Consistency of the blind source separation computed with five common algorithms for magnetoencephalogram background activity

Consistency of the Blind Source Separation Computed With Five Common Algorithms for Magnetoencephalogram Background Activity

Javier Escudero[1,*], Roberto Hornero[2], Daniel Abásolo[2]

Abstract

Blind source separation (BSS) is widely used to analyse brain recordings like the magnetoencephalogram (MEG). However, few studies have compared different BSS decompositions of real brain data. Those comparisons were usually limited to specific applications. Therefore, we aimed at studying the consistency (i.e., similarity) of the decompositions estimated for real MEGs from 26 subjects using five widely used BSS algorithms (AMUSE, SOBI, JADE, extended-Infomax and FastICA) for five epoch lengths (10 s, 20 s, 40 s, 60 s and 90 s). A statistical criterion based on Factor Analysis was applied to calculate the number of components into which each epoch would be decomposed. Then, the BSS techniques were applied. The results indicate that the pair of algorithms 'AMUSE–SOBI', followed by 'JADE–FastICA', provided the most similar separations. On the other hand, the most dissimi-

*Corresponding author

Email address: javier.escudero@ieee.org javier.escudero@plymouth.ac.uk (Javier Escudero)

[1]J. Escudero was with the Biomedical Engineering Group, University of Valladolid. Now, he is with the Research Group of Signal Processing and Multimedia Communications, School of Computing and Mathematics, PL4 8AA, Plymouth, University of Plymouth.

[2]Biomedical Engineering Group, E.T.S.I. Telecomunicación, University of Valladolid, 47011, Valladolid, Spain.

lar outcomes were computed with 'AMUSE–JADE' and 'SOBI–JADE'. The BSS decompositions were more similar for longer epochs. Furthermore, additional analyses of synthetic signals supported the results of the real MEGs. Thus, when selecting BSS algorithms to explore brain signals, the techniques offering the most different decompositions, such as AMUSE and JADE, may be preferred to obtain complementary, or at least different, perspectives of the underlying components.

Keywords: Algorithm comparison, Blind Source Separation (BSS), Consistency, Independent Component Analysis (ICA), Magnetoencephalogram (MEG)

---

1. Introduction

The electroencephalogram (EEG) and the magnetoencephalogram (MEG) are the only techniques that measure the synchronous oscillations of the cortex directly and non-invasively. Whereas the former records the electrical brain activity, the latter reflects the corresponding magnetic fields [1]. These signals have slightly different characteristics. For instance, MEG is only affected by current flows oriented parallel to the scalp and it is less distorted than the EEG by extra-cerebral tissues [1]. Despite these subtle differences, similar problems are faced when analysing both recordings. Firstly, the signals acquired at a particular sensor are a weighted linear mixture of the underlying brain activity [2]. Therefore, the isolation and analysis of the electromagnetic activity generated by a specific source of interest is a complex task [2]. Moreover, the brain activity is usually recorded together with undesired signals (i.e., artefacts) of physiological or environmental origin [2, 3].

Blind Source Separation (BSS) is useful to overcome some difficulties encountered in EEG and MEG analysis [2, 3]. The BSS estimates the constituent sources (or components) of the observations assuming a linear mixture model [3]. Although the components and the mixing system are unknown, they can be estimated thanks to a minimal set of assumptions that includes the statistical independence of the sources [2–5].

BSS has been widely applied to EEG and MEG data [2, 3, 5]. For instance, diverse methodologies have been used to detect and remove the artefacts [5–9]. BSS is also helpful to isolate brain activity related to specific brain functions [3, 4, 10] or to improve the discrimination of demented patients against controls [11–13].

There is a wide variety of BSS techniques available and not all algorithms are based on the same principles. For a review see, for example, [3–5]. Theoretical relationships exist among some of the metrics used in the algorithms. However, it may be difficult to select a priori the most appropriate algorithm for a particular application [6, 8]. These methods are data-driven and, by their own nature, exploratory [5].

In order to try to clarify the relationships between BSS techniques, a few studies have compared some algorithms (see [10] and references therein). However, most of these analyses were based on synthetic (i.e., artificial) signals. For instance, three Higher-Order Statistics (HOS) algorithms were compared in [14]. However, basic hypotheses in HOS-BSS were violated in the experimental design: some synthetic sources were sub-Gaussian and, in some cases, moving sources were simulated [14]. This can limit the reliability of the results. Moreover, the analysis focused on acoustic signals and

the extension of those results to brain recordings is not straightforward [14]. Computational and statistical comparisons among HOS methods were also performed with super-Gaussian synthetic signals [4]. The main conclusions supported the robustness of HOS techniques under slight violations of the assumptions. Additional analysis suggested that different techniques may reveal different components when applied to real signals [4].

Diverse studies have compared BSS algorithms in artefact removal from EEGs [8, 15–17]. The independence of the extracted components was checked in the removal of ocular artefacts [18, 19]. However, the most commonly used algorithms were left out of this analysis and the evaluation was done in terms of mutual information [18, 19]. This might bias the analysis in favour of those algorithms directly based on this metric. Moreover, the significance of the differences among algorithms was not tested [19]. Other analyses have evaluated the performance of BSS algorithms regarding the quality of their artefact removal [7]. Recently, an extensive study focused on EEG data has been published [10]. Nevertheless, it was entirely based on synthetic data [10]. The outputs of three common BSS algorithms have also been compared against a new BSS approach based on the short-time Fourier transform [20]. This study suggests that, in the case of spontaneous activity, HOS methods tend to focus on the extraction of artefacts whereas a Second-Order Statistics (SOS) approach failed since it tended to extract components with very similar spectra [20]. However, this analysis was mainly carried out in the specific framework of the study of the phase differences between components with data from only one subject [20].

To sum up, most comparisons among BSS algorithms were carried out

with simulated signals only or in a very particular context, such as artefact removal [6, 8, 16, 17]. This may limit the application of the results to other settings. Moreover, a detailed study on the similarity of the decompositions for real brain recordings computed with different algorithms is lacking [2].

