of interest during the early steps. Besides when the number of all sources largely encompasses the number of sources of interest, the computational complexity of the deflationary algorithms is greatly reduced.

In this paper, we propose an efficient Jacobi-like Deflationary ICA algorithm, called JDICA, based on second and Fourth Order (FO) statistics. The deflation procedure of our algorithm is inspired by [4]. The gradient-based ICA algorithm (called DelL$^{\mathbb{R}}$ throughout this paper) proposed in [4], estimates the sources one by one using a smart projection-based deflation scheme. According to its gradient-based structure, the step size must be precisely chosen to guarantee acceptable results, especially with noisy data. A multi-initialization procedure can even be necessary in some practical contexts. In order to overcome these drawbacks, we propose a Jacobi-like algorithm to maximize the contrast function computed from the FO cumulants of the whitened observations.

We have examined the effectiveness of JDICA in denoising of simulated interictal epileptic data when a low number of electrodes is available as for children. The comparison in terms of performance and numerical complexity with classical deflationary ICA algorithms, namely FastICA [6], RobustICA [10] and DelL$^{\mathbb{R}}$ shows that JDICA offers a better accuracy than DelL$^{\mathbb{R}}$ and a lower numerical complexity than FastICA and RobustICA.

II. METHODOLOGY

We assume that we have some realizations of the real-valued random vector $\boldsymbol{x}$ (1). Since JDICA, like a large group of ICA algorithms, needs a prewhitening step [4] without loss of generality, we assume that vector $\boldsymbol{x}$ denotes the prewhitened observation random vector and matrix $\boldsymbol{A} = [\boldsymbol{a}_1, \ldots, \boldsymbol{a}_P]$ is a $(P \times P)$ real-valued orthogonal mixing matrix. The aim of our method is then to estimate the P columns $\boldsymbol{a}_p$ of $\boldsymbol{A}$ and the P corresponding sources such that $s_p = \boldsymbol{a}_p^\top \boldsymbol{x}$. More particularly, vector $\boldsymbol{a}_p$ can be identified by maximizing the following contrast function:

$$\mathfrak{F}(\boldsymbol{g}^{(p)}) = \frac{1}{4}[C_4(y_p)]^2 = \frac{1}{4}[C_4(\boldsymbol{g}^{(p)^T}\boldsymbol{x})]^2 \qquad (2)$$

with respect to $\boldsymbol{g}^{(p)}$ where $C_4(y_p)$ is the FO marginal cumulant of $y_p = \boldsymbol{g}^{(p)^\top}\boldsymbol{x}$. The advantage of defining such a contrast function is that the arguments of the local maxima of $\mathfrak{F}$ on the unit sphere are the vectors $\{\pm\boldsymbol{a}_p\}_{p=1,\ldots,P}$[4]. This property ensures our maximization (2) to converge to one column of the matrix

A. Consequently one of the sources is extracted. Thus a projection deflation procedure is applied to subtract the contribution of the extracted source from the mixture. These two steps require a particular parametrization of the elements of the unit sphere which is given by:

Lemma 1: Each unit norm column vector $\boldsymbol{g} \in \mathbb{R}^P$ whose last component g_P is strictly positive can be represented as the last column of an orthogonal matrix given by:

$$\boldsymbol{G}(\boldsymbol{t}) = \boldsymbol{G}_{P-1}(t_{P-1}) \ldots \boldsymbol{G}_2(t_2)\boldsymbol{G}_1(t_1) \quad (3)$$

where the $P - 1$ real-valued elements of $\boldsymbol{t} = [t_1, \ldots, t_{P-1}]^\top$ correspond to tangents of uniquely defined angles belonging to $]-\pi/2, \pi/2[$ and $\boldsymbol{G}_p(t_p)$ is a Givens rotation of size $(P \times P)$ derived from an identity matrix for which the (p,p)-th, (P,P)-th, (p,P)-th, (P,p)-th components are replaced with $(1+t_p^2)^{-1/2}$, $(1+t_p^2)^{-1/2}$, $t_p(1+t_p^2)^{-1/2}$ and $-t_p(1+t_p^2)^{-1/2}$, respectively.

