Blind separation of skewed signals in instantaneous mixtures

Nikolaos Mitianoudis, Tania Stathaki
Communications and Signal Processing Group
Imperial College
Exhibition Road, SW7 2AZ London, UK
Email: n.mitianoudis@imperial.ac.uk

Mike Davies
Centre for Digital Music
Queen Mary London
Mile End Road, E1 4NS, London, UK

***Abstract*— The problem of source separation of instantaneous mixtures has been addressed thoroughly in literature in the past. The assumption of statistical independence between the source signals, led to the introduction of Independent Component Analysis (ICA). A number of methods, based on the ICA framework, can identify nonGaussian sources in instantaneous mixtures with robust convergence and performance. However, in several biomedical applications, there is a need to identify and separate signals that, apart from being nonGaussian, are not symmetric. In this article, the authors present a method for blind identification and separation of skewed (non-symmetric) signals in a linear instantaneous mixture.**

I. Introduction

Assume a set of M sensors monitoring a phenomenon via the signals $\underline{x}(n) = [x_1(n), x_2(n), \ldots, x_M(n)]^T$. Let us also assume that there are a number of underlying N factors (sources) $\underline{s}(n) = [s_1(n), s_2(n), \ldots, s_N(n)]^T$ that trigger the phenomenon, observed by the sensors. We will assume that the contribution of each factor is transmitted with insignificant delay to the observing sensors, i.e. instantaneously. In addition, possible corruption by additive noise is considered insignificant. The following model connects the observed signals with the source signals via instantaneous mixing.

$$\underline{x} = A\underline{s} \tag{1}$$

where A is a mixing matrix denoting instantaneous transmission. In this study, we will assume equal number of sources and sensors ($N = M$). Although this linear instantaneous model may seem unrealistic, there are a lot of real-life applications, where it can serve as a very good approximation. In biomedical signal processing, there are a number of monitoring signals that are considered instantaneous mixtures of input sources, such as the *electrocardiograph* (ECG) and the *electroencephalogram* (EEG) signals. The general non-linear source separation problem is a more demanding problem, which usually can not be addressed by the methods proposed for linear mixtures.

A number of blind source separation approaches have been proposed in the past [4]. Introducing the assumption of statistical independence between the source signals, led to the development of *Independent Component Analysis* (ICA). Using this framework, one can separate nonGaussian sources (in fact only one is allowed to be Gaussian) with a number of different methodologies. Some approaches perform separation by minimising the *Kullback-Leibler* (KL) divergence between the separated sources and *several probabilistic priors* on the source signals. Other approaches minimise the *mutual information* conveyed by the separated sources or perform approximate diagonalisation of a *cumulant tensor* of the mixtures. Finally, some methods perform separation by estimating the directions of the most nonGaussian components using *kurtosis* or *negentropy*, as nonGaussianity measures. For more on these techniques, one can refer to tutorial books on ICA, such as [1], [4].

However, in some applications it is necessary to identify signals with other statistical properties, apart from nonGaussianity. One of these characteristics might be the symmetry of the distribution. Several sources of interest in biomedical signals are skewed. In [9], Stetson used ICA and skewness as a criterion to identify the arterial pulse signal from noisy mixtures. In [8], Sanei and Shoker proposed the use of ICA and support-vector machines together with a number of features (among which was skewness) to identify the eye-blinking artifact in EEG signals. Consequently, this asymmetry can be used as a tool to identify certain signals in biomedical applications.

In this paper, we derive a FastICA [5] type of algorithm for skewness explicitly. In a similar manner to optimising *kurtosis* [5], one can optimise third-order moments, like *skewness*, to separate non-symmetrical signals. The proposed algorithm shows promising results on artificial data in terms of convergence and separation quality. We also present some preliminary results on Ventricular Activity separation from ECG data and on eye-blinking identification in EEG data.