Thus, it is important to study the consistency (i.e., similarity) of the separations estimated from real electromagnetic recordings. This could lead to further understanding of the relationships among BSS techniques and to more informed decisions about which algorithms could offer complementary perspectives in one particular study. By offering information about which BSS methods provide more similar results, the search for appropriate techniques for the problem at hand would be facilitated. To achieve this goal, real MEG background activity will be decomposed using five widely used BSS algorithms in the analysis of EEGs and MEGs: algorithm for multiple unknown signals extraction (AMUSE), second-order blind identification (SOBI), joint approximate diagonalisation of eigenmatrices (JADE), Lee-Sejnowski's extended-Infomax algorithm and Hyvärinen-Oja's FastICA algorithm. The results obtained from the real MEG activity will be complemented by measuring the quality of the BSS in a dataset of synthetic signals.

2. Subjects and MEG Recording

Twenty-six healthy elderly subjects without past or present mental disorders participated in this study (9 men and 17 women). Their mean age was 71.77 ± 6.38 years (mean $\pm$ standard deviation, SD). These subjects are part of a larger database acquired to study the effects of Alzheimer's disease in

the MEG (see, for instance, [12, 13]). We limited the analyses to the control subjects to avoid any bias in the results due to that dementia. All subjects gave their informed consent to participate in the current research, which was approved by the local ethics committee.

The MEG recording process was carried out in a magnetically shielded room with a 148-channel whole-head magnetometer (MAGNES 2500 WH, 4D Neuroimaging) located in the MEG Centre Dr. Pérez-Modrego at the Complutense University of Madrid (Spain). During this procedure, the subjects lay on a patient bed with eyes closed in a relaxed state. They were asked to stay awake and not to move eyes and head. For each subject, five minutes of MEG recording were acquired at a sampling rate of 678.19 Hz. Then, the data were decimated to a sampling frequency of $f_s = 169.55$ Hz. Afterwards, the recordings were processed with a band-pass FIR filter with cut-off frequencies at 0.5 Hz and 60 Hz. Finally, the MEGs were divided into epochs of 10 s, 20 s, 40 s, 60 s and 90 s (1695, 3390, 6780, 10170 and 15255 samples, respectively).

3. Blind Source Separation (BSS)

3.1. Linear Mixing Model for BSS

BSS techniques attempt to represent a set of m measured time-varying signals, $\mathbf{x}(t) = [x_1(t), x_2(t), \ldots, x_m(t)]^{\mathrm{T}}$, where $^{\mathrm{T}}$ denotes transposition, as a linear mixture of l latent underlying components (or sources), $\mathbf{s}(t) = [s_1(t), s_2(t), \ldots, s_l(t)]^{\mathrm{T}}$, given by a full-rank $m \times l$ mixing matrix, $\mathbf{A}$ [3–5]. A vector $\mathbf{n}(t) = [n_1(t), n_2(t), \ldots, n_m(t)]^{\mathrm{T}}$ can also be included in the model to account for measurement noise [3, 9, 21, 22]. Hence, the BSS model can

be represented as:

$$\mathbf{x}(t) = \mathbf{A}\mathbf{s}(t) + \mathbf{n}(t). \quad (1)$$

In EEG and MEG analysis, $\mathbf{x}(t)$ denotes the recordings, whereas $\mathbf{s}(t)$ represents either neural activity or interference signals of diverse origins [3].

Since only the observations $\mathbf{x}(t)$ are available, several assumptions are needed to estimate $\mathbf{A}$ and $\mathbf{s}(t)$ [3, 4]. In addition to linearity, it is hypothesized that $m \geq l$ and that the mixture is stationary. Moreover, the components are assumed to be mutually independent or, alternatively, decorrelated at any time delay [3, 5]. All these hypotheses have been validated for brain signals [2–5].

3.2. BSS Algorithms

Five BSS algorithms commonly used in the analysis of EEGs and MEGs were compared: AMUSE, SOBI, JADE, extended-Infomax and FastICA [2–4, 6–8, 12, 13].

AMUSE [23] and SOBI [24] are time-structure based methods, also known as SOS-BSS. They assume that the sources have no spatial-temporal correlations [3]. Thus, these techniques try to diagonalize a set of cross-covariance matrices computed from $\mathbf{x}(t)$. AMUSE only considers two time delays – usually $\tau = 0$ and $\tau = 1$ sample, which corresponds to $\tau = 0.0059$ s at $f_s = 169.55$ Hz [23]. As a result, it orders the components by decreasing linear predictability, a criterion closely related to the signal spectral content [12, 13]. On the other hand, SOBI uses iterative procedures to simultaneously diagonalise multiple temporal lags [24]. Similarly to [15], SOBI was applied with 50 consecutive time lags from $\tau = 1$ sample to $\tau = 50$ samples

($\tau = 0.2949$ s at f_s). This choice was supported by the fact that this set of delays covered a wide time interval without extending beyond the support of the average autocorrelation function of the MEG recordings.

On the other hand, JADE [25], extended-Infomax [26] and FastICA [4] rely on HOS, that is, statistical parameters like negentropy or kurtosis. They look for non-Gaussian sources assuming that $\mathbf{x}(t)$ are observations of random variables where the temporal order is irrelevant [3, 4]. In this study, FastICA was applied with the non-linearity $\tanh(\cdot)$ and the deflationary approach [4]. This function was selected for being a good general-purpose function [4]. The extended version of Infomax was used in order to estimate both sub- and super-Gaussian sources [26]. This version of the algorithm has been widely applied to EEG and MEG [8, 15, 16]. The number of each type of components was automatically determined [26]. JADE has no input parameters [4, 6, 25].

All these BSS algorithms are contained in the EEGLAB [27], FastICA [28] and ICALAB toolboxes [29].

3.3. Preprocessing and Model Order Selection

The implementation of most BSS algorithms assumes a noiseless mixture where $m = l$ [3, 4]. However, EEG and MEG are affected by measurement noise whose power may not be negligible [9, 21, 22, 30, 31]. Furthermore, the number of channels in current EEG and MEG systems can be much larger than that of meaningful BSS components (i.e., $m > l$) [21, 30]. Hence, a suitable preprocessing is important to reduce the importance of the measurement noise and the dimensionality of the input signals of the BSS algorithms [3, 9, 21].