Proof derives from [4, lemma 2.2] by expressing $\cos(\theta)$ and $\sin(\theta)$ as a function of $t = \tan(\theta)$. This parametrization differs from that of [4] and allows us both to reformulate the contrast (2) as a rational function and to consider other optimization strategies such as a Jacobi-like procedure.

To extract the first source, we then propose to compute a matrix $\boldsymbol{G}^{(1)}(\boldsymbol{t})$ such that its last column, $\boldsymbol{g}^{(1)}(\boldsymbol{t})$, maximizes the contrast function (2) with respect to $\boldsymbol{t}$. Our Jacobi-like optimization procedure consists of decomposing $\boldsymbol{G}^{(1)}(\boldsymbol{t})$ as a product of $P - 1$ elementary Givens rotations $\boldsymbol{G}_p^{(1)}(t_p)$ and of sequentially identifying the $P - 1$ corresponding parameters t_p. The $(P - 1)$-dimensional optimization problem is thus replaced with $P - 1$ sequential mono-dimensional optimization problems. In practice, several sweeps of the $P - 1$ parameters are necessary to achieve convergence. More precisely, let us consider the p-th mono-dimensional maximization problem of a sweep of our Jacobi-like procedure. It consists in computing matrix $\boldsymbol{G}^{(1+)}(t_p)$ defined by $\boldsymbol{G}^{(1+)}(t_p) = \boldsymbol{G}_p^{(1)}(t_p)\boldsymbol{G}^{(1-)}$ such that its last column, $\boldsymbol{g}^{(1+)}(t_p)$, maximizes the contrast function (2), where $\boldsymbol{G}^{(1-)}$ is the product of all the elementary Givens rotations estimated previously.

Denoting the last column of $\boldsymbol{G}^{(1-)}$ by $\boldsymbol{g}^{(1-)} = [g_1^{(1-)}, \ldots, g_P^{(1-)}]^\top$, the last column of $\boldsymbol{G}^{(1+)}$ can be written as:

$$\begin{aligned}
\boldsymbol{g}^{(1+)}(t_p) &= [g_1^{(1+)}(t_p), \ldots, g_P^{(1+)}(t_p)]^\top \\
&= [g_1^{(1-)}, \ldots, g_{p-1}^{(1-)}, g_p^{(1+)}(t_p), g_{p+1}^{(1-)}, \ldots, g_P^{(1+)}(t_p)]^\top
\end{aligned} \quad (4)$$

where:

$$g_p^{(1+)}(t_p) = \frac{1}{\sqrt{1+t_p^2}} g_p^{(1-)} + \frac{t_p}{\sqrt{1+t_p^2}} g_P^{(1-)} \quad (5)$$

$$g_P^{(1+)}(t_p) = \frac{-t_p}{\sqrt{1+t_p^2}} g_p^{(1-)} + \frac{1}{\sqrt{1+t_p^2}} g_P^{(1-)} \quad (6)$$

It appears that only the p-th and P-th components of $\boldsymbol{g}^{(1+)}(t_p)$ depend on t_p. Then, we set the derivative of the contrast function with respect to t_p equal to zero to find the appropriate t_p value:

$$\frac{\partial C_4(y_1)^2/4}{\partial t_p} = \frac{1}{2} C_4(y_1) \frac{\partial C_4(y_1)}{\partial t_p} = 0 \quad (7)$$

which results in simply vanishing $\partial C_4(y_1)/\partial t_p$.

Now using the multi-linearity property of cumulants, it is shown that $C_4(y_1)$ can be written as follows:

$$\begin{aligned}
C_4(y_1) =\ & d_{40}(g_p^{(1+)}(t_p))^4 + d_{31}(g_p^{(1+)}(t_p))^3(g_P^{(1+)}(t_p)) \\
& + d_{22}(g_p^{(1+)}(t_p))^2(g_P^{(1+)}(t_p))^2 + d_{13}(g_p^{(1+)}(t_p))(g_P^{(1+)}(t_p))^3 \\
& + d_{04}(g_P^{(1+)}(t_p))^4 + d_{30}(g_p^{(1+)}(t_p))^3 + d_{21}(g_p^{(1+)}(t_p))^2(g_P^{(1+)}(t_p)) \\
& + d_{12}(g_p^{(1+)}(t_p))(g_P^{(1+)}(t_p))^2 + d_{03}(g_P^{(1+)}(t_p))^3 + d_{20}(g_p^{(1+)}(t_p))^2 \\
& + d_{11}(g_p^{(1+)}(t_p))(g_P^{(1+)}(t_p)) + d_{02}(g_P^{(1+)}(t_p))^2 + d_{10}(g_p^{(1+)}(t_p))
\end{aligned}$$

$$+ d_{01}(g_P^{(1+)}) + d_{00} \quad (8)$$

where the coefficients d_{ij} are defined in appendix. Consequently, by computing the derivative of (8), we obtain:

$$\sum_{m=0}^{4} e_m t_p^m + \sqrt{1+t_p^2}\left(\sum_{n=0}^{3} f_n t_p^n\right) = 0 \quad (9)$$

where the coefficients e_m and f_n are given in appendix. Equation (9) can be simplified to an 8-th degree polynomial equation as follows:

$$\sum_{m=0}^{8}\sum_{n=0}^{m} e_n e_{m-n} t_p^m - (1+t_p^2)\sum_{m'=0}^{6}\sum_{n'=0}^{m'} f'_{n'} f_{m'-n'} t_p^{m'} = 0 \quad (10)$$

By rooting (10), 8 solutions $\hat{t}_p$ are obtained. Then we calculate the contrast function (2) for all real-valued roots and we choose the root $\hat{t}_p^{(opt)}$ which maximizes it. Eventually, we calculate the matrices $\boldsymbol{G}_p^{(1)}(\hat{t}_p^{(opt)})$ and $\boldsymbol{G}^{(1+)}(\hat{t}_p^{(opt)})$. This procedure is performed iteratively for all $p \in \{1, \ldots, P - 1\}$ and for several sweeps until convergence. At this stage, the first column $\hat{\boldsymbol{a}}_1$ of the estimated mixing matrix is given by the last update of $\boldsymbol{g}^{(1+)}(\hat{t}_{P-1}^{(opt)})$ and the first source is estimated by $\hat{s}_1 = \hat{\boldsymbol{a}}_1^\top \boldsymbol{x}$.

After estimating the first source, we remove its contribution from the observations by projecting the observations onto the subspace orthogonal to that spanned by $\hat{\boldsymbol{a}}_1$ by computing $\boldsymbol{x}^{(1)} = \Pi^{(1)}\boldsymbol{x}$ where $\Pi^{(1)}$ is a $(P - 1 \times P)$ projection matrix built by stacking vertically the $P - 1$ first rows of the last update of $\boldsymbol{G}^{(1+)}(\hat{t}_p^{(opt)})^\top$. Now to estimate the other sources, the same procedure should be done by using equations (4) to (10). The only difference is that the vector of observations $\boldsymbol{x}$ should be replaced by the observation $\boldsymbol{x}^{(p-1)}$ of reduced dimension $(P - p + 1)$ in order to extract the p-th source.

Note that the estimation of FO cumulants is not required at each iteration of our Jacobi-like procedure. The N^4 FO cumulants $C_{n_1,n_2,n_3,n_4,\boldsymbol{x}}$ of vector $\boldsymbol{x}$ can be estimated at the beginning of the procedure and sorted in a $(N^2 \times N^2)$ matrix, $\boldsymbol{Q}_{\boldsymbol{x}}$, called quadricovariance [1]. The FO cumulants $C_{n_1,n_2,n_3,n_4,\boldsymbol{x}^{(p)}}$ of vector $\boldsymbol{x}^{(p)}$ can then be derived using the following formula $\boldsymbol{Q}_{\boldsymbol{x}^{(p)}} = \boldsymbol{H}\boldsymbol{Q}_{\boldsymbol{x}}\boldsymbol{H}^\top$ where:

$$\boldsymbol{H} = \prod_{i=1}^{p}(\Pi^{(i)} \otimes \Pi^{(i)}) \quad (11)$$

with $\otimes$ the Kronecker product operator.