II. Source separation of skewed signals

A. Definition of skewness

In statistics, *skewness* is a measure of symmetry, or more precisely, the lack of symmetry. A data set is symmetric if it looks the same to the left and right of the center point (sample mean). Skewness can also be considered a third-order moment. Assuming that u is a random variable of non-zero mean, then skewness can be defined as follows:

$$\text{skew}(u) = \frac{\mathcal{E}\{(u - \mathcal{E}\{u\})^3\}}{\mathcal{E}\{(u - \mathcal{E}\{u\})^2\}^{3/2}} \quad (2)$$

where $\mathcal{E}\{\cdot\}$ represents the expectation operator. Skewness takes positive values when the signal is asymmetrical to the right and negative values, when it is skewed to the left. As we are not interested in the sign of skewness, it would be appropriate to optimise the absolute value of skewness. As some of the input signals will be skewed, the mean should not be considered negligible and should be taken into account in our analysis.

B. Principal Component Analysis

The first task will be to "prewhiten" the data. A Principal Component Analysis (PCA) step will orthogonalise (decorrelate) and normalise the data to unit variance [4]. The prewhitening matrix V is formed by the eigenvectors of the covariance matrix $C = \mathcal{E}\{(\underline{x} - \mathcal{E}\{\underline{x}\})(\underline{x} - \mathcal{E}\{\underline{x}\})^T\}$. Assuming that H is a matrix containing all the eigenvectors of C and D a diagonal matrix containing the eigenvalues of C. The eigenvalue at the i-th diagonal element should correspond to the eigenvector at the i-th column of H. Then, the prewhitening matrix V and the "whitened" data $\underline{z}$ are given by the following equations:

$$V = D^{-0.5} H^T \quad (3)$$

$$\underline{z} = V\underline{x} \quad (4)$$

where $\mathcal{E}\{(\underline{z} - \mathcal{E}\{\underline{z}\})(\underline{z} - \mathcal{E}\{\underline{z}\})^T\} = I$. As it is shown in source separation literature [4], decorrelation is not a sufficient condition to separate independent signals, as it will only decorrelate (orthogonalise) the data. Therefore, in order to isolate L skewed components, we need to estimate L projection operators $\underline{w}_i$ that will isolate the skewed components u_i, existing in the mixture ($L \leq N$).

$$u_i = \underline{w}_i^T \underline{z} \quad \forall\ i = 1, \ldots, L \quad (5)$$

C. Separation optimising skewness

In this section, we seek to identify a single skewed component u in the mixture with the help of skewness. In order to estimate u, we need to optimise the absolute value of skewness to cater for both directions of skew. In this effort, we will optimise the nominator only. Restricting the projection operator $\underline{w}$ to perform rotation only, we have to impose the constraint that $||\underline{w}||^2 = 1$, where $||\cdot||$ represents the $\mathcal{L}_2$-norm. In this case, we have

$$\begin{aligned} \mathcal{E}\{(u - \mathcal{E}\{u\})^2\} &= \underline{w}^T \mathcal{E}\{(\underline{z} - \mathcal{E}\{\underline{z}\})(\underline{z} - \mathcal{E}\{\underline{z}\})^T\}\underline{w} \\ &= \underline{w}^T I \underline{w} = ||\underline{w}||^2 = 1 \end{aligned} \quad (6)$$

Consequently, the denominator in the definition of skewness remains 1, as long as we prewhiten the input data and keep the projection vector $\underline{w}$ normalised to unit variance. As a result, we need to optimise the following cost function:

$$J(\underline{w}) = |G(\underline{w})| = \left|\mathcal{E}\{(u - \mathcal{E}\{u\})^3\}\right| \quad (7)$$

Expanding $J(\underline{w})$, we get the following simplified expression

$$J(\underline{w}) = |\mathcal{E}\{u^3\} - 3\mathcal{E}\{u^2\}\mathcal{E}\{u\} + 2\mathcal{E}\{u\}^3| \quad (8)$$

The optimisation problem to be solved is set as follows:

$$\max_{w} J(\underline{w}) \quad (9)$$

$$\text{subject to } ||\underline{w}||^2 = 1 \quad (10)$$

The first step is to estimate the derivative of $\partial J / \partial \underline{w}$.