The preprocessing applied before a BSS algorithm is usually based on

Principal Component Analysis (PCA) [3]. Nevertheless, this approach has some drawbacks as it implies a certain degree of arbitrariness in the estimation of l. Moreover, it is not clear that the external noise is weak enough at all sensors [3, 9, 21]. In contrast to PCA, we apply a preprocessing based on factor analysis (FA) that can deal with different noise power at each sensor. Moreover, the model order (l) has been estimated with a statistical criterion: the Minimum Description Length (MDL) [30]. The preprocessing variables are computed for the range of possible l values and, for each of them, the value of the statistical criterion MDL is computed. Then, the optimum l is selected as the one providing the minimum MDL. A detailed description of FA and the MDL can be found in [30] or [9]. This preprocessing was evaluated in [9] using synthetic data. The results suggested that it provided more accurate estimations of l than other commonly used PCA-based approaches. Furthermore, other studies have found that FA is more parsimonious when estimating the value of l in real EEGs and MEGs than classical PCA schemes [21, 22].

3.4. Comparison of BSS Algorithms

A completely accurate quantification of the performance provided by a BSS algorithm q can only be achieved if either the original mixing matrix, $\mathbf{A}$, or set of sources, $\mathbf{s}(t)$, is known [11, 32, 33]. This is the case when analysing synthetic signals. For real EEG and MEG recordings, these data are not available. However, the consistency of various BSS algorithms can still be precisely computed [29]. In order to do so, two different BSS algorithms (algorithm q and algorithm r) must be applied to the same input data in order to estimate the corresponding mixing matrices: $\mathbf{A}^q$ and $\mathbf{A}^r$ [29]. Then,

the columns of these matrices are normalized to unit length vectors and a matrix $\mathbf{P}^{qr}$, whose size is $l \times l$, is computed as:

$$\mathbf{P}^{qr} = (\mathbf{A}^q)^{-1} \mathbf{A}^r. \tag{2}$$

If the two algorithms q and r provide exactly the same separation, $\mathbf{P}^{qr}$ will be a generalized permutation matrix [4]. Similarly, the closer $\mathbf{P}^{qr}$ is to a permutation matrix, the more consistent the separations of the algorithms q and r are [29].

In order to measure the degree to which $\mathbf{P}^{qr}$ is close to a permutation matrix, we define the metric F as $F = (F_1 + F_2)/2$, with F_1 and F_2 computed as in [11]:

$$F_1 = \frac{1}{l} \sum_{i=1}^{l} \left[\frac{1}{l-1} \left(\sum_{j=1}^{l} \frac{|p_{ij}|}{\max_k |p_{ik}|} - 1 \right) \right], \tag{3}$$

and

$$F_2 = \frac{1}{l} \sum_{j=1}^{l} \left[\frac{1}{l-1} \left(\sum_{i=1}^{l} \frac{|p_{ij}|}{\max_k |p_{kj}|} - 1 \right) \right], \tag{4}$$

where p_{ij} denotes an element of $\mathbf{P}^{qr}$ and l is the number of components.

F_1 measures the average coupling of other sources into one particular component, whereas F_2 accounts for the fact that two or more estimated components represent exactly the same original source [11]. It is worth noting that F_1 and F_2 are normalized so that their values do not depend on the dimensions of $\mathbf{P}^{qr}$. Since F_1 and F_2 are bounded between 0 (for a perfect generalized permutation matrix) and 1, F also ranges between 0 and 1. Hence, the lower the value of F for a pair of algorithms, the more consistent they are (i.e., the outcomes of both algorithms are more similar).

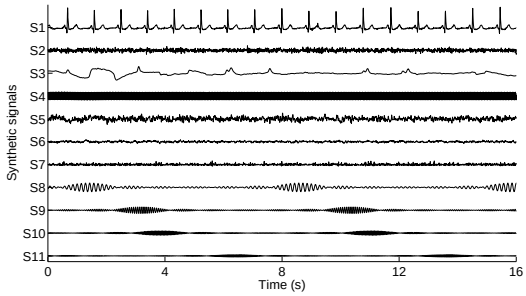
3.5. Synthetic signals

For the sake of completeness, synthetic signals are generated to evaluate the quality of the separations computed with AMUSE, SOBI, JADE, extended-Infomax and FastICA. However, it should be borne in mind that the reliability of any results computed from simulated data is limited.

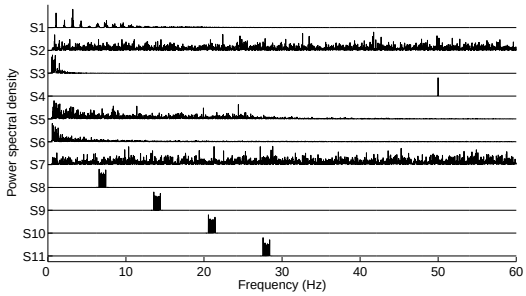
The synthetic signals used in this study were developed in [9]. They are composed by 11 inner components. These signals have the same sample frequency and were processed with the same filter as the real MEG recordings. Fig. 1 depicts one example of each synthetic source including their time plot, power spectral density and histogram. Additional details can be found in [9]:

1. S1 corresponds to a real electrocardiogram representing the cardiac artefact.
2. S2 is an inner white Gaussian noise source.
3. S3 is a real electrooculogram illustrating ocular activity.
4. S4 is a sine wave at 50 Hz.
5. S5 is a real MEG channel selected to have minimal artefactual activity.
6. S6 is a $1/f$ noise source.
7. S7 is a white exponential noise source.
8. S8 to S11 represented rhythmic activity with main frequencies are 7 Hz, 14 Hz, 21 Hz and 28 Hz.

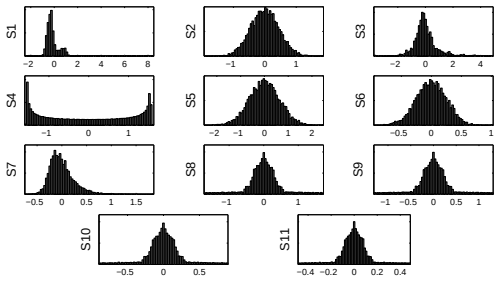
The synthetic signals allow to evaluate how close the mixing matrix computed with the BSS is to the actual one. In order to do so, the metric F is calculated from a matrix $\mathbf{P}^{qr}$ where $\mathbf{A}^q$ refers to the known synthetic mixing matrix and $\mathbf{A}^r$ is estimated with a BSS technique.



(a) Time plot.



(b) Power spectral density.



