# ICA AND BSS APPLIED TO BIOMEDICAL SIGNALS: AN OVERVIEW OF THEIR USE IN BIOMEDICINE

C.J. James* and C.W. Hesse

Signal Processing and Control Group, ISVR, University of Southampton, Southampton, UK

* C.James@soton.ac.uk

**Abstract: Independent Component Analysis (ICA) and Blind Source Separation (BSS) are increasing in popularity in the field of biomedical signal processing. These are generally used to separate digital recordings of multi-channel biosignals into their constituent components. The wider use of ICA/BSS in the field of biomedical signal processing has been facilitated through a number freely available toolboxes that implement popular flavours of ICA/BSS, such as the FastICA algorithm. However, the application of these techniques in a blind fashion could lead to many erroneous conclusions drawn from data-sets for which 'standard' ICA/BSS techniques are not immediately applicable. In this work we set out to summarise some of the different techniques of ICA/BSS in the literature, in view of the possible issues that may arise when these are applied blindly to biomedical signal measurements.**

## Introduction

The ultimate aim in processing electromagnetic (EM) biosignals is that of extracting information underlying those measurements made over time. Such a set of multi-channel measurements are usually recorded using a known spatial distribution of the recording sensors with respect to the human body (such as across the scalp, chest, limbs, etc.), giving rise to a set of temporally and spatially correlated measurements. Signal(s) of interest are seldom recorded in isolation and are generally mixed with other ongoing 'background' activity and sensor noise, and are almost certainly contaminated by artifacts of either physiological or environmental origins; furthermore, the signal-to-noise (SNR) ratio of the desired signal is generally quite poor. When viewing measured biomedical signals, a clinician, through training and experience generally looks for distinct patterns of activity with particular spatial distributions – exactly what the clinician is looking for depends on the application domain. One viewpoint is that the recorded data contains measurements of a finite set of separate, overlapping (in both space and time) activities which can be both physiological and/or artifactual in nature. A clinician would then attempt to 'unmix' these sources visually using human reasoning, in order to be able to arrive at a diagnosis or prognosis. It would be of great benefit to clinicians if it were possible to automate the analysis of biomedical signals to do the following:

1. unmix and isolate sets of EM biosignal measurements into their constituent components or sources;
2. provide information as to the number of distinct sources underlying the measurements;
3. provide the spatial distribution of each source along with the time-series of the source itself;
4. track changes in the number, spatial distribution and morphology of the sources over time.

Within this context, the technique of Independent Component Analysis (ICA) – as a subset of the technique of Blind Source Separation (BSS) – provides a tool for giving a solution to the requirements listed above. ICA essentially extracts a set of underlying sources or components from a set of random variables, measurements or signals. The technique typically uses a large set of observed multivariate data to define a generative model for the observed data. The components are assumed to be mixed, either linearly or nonlinearly, and the components themselves – along with the mixing system – are assumed to be unknown.

For ICA the fundamental assumption is that the sources are *mutually independent*. ICA de-mixes or extracts these sources by exploiting this independence of the sources underlying the measured data and is a more powerful technique than classical methods such as Principal Component Analysis (PCA).

## Independent Component Analysis

The basic BSS problem that ICA attempts to solve assumes a set of $m$ measured data points at time instant $t$, $\mathbf{x}(t) = [x_1(t), x_2(t), \ldots, x_m(t)]^T$ to be a combination of $n$ unknown underlying sources $\mathbf{s}(t) = [s_1(t), s_2(t), \ldots, s_n(t)]^T$. The mixing of the sources is generally assumed to be linear, and the mixing matrix describing the linear combination of the $\mathbf{s}(t)$ is given by the full rank $m \times n$ matrix $\mathbf{A}$ such that

$$\mathbf{x}(t) = \mathbf{A}\mathbf{s}(t), \tag{1}$$

it is also generally assumed that the number of underlying sources is less than or equal to the number of measurement channels ( $n \leq m$ ).