III. NUMERICAL COMPLEXITY

In this section, we analyze the numerical complexity of the proposed algorithm in terms of real-valued floating point operations (flops). A flop corresponds to a multiplication followed by an addition, but in practice only the number of multiplications is computed. In the following computations, P, N and T are the number of sources, the number of observation channels and the number of time samples, respectively. $f_4(P) = P(P+1)(P+2)(P+3)/24$ is equal to the number of free entries in a fourth order cumulant tensor of dimension P enjoying all symmetries. $B = \min(T\frac{N^2}{2} + 4\frac{N^3}{3} + PNT, 2TN^2)$ is the number of flops required to perform spatial whitening. R is the complexity required to compute the roots of a real 8-th degree polynomial by using the companion matrix technique (we may take $R = 972$ flops). As a result the proposed ICA algorithm requires $B + 2T + 2P + NP^2 + 3Tf_4(P) + \sum_{p=2}^{P-1}(2p^2(p-1)^2(p^2-p+1)) + \sum_{p=2}^{P}(p^2T + pP + (p-1)P^2) + \sum_{p=2}^{P} It_{P-p+1}(R+4p^3/3-7p^2+62p/3+195 + \min((p-2)T+8(4T+8), 4p+8(2p^4+p^2+7)))$ flops to extract all P sources.
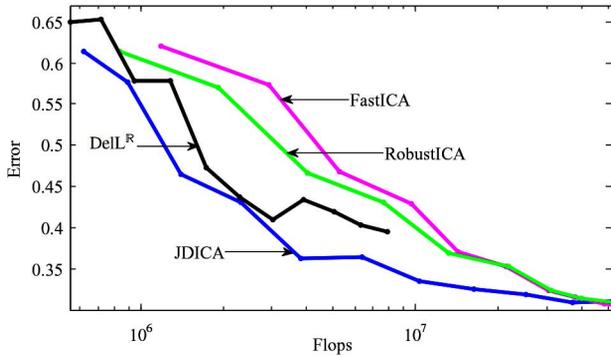
Fig. 1. Average Error as a function of flops obtained by varying the number of estimated sources $\hat{P}$ with $\text{SNR} = -5\text{dB}$.

IV. PERFORMANCE ANALYSIS ON SIMULATED DATA

A. Data Generation

The simulated epileptic EEG was generated using a realistic model developed in our team [3]. We built a mesh of the cortical surface from a 3D MRI T1 image of a subject (BrainVisa, SHFJ, Orsay, France). This mesh is composed of 40500 triangles of mean surface $5mm^2$. A current dipole is placed at the barycenter of each triangle and oriented orthogonally to the triangle surface, leading to a field of current dipoles. From this mesh, P_e distributed sources, called "patches", generating interictal spikes, are defined. Each patch is composed of 100 dipole sources to which we assigned hyper-synchronous spike-like activities generated from a model of neuronal populations [3]. From this setup and considering 12 electrodes, namely Fp1, Fp2, C3, C4, T3, T4, O1, O2, F7, F8, T5 and T6, the forward problem was then calculated using a realistic head model made of three nested homogeneous volumes shaping the brain, the skull and the scalp (ASA, ANT, Enschede, Netherlands). The aforementioned electrodes are commonly used to record EEG in pediatric patients. The epileptic activity at the level of these electrodes, namely the signal of interest, was then obtained by solving the forward problem using a realistic head model and the Boundary Element Method (BEM). In this paper, we considered a single patch localized in the left superior temporal gyrus and 50 Monte Carlo simulations were generated. In addition a 12-dimensional signal of non-interest extracted from real 12-channel EEG and composed of muscle activity, background EEG and instrument noise was added to each trial with a specified Signal-to-Noise Ratio (SNR).