(c) Histogram.

Figure 1: Example of synthetic sources. The time plot (a), power spectral density (b) and histogram (c) are shown.

To avoid any influence of the preprocessing, the number of mixtures is set to the number of components ($m = l = 11$). The synthetic mixing matrix is created with a Gaussian process with zero mean and SD equal to one [9]. In order to study the influence of the synthetic data length, epochs of 2 s, 4 s, 8 s, 16 s, 32 s and 64 s are considered. For each length, 100 different instances of the synthetic signals are created with random delays or phases.

3.6. Statistical Analysis

Boxplots are used to visually summarise distributions of data. This diagram is composed of a box with three horizontal lines at the lower quartile, median and upper quartile values. The confidence interval of the median is indicated with a couple of notches. The boxplot also has two whiskers to show the extent of the rest of the data, which is estimated as 1.5 times the interquartile range. Values beyond the end of the whiskers are considered outliers and are marked with a '+'.

For the real MEG signals, a one-way ANalysis Of VAriance (ANOVA) is used to test whether the means of several groups are all equal. This procedure offers the possibility of partitioning the observed covariance in the data into components due to diverse explanatory variables (e.g., 'Pair of algorithms'). Additionally, a quantitative predictor (i.e., covariate) can be removed from the samples by a regression in order to account for some variability and increase statistical power. In this case, the number of components (l) can be taken as a covariate. The Scheffé's correction will be applied in the post-hoc multiple comparison procedure. In the case of the synthetic data, it is not necessary to consider l as a covariate in the ANOVA since it has a constant value.

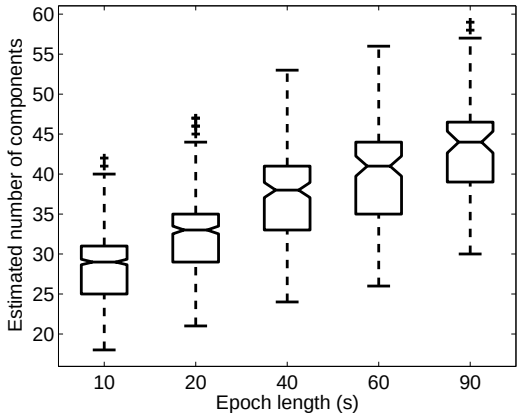


Figure 2: Boxplots showing the number of components (l) estimated for the real MEG epoch lengths.

4. Results

Our main objective was to study the consistency of real MEG data decompositions estimated with five common BSS algorithms for five epoch lengths: 10 s, 20 s, 40 s, 60 s and 90 s. Firstly, the value of l was estimated for each case with the MDL [9, 30]. Fig. 2 shows the boxplots representing the distributions of l for every epoch length. As it can be expected, l tended to increase with the epoch length of the MEG signal.

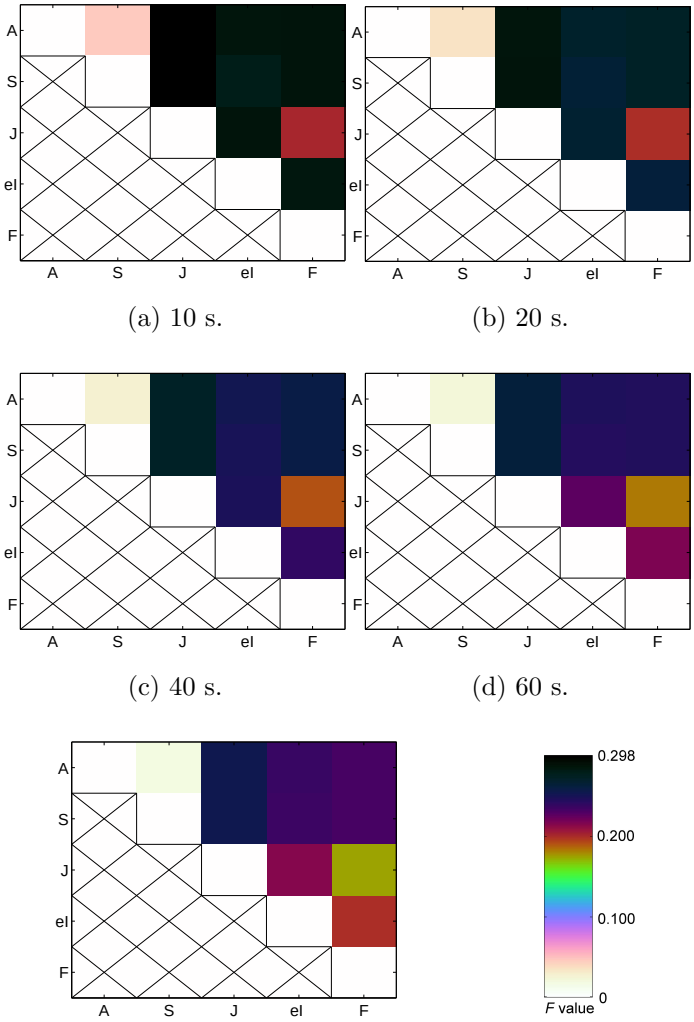
Secondly, the MEG recordings were preprocessed with the optimal l value estimated for each epoch. These preprocessed signals were decomposed with the five BSS methods. Then, the matrices $\mathbf{P}^{qr}$ were computed for each epoch and pair of algorithms and characterized with the metric F. In order to reduce the amount of data to be analysed, we studied only one matrix, $\mathbf{P}^{qr}$, instead of both $\mathbf{P}^{qr}$ and $\mathbf{P}^{rq}$. This decision was supported by the fact that the average absolute differences for the F metric between $\mathbf{P}^{qr}$ and $\mathbf{P}^{rq}$ were

always lower than 1.2%.

For each length, the F values obtained for every pair of algorithms were averaged. These results are depicted in Fig. 3, where all subplots are represented with the colour scale used to represent the data [34]. Lower F values are related to more consistent (more similar) pairs of algorithms. For all epoch lengths, Fig. 3 suggests that the most consistent pair of algorithms was 'AMUSE–SOBI' (SOS-based methods), followed by the pair 'JADE–FastICA', which involve HOS. Moreover, Fig. 3 shows that the general level of consistency improved as the length of the real MEG epochs increased. This suggests that the separations provided by different algorithms tended to converge as larger signals were decomposed.

