

# A FAST ALGORITHM FOR BLIND EXTRACTION OF SMOOTH COMPLEX SOURCES WITH APPLICATION IN EEG CONDITIONING

*Soroush Javidi and Danilo P. Mandic*

Department of Electrical and Electronic Engineering, Imperial College, London, UK  
{soroush.javidi,d.mandic}@imperial.ac.uk

## ABSTRACT

A fast complex-valued blind source extraction algorithm designed for the extraction of smooth sources, is introduced. This is achieved based on a smoothness constrained cost and by considering the augmented statistics of the latent source signals. The methodology, based on the  $\mathbb{C}\mathbb{R}$  calculus and complex FastICA framework, is shown to be capable of extraction of both complex circular and noncircular smooth sources for real-time brain computer interfacing. The performance of the algorithm is verified on benchmark signals, followed by a biomedical application in the extraction of artifacts from recorded EEG signals in the order of smoothness.

## 1. INTRODUCTION

Blind source extraction (BSE) belongs to the class of blind source separation (BSS) techniques, whereby the aim is not only to recover latent sources from an observed mixture, but to perform the recovery in a certain order. This way, we extract a subset of the sources based on a desired signal property [1], which is particularly useful for large datasets. Several classes of BSE algorithms based on the temporal structure and higher-order statistics of sources have already been explored in the real domain [2, 3]. In the complex domain, such BSE algorithms have been recently introduced; they were designed so as to cater for the generality of complex signals, both circular and noncircular. Performances of these algorithms in the extraction of synthetic and real-world signals have been demonstrated on the extraction of complex circular and noncircular brain electrical activity sources [4].

The unique property of complex statistics is that the complete second-order information about a signal  $\mathbf{z}(k) \in \mathbb{C}^N$  is based on the covariance  $\mathcal{C}_{\mathbf{z}\mathbf{z}} = E\{\mathbf{z}(k)\mathbf{z}^H(k)\}$  and pseudo-covariance  $\mathcal{P}_{\mathbf{z}\mathbf{z}} = E\{\mathbf{z}(k)\mathbf{z}^T(k)\}$  matrices [5, 6]. Signals for which the pseudo-covariance vanishes, that is  $\mathcal{P}_{\mathbf{z}\mathbf{z}} = \mathbf{0}$ , are termed proper or second-order circular, while most complex signals encountered in signals processing applications are improper or second-order noncircular. However, in the literature complex-valued signals have been, either explicitly or implicitly, traditionally considered proper.

The resulting algorithms are optimal for only a subset of complex signals (proper signals), while performing sub-optimally for the wider class of improper signals.

The recent resurgence in real-time signal processing algorithms in  $\mathbb{C}$  has been made possible by so called augmented complex statistics which allow for the consideration of the full second-order information available within the signal [7]. Furthermore, the concept of  $\mathbb{C}\mathbb{R}$  calculus [8, 9] has aided the analysis of cost functions which would otherwise not satisfy the stringent Cauchy-Riemann conditions of the standard complex calculus.

Smoothness is a fundamental signal property, which can be modelled based on the gradients of data vectors. It can also aid BSS and BSE as, for instance, in electroencephalography (EEG) analysis, artifacts coming from eye muscles are smoother than the background EEG. The BSE for real-valued smooth signals in the time-domain was introduced in [10], and an implementation based in the frequency domain was recently proposed in [11]. Processing in the time domain has its merits in retaining the signals in their original form and avoiding extra computations. In addition, processing in the Fourier domain will result in an intermediary smoothing step due to the low-pass filtering performed by the Fourier Transform.

To this end, we propose a blind extraction algorithm for complex-valued signals in time domain. In a similar manner to [10], we introduce a fast converging algorithm by using a fixed-point type update based on the existing complex FastICA algorithm [12, 13]. Such an extraction algorithm can thus be seen as a constrained version of the complex FastICA algorithm, and as shown in the derivation, it simplifies into the unconstrained complex FastICA where the smoothness constraint is removed.

The performance is verified on the removal of artifacts from real-world EEG recordings. It is shown that several types of eye movement artifacts can be successfully removed using the proposed algorithm, this making it attractive for brain computer interface (BCI) applications. Thus by removing or rejecting the artifact sources, further processing on the remaining pure EEG signals is made possible in real-time.