B. Results

We compared the performance of the proposed JDICA algorithm with three deflationary ICA algorithms, namely FastICA, RobustICA, DelL$^{\mathbb{R}}$. Note that, unlike the three other algorithms, RobustICA does not require any prewhitening. The performance was computed as a function of computational complexity using the Normalized Mean-Squared Error (NMSE) as defined in [1]. In our experiment, the data length is fixed to 5120 samples and the SNR value is -5 dB. By varying the number of estimated sources $\hat{P}$ in the range of 2 to 12, we make vary the computational complexity of each algorithm. Fig. 1 shows then the average Error as a function of flops at the output of the four algorithms. This figure illustrates that the JDICA algorithm offers the best compromise between performance and numerical complexity when a low number of electrodes is used even if RobustICA converges faster. It implies that one iteration of RobustICA requires more flops than one sweep of JDICA.
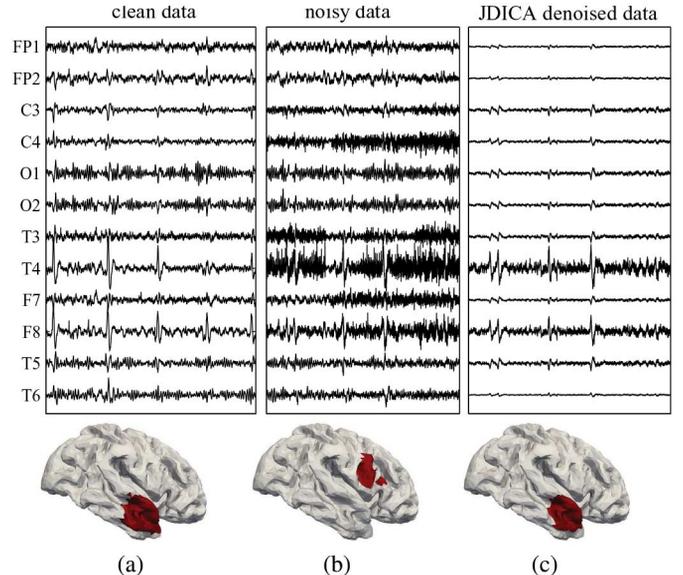
Fig. 2. Denoising of real interictal spikes data (a) a noise-free interictal spikes, (b) an epoch including spikes hidden in muscle activity and (c) EEG denoised by JDICA. The source localization results at the output of 4-ExSo-MUSIC are depicted at the bottom of each column.

V. APPLICATION TO REAL DATA

In this section we evaluate JDICA in the case of real data. The JDICA algorithm was applied to denoise interictal spikes obtained from a patient suffering from drug-resistant partial epilepsy. Scalp-EEG data were acquired from 12 electrodes at a sampling frequency of 256 Hz. These data were reviewed in order to isolate an epoch of clean data containing interictal spikes (Fig. 2(a)) and an epoch of noisy EEG containing spikes hidden by muscle activity of high amplitude (Fig. 2(b)).

The same procedure as for simulated data was applied to reconstruct the denoised EEG signals by using JDICA (Fig. 2(c)). Since we do not know the ground truth to evaluate the performance of the proposed method, a source localization process was performed on the original clean signal (considered as a reference), on the noisy data, as well as on data denoised by JDICA. The recent 4-ExSo-MUSIC algorithm [2] was used to achieve source localization. As shown in Fig. 2, the epileptic spikes maximal at temporal and frontotemporal electrodes (T4, F8) on clean data are retrieved at the same electrodes on denoised data. In addition, the muscle activity visible on noisy data is strongly reduced by the JDICA procedure at F8 and T4 and almost entirely removed at other channels. Source localization (bottom of Fig. 2) of clean (2a) and of denoised spikes (2c) is similar (right anterior temporal) and consistent with the patient pathology. For noisy data, the spike source is incorrectly localized.

VI. CONCLUSION

In this paper, we proposed a new deflationary ICA algorithm based on a Jacobi-like optimization procedure to separate independent sources. We examined the effectiveness of the proposed algorithm in denoising of simulated pediatric epileptic data. The comparison in terms of performance and numerical complexity with the FastICA, RobustICA and DelL$^{\mathbb{R}}$ algorithms shows that the proposed algorithm offers the best trade-off between performance and numerical complexity when a low number of electrodes is available, such as in pediatric patients. We also examined the feasibility of JDICA in the case of real interictal data and showed that the JDICA algorithm is able to properly denoise real data as well as simulated ones. As a part of our future work, we will examine the proposed algorithm with higher number of electrodes which may lead to different results.