ISSN: 1727-1983 

The task of the ICA algorithms is to recover the original sources $\mathbf{s}(t)$ from just the observations $\mathbf{x}(t)$ and this generally translates to that of finding a separating or de-mixing matrix **W** such that

$$\hat{\mathbf{s}}(t) = \mathbf{W}\mathbf{x}(t), \quad (2)$$

given the set of observed values in $\mathbf{x}(t)$ and where $\hat{\mathbf{s}}(t)$ are the resulting *estimate* of the underlying sources.

***The Basic Assumptions:*** In reality the basic mixing model assumed in eqn (1) is simplistic and used for ease of implementation. The more general mixing model, which makes no assumptions on the linearity of the mixing and allows additive noise, may be a more realistic model for a system *in general*, so

$$\mathbf{x}(t) = \mathbf{f}\{\mathbf{s}(t)\} + \mathbf{n}(t) \quad (3)$$

where **f** can be any unknown function and $\mathbf{n}(t)$ is additive sensor noise corrupting the measurements $\mathbf{x}(t)$ (generally i.i.d. spatially and temporally white noise).

In (3), the BSS problem is now that of obtaining a demixing matrix by inverting **f** whilst having no information on the properties of either **s** or **n** (or **f**); it can be appreciated that without making any assumptions about the nature of the data, noise or mixing process, the BSS problem will remain intractable. This is why basic assumptions are made when formulating ICA algorithms in order to make the problem more tractable. Within a biomedical signal analysis context, most of these basic assumptions still make the technique attractive and viable.

The following are some of the more common assumptions made when applying ICA algorithms:

***i. Linear mixing:*** this reduces eqn (3) to

$$\mathbf{x}(t) = \mathbf{A}\mathbf{s}(t) + \mathbf{n}(t) \quad (4)$$

where **A** is the linear mixing matrix. In a biomedical signals context, linear mixing assumes mixing of the sources using simple linear superposition of the attenuated sources at the measurement channel – for the most part a reasonable assumption to make. For the most part assuming instantaneous mixing is perfectly legitimate as this assumes that transmission through the mixing medium is instantaneous – this holds for such applications as fMRI and EM brain signals. Quantities such as sound signals measured through microphones then do become an issue, however, as this assumes convolutive mixing.

***ii. Noiseless mixing:*** The majority of the traditional ICA models assume that the observations $\mathbf{x}(t)$ are noiseless (or that the noise term $\mathbf{n}(t)$ is negligible), i.e. eqn (4) reduces to eqn (1). Whilst this is probably less realistic in practical terms, it allows ICA algorithms to separate sources of interest even if the separate sources themselves remain contaminated by the measurement noise.

***iii. Square mixing:*** So far it has been assumed that the mixing matrix **A** may be non-square ($m \times n$), and in the case of physiologic signal analysis it is likely that the number of underlying sources $n$ is *less* than the number of measurement channels $m$ in use. However most classic ICA algorithms assume a *square mixing* matrix, i.e. $m=n$, this makes the BSS problem more tractable. Within a biomedical signal analysis perspective the square-mixing assumption is sometimes less than desirable, particularly is situations where high density measurements are made over relatively short periods of time such as in most magnetoencephalogram (MEG) and functional magnetic resonance imaging (fMRI) recordings. The probability of there being as many sources as measurement channels in these situations is less likely. For this reason most researchers apply data-reduction techniques to the data prior to ICA [1] although this may be ill-advised in certain situations.

***iv. Stationary mixing:*** It is assumed that the statistics of the mixing matrix **A** do not change with time – i.e. the assumption of stationarity of the mixing matrix. In biomedical signal analysis terms this means that the physics of the mixing of the sources as measured by the sensors is not changing – this may not be the case in situations where, for example, electrocardiogram (ECG) is recorded on chest electrodes and the electrodes move over time due to breathing, etc.. However, in EM brain signal recordings the assumption of a stationary mixing matrix can be interpreted as the fixed biophysical structure of the brain itself whilst the sources distributed within this structure change their intensity over time, which is perfectly plausible.