## 2. SMOOTHNESS-BASED BLIND SOURCE EXTRACTION

### 2.1. The Concept of Smoothness in $\mathbb{C}$

The mathematical concept of a smooth function is based on differentiability. Consider the Sobolev space  $W^{p,q} \subset \mathbb{R}^N$  defined as the space where the  $p$ -th power of a function  $f \in W^{p,q}$  and also its first  $q$ -th derivatives are integrable [14]. The norm is then defined as

$$\|f\|_{W^{p,q}} = \left( \sum_{i=0}^q \|D^{(i)}f\|_p^p \right)^{1/p} \quad (1)$$

where  $D^{(i)}f$  denotes the  $i$ th derivative. Due to the duality between  $\mathbb{C}$  and  $\mathbb{R}^2$  [9], the above definition can also be adopted for complex-valued functions. We shall utilise the Sobolev norm for the space  $W^{2,1}$ , where only the second power of the function and its first derivative are considered. Taking an arbitrary upper bound of the ratio between the Sobolev and Euclidean norms of the function  $f$  and after some simplification, we arrive at

$$\frac{\|D^{(1)}f\|_2^2}{\|f\|_2^2} \leq \rho_s \quad (2)$$

where  $\rho_s$  is the upper bound of the ratio, also referred to as the smoothness factor. For a discrete signal  $z(k)$ , a more simplified form is given by

$$E\{|\Delta z(k)|^2\} - \rho_s E\{|z(k)|^2\} \leq 0 \quad (3)$$

where  $\Delta z(k) = z(k) - z(k-1)$ ; a geometric interpretation is given in Fig. 1. Thus, in a similar fashion to the real-valued definition, (3) models a complex-valued signal with a slow varying temporal profile as a smooth signal. Intuitively, a complex-valued signal  $z(k)$  is smooth if the variance of the difference between consecutive samples is less than a pre-defined fraction of the variance of the signal itself. This can also be viewed as measuring the variation in the gradient of the signal<sup>1</sup>.

Alternatively, consider the definition (2) for  $z(k) = z_r(k) + jz_i(k)$ , expressed in its dual form  $\omega_k = [z_r, z_i]^T \in \mathbb{R}^2$ . Then,

$$E\{\langle \Delta \omega_k, \Delta \omega_k \rangle\} - \rho_s E\{\langle \omega_k, \omega_k \rangle\} \leq 0 \quad (4)$$

$$E\{\Delta z_r^2(k)\} + E\{\Delta z_i^2(k)\} - \rho_s (E\{z_r^2(k)\} + E\{z_i^2(k)\}) \leq 0$$

where symbol  $\langle \cdot, \cdot \rangle$  denotes the inner product.

Notice that the smoothness definition based on the Sobolev norm of  $W^{2,1}$  corresponds to the covariances  $\mathcal{C}_{zz}(0)$  and  $\mathcal{C}_{zz}(1)$ , that is, the covariances of lag zero and one.

<sup>1</sup>In  $\mathbb{C}$  relationships such as  $>$  and  $<$  do not apply and we need to resort to the duality between  $\mathbb{R}^2$  and  $\mathbb{C}$ , and to use so called lexicographic ordering.

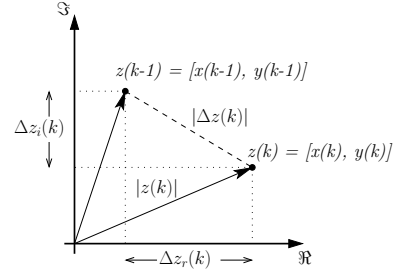


Fig. 1. Geometric interpretation of the smoothness definition given in (3)

### 2.2. The BSE Problem

Consider an observation  $\mathbf{x}(k) \in \mathbb{C}^N$  formed from the linear weighted combination of latent sources  $\mathbf{s}(k) \in \mathbb{C}^{N_s}$ , given by

$$\mathbf{x}(k) = \mathbf{A}\mathbf{s}(k) \quad (5)$$

where  $\mathbf{A} \in \mathbb{C}^{N \times N_s}$  is the mixing matrix, and  $N_s$  is the number of sources. The sources are assumed independent and the observation mixture is whitened prior to processing. The aim is to find a demixing vector  $\mathbf{w}$  that will recover one of the sources, given by

$$y(k) = \mathbf{w}^H(k)\mathbf{x}. \quad (6)$$

Following the standard BSS methodology [15, 12], this can be achieved by maximising the non-Gaussianity of  $y(k)$  reflected in the cost function

$$\mathcal{J}_N(\mathbf{w}, \mathbf{w}^*) = E\{G(|\mathbf{w}^H \mathbf{x}|^2)\} \quad (7)$$

where  $G$  is a nonlinearity used to approximate the associated negentropy, and for generality,  $\mathcal{J}_N$  is expressed using both the conjugate coordinates  $\mathbf{w}$  and  $\mathbf{w}^*$ .