APPENDIX

$$d_{40} = C_{p,p,p,p,\boldsymbol{x}}, \quad d_{31} = 4C_{P,p,p,p,\boldsymbol{x}}, \quad d_{22} = 6C_{P,P,p,p,\boldsymbol{x}}, \quad d_{13} = 4C_{P,P,P,p,\boldsymbol{x}}, \quad d_{04} = C_{P,P,P,P,\boldsymbol{x}}$$

$$d_{30} = 4 \sum_{\substack{i_1=1 \\ i_1 \neq p}}^{P-1} C_{p,p,p,i_1,\boldsymbol{x}} g_{i_1}^{(1-)}, \quad d_{21} = 12 \sum_{\substack{i_1=1 \\ i_1 \neq p}}^{P-1} C_{P,p,p,i_1,\boldsymbol{x}} g_{i_1}^{(1-)}$$

$$d_{12} = 12 \sum_{\substack{i_1=1 \\ i_1 \neq p}}^{P-1} C_{P,P,p,i_1,\boldsymbol{x}} g_{i_1}^{(1-)}, \quad d_{03} = 4 \sum_{\substack{i_1=1 \\ i_1 \neq p}}^{P-1} C_{P,P,P,i_1,\boldsymbol{x}} g_{i_1}^{(1-)}$$

$$d_{20} = 6 \sum_{\substack{i_1=1 \\ i_1 \neq p}}^{P-1} \left(C_{p,p,i_1,i_1,\boldsymbol{x}} (g_{i_1}^{(1-)})^2 + 2 \sum_{\substack{i_2=1 \\ i_2 \neq p}}^{i_1-1} C_{p,p,i_1,i_2,\boldsymbol{x}} g_{i_1}^{(1-)} g_{i_2}^{(1-)} \right)$$

$$d_{11} = 12 \sum_{\substack{i_1=1 \\ i_1 \neq p}}^{P-1} \left(C_{P,P,p,i_1,\boldsymbol{x}} (g_{i_1}^{(1-)})^2 + 2 \sum_{\substack{i_2=1 \\ i_2 \neq p}}^{i_1-1} C_{P,p,i_1,i_2,\boldsymbol{x}} g_{i_1}^{(1-)} g_{i_2}^{(1-)} \right)$$

$$d_{02} = 6 \sum_{\substack{i_1=1 \\ i_1 \neq p}}^{P-1} \left(C_{P,P,i_1,i_1,\boldsymbol{x}} (g_{i_1}^{(1-)})^2 + 2 \sum_{\substack{i_2=1 \\ i_2 \neq p}}^{i_1-1} C_{P,P,i_1,i_2,\boldsymbol{x}} g_{i_1}^{(1-)} g_{i_2}^{(1-)} \right)$$

$$d_{10} = 4 \sum_{\substack{i_1=1 \\ i_1 \neq p}}^{P-1} \left(C_{P,i_1,i_1,i_1,\boldsymbol{x}} (g_{i_1}^{(1-)})^3 + 3 \sum_{\substack{i_2=1 \\ i_2 \neq i_1,p}}^{P-1} C_{k,i_1,i_1,i_2,\boldsymbol{x}} (g_{i_1}^{(1-)})^2 g_{i_2}^{(1-)} \right.$$
$$\left. + 6 \sum_{\substack{i_2=1 \\ i_2 \neq p}}^{i_1-1} \sum_{\substack{i_3=1 \\ i_3 \neq p}}^{i_2-1} C_{p,i_1,i_2,i_3,\boldsymbol{x}} g_{i_1}^{(1-)} g_{i_2}^{(1-)} g_{i_3}^{(1-)} \right)$$