For each epoch length, a one-way ANOVA with the Scheffé's multiple comparison procedure, 'Pair of algorithms' as the grouping factor and 'Number of estimated components' (l) as a covariate was used to statistically evaluate the differences in the F values.

For epochs of 10 s, there were significant differences in the F values as a consequence of the factor 'Pair of Algorithms', the covariate 'Number of estimated components' and their interaction ($p \ll 0.0001$ in all cases). The slopes of the regression of F against l were significantly different from 0 ($p < 0.05$) for the pairs of algorithms 'AMUSE–SOBI', 'AMUSE–extended-Infomax' and 'SOBI–extended-Infomax'. In the first case, F slightly increased with n (a larger number of components made the decompositions more different). For the other two pairs, more components produced lower F. Finally, the post-hoc multiple comparison procedure confirmed that the level of consistency of the 'AMUSE–SOBI' pair was significantly lower from



(a) 10 s. (b) 20 s.

(c) 40 s. (d) 60 s.

(e) 90 s. (f) Colour scale.

Figure 3: Average F values for each pair of BSS algorithms (A: AMUSE, S: SOBI, J: JADE, eI: extended-Infomax, F: FastICA) and length of the real MEG epochs: (a) 10 s, (b) 20 s, (c) 40 s, (d) 60 s and (e) 90 s. For the sake of clarification, the zero-level of the F metric for redundant pairs has been included in the diagonal of each subplot. The colour scale [34] used to represent the F values appears in (f).

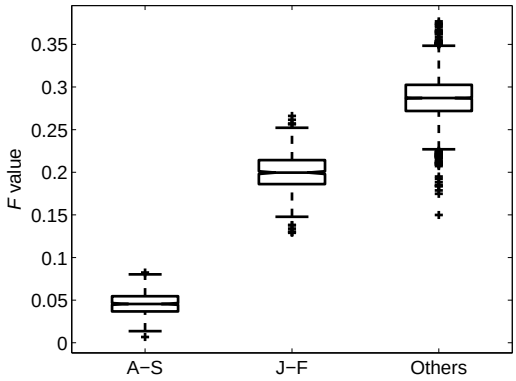


Figure 4: Boxplots showing the F values for the pairs of algorithms 'AMUSE–SOBI' (A-S) and 'JADE–FastICA' (J-F) and jointly for the other 8 pairs of algorithms (Others) for real MEG epochs of 10 s.

that of 'JADE–FastICA', and that the F values for these two pairs also differed significantly from the other eight pairs. Of note is that 'AMUSE–JADE' and 'SOBI–JADE' offered the most different separations. To illustrate this statistical differences among pairs of algorithms, Fig. 4 shows the boxplots of the F values for pairs 'AMUSE–SOBI', 'JADE–FastICA' and the rest of pairs. It can be observed that 'AMUSE–SOBI' has the lowest F values, followed by 'JADE–FastICA'.

The results obtained for epochs of 20 s were very similar to those previously reported for 10 s, with the same level of significant differences in the grouping factor and covariate. In this case, the regression of F against l was significantly positive for 'JADE–FastICA' as well as for the cases reported for the 10 s case.

The case of epoch length equal to 40 s presents slight deviations from the previous results. The F values varied significantly with 'Pair of Algorithms'

($p \ll 0.0001$), 'Number of estimated components' ($p = 0.0100$) and their interaction ($p \ll 0.0001$). The slopes of the regression for 'AMUSE–extended-Infomax', 'AMUSE–FastICA', 'SOBI–extended-Infomax' and 'SOBI–FastICA' decreased with l, whereas the pair 'JADE–FastICA' offered less similar separations for larger l. Likewise the previous cases, the post-hoc multiple comparison procedure indicated that 'AMUSE–SOBI' and 'JADE–FastICA' offered the most similar decompositions between algorithms. The analysis also suggested that the outcomes of 'AMUSE–JADE' and 'SOBI–JADE' were the most dissimilar.

When epochs of 60 s were studied, the F values only presented significant differences for 'Pair of Algorithms' and its interaction with l ($p \ll 0.0001$ in both cases). The slopes of F against l that are significantly different from zero are identical to those indicated in the analysis made for epochs of 40 s. The multiple comparison procedure suggested that 'AMUSE–SOBI' and 'JADE–FastICA', in that order, were the most consistent pairs of algorithms. There was also a tendency for 'AMUSE–JADE' and 'SOBI–JADE' to compute the least similar decompositions.

The decompositions of 90 s were also analysed. Only the 'Pair of Algorithms' and its interaction with l had significant p values ($p \ll 0.0001$). The pairs 'JADE–FastICA' and 'JADE–extended-Infomax' offered more different separations when l increased ($p < 0.05$). 'AMUSE–extended-Infomax', 'AMUSE–FastICA' and 'SOBI–extended-Infomax' had regression slopes significantly ($p < 0.05$) lower than zero. For this epoch length, 'AMUSE–JADE' and 'SOBI–JADE' computed the least consistent separations. On the other hand, four pairs of algorithms had significantly different population marginal

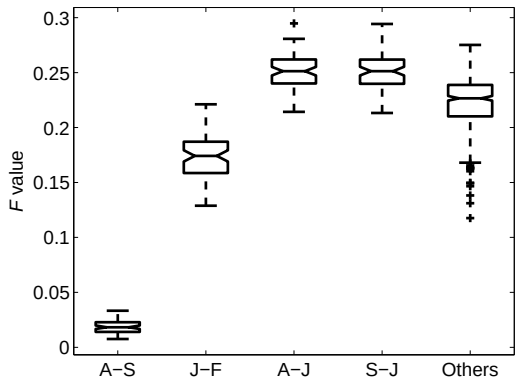
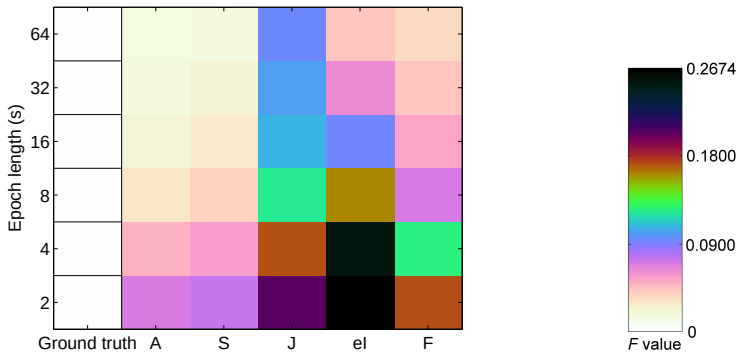


Figure 5: Boxplots showing the F values for the pairs of algorithms 'AMUSE–SOBI' (A-S), 'JADE–FastICA' (J-F), 'AMUSE–JADE' (A-J) and 'SOBI–JADE' (S-J) and jointly for the other 6 pairs of algorithms (Others) for real MEG epochs of 90 s.

means for the F values from the rest of pairs: 'AMUSE–SOBI', 'JADE–FastICA', 'extended-Infomax–FastICA' and 'JADE–extended-Infomax' (from more consistent to more dissimilar). The results for the most consistent and dissimilar pairs of algorithms are illustrated with boxplots in Fig. 5.