To ensure that components with certain smoothness characteristics are extracted, we aim to impose further constraints on  $\mathcal{J}_N$ . Based on the definition in (3), for the removed source  $y(k)$  given in (6), the smoothness measure becomes

$$\mathcal{J}_S(\mathbf{w}, \mathbf{w}^*) = \beta (E\{|\mathbf{w}^H \Delta \mathbf{x}(k)|^2\} - \rho_s E\{|\mathbf{w}^H \mathbf{x}(k)|^2\}) \quad (8)$$

The constant  $\beta = \{-1, 1\}$  gives us a degree of freedom in dealing with smooth sources, for instance  $\beta = -1$  the extraction of the most non-smooth source will be achieved.

Thus, the optimisation problem of BSE of latent sources based on the smoothness constraint (S-cBSE) can be stated as

$$\mathbf{w}_{opt} = \arg \max_{\|\mathbf{w}\|_2=1} \mathcal{J}_N(\mathbf{w}, \mathbf{w}^*)$$

$$\text{subject to } \mathcal{J}_S(\mathbf{w}, \mathbf{w}^*) \leq 0. \quad (9)$$

After every step, the demixing  $\mathbf{w}$  is normalised, to avoid spurious solutions.

To solve the optimisation problem (9), we will employ the method of Lagrangian multipliers. The extrema of the Lagrangian can be found using the Newton method, resulting in faster convergence to the solution; this method has been shown to be stable for a related unconstrained problem in the complex domain [13], and a detailed proof of the derivation is given in the Appendix. The Newton based optimisation of the Lagrangian is performed as

$$\begin{aligned} \Delta \mathbf{w} &= \left( \mathbf{H}_{\mathbf{w}\mathbf{w}^*} - \mathbf{H}_{\mathbf{w}^*\mathbf{w}^*} \mathbf{H}_{\mathbf{w}^*\mathbf{w}}^{-1} \mathbf{H}_{\mathbf{w}\mathbf{w}} \right)^{-1} \\ &\quad \times \left( \mathbf{H}_{\mathbf{w}^*\mathbf{w}^*} \mathbf{H}_{\mathbf{w}^*\mathbf{w}}^{-1} \frac{\partial \mathcal{L}}{\partial \mathbf{w}} - \frac{\partial \mathcal{L}}{\partial \mathbf{w}^*} \right) \\ \Delta \lambda &= \nabla_{\lambda} \mathcal{L} \\ \mathbf{w}(k+1) &\leftarrow \mathbf{w}(k+1) / \|\mathbf{w}(k+1)\|_2 \end{aligned} \quad (10)$$

where  $\mathcal{L}(\mathbf{w}, \mathbf{w}^*, \lambda)$  is the Lagrangian function,  $\lambda$  is the Lagrangian multiplier and  $\mathbf{H}$  matrices are the Hessians of  $\mathcal{L}$ .

To extract successive smooth (non-smooth) sources, a deflationary orthogonalisation process using the Gram-Schmidt method is performed after each iteration of the extraction algorithm in (10). While this allows for unambiguous extractions, errors in the extraction and thus deflation process can accumulate, resulting in decreased performance over consecutive extractions<sup>2</sup>. The deflation procedure for the  $i$ th demixing vector can be compactly written as

$$\mathbf{w}_i(k+1) \leftarrow \mathbf{w}_i(k+1) - \mathbf{W}\mathbf{W}^H \mathbf{w}_i(k+1) \quad (11)$$

where  $\mathbf{W} = [\mathbf{w}_1(k+1), \dots, \mathbf{w}_i(k+1)]$ .

### 3. PERFORMANCE BENCHMARKING

To illustrate the performance of the proposed algorithm, sub-Gaussian and super-Gaussian complex-valued sources with different degrees of noncircularity were used. The smoothness degree of the sources

$$\rho_s = \frac{E\{|\Delta z(k)|^2\}}{E\{|z(k)|^2\}} \quad (12)$$

was measured using (3), while the circularity was measured as the ratio of the absolute value of the pseudo-variance  $\tau_z^2 = E\{z^2\}$  to the variance  $\sigma_z^2 = E\{|z|^2\}$  of the source, as described in [16], and expressed as

$$r = \frac{|\tau_z^2|}{\sigma_z^2}, \quad r \in [0, 1]. \quad (13)$$

The value  $r = 0$  denotes a second-order circular source, while  $r = 1$  indicates a highly noncircular source.