$$d_{01} = 4 \sum_{\substack{i_1=1 \\ i_1 \neq p}}^{P-1} \left(C_{P,i_1,i_1,i_1,\boldsymbol{x}} (g_{i_1}^{(1-)})^3 + 3 \sum_{\substack{i_2=1 \\ i_2 \neq i_1,p}}^{P-1} C_{P,i_1,i_1,i_2,\boldsymbol{x}} (g_{i_1}^{(1-)})^2 g_{i_2}^{(1-)} \right.$$
$$\left. + 6 \sum_{\substack{i_2=1 \\ i_2 \neq k}}^{i_1-1} \sum_{\substack{i_3=1 \\ i_3 \neq p}}^{i_2-1} C_{P,i_1,i_2,i_3,\boldsymbol{x}} g_{i_1}^{(1-)} g_{i_2}^{(1-)} g_{i_3}^{(1-)} \right)$$

$$d_{00} = \sum_{\substack{i_1=1 \\ i_1 \neq p}}^{P-1} \sum_{\substack{i_2=1 \\ i_2 \neq p}}^{P-1} \sum_{\substack{i_3=1 \\ i_3 \neq p}}^{P-1} \sum_{\substack{i_4=1 \\ i_4 \neq p}}^{P-1} C_{i_1,i_2,i_3,i_4,\boldsymbol{x}} g_{i_1}^{(1-)} g_{i_2}^{(1-)} g_{i_3}^{(1-)} g_{i_4}^{(1-)}$$

$$e_0 = d_{11}(g_P^{(1-)})^2 - d_{11}(g_p^{(1-)})^2 + d_{13}(g_P^{(1-)})^4 - d_{31}(g_p^{(1-)})^4 + (2d_{22}-4d_{04})g_p^{(1-)}(g_P^{(1-)})^3 + (4d_{40}-2d_{22})(g_p^{(1-)})^3 g_P^{(1-)}$$
$$+ (3d_{31}-3d_{13})(g_p^{(1-)})^2(g_P^{(1-)})^2 + (2d_{20}-2d_{02})g_p^{(1-)}g_P^{(1-)}$$

$$e_1 = (2d_{02}-2d_{20})(g_P^{(1-)})^2 + (2d_{20}-2d_{02})(g_p^{(1-)})^2 + (2d_{22}-4d_{04})(g_P^{(1-)})^4 + (2d_{22}-4d_{40})(g_p^{(1-)})^4 - 4d_{11}g_p^{(1-)}g_P^{(1-)}$$
$$+ (6d_{31}-10d_{13})g_p^{(1-)}(g_P^{(1-)})^3 + (6d_{13}-10d_{31})(g_p^{(1-)})^3 g_P^{(1-)} + (12d_{04}-12d_{22}+12d_{40})(g_p^{(1-)})^2(g_P^{(1-)})^2$$

$$e_2 = (3d_{31}-3d_{13})(g_P^{(1-)})^4 + (3d_{31}-3d_{13})(g_P^{(1-)})^4 + (12d_{04}-12d_{22}+12d_{40})g_p^{(1-)}(g_P^{(1-)})^3 + (12d_{22}-12d_{04}-12d_{40})(g_p^{(1-)})^3 g_P^{(1-)}$$
$$+ (18d_{13}-18d_{31})(g_p^{(1-)})^2(g_P^{(1-)})^2$$

$$e_3 = (2d_{02}-2d_{20})(g_p^{(1-)})^2 + (4d_{04}-2d_{22})(g_P^{(1-)})^4 + (2d_{20}-2d_{02})(g_P^{(1-)})^2 + (4d_{40}-2d_{22})(g_p^{(1-)})^4 - 4d_{11}g_p^{(1-)}g_P^{(1-)}$$
$$+ (6d_{13}-10d_{31})g_p^{(1-)}(g_P^{(1-)})^3 + (6d_{31}-10d_{13})(g_p^{(1-)})^3 g_P^{(1-)} + (12d_{22}-12d_{04}-12d_{40})(g_p^{(1-)})^2(g_P^{(1-)})^2$$