$$\begin{aligned} \frac{\partial J}{\partial \underline{w}} &= \text{sgn}\left(G(\underline{w})\right) \\ &\cdot \quad (3\mathcal{E}\{\underline{z}(\underline{w}^T\underline{z})^2\} - 6\mathcal{E}\{\underline{z}(\underline{w}^T\underline{z})\}\mathcal{E}\{\underline{w}^T\underline{z}\} \\ &- \quad 3\mathcal{E}\{(\underline{w}^T\underline{z})^2\}\mathcal{E}\{\underline{z}\} + 2 \cdot 3\mathcal{E}\{(\underline{w}^T\underline{z})\}^2\mathcal{E}\{\underline{z}\}) \end{aligned} \quad (11)$$

$$\begin{aligned} \frac{\partial J}{\partial \underline{w}} &= \text{sgn}(G(\underline{w}))[3(\mathcal{E}\{\underline{z}(\underline{w}^T\underline{z})^2\} - \mathcal{E}\{(\underline{w}^T\underline{z})^2\}\mathcal{E}\{\underline{z}\}) \\ &- \quad 6(\mathcal{E}\{\underline{z}(\underline{w}^T\underline{z})\}\mathcal{E}\{\underline{w}^T\underline{z}\} - \mathcal{E}\{\underline{w}^T\underline{z}\}^2\mathcal{E}\{\underline{z}\})] \end{aligned} \quad (12)$$

One can perform gradient ascent optimisation to find the optima of the cost function, as shown below.

$$\underline{w}^+ \leftarrow \underline{w} + \eta \frac{\partial J(\underline{w})}{\partial \underline{w}} \quad (13)$$

$$\underline{w}^+ \leftarrow \underline{w}^+ / ||\underline{w}^+|| \quad (14)$$

The new estimate for the projection operator $\underline{w}^+$ is estimated in terms of the previous estimate $\underline{w}$ and the derivative of the cost function, weighted by the learning rate η. The second step normalises the unmixing vector to unit L_2-norm. These two steps are repeated until convergence, i.e. $|\underline{w}^T\underline{w}^+| \rightarrow 1$. However, gradient optimisation methods suffer from convergence speed problems, as their convergence is controlled by the learning rate η. In practice, a bad choice for the learning rate η can inhibit or delay the convergence of the optimisation considerably.

However, one can form a "fixed-point" rule to accelerate the convergence of the algorithm [5]. At a stable point of the algorithm, the gradient must point to the direction of $\underline{w}$. This implies that the gradient should be equal to $\underline{w}$ multiplied by some scalar constant. Only in this case adding $\underline{w}$ to the gradient is not changing its direction and we have convergence. In addition, the arbitrary scaling (or sign) is effectively removed by normalising to unit norm. Hence, the "fixed-point" can be found at those points where the new estimate is equal to the gradient.

$$\underline{w}^+ \propto \frac{\partial J(\underline{w})}{\partial \underline{w}} \quad (15)$$

To reduce the computational cost, one can remove the $\text{sgn}(\cdot)$ expression from $\partial J / \partial \underline{w}$ [5]. Effectively, the correct sign identifies signals skewed to the left or to the right. However, this information might as well be lost in the scale ambiguity of

the linear instantaneous model. As a result, there is no practical need to maintain this costly term in the update algorithm. Consequently, the following update algorithm is proposed:

$$\begin{aligned} \underline{w}^+ &\leftarrow \mathcal{E}\{\underline{z}(\underline{w}^T\underline{z})^2\} - \mathcal{E}\{(\underline{w}^T\underline{z})^2\}\mathcal{E}\{\underline{z}\} \\ &- 2\mathcal{E}\{\underline{w}^T\underline{z}\}(\mathcal{E}\{\underline{z}(\underline{w}^T\underline{z})\} - \mathcal{E}\{\underline{w}^T\underline{z}\}\mathcal{E}\{\underline{z}\}) \quad (16) \end{aligned}$$

$$\underline{w}^+ \leftarrow \underline{w}^+/||\underline{w}^+|| \quad (17)$$

The algorithm proposed above, extracts only one component, which corresponds to the local optimum that is closest to the random initial guess for $\underline{w}$. The same rule can be used to get other skewed components that exist in the linear mixture. The above rule is randomly re-initialised to trace other skewed components. However, the algorithm should not converge to the same component. As all solutions lie in a orthogonal structure, due to prewhitening, we must always search for a solution in planes that are orthogonal to the planes defined by the estimated components. In other words, the new components should always be orthogonal to the already estimated components. Hence, for the $i-th$ component, the update for $\underline{w}_i^+$ should always be orthogonal to the space spanned by the vectors $\underline{w}_1, \underline{w}_2, \ldots, \underline{w}_{i-1}$ [5].

$$\underline{w}_i^+ \leftarrow \underline{w}_i^+ - BB^T\underline{w}_i^+ \quad (18)$$

where $B = [\underline{w}_1\ \underline{w}_2 \ldots \underline{w}_{i-1}]$.