<sup>2</sup>In practical applications, this usually does not pose a problem, as we are interested in only 1-2 smooth sources (artifacts).

The performance of the algorithm was measured using the Performance Index (PI) given by [1]

$$PI = 10 \log_{10} \left( \frac{1}{M} \left( \sum_{i=1}^M \frac{|u_i|^2}{\max\{|u_1|^2, \dots, |u_M|^2\}} \right) \right). \quad (14)$$

where  $\mathbf{u} = \mathbf{A}^H \mathbf{w} = [u_1, \dots, u_M]$ .

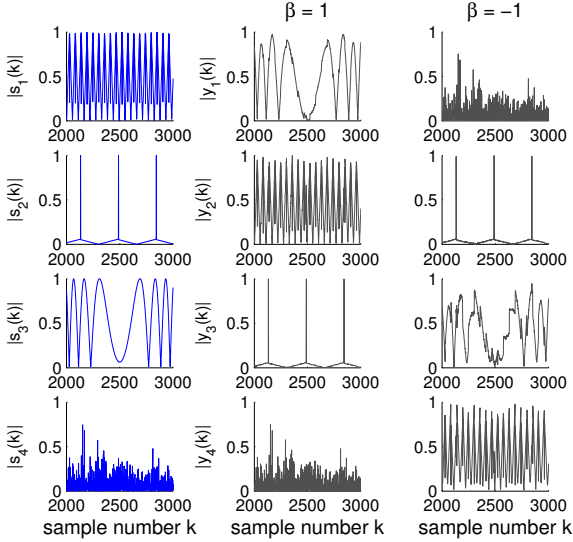
Four complex-valued sources of 5000 samples were mixed using a randomly generated  $4 \times 4$  mixing matrix to form the observed mixtures. The magnitude of the sources are shown in Fig. 2 and the signal properties given in Table 1. The mixture was whitened and the latent sources were extracted using the S-cBSE algorithm (10). In the first experiment, the value of  $\beta = 1$ ,  $\rho_s = 0.9$ ,  $\lambda = 1$  and  $\mu_\lambda = 0.01$ . As the signals were synthetically generated, the value of  $\rho_s$  was chosen based on measurements of the signal smoothness. The nonlinearity  $G(z) = \log \cosh(z)$  ensured that the negentropy of both sub-Gaussian and super-Gaussian sources were sufficiently approximated for maximisation.

Fig. 2 shows the sources which were successfully extracted based on the smoothness criterion. For comparison, the measured smoothness factors for the extracted sources are given in Table 1. Notice that as  $\{\rho_{s_3}, \rho_{s_1}\} \leq 0.9$  it is expected that only sources  $s_3(k)$  and  $s_1(k)$  were to be extracted, however, the algorithm also successfully extracted the subsequent sources  $s_2(k)$  and  $s_4(k)$ . This can be attributed to the strong non-Gaussianity condition in (7), which is sufficient for successful extraction. The plot of performance index at each iteration (Fig. 3) shows that the algorithm achieved convergence with a PI of around -30dB for the source estimates  $y_1(k)$ ,  $y_2(k)$  and  $y_4(k)$  in under 10 iterations, while source estimate  $y_3(k)$  achieved a PI of under -35dB in 19 iterations. Alternatively, expressed in terms of the signal-to-interference ratio (SIR), the values for the consecutive extractions were respectively 29.81dB, 23.23dB, 21.76dB and 25.68dB. For comparison, the performance of the S-cBSE algorithm in extracting the first desired source based on the standard complex FastICA [12] is shown by the dotted line. Notice that the algorithm is not suitable for the extraction of noncircular sources.

In the next experiment, the objective was to extract the non-smooth sources, for which  $\beta = -1$  and  $\rho_s = 2$ . The values of the other parameters were set empirically to  $\lambda = 20$  and  $\mu_\lambda = 1$  and the nonlinear function  $G$  was kept as before. The sources were extracted in the order of increasing smoothness, with the performance indices over the extraction process plotted in Fig. 3. The PI value for the source estimate  $y_1(k)$  was around -30dB while  $y_2(k)$  achieved a limit cycle with a varying PI of around -22dB to -30dB. Source estimate  $y_3(k)$  initially converged but diverged after 3 iterations and  $y_4(k)$  only achieved a PI of around -20dB. While source  $s_4(k)$  was the only non-smooth signal according to the value set for  $\rho_s$ , source  $s_2(k)$  was also success-