Finally, synthetic data were used to complement the previous analyses. AMUSE, SOBI, JADE, extended-Infomax and FastICA were used to estimate the mixing matrix. The metric F was computed from a matrix $\mathbf{P}^{qr}$ where q represented the known synthetic mixing matrix and r referred to one of the BSS methods. Hence, F indicated the quality of the decomposition, with lower F values accounting for more accurate estimations. For each synthetic epoch length, the F values of the 100 instances of the simulated data were averaged. These results are illustrated in Fig. 6. Similarly to the real recordings, Fig. 6 indicates that the accuracy of the separation increased



(a) F values for the synthetic data. (b) Colour scale.

Figure 6: Average F values for the decompositions of the synthetic data computed with AMUSE (A), SOBI (S), JADE (J), extended-Infomax (eI) and FastICA (F) for epochs of 64 s, 32 s, 16 s, 8 s, 4 s and 2 s. The zero-level of the F metric for a perfect decomposition is also plotted (Ground truth). The colour scale [34] used to represent the F values appears in (b).

with the epoch length and that AMUSE and SOBI offered similar levels of accuracy in the BSS.

A Scheffé-corrected one-way ANOVA with 'Algorithm' as factor was applied to assess the differences in the F values of the synthetic signals. For all epoch lengths, the differences in the separation accuracy were significant ($p \ll 0.0001$) and showed the same pattern in the multiple comparison analysis. AMUSE and SOBI provided the best estimations of the synthetic mixing matrix. Their level accuracy was significantly different from that of HOS-BSS techniques. As for this type of methods, FastICA calculated more accurate separations than JADE and extended-Infomax and the quality of JADE did not improve with the signal length as much as in the other cases. The differences among the three HOS-BSS approaches were significant for all

signal lengths.

5. Discussion and Conclusions

We compared five BSS algorithms in terms of the similarity between their decompositions for the real signals recorded from 26 subjects. We only evaluated one matrix, $\mathbf{P}^{qr}$, for each pair of algorithms instead of both $\mathbf{P}^{qr}$ and $\mathbf{P}^{rq}$ as the average differences for the F metric between $\mathbf{P}^{qr}$ and $\mathbf{P}^{rq}$ were always lower than 1.20%. By taking this decision, we tried to reduce the surplus complexity and redundancy of the problem. The results indicated that 'AMUSE–SOBI' and 'JADE–FastICA', in that order, are the pairs of algorithms that provide more similar BSS decompositions, while 'AMUSE–JADE' and 'SOBI–JADE' provided the most different outcomes. The separations tended to be more simillar for longer epochs. These results were supported by a complementary analysis of synthetic signals.

The preprocessing does not constitute a BSS algorithm itself. It relies on the classical projection technique of FA [9, 30, 31]. However, this preprocessing is important for several reasons [4, 5, 9, 21, 30]:

1. The number of inner meaningful components in real EEGs and MEGs may be less than the number of available channels.
2. A dimensionality reduction may sometimes be necessary to avoid "overfitting".
3. A dimensionality reduction may help to reduce the importance of the external noise.

The preprocessing included the estimation of the optimum number of components (l).

The dependence of the preprocessing on the signal length was studied for real epochs of 10 s, 20 s, 40 s, 60 s and 90 s. As it was expected, l increased with the epoch length. This means that longer signals tend to be composed of more inner sources or, at least, need to consider more components to obtain an optimum decomposition. These results are supported by other contributions about the model order selection in EEG and MEG [21, 22]. These studies investigated the performance of diverse approaches based on PCA and FA to estimate the number of BSS components in real EEG and MEG. Those results indicated that probabilistic PCA and FA models yield estimations of the dimensionality that are more reliable and independent of the signal power than commonly used PCA approaches [22]. The estimated values of l were about one third of the measurement space dimension [22]. In our case, the number of components was usually lower than one third of channels, specially for the shorter epochs. This suggested that, in the case of MEG equipment, more channels do not necessarily reflect more brain signals [22]. What is more, the data dimension reduction is supported by the statistical properties of the signal and the FA models may offer an appropriate description of the brain recordings [21, 22].

The visual representation of the results provided by Fig. 3 clearly indicated that the pair 'AMUSE–SOBI' provided the most similar decomposition of the analysed MEGs. This result was confirmed by the fact that these techniques computed the most accurate decompositions of the synthetic data. The principle beneath these two techniques is the simultaneous diagonalisation of several time-delayed cross-covariance matrices [3, 4]. We also found that the decompositions of JADE and FastICA are characterized by a high

degree of similarity for the real signals. This might be explained by the fact that the theoretical principles of both algorithms could be related [4].

Additionally, the algorithms tended to estimate more similar decompositions of the real MEGs as longer epochs were considered. This may be due to the fact that, although l increased with the epoch length, the number of data samples considered increased more rapidly than the number of elements to be estimated in $\mathbf{A}$ [2]. This fact is illustrated in Table 1, which depicts the number of available data samples and the median of the number of elements to be estimated in $\mathbf{A}$ for each epoch length. It is clear that, the longer the epoch, the larger the number of samples available to estimate each element of $\mathbf{A}$. Hence, the decompositions may be considered more reliable for longer epochs [2], which could explain the 'relative similarity' of the BSS outcomes for long real MEG epochs. This result was supported by the analysis of synthetic data. Table 1 also depicts the number of synthetic samples available to compute the BSS. As the number of synthetic components was fixed, longer signals offered more accurate decompositions. Yet, it should be noticed that, for the synthetic data, the quality of the JADE separation did not improve as much as with the other techniques.