***v. Statistical Independence of the sources:*** The most important assumption made in applying ICA is that the sources are *mutually independent*. Statistical independence is a stronger assumption than uncorrelatedness, and while statistically independent sources are necessarily uncorrelated, the converse does not follow. This means, for example, that independent variables are uncorrelated and have no higher order correlations. In the case of time-series data it is assumed that each source is generated by a random process which is independent of the random processes generating other sources. Thus, the BSS problem can be made further tractable by allowing the introduction of a set of algorithms that can take advantage of this independence of the sources. The assumption of independence of the sources can be quite obvious in some situations as, for example, when used in artifact rejection separating brain signals from, say, 50Hz line noise or ocular artifact [2]; or when separating fetal


ISSN: 1727-1983 

electrocardiogram (FECG) from maternal ECG (MECG) through trans-abdominal recordings [3].

The utility of ICA in light of the above assumptions, when placed in a physiologic analysis setting, should be assessed on an individual biomedical signals application basis. However, for the most part, ICA can still be a very useful technique even with these simplifying assumptions being made.

### A. ICA through Higher Order Statistics

Some of the most commonly cited ICA algorithms perform BSS of statistically independent sources based on techniques involving higher-order statistics (HOS). Several such implementations can be found in the literature; [4,5,6,7,8]. For these methods the measurements $\mathbf{x}(t)$ of eqn (1) are assumed to be observations of random variables, where temporal ordering is irrelevant and which are generated as a linear mixture of statistically independent sources.

When seeking statistical independence in sources, it is not enough to obtain uncorrelatedness between the sources, which is what PCA does – although decorrelating the measured data is generally a useful first step – statistical independence is based on HOS. It turns out that it is possible to obtain an estimate $\hat{\mathbf{s}}(t)$ of the sources $\mathbf{s}(t)$ iff the sources $\mathbf{s}(t)$ are non-Guassian. In practice it is enough to try and make the estimates $\hat{\mathbf{s}}(t)$ as non-Gaussian as possible as, according to the central limit theorem, sums of non-Gaussian random variables are closer to Gaussian than the originals. In this way looking for independent sources is equivalent to looking for non-Gaussian sources. As a consequence, this highlights a potential limitation of the method when used for biomedical signal processing, as ICA using this technique can only resolve independent sources which have non-Gaussian distributions (or at most only one source having a Gaussian distribution).

Three of the most popular and widely referenced techniques for implementing ICA using a HOS approach are:

*(a) Non-Gaussianity through Kurtosis – FastICA:* FastICA is one of the more referenced ICA techniques in the literature [7] and it is distributed as a freely downloadable set of Matlab® functions from the internet [9]. This algorithm attempts to separate underlying sources from the given measurement set based on their 'non-Gaussianity'. The simple premise behind FastICA is that the fast fixed-point iterative algorithm undertakes to find projections that maximize the non-Gaussianity of components by their Kurtosis (the 4th order cumulant given to a random variable).

*(b) Non-Gaussianity through Negentropy – Infomax:* Another algorithm that implements ICA through attempting to discover non-Gaussianity of the sources is the Bell-Sejnowski algorithm [6,10], where non-Gaussianity is measured using negentropy, which is based on the information-theoretic quantity of differential entropy. For random variables with equal variance but different distributions, Gaussian random variables have the largest entropy, i.e. contain the least information. Thus, negentropy (or differential entropy) is defined as the difference between the entropy of a Gaussian random variable with the same variance as the observed random variable, and the entropy of the observed variable. Negentropy is zero when the observed random variable is also Gaussian, and positive when the observed variable is non-Gaussian.

*(c) Joint Approximate Diagonalisation of Eigenmatrices – JADE:* This approach is known as ICA by tensorial methods using higher-order cumulant tensors [11]. The covariance matrix is the second-order cumulant tensor, and the fourth-order tensor is defined by the fourth order cumulants. By performing an eigenvalue decomposition of the covariance matrix of the data, $\mathbf{C}_\mathbf{x}$, the data are transformed such that the second-order correlations are zero. Similarly, as a generalisation of this principle, fourth-order cumulant tensors can be used to make the fourth-order cumulants zero or as close to zero as possible. As with HOS methods described previously, reducing the fourth-order cumulants to zero in this way implies statistical independence of the sources and JADE is the algorithm that implements this. As the name implies, JADE involves the joint diagonalisation of a number of matrices (i.e. attempts to make all off diagonals zero or close to zero as possible) [12,13].