**Table 1.** Source properties for extraction simulations

Source	$r$	$\rho_s$	$\beta = 1$ $\beta = -1$	
			$\hat{\rho}_s$	$\hat{\rho}_s$
$s_1(k)$	0.9997	0.1154	0.1200	0.0193
$s_2(k)$	0.9865	1.4771	1.4745	1.4782
$s_3(k)$	0.9998	0.0148	0.0150	0.1136
$s_4(k)$	0.9995	2.0214	2.0219	2.0204

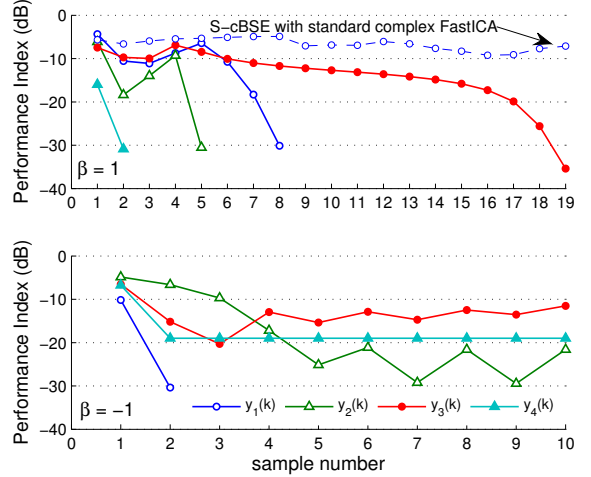


**Fig. 2.** Performance of the algorithm (10) in the extraction of smooth ( $\beta = 1$ ) and non-smooth ( $\beta = -1$ ) sources

fully extracted due to the close proximity to the smoothness criterion. However, note that sources  $s_1(k)$  and  $s_3(k)$  were not successfully extracted due to the disparity between the values of  $\rho_{s_1}$  and  $\rho_{s_3}$  to  $\rho_s = 2$  as set for this experiment. The SIR for the consecutive extractions were respectively 23.87dB, 27.45dB, 3.93dB and 3.87dB.

#### 4. ARTIFACT EXTRACTION FROM EEG

The S-cBSE algorithm was next utilised to extract power line noise, biological eye blink (electrooculogram, EOG), and eye muscle activity (electromyogram, EMG) artifacts, common in EEG recordings. The aim is to condition the contaminated recording so that further processing, such as those necessary in real-time BCI, can be performed on pure EEG signals. To this end, an EEG recording was made where the electrodes were placed according to the 10-20 systems at positions Fp1, Fp2, C3, C4, O1, O2 and the ground placed at Cz. The electrical activity from the EOG and EMG artifacts were recorded using the vEOG and hEOG channels, with electrodes placed around the eye. The recording lasted 30s and the data were sampled at a rate of 256Hz. In the first study, the participants were asked to blink at random intervals while looking straight. In the second study,



**Fig. 3.** Performance of the algorithm (10) in the extraction of smooth ( $\beta = 1$ ) sources and non-smooth ( $\beta = -1$ ) sources. Performance of the S-cBSE algorithm based on the standard complex FastICA [12] for the first extracted source with  $\beta = 1$  is shown by the dotted line.

the instructions were to move the eyes in a vertical motion at random intervals.

The recorded EEG channels were combined into temporal complex-valued mixtures such that the real and imaginary components comprised symmetric EEG channels. In this manner, the cross-information due to the phase and magnitude relationship between pairs of symmetric electrodes was utilised by the extraction algorithm [17]. The complex EEG mixtures  $x(k)$  were generated as

$$\begin{aligned}
 x_1(k) &= \text{Fp1}(k) + j\text{Fp2}(k) \\
 x_2(k) &= \text{C3}(k) + j\text{C4}(k) \\
 x_3(k) &= \text{O1}(k) + j\text{O2}(k).
 \end{aligned} \tag{15}$$