If we are not interested in preserving the mean of the original signals (which will be also affected by the scale ambiguity), we can simplify the proposed update rule and reduce the computational cost of the algorithm. We can normalise the input data to zero mean, using the following prewhitening step.

$$\underline{z} = V(\underline{x} - \mathcal{E}\{\underline{x}\}) \quad (19)$$

Consequently, $\mathcal{E}\{\underline{z}\} = 0$ and $\mathcal{E}\{\underline{w}^T\underline{z}\} = 0$ and therefore the update rule in (16) can be simplified to the following:

$$\underline{w}^+ \leftarrow \mathcal{E}\{\underline{z}(\underline{w}^T\underline{z})^2\} \quad (20)$$

This simplification decreases the computational cost of the algorithm significantly, however, we lose possible bias information of the input signals.

One benefit of the proposed algorithm is that it might be a more appropriate tool for biomedical applications, where skewness is more important to nonGaussianity for certain categories of signals. Thus, one can gain all the benefits of estimating lower-order moments. Also, skewness, along with kurtosis, can be defined as a cumulant, unlike general nonlinearities that can be used in ICA. Therefore, it should be possible to make the proposed algorithm "blind" to additive noise of known covariance [3].

III. Experiments

For this experiment, we created four artificial mixtures of two symmetrical and two skewed signals, in order to test the algorithm's performance. We used a uniformly distributed source from -1 to 1 and a Gaussian source with zero mean and unit variance as symmetrical signals. The skewed signals were two Weibull distributed signals with different parameter values, one skewed to the left and one skewed to the right. The distributions of the four input signals are shown in figure 1. We used 5000 points of these randomly generated sources, which were mixed with the following example mixing matrix A.

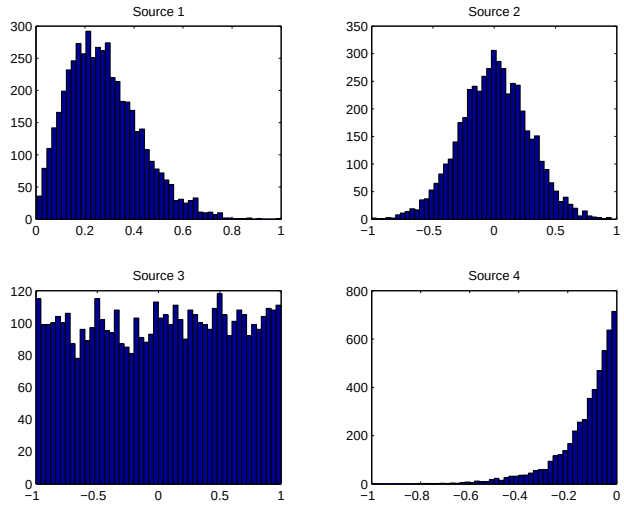

Fig. 1. Histograms of the four input sources used in the first experiment. Source 1 is skewed to the left and Source 4 is skewed to the right. Both skewed signals have non-zero mean.

$$A = \begin{bmatrix} 0.40 & 0.25 & 0.1 & 0.35 \\ 0.17 & 0.25 & 0.45 & 0.13 \\ 0.15 & 0.10 & 0.20 & 0.65 \\ 0.23 & 0.57 & 0.10 & 0.10 \end{bmatrix} \quad (21)$$

In figure 2, we can see the convergence of the update rule in (16). The algorithm converged almost after $5-10$ iterations using random initialisation. We also measured the Signal-to-Noise Ratio (SNR), comparing the original input signal s_i with the corresponding separated signal u_j, using the definition in (22). To perform accurate measurement, the scale and permutation ambiguity of the model should be taken into account and the signals should be normalised accordingly.

$$SNR_{(\text{dB})} = 10\log_{10}\frac{\mathcal{E}\{s_i^2\}}{\mathcal{E}\{(s_i - u_j)^2\}} \quad (22)$$

The algorithm managed to identify and separate both signals with promising performance. The histograms of the two identified sources are depicted in figure 3. The algorithm has isolated the two skewed signals from the mixture. The quality of separation was also promising, giving $SNR_1 = 25.4060dB$ and $SNR_2 = 40.4802dB$ for the two sources. If we are searching for more than 2 sources, then the algorithm will not be able to unmix the other signals. Instead, the separation results are Gaussian-like signals, i.e. the mixture of the two symmetrical signals.