### B. ICA through Time Structure Analysis

A completely different approach to ICA is given by considering the time structure of the sources. The assumption of independence of the sources has a very important and useful consequence: the source waveforms have no spatial temporal or spatial time-frequency correlations. The basic approach here is to capture the dependency structure of the observed signals using a set of square matrices (in the form of a *stack* of matrices), and then find the de-mixing matrix which is the joint diagonaliser of the stack, i.e.

$$\mathbf{C}_\mathbf{x}^k = \mathbf{A}\mathbf{C}_\mathbf{s}^k\mathbf{A}^\mathrm{T}, \tag{6}$$

where $\mathbf{C}_\mathbf{x}^k$ is the $k$-th covariance matrix of the data $\mathbf{x}(t)$, $\mathbf{C}_\mathbf{s}^k$ the corresponding covariance matrix of the sources $\mathbf{s}(t)$, and $\mathbf{A}$ is the mixing matrix. Conversely, the source covariances are obtained from the data covariance through the inversion

$$\mathbf{C}_\mathbf{s}^k = \mathbf{W}\mathbf{C}_\mathbf{x}^k\mathbf{W}^\mathrm{T}, \tag{7}$$


ISSN: 1727-1983 

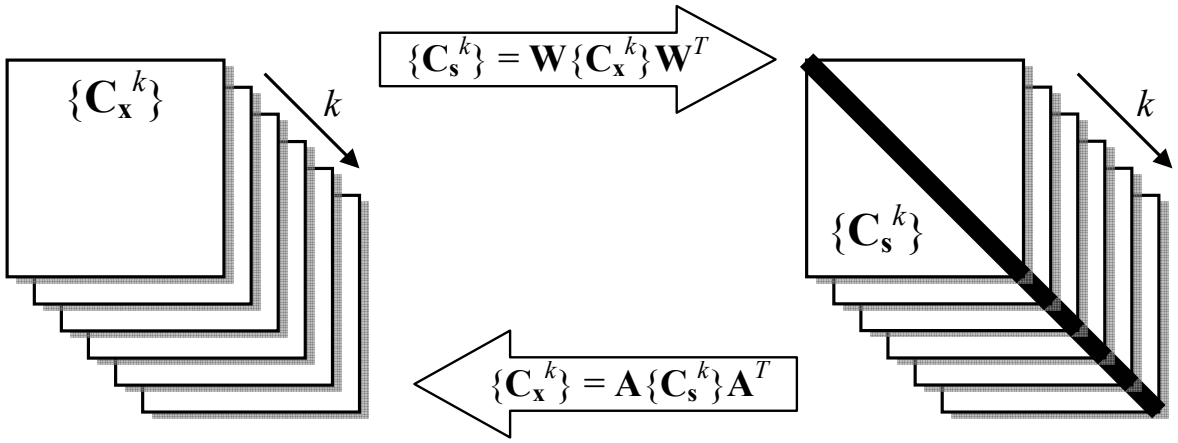


Figure 1: The relationship between the two covariance matrix stacks of $\mathbf{C}_\mathbf{x}^k$ and $\mathbf{C}_\mathbf{s}^k$ in time structure based ICA. The mixing matrix **A** can be seen to link the covariance stack of the sources to the covariance stack of the measurements and vice-versa with the unmixing matrix **W**.

where **W** is the unmixing matrix. These two equations hold in general, regardless of the nature of the matrices in the stack. The relationship between the two covariance matrix stacks of $\mathbf{C}_\mathbf{x}^k$ and $\mathbf{C}_\mathbf{s}^k$ are shown pictorially in Figure 1, where the mixing matrix **A** links the covariance stack of the sources to the covariance stack of the measurements and vice-versa with the unmixing matrix **W**.