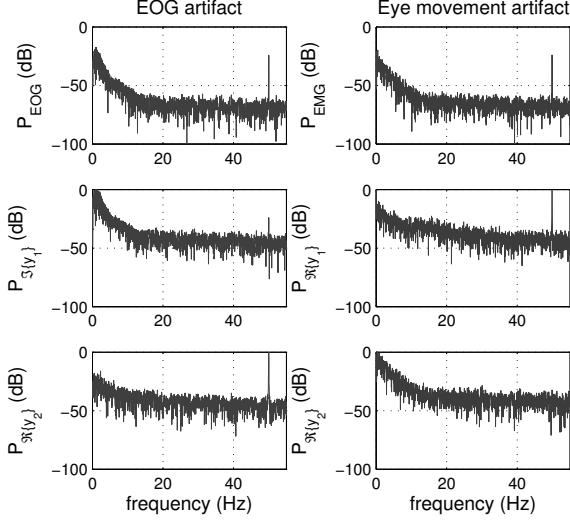
In the EOG study, the algorithm (10) was used to extract two independent sources, and was initialised respectively with  $\beta_1 = 1, \beta_2 = -1$ , and  $\rho_{s,1} = 0.01, \rho_{s,2} = 0.9$  for the first and second extraction steps, while the value of  $\lambda = 80$  and  $\mu_\lambda = 1$  for both steps. These values were deduced from prior information about both artifacts; the periodic power line noise was non-smooth<sup>3</sup>, while the intermittent EOG activity was smooth in comparison to the pure EEG data.

The real and imaginary components of the complex-valued extracted signal  $y(k)$  represent the actual real-valued latent sources. After the completion of each extraction stage, the smoothness of the real and imaginary components were measured, and the component matching the criterion was removed. The smoothness values for the extracted signals  $y_1(k)$  and  $y_2(k)$ , and their respective real and imaginary

<sup>3</sup>This can be attributed to the low sampling rate, a limitation of the recording hardware.

**Table 2.** Smoothness properties for extracted EEG artifacts. The rejected components are shown in bold font.

Dataset	Source	$\rho_s$	$\rho_s(\Re, \Im)$
'EOG'	$y_1(k)$	0.0274	0.2706, <b>0.0085</b>
	$y_2(k)$	1.2910	<b>1.3179</b> , 0.8494
'EMG'	$y_1(k)$	0.7333	<b>0.7748</b> , 0.2323
	$y_2(k)$	0.1438	<b>0.0142</b> , 0.1242



**Fig. 4.** *Left:* Power spectrum of the recorded EOG and the extracted artifacts, *Right:* Power spectrum of the EMG due to eye movement and the extracted artifacts.

components are given in Table 2.

A qualitative assessment of the extraction was performed by comparing the power spectrum of the reference biological artifact and the power spectrum of the extracted artifacts, such as the EOG shown in the left column of Fig. 4. The power spectrum of the raw EOG illustrates the presence of frequencies from 0Hz-5Hz and the power line activity at 50Hz. The power spectrum of  $\Im\{y_1(k)\}$  shows that the algorithm successfully extracted the EOG source, while attenuating the 50Hz frequency. The 50Hz source was contained within the real component of the second extracted source  $\Re\{y_2(k)\}$ , as seen from the corresponding power spectrum.

For the EMG study, the S-cBSE algorithm was initialised such that  $\beta_1 = -1, \beta_2 = 1$  and  $\rho_{s,1} = 0.9, \rho_{s,2} = 0.05$ . The parameters  $\lambda_1 = 1, \lambda_2 = 10$  and  $\mu_\lambda = 1$  for both extraction steps. The smoothness factor of the extracted sources and their respective components are given in Table 2, and the power spectrum associated with the recorded eye muscle activity and the extracted components is given in Fig. 4. Observe that the real component of  $y_1(k)$  contained the power line activity, while the real component of  $y_2(k)$  represented the EMG activity.

## 5. CONCLUSIONS AND FUTURE WORK

We have introduced an algorithm for smoothness based complex blind source extraction (S-cBSE). The concept of smoothness has been defined for general complex-valued signals and was employed to define a constrained cost function, based on the maximisation of non-Gaussianity. The fast convergence of the algorithm is inherited from FastICA, confirmed on benchmark data. Further, an application in the extraction of power line noise and biological artifacts from contaminated EEG recordings has been addressed.

As illustrated in the simulations, the algorithm responds to the choice of the smoothness factor  $\rho_s$ . Extensions of this work will include stricter definition of smoothness by considering different norms in the Sobolev space. In addition, while the algorithm is unaffected by noise represented as a source signal, the performance degrades in the presence of additive noise. It is thus desirable to enhance robustness through bias removal within the cost function.