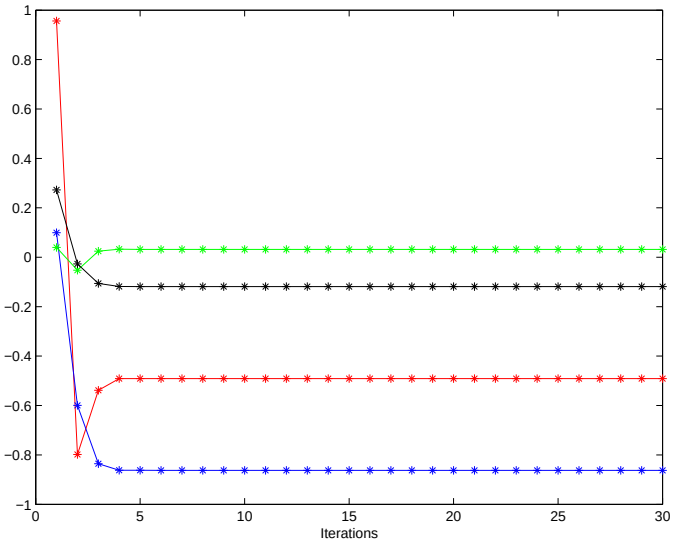

Fig. 2. Convergence of the four coefficients of the vector $\underline{w}$ using the fixed-point algorithm in (16) and random initialisation.

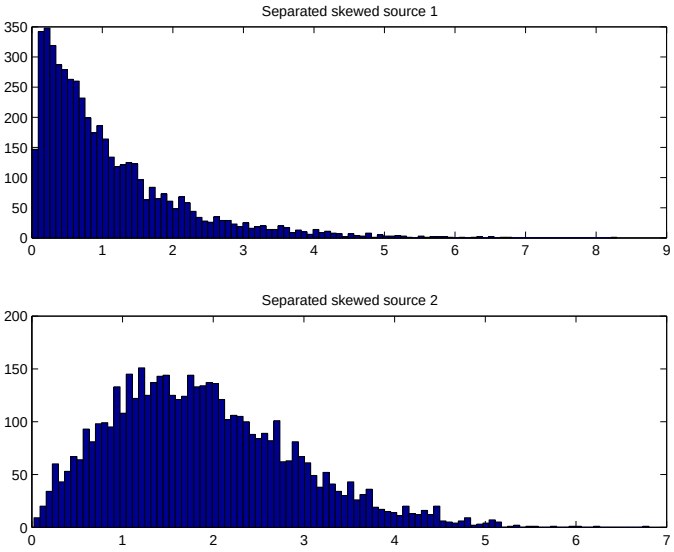

Fig. 3. Histograms of the two skewed separated sources. Comparing to the original histograms in figure 1, we can observe the scale and permutation ambiguity in the estimation of the linear generative model in (1).

IV. Application to Biomedical Signals

The application of Independent Component Analysis (ICA) on biomedical signal processing has been highlighted in the literature during the last years [4]. In medical applications, a number of sensors are used to observe the function of body organs. Usually, these sensors capture the activity of the organs by measuring the electric potential at several points of the body. The medical doctor examines these measurements over time series and can infer about the medical condition.

Let us briefly explore how the linear generative model of (1) can be applied in this case. In fact, these sensors capture the overall picture of a more complicated phenomenon that can be usually analysed into contributions from different underlying components. The contribution of each of these components to the signal, observed by each sensor, may be different in terms of amplitude and delay. However, as we are referring to electrical current measurements and noting that the sensors are usually placed relatively close, it is common to assume that the contribution of each component arrives with insignificant delay. Consequently, we can consider the mixing to be instantaneous.

Assume that the individual components are also statistically independent, ICA can be employed to separate these components. A number of approaches have employed ICA for separation for several types of biomedical signals. However, there are several components that can be identified using other statistical characteristics than statistical independence, i.e. non-Gaussianity. In this paper, we will use skewness to identify several known components in electrocardiograms (ECG) and electroencephalograms (EEG) that are not symmetric.