$$e_4 = d_{11}(g_p^{(1-)})^2 + d_{13}(g_p^{(1-)})^4 - d_{11}(g_P^{(1-)})^2 - d_{31}(g_P^{(1-)})^4 + (4d_{04}-2d_{22})(g_p^{(1-)})^3 g_P^{(1-)} + (2d_{22}-4d_{40})g_p^{(1-)}(g_P^{(1-)})^3$$
$$+ (3d_{31}-3d_{13})(g_p^{(1-)})^2(g_P^{(1-)})^2 + (2d_{02}-2d_{20})g_p^{(1-)}g_P^{(1-)}$$

$$f_0 = d_{10}g_P^{(1-)} - d_{01}g_p^{(1-)} + d_{12}(g_P^{(1-)})^3 - d_{21}(g_p^{(1-)})^3 + (2d_{21}-3d_{03})g_p^{(1-)}(g_P^{(1-)})^2 + (3d_{30}-2d_{12})(g_p^{(1-)})^2 g_P^{(1-)}$$

$$f_1 = (2d_{12}-3d_{30})(g_P^{(1-)})^3 + (2d_{21}-3d_{03})(g_p^{(1-)})^3 + (6d_{03}-7d_{21})(g_p^{(1-)})^2 g_P^{(1-)} + (6d_{30}-7d_{12})g_p^{(1-)}(g_P^{(1-)})^2 - d_{10}g_p^{(1-)} - d_{01}g_P^{(1-)}$$

$$f_2 = d_{10}g_P^{(1-)} - d_{01}g_p^{(1-)} + (2d_{21}-3d_{03})(g_p^{(1-)})^3 + (6d_{03}-7d_{21})g_p^{(1-)}(g_P^{(1-)})^2 + (7d_{12}-6d_{30})(g_p^{(1-)})^2 g_P^{(1-)} + (3d_{30}-2d_{12})(g_P^{(1-)})^3$$

$$f_3 = (2d_{12}-3d_{30})g_p^{(1-)}(g_P^{(1-)})^2 - d_{10}g_p^{(1-)} - d_{12}(g_p^{(1-)})^3 - d_{21}(g_P^{(1-)})^3 - d_{01}g_P^{(1-)} + (2d_{21}-3d_{03})(g_p^{(1-)})^2 g_P^{(1-)}$$

REFERENCES

[1] 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 Bull. Polish Acad. Sci.*, vol. 60, no. 3, pp. 407–418, 2012.

[2] G. Birot, L. Albera, F. Wendling, and I. Merlet, "Localization of extended brain sources from EEG/MEG: The ExSo-MUSIC approach," *NeuroImage*, vol. 56, no. 1, pp. 102–113, May 2011.

[3] D. Cosandier-Rimélé, I. Merlet, J. M. Badier, P. Chauvel, and F. Wendling, "The neuronal sources of EEG: Modeling of simultaneous scalp and intracerebral recordings in epilepsy," *NeuroImage*, vol. 42, no. 1, pp. 135–146, Apr. 2008.

[4] N. Delfosse and P. Loubaton, "Adaptive blind separation of independent sources: A deflation approach," *Signal Process.*, vol. 45, no. 1, pp. 59–83, 1995.

[5] A. Delorme, J. Palmer, J. Onton, R. Oostenveld, and S. Makeig, "Independent EEG sources are dipolar," *PloS One*, vol. 7, no. 2, p. e30135, 2012.

[6] A. Hyvärinen and E. Oja, "A fast fixed-point algorithm for independent component analysis," *Neural Comput.*, vol. 9, no. 7, pp. 1483–1492, 1997.

[7] J.-X. Mi, "A novel algorithm for Independent Component Analysis with reference and methods for its applications," *PloS One*, vol. 9, no. 5, p. e93984, 2014.

[8] A. Nandi and V. Zarzoso, "Fourth-order cumulant based blind source separation," *IEEE Signal Process. Lett.*, vol. 3, no. 12, pp. 312–314, 1996.

[9] B. Stoll and E. Moreau, "A generalized ICA algorithm," *IEEE Signal Process. Lett.*, vol. 7, no. 4, pp. 90–92, 2000.

[10] V. Zarzoso and P. Comon, "Robust independent component analysis by iterative maximization of the kurtosis contrast with algebraic optimal step size," *IEEE Trans. Neural Netw.*, vol. 21, no. 2, pp. 248–261, Feb. 2010.