The statistical analysis carried out for every epoch length pointed out the statistical significance of the similarity between the decomposition computed by 'AMUSE–SOBI' and 'JADE–FastICA'. On the other hand, the pairs 'AMUSE-JADE' and 'SOBI-JADE' used to provide the most dissimilar separations of the MEG signals. A relatively consistent pattern was that a larger number of components made the outcomes of the separations calculated by 'AMUSE-SOBI' and 'JADE-FastICA' slightly more different. Surprisingly,

Table 1: Ratios of the number of data samples for each epoch length divided by the median value of the number of elements in $\mathbf{A}$ for the real MEG recordings and for the synthetic signals.

Real MEG recordings				
Epoch length	Data samples	Median of l	Elements in $\mathbf{A}$ (l^2)	Ratio
10 s	1695	29	841	2.02
20 s	3390	33	1089	3.11
40 s	6780	38	1444	4.67
60 s	10170	41	1681	6.05
90 s	15255	44	1936	7.88
Synthetic signals				
Epoch length	Data samples	Value of l	Elements in $\mathbf{A}$ (l^2)	Ratio
2 s	339	11	121	2.80
4 s	678	11	121	5.60
8 s	1356	11	121	11.21
16 s	2713	11	121	22.42
32 s	5425	11	121	44.83
64 s	10851	11	121	89.68

the pairs 'AMUSE-extended-Infomax' and 'SOBI-extended-Infomax' computed BSS decompositions that were slightly more similar for larger values of l.

A few studies have compared the performance of several BSS techniques from different perspectives. For instance, synthetic signals have been used to evaluate whether the BSS improved the automatic detection of artefacts in the EEG [7] or to assess the quality of the BSS decomposition [10]. Our results from simulated data agree with those of [10] in the sense that SOS-BSS techniques seemed to calculate more accurate decompositions. SOS-BSS methods also performed better than HOS-BSS techniques in a detailed analysis of the ocular artefact rejection for EEG [16, 17]. However, in the artefact detection problem, Infomax performed better than FastICA and SOBI in [7]. Artificially mixed EEG signals have also been analysed in [8] to compare the relative performance of a few BSS algorithms to isolate the artefacts. The results varied depending on which type of contamination was considered [8]. Real MEG signals were decomposed in [6] to evaluate their ability to extract artefacts by comparing the contaminated components of different algorithms with reference signals. Some of these previous studies used small datasets or small numbers of channels and the evaluation of the algorithms was frequently based on subjective criteria [10]. Moreover, it must be noted that different sets of synthetic data could produce different results. In contrast, we analysed the decompositions of real MEG recordings from 26 subjects globally to gain insight into the similarities between some of the most commonly used BSS algorithms. Instead of comparing a manually selected subset of components [6], the entire decomposition was assessed since

the metric F [11] was computed from the mixing matrices $\mathbf{A}^q$ and $\mathbf{A}^r$.

It is important to note that the real sources are unknown, hence the term blind [5]. Therefore, assessing the performance of the BSS analysis is not straightforward at all since the separation cannot be absolutely validated for real data [3, 5]. Thus, the analyses of the real signals were exploratory and only aimed at measuring the similarity between the results of the BSS algorithms and not at evaluating the actual quality of the separation. This can only be achieved with some kind of synthetic signals [32, 33]. Yet, our complementary analysis of the simulated data supported the results derived from the real MEG recordings. This suggests that the results for each pair of algorithms are indeed due to the methodology of the BSS techniques and not to this particular application. Our study is also limited by the fact that only real signals of MEG background activity were studied. Additionally, only recordings from elderly people were analysed. Thus, the results might be difficult to generalise to younger subjects.

To sum up, this study evaluated the degree to which diverse BSS techniques provide similar decompositions for real MEG background activity. The most similar separations were computed with 'AMUSE–SOBI', followed by 'JADE–FastICA'. The pairs 'AMUSE–JADE' and 'SOBI–JADE' used to provide the most dissimilar outcomes. Finally, the overall level of similarity increased as longer signals were decomposed. These results were supported by a study based on synthetic signals. Since diverse BSS methods may offer relatively different perspectives when applied to real signals [4], these results should be taken into account when deciding which BSS algorithms are to be applied to brain signals. For instance, if only two BSS are to be

selected for an exploratory analysis, the algorithms AMUSE and JADE will provide relatively different perspectives of the data and minimise the amount of redundant information.

6. Acknowledgement

The authors would like to thank Dr. Alberto Fernández ("Centro de Magnetoencefalografía Dr. Pérez-Modrego"), who collected the MEG recordings. This study was partially supported by the "Ministerio de Ciencia e Innovación" and FEDER grant TEC2008-02241.

References

[1] R. Hari, Magnetoencephalography in clinical neurophysiological assessment of human cortical functions, in: E. Niedermeyer, F. Lopes da Silva (Eds.), Electroencephalography: Basic Principles, Clinical Applications, and Related Fields, Lippincott Williams and Wilkins, Philadelphia, U.S.A., 2004, pp. 1165–1197.

[2] J. Onton, M. Westerfield, J. Townsend, S. Makeig, Imaging human EEG dynamics using independent component analysis, Neuroscience and Biobehavioral Reviews 30 (6) (2006) 808–822.

[3] C. James, C. Hesse, Independent component analysis for biomedical signals, Physiological Measurement 26 (1) (2005) R15–R39.

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

[5] R. Vigário, E. Oja, BSS and ICA in Neuroinformatics: From Current Practices to Open Challenges, IEEE Reviews in Biomedical Engineering 1 (2008) 50–61.

[6] H. Zavala Fernández, T. Sander, M. Burghoff, R. Orglmeister, L. Trahms, Comparison of ICA Algorithms for the Isolation of Biological Artifacts in Magnetoencephalography, Lecture Notes in Computer Science 3889 (2006) 511–518.

[7] A. Delorme, T. Sejnowski, S. Makeig, Enhanced detection of artifacts in EEG data using higher-order statistics and independent component analysis, NeuroImage 34 (4) (2007) 1443–1449.