A. Electrocardiogram (ECG) Signals

Electrocardiogram (ECG) is a test that measures the electrical activity of the heart. It is used to measure the rate and regularity of heartbeats, as well as the size and position of the chambers, the presence of any damage to the heart and the effects of drugs or devices used to regulate the heart (e.g. a pacemaker). In literature, it has been shown that the electric potential in one part of the body surface can be obtained by adding partial contributions of the heart potentials, each one scaled by a transfer coefficient [7]. As a result, the instantaneous mixture model of (1) can be used to model the ECG monitoring system. The voltages for the 12-lead ECG can be expressed as an instantaneous mixture of the heart potentials.

In figure 4, we can see twelve signals that are given from a 12-lead ECG. The ECG can detect several parts of these signals that are associated with Ventricular Activity (VA) and also a number of more subGaussian signals that are associated with Atrial Activity (AA) [7]. A number of methods have been proposed for the identification of VA or AA signals, based on either ICA or other methods. However, one can observe that VA signals are not symmetrical, compared to AA signals or other noise signals. Therefore, one can optimise skewness, as shown earlier on, to identify and separate signals associated with VA.

For this experiment, we used a 12-lead set of signals from the PhysioNet database [11], sampled at 1 KHz (see figure 4). The signals were filtered initially using a notch filter to remove mains interference and then with a band-pass filter with cut-off frequencies of 0.5 and 60 Hz to remove baseline wandering and thermal noise [7]. We applied the algorithm in (16), using 5000 input samples of each electrode, to estimate 4 skewed components in the mixture. The separated signals are depicted in figure 5. The algorithm managed to identify three components that are associated with VA. The algorithm also identified a fourth non-symmetrical component, but can not be clearly associated with VA.

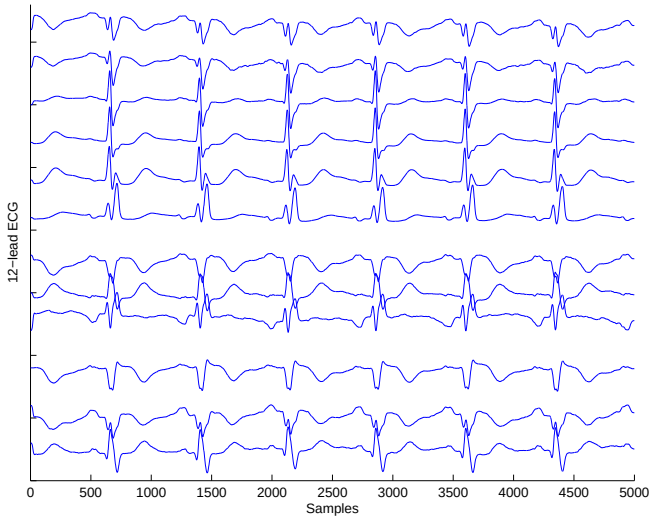

Fig. 4. A twelve-lead input ECG signal from the Physionet database [11].

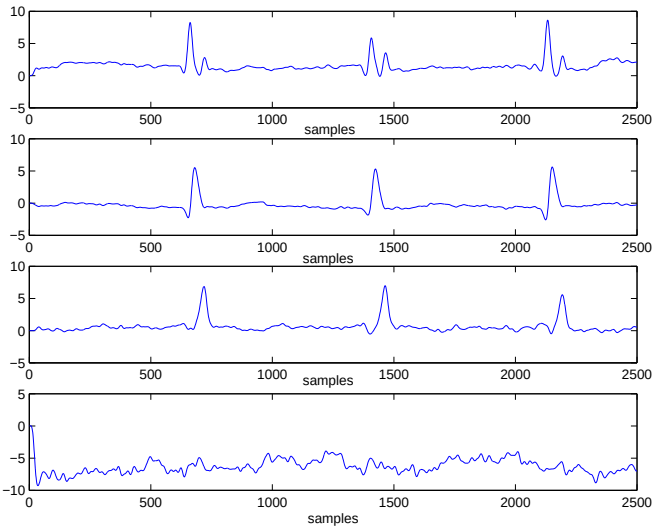

Fig. 5. Trying to separate four skewed components, the algorithm identified the first three, which can be associated with VA.