## 6. APPENDIX

### 6.1. Derivation of the S-cBSE Algorithm

First, note that due to the whiteness of  $\mathbf{x}(k)$ , the cost  $\mathcal{J}_S$  in (8) can be expanded as

$$\begin{aligned} \mathcal{J}_S &= \mathbf{w}^H E\{\Delta\mathbf{x}\Delta\mathbf{x}^H\}\mathbf{w} - \rho_s \mathbf{w}^H E\{\mathbf{x}\mathbf{x}^H\}\mathbf{w} \\ &= \mathbf{w}^H \underbrace{[\mathcal{C}_{\Delta\mathbf{x}\Delta\mathbf{x}} - \rho_s \mathbf{I}]}_{\triangleq \mathbf{B}} \mathbf{w} \end{aligned} \quad (16)$$

where  $\mathbf{B} = \mathbf{B}^H$  and  $\mathbf{I}$  is the identity matrix.

To solve the constrained optimisation problem (9), consider the Lagrangian function  $\mathcal{L}(\mathbf{w}, \mathbf{w}^*, \lambda) : \mathbb{C}^N \times \mathbb{C}^N \times \mathbb{R} \mapsto \mathbb{R}$  given by

$$\mathcal{L}(\mathbf{w}, \mathbf{w}^*, \lambda) = \mathcal{J}_N(\mathbf{w}, \mathbf{w}^*) + \lambda \mathcal{J}_S(\mathbf{w}, \mathbf{w}^*) \quad (17)$$

where  $\lambda \in \mathbb{R}$  is the Lagrange multiplier. For the inequality constraint  $\mathcal{J}_S$ , the Karush-Kuhn-Tucker conditions are to be considered and satisfied. However, we will follow the method in [10] and transform the smoothness inequality constraint into the equality constraint  $\hat{\mathcal{J}}_S = \max(\mathcal{J}_S, 0) = 0$ , resulting in a simpler solution.

The Newton method is then used to find the extrema of the Lagrangian, defined in augmented complex form as [18]

$$\Delta \mathbf{w}^a = -\mathbf{H}_{\mathbf{w}\mathbf{w}}^{a-1} \left( \frac{\partial \mathcal{L}}{\partial \mathbf{w}^{a*}} \right) \quad (18)$$

where  $\mathbf{w}^a = [\mathbf{w}^T, \mathbf{w}^H]^T$  denotes an augmented complex column vector and  $\mathbf{H}_{\mathbf{w}\mathbf{w}}^a$  is the augmented Hessian matrix, given by<sup>4</sup>

$$\mathbf{H}_{\mathbf{w}\mathbf{w}}^a = \begin{bmatrix} \mathbf{H}_{\mathbf{w}\mathbf{w}^*} & \mathbf{H}_{\mathbf{w}^*\mathbf{w}^*} \\ \mathbf{H}_{\mathbf{w}\mathbf{w}} & \mathbf{H}_{\mathbf{w}^*\mathbf{w}} \end{bmatrix} \quad (19)$$

<sup>4</sup>For the calculation of the S-cBSE algorithm based on the standard complex FastICA, the block off-diagonal elements of  $\mathbf{H}_{\mathbf{w}\mathbf{w}}^a$  are assumed zero, and form a quasi-Newton Hessian matrix.

Expanding the augmented Newton update and solving for  $\Delta \mathbf{w}$  results in the Newton step given in (10) (see also [9]), where the individual gradient components, calculated using  $\mathbb{C}\mathbb{R}$  calculus, are given by

$$\begin{aligned}\frac{\partial \mathcal{L}}{\partial \mathbf{w}^*} &= E\{g(|y|^2)y^* \mathbf{x}\} + \lambda \epsilon \beta \mathbf{B} \mathbf{w} \\ \frac{\partial \mathcal{L}}{\partial \mathbf{w}} &= \left( \frac{\partial \mathcal{L}}{\partial \mathbf{w}^*} \right)^*,\end{aligned}\quad (20)$$

and the Hessian components are given by

$$\begin{aligned}\mathbf{H}_{\mathbf{w}^* \mathbf{w}^*} &= \frac{\partial}{\partial \mathbf{w}^*} \left( \frac{\partial \mathcal{L}}{\partial \mathbf{w}^*} \right)^T = E\{g'(|y|^2)y^{*2} \mathbf{x} \mathbf{x}^T\} \\ &\approx E\{g'(|y|^2)y^{*2}\} E\{\mathbf{x} \mathbf{x}^T\} \\ \mathbf{H}_{\mathbf{w}^* \mathbf{w}} &= \frac{\partial}{\partial \mathbf{w}^*} \left( \frac{\partial \mathcal{L}}{\partial \mathbf{w}} \right)^T = E\{g'(|y|^2)|y|^2 + g(|y|^2)\} \mathbf{I} \\ &\quad + \lambda \epsilon \beta \mathbf{B} \\ \mathbf{H}_{\mathbf{w} \mathbf{w}} &= (\mathbf{H}_{\mathbf{w}^* \mathbf{w}^*})^* \quad \mathbf{H}_{\mathbf{w} \mathbf{w}^*} = (\mathbf{H}_{\mathbf{w}^* \mathbf{w}})^*,\end{aligned}\quad (21)$$

with  $\epsilon = (\text{sgn}(\mathcal{J}_S) + 1)/2$ , and  $g$  and  $g'$  denote the first and second derivative of the nonlinearity  $G$ . As in [15], for whitened data we make use of the approximation  $E\{f(x)xx\} \approx E\{f(x)\}E\{xx\}$ .