The index $k$ is an index into the matrix stack and will have different interpretations depending on what quantities are being measured. For example, when the temporal dependancy is captured through temporal corellation measured at different lags, $k$ is an index into the cross-covariances at each lag, starting from $k$=0,1,2, etc. … until a maximum number of lags is reached. So, for a maximum of $L$ lags, there will be $L+1$ matrices in the stack ($k$=0,1,2,..., $L$).

There are two ways of estimating the mixing matrix **A** with reference to a stack of correlation matrices. The most common approach is to estimate the de-mixing matrix **W** first – this is known as an *inverse* estimation method. Since $\mathbf{C_s}$ is supposed to be diagonal, we can optimise the coefficients of **W** in such a way as to make the matrix given by $\mathbf{W}\mathbf{C}_\mathbf{x}^k\mathbf{W}^\mathrm{T}$ as diagonal as possible. The diagonality of a matrix can be measured, for example, by the sum of the squared off-diagonal elements. Previous methods employed orthogonal constraints [12,13,14], however more advanced methods allow non-orthogonal diagonalisation [15] but require a square mixing matrix. Conversely, *forward* estimation methods (e.g. [16]) have the advantage that they allow non-orthogonal and non-square mixing. However, these methods are not quite as efficient as some of the inverse methods, and still require some estimate of the number of sources.

**ICA applied to Neurophysiology**

The rationale behind this paper is that of illustrating the variety of specialisations that ICA allows, hopefully highlighting the many more potential uses of ICA in biomedicine than is currently presented in the literature. For the most part the use of ICA as a 'black-box' method may result in situations, such as the violation of some assumption, which implies that ICA is inferior where, in fact, it could have been better had it been used appropriately.

In our implementations of ICA for *EM brain signal analysis* we make assumptions that are in keeping with the general assumptions governing the application of ICA. In particular we assume that:

1. The measured neurophysiological signals are a linear summation of the electrical/magnetic activity from various brain regions.
2. The EM field distribution is spatially fixed and only the electrical 'strength' is changing within these regions.
3. Any activity of interest is independent of the ongoing background EM brain activity. This certainly holds true for most artifacts and to activity such as epileptic seizure activity (at least early on in the evolution of a seizure).

Whilst ICA is not necessarily suited for use in all problems in this domain, one of the biggest reasons for using ICA in EM brain signal analysis is the fact that multi-source activity can be naturally separated into *neurophysiologically meaningful* components. Standard signal processing techniques such as matched and/or adaptive filters *can* be used to detect and extract activity of interest, but these generally require much more detailed *a priori* knowledge about the characteristics of


ISSN: 1727-1983 

each of the signals in question. Furthermore, such techniques are never as discriminative as ICA can be, because there are usually residuals in performing unmixing in this way. ICA also unmixes signals by making very basic assumptions about the data (that of independence being foremost) and it makes little difference if the signals are artifactual in origin or brain-signals, for example, for the technique to work – standard techniques are usually not as flexible. Further information on our specific applications of ICA to EM brain signal analysis can be found in [17].

## Conclusions

This paper describes the technique of ICA as a method for performing BSS in the context of biomedical signal processing. The generic technique of ICA is introduced along with the fundamental assumptions that are generally made in order to make the problem more tractable. In essence, ICA techniques make assumptions based on the mixing of the independent sources and based on the statistical independence of those same sources. The mixing assumptions such as those of linearity, stationarity and square mixing are made in order to allow specific embodiments of ICA to be easily formulated and may be relaxed at will – depending on the algorithm in use. The same holds, for example, for the assumption of noiseless mixing. It has been shown that although these assumptions might make ICA seem restrictive in its potential applications, it has found many applications in the biomedical signal processing field as shown in the literature.