Hence, the proposed algorithm can be used as a method of isolating VA signals. It can also be used as a preprocessing step to identify the AA signal. Using the proposed algorithm, one can separate VA signals and then use second-order methods to isolate the AA signal, as proposed by Rieta et al [7].

B. Electroencephalogram (EEG) Signals

The electroencephalogram (EEG) is a medical test used to measure/monitor the electrical activity of the brain via electrodes applied to the scalp. This safe and painless procedure can help diagnose a number of conditions, including epilepsy, sleep disorders and brain tumours. This method can provide direct information about the neural dynamics on a millisecond scale [6], [10]. The EEG are very sensitive, low-amplitude signals (μV) and are usually contaminated by noise. As a result, a number of methods were proposed in literature to

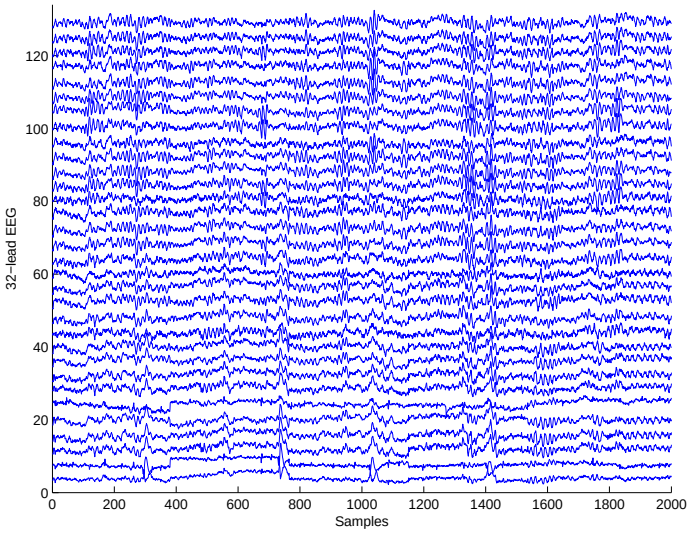

Fig. 6. A thirty two-lead input EEG signal from the EEGLAB test data [2].

isolate meaningful signals from artifacts and noise signals in an EEG recording.

The application of ICA to the study of EEG signals is valid only if some conditions are at least approximately satisfied [4]. First of all, we have to assume the existence of independent components (source signals), i.e. hidden neural centres that can serve as sources. This assumption can be justified statistically, however, in general there is no physical evidence confirming its validity in full. Secondly, the assumption of linear instantaneous mixing should also hold. As most EEG signals lie below 1 KHz, the propagation of the signals can be considered immediate and hence, there is no need to introduce time delays in the model. Finally, the mixing has to be stationary, i.e. the mixing matrix A should not be changing. Although, the underlying source signals are documented to be non-stationary, the mixing matrix should remain stationary, as long as the electrodes in the cask are not moving. In practice, we are bound to have slight movement during the experiment, however, we will assume that the change in the mixing matrix is insignificant.

Using ICA, one can identify hidden underlying sources in EEG [6], [10]. The estimated independent components can be localised and projected on the sensors' space and highlight activity in certain parts of the brain. However, EEG signals are contaminated by certain artifacts, i.e. signals that are not generated by brain activity, but by some external disturbances (e.g. muscle activity, heart beat and eye blinking). A number of methods have been used to remove the possible artifacts present in EEG [6], [8], [10]. The eye blinking signal is an artifact, reported to be usually skewed to one side [8]. Consequently, one might use the update algorithm described earlier to identify and remove the eye blinking signal from EEG recordings.

We performed some preliminary experiments to test the previous argument. We used the test EEG data set that is provided with the EEG software package EEGLAB [2]. This