[8] S. Fitzgibbon, D. Powers, K. Pope, C. Clark, Removal of EEG Noise and Artifact Using Blind Source Separation, Journal of Clinical Neurophysiology 24 (3) (2007) 232–243.

[9] J. Escudero, R. Hornero, D. Abásolo, A. Fernández, M. Lopez-Coronado, Artifact Removal in Magnetoencephalogram Background Activity With Independent Component Analysis, IEEE Transactions on Biomedical Engineering 54 (11) (2007) 1965–1973.

[10] M. Klemm, J. Haueisen, G. Ivanova, Independent component analysis: comparison of algorithms for the investigation of surface electrical brain activity, Medical & Biological Engineering & Computing 47 (4) (2009) 413–423.

[11] C. Melissant, A. Ypma, E. Frietman, C. Stam, A method for detection of

Alzheimer's disease using ICA-enhanced EEG measurements, Artificial Intelligence in Medicine 33 (3) (2005) 209–222.

[12] J. Escudero, R. Hornero, J. Poza, D. Abásolo, A. Fernández, Assessment of classification improvement in patients with Alzheimer's disease based on magnetoencephalogram blind source separation, Artificial Intelligence in Medicine 43 (1) (2008) 75–85.

[13] J. Escudero, R. Hornero, D. Abásolo, A. Fernández, Blind source separation to enhance spectral and non-linear features of magnetoencephalogram recordings. Application to Alzheimer's disease, Medical Engineering and Physics 31 (7) (2009) 872–879.

[14] Y. Li, D. Powers, J. Peach, Comparison of blind source separation algorithms, Advances in Neural Networks and Applications (2001) 18–23.

[15] M. Crespo-Garcia, M. Atienza, J. Cantero, Muscle artifact removal from human sleep EEG by using independent component analysis, Annals of Biomedical Engineering 36 (3) (2008) 467–475.

[16] S. Romero, M. Mañanas, M. Barbanoj, A comparative study of automatic techniques for ocular artifact reduction in spontaneous EEG signals based on clinical target variables: A simulation case, Computers in Biology and Medicine 38 (3) (2008) 348–360.

[17] S. Romero, M. Mañanas, M. Barbanoj, Ocular Reduction in EEG Signals Based on Adaptive Filtering, Regression and Blind Source Separation, Annals of Biomedical Engineering 37 (1) (2009) 176–191.

[18] V. Krishnaveni, S. Jayaraman, P. Kumar, K. Shivakumar, K. Ramadoss, Comparison of Independent Component Analysis Algorithms for removal of ocular artifacts from Electroencephalogram, Measurement Science Review Journal 5 (2) (2005) 67–79.

[19] V. Krishnaveni, S. Jayaraman, K. Ramadoss, Application of mutual information based least dependent component analysis (MILCA) for removal of ocular artifacts from electroencephalogram, International Journal of Biological and Medical Sciences 1 (1) (2006) 63–74.

[20] A. Hyvärinen, P. Ramkumar, L. Parkkonen, R. Hari, Independent component analysis of short-time Fourier transforms for spontaneous EEG/MEG analysis, NeuroImage 49 (1) (2010) 257–271.

[21] C. Hesse, Model order estimation for blind source separation of multichannel magnetoencephalogram and electroencephalogram signals, in: Proceedings of the 30th Annual International Conference of the IEEE Engineering in Medicine and Biology Society, Vancouver, BC, Canada, 2008, pp. 3348–3351.

[22] C. Hesse, On Estimating the Signal Subspace Dimension of High-Density Multichannel Magnetoencephalogram Measurements, in: Proceedings of the 29th Annual International Conference of the IEEE Engineering in Medicine and Biology Society, Lyon, France, 2007, pp. 6227–6230.

[23] L. Tong, R. Liu, V. Soon, Y. Huang, Indeterminacy and identifiability of blind identification, IEEE Transactions on Circuits and Systems 38 (5) (1991) 499–509.

[24] A. Belouchrani, K. Abed-Meraim, J. Cardoso, E. Moulines, A blind source separation technique using second-order statistics, IEEE Transactions on Signal Processing 45 (2) (1997) 434–444.

[25] J. Cardoso, A. Souloumiac, Blind Beamforming for Non Gaussian Signals, IEE proceedings. Part F. Radar and Signal Processing 140 (6) (1993) 362–370.

[26] T. Lee, M. Girolami, T. Sejnowski, Independent component analysis using an extended infomax algorithm for mixed subgaussian and supergaussian sources, Neural Computation 11 (2) (1999) 417–441.

[27] A. Delorme, S. Makeig, EEGLAB: an open source toolbox for analysis of single-trial EEG dynamics including independent component analysis, Journal of Neuroscience Methods 134 (1) (2004) 9–21.

[28] H. Gävert, J. Hurri, J. Särelä, A. Hyvärinen, FastICA Toolbox, website, [online] `http://www.cis.hut.fi/projects/ica/fastica/` (March 2010).
URL `http://www.cis.hut.fi/projects/ica/fastica/`

[29] A. Cichocki, S. Amari, K. Siwek, T. Tanaka, A. Huy Phan, ICALAB for Signal Processing, website, [online] `http://www.bsp.brain.riken.jp/ICALAB/ICALABSignalProc/` (March 2010).
URL `http://www.bsp.brain.riken.jp/ICALAB/ICALABSignalProc/`

[30] S. Ikeda, K. Toyama, Independent component analysis for noisy data–MEG data analysis, Neural Networks 13 (10) (2000) 1063–1074.

[31] J. Cao, N. Murata, S. Amari, A. Cichocki, T. Takeda, A robust approach to independent component analysis of signals with high-level noise measurements, IEEE Transactions on Neural Networks 14 (3) (2003) 631–645.

[32] A. Mansour, M. Kawamoto, N. Ohnishi, A survey of the performance indexes of ICA algorithms, in: Proc. of the 21st IASTED Int. Conf.: Modelling, Identification and Control (MIC02), 2002, pp. 660–666.

[33] E. Vincent, R. Gribonval, C. Févotte, Performance measurement in blind audio source separation, IEEE Trans. on Audio, Speech and Language Processing 14 (4) (2006) 1462–1469.

[34] J. McNames, An effective color scale for simultaneous color and gray-scale publications, IEEE Signal Processing Magazine 23 (2006) 82–96.