Of the many possible algorithms devised towards solving the BSS problem, ICA is popularly solved through the use of HOS techniques – basically through separating statistically independent sources based on their non-Gaussianity. We show that another, more appealing (with biomedical signal analysis in mind) viewpoint is that of using spatio-temporal and spatial-time frequency based ICA techniques. The main difference between the two being that in the latter technique the information inherent in the time-sequence of the measured data points is made use of – whereas in the former it is not. It can be seen that in the biomedical signal processing field where the analysis of information is generally based on the frequency content of recordings and on waveform morphology, such ICA techniques prove invaluable.

## Acknowledgements

This work is funded by EPSRC Grant #GR/S13132/01.

## References

[1] HYVÄRINEN, A., KARHUNEN, J. AND OJA, E. (2001): 'Independent Component Analysis', (John Wiley and Sons, New York), pp. 125–144

[2] JAMES, C.J. AND GIBSON, O.J. (2003): 'Temporally Constrained ICA: an Application to Artifact Rejection in Electromagnetic Brain Signal Analysis', *IEEE Trans. Biomed. Eng.*, **50**, 9, pp. 1108–1116

[3] DE LATHAUWER, L., DE MOOR, B. AND VANDEWALLE, J. (2000): 'Fetal electrocardiogram extraction by blind source subspace separation', *IEEE Trans. Biomed. Eng.,* **47**, 5, pp. 567– 572

[4] MAKEIG, S., JUNG, T.P., BELL. A.J., GHAHREMANI. A AND SEJNOWSKI, T.J. (1997): 'Blind separation of auditory event-related brain responses into independent components', *Proc. Natl. Acad. Sci., USA*, **94**, pp. 10979–10984

[5] COMON, P. (1994): 'Independent Component Analysis, a new concept?', *Signal Processing*, **36**, pp. 287–314

[6] BELL, A.J. AND SEJNOWSKI. T.J. (1995): 'An Information-Maximization Approach to Blind Separation and Blind Deconvolution', *Neural Computation*, **7**, pp. 1129–1159

[7] HYVÄRINEN, A. AND OJA, E. (1997): 'A fast fixed-point algorithm for independent component analysis', *Neural Computation*, **9**, pp. 483–1492

[8] JAMES, C.J. AND LOWE, D. (2001): 'ICA in Electromagnetic Brain Signal Analysis', Proc. 4th Intl. Conf. Neural Networks and Expert Systems in Medicine & Healthcare – NNESMED 2001*, Greece*, pp. 197–202

[9] FastICA MATLAB Package [Online] Available: http://www.cis.hut.fi/projects/ica/fastica

[10] EEGLAB MATLAB Package [Online] Available: http://www.sccn.ucsd.edu/~scott/ica.html

[11] HYVÄRINEN, A., KARHUNEN, J. AND OJA, E. (2001): 'Independent Component Analysis', (John Wiley and Sons, New York), pp. 229–237

[12] CARDOSO, J.F. AND SOULOUMIAC, A. (1993): 'Blind beamforming for non Gaussian signals', *IEE Proceedings-F*, **140**, 6, pp. 362–370

[13] CARDOSO, J.F. AND SOULOUMIAC, A. (1996): 'Jacobi angles for simultaneous diagonalization', *SIAM J. Matrix Anal. Applicat.*, **17**, 1, pp. 161–164

[14] PHAM, D.H. (2001): 'Joint approximate diagonalization of positive definite matrices', *SIAM J. Matrix Anal. Applicat.*, **22**, 4, pp. 1136–1152

[15] ZIEHE, A., LASKOV, P., MÜLLER, K.R. AND NOLTE, G. (2003): 'A linear least-squares algorithm for joint diagonalization', *Proc. Int. Conf. on Independent Component Analysis and Blind Signal Separation (ICA2003)*, Nara, Japan, pp.469–474

[16] YEREDOR, A. (2002): 'Non-orthogonal joint diagonalization in the least-squares sense with application in blind source separation', *IEEE Trans. Signal Processing*, **50**, 7, pp. 1545–1553

[17] JAMES, C.J. AND HESSE, C.W. (2005): 'Independent Component Analysis for Biomedical Signals', *Physiol. Meas.,* **26**, pp. 15–39


ISSN: 1727-1983 