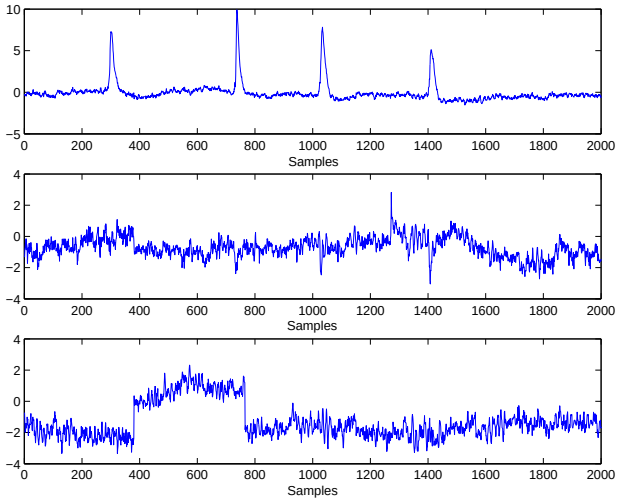

Fig. 7. Identifying three skewed components from the 32-lead EEG. The first choice for the algorithm was an artifact associated with eye-blinking.

consists of a 32-lead EEG recording, sampled at 500 Hz (see figure 6). The whole available data set ($\sim$ 61 secs) was used in the update rule of (16). The algorithm was requested to identify three skewed components. In figure 7, the first identified component is the one usually associated with eye blinking [6], [10]. This implies that skewness is a proper criterion to identify the eye blinking artifacts. This preliminary effort demonstrates a valid application of the proposed method in the area of EEG analysis.

The algorithm's convergence in the case of real biomedical data was comparable to the performance with artificial data. In either case, the algorithm required not more than 30 iterations for convergence.

V. Conclusion

In this paper, we have proposed an algorithm to separate skewed sources in an instantaneous mixture. After a prewhitening step, the algorithm optimises a third-order moment, i.e. skewness, using a "fixed-point" optimisation scheme to perform identification and separation of the skewed sources. The algorithm's convergence and performance in an artificial experiment seems very promising. In addition, we explored the use of the proposed technique in several biomedical applications. More specifically, we have given some preliminary examples of Ventricular Activity signal separation in ECG signals and eye blinking removal from EEG signals.

In the future, the authors would like to quantify the convergence of the proposed technique and investigate more solidly its performance in the case of ECG and EEG signals.

Acknowledgment

The authors would like to thank Dr. Stella S. Daskalopoulou for the insightful advice and guidance on the medical part of the paper.

References

[1] A. Cichocki, S.I. Amari, *Adaptive Blind Signal and Image Processing. Learning algorithms and applications*, John Wiley & Sons, 2002.

[2] A. Delorme, S. Makeig, *EEGLAB: an open source toolbox for analysis of single-trial EEG dynamics*, Journal of Neuroscience Methods, 134:9-21, (2004).

[3] A. Hyvärinen. Independent component analysis in the presence of Gaussian noise by maximizing joint likelihood. *Neurocomputing*, 22:49–67, 1998.

[4] A. Hyvärinen, J. Karhunen, E. Oja, *Independent Component Analysis*, John Wiley & Sons, 2001.

[5] A. Hyvärinen, E. Oja, *A Fast Fixed-Point Algorithm for Independent Component Analysis*, Neural Computation, 9(7):1483-1492, 1997.

[6] T.-P. Jung, C. Humphries, T.-W. Lee, S. Makeig, M. J. McKeown, V. Iragui, T. Sejnowski, *Extended ICA removes artifacts from electroencephalographic recordings*, Psychophysiology, 37, 2000, 163178. Cambridge University Press.

[7] J.J. Rieta, F. Castells, C. Sanchez, J. Igual, *ICA applied to atrial fibrillation analysis*, 4-th Int. Symposium on Independent Component Analysis and Blind Signal Separation, April 2003, Nara, Japan.

[8] S. Sanei, L. Shoker, *Artefact Removal from EEGs Using a Hybrid BSS-SVM Algorithm*, Invited Talk, IEE Biomedical Signal Processing Workshop, London, Dec. 1, 2004.

[9] P.F. Stetson, *Independent Component Analysis of Pulse Oximetry Signals based on Derivative Skew*, 5-th Int. Symposium on Independent Component Analysis and Blind Signal Separation, September 2004, Granada, Spain.

[10] R. Vigário, V. Jousmäki, M. Hamäläinen, R. Hari, E. Oja, *Independent Component Analysis for identification of artifacts in Magnetoencephalographic recordings*, Advances in Neural Information Processing System 10 (Proc. NIPS 97), pp. 229 - 235. MIT Press.

[11] PhysioBank website, *The PTB Diagnostic ECG Database*, http://www.physionet.org/physiobank/database/ptbdb/