The value of  $\lambda$  is updated using a gradient ascent method at each iteration, given in (10). If the value of  $\lambda = 0$ , we arrive at the unconstrained problem, for which the solution is given in [13], as a generalised complex FastICA algorithm.

## 7. REFERENCES

- [1] A. Cichocki and S. Amari, *Adaptive Blind Signal and Image Processing, Learning Algorithms and Applications*, Wiley, 2002.
- [2] Q. Shi, R. Wu, and S. Wang, "A novel approach to blind source extraction based on skewness," in *ICSP 2006*, Nov. 2006, vol. 4, pp. 3187–3190.
- [3] W. Liu, D. P. Mandic, and A. Cichocki, "Blind second-order source extraction of instantaneous noisy mixtures," *IEEE Transactions of Circuits and Systems II*, vol. 53, no. 9, pp. 931–935, 2006.
- [4] S. Javidi, D. P. Mandic, and A. Cichocki, "Complex blind source extraction from noisy mixtures using second order statistics," *IEEE Transactions on Circuits and System I*, 2010 (accepted).
- [5] B. Picinbono and P. Bondon, "Second-order statistics of complex signals," *IEEE Transactions on Signal Processing*, vol. 45, no. 2, pp. 411–420, 1997.
- [6] F.D. Neeser and J.L. Massey, "Proper complex random processes with applications to information theory," *IEEE Transactions on Information Theory*, vol. 39, no. 4, pp. 1293–1302, 1993.
- [7] D. P. Mandic and S. L. Goh, *Complex Valued Non-linear Adaptive Filters: Noncircularity, Widely Linear and Neural Models*, Wiley, 2009.
- [8] W. Wirtinger, "Zur formalen theorie der funktionen von mehr komplexen veränderlichen," *Mathematische Annalen*, vol. 97, no. 1, pp. 357–375, Dec. 1927.
- [9] K. Kreutz-Delgado, "The complex gradient operator and the  $\mathbb{C}\mathbb{R}$ -calculus," *Dept. of Electrical and Computer Engineering, UC San Diego, Course Lecture Supplement No. ECE275A*, pp. 1–74, 2006.
- [10] N. Mitianoudis, T. Stathaki, and A. G. Constantinides, "Smooth signal extraction from instantaneous mixtures," *IEEE Signal Processing Letters*, vol. 14, no. 4, pp. 271–274, 2007.
- [11] L. T. Duarte, B. Rivet, and C. Jutten, "Blind extraction of smooth signals based on a second-order frequency identification algorithm," *IEEE Signal Processing Letters*, vol. 17, no. 1, pp. 79–82, 2010.
- [12] E. Bingham and A. Hyvärinen, "A fast fixed point algorithm for independent component analysis of complex valued signals," *Journal of Neural Systems*, vol. 10, pp. 1–8, 2000.
- [13] M. Novey and T. Adalı, "On extending the complex FastICA algorithm to noncircular sources," *IEEE Transactions on Signal Processing*, vol. 56, no. 5, pp. 2148–2154, 2008.
- [14] R. A. Adams and J. J. F. Fournier, *Sobolev spaces*, Academic Press, 1975.
- [15] A. Hyvärinen, J. Karhunen, and E. Oja, *Independent Component Analysis*, Wiley, 2001.
- [16] E. Ollila, "On the circularity of a complex random variable," *IEEE Signal Processing Letters*, vol. 15, pp. 841–844, 2008.
- [17] D. P. Mandic, S. Javidi, G. Soudretis, and S. L. Goh, "Why a complex valued solution for a real domain problem," in *IEEE Workshop on Machine Learning for Signal Processing*, Aug. 2007, pp. 384–389.
- [18] A. van den Bos, "Complex gradient and Hessian," *IEE Proceedings of Vision, Image and Signal Processing*, vol. 141, no. 6, pp. 380–383, 1994.