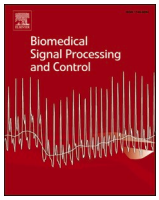
Biomedical Signal Processing and Control 72 (2022) 103292

Contents lists available at ScienceDirect

Biomedical Signal Processing and Control

journal homepage: www.elsevier.com/locate/bspc



Which BSS method separates better the EEG Signals? A comparison of five different algorithms

Christos Stergiadis, Vasiliki-Despoina Kostaridou, Manousos A. Klados *

Department of Psychology, University of York Europe Campus, CITY College 24, Proxenou Koromila Street, 546 22 Thessaloniki, Greece.

ARTICLE INFO

Keywords:
EEG
BSS
ICA
Signal processing
EEG decomposition

ABSTRACT

A very common strategy for rejecting electroencephalographic (EEG) artifacts, includes the decomposition of filtered EEG signals using a Blind Source Separation (BSS) algorithm, the identification and removal of artifactual components and the reconstruction of the cleaned EEG signals. In this pipeline, the performance of the BSS algorithm, which is defined as its ability to separate properly the independent sources (like the EEG from artifactual sources), is very crucial for rejecting most of the artifacts, while maintaining the most part of EEG intact. The overwhelming majority of the published papers uses the extended INFOMAX version of Independent Component Analysis (ICA) for artifact rejection purposes. But is this the most efficient algorithm to separate EEG signals into independent components? This study comes to shed light to the aforementioned question by assessing the performance of the five most common BSS algorithms. The normalized entropy of the brain-related components, their correlation between independent components with the original sources and the amount of the overall mutual information reduction (MIR) achieved by each decomposition were computed in datasets with systematically varying numbers of electrodes (ranging from 19 to 99), from 26 real human scalp EEG recordings. Additionally, 54 different datasets containing artificially contaminated EEG signals were also examined for the same purpose, on the basis of the Euclidean distance and the correlation, between the generated Independent Components (ICs) and the original vertical and horizontal eye signals, which were used for the contamination. The results support that the Adaptive Mixture ICA was the best performing BSS method.

1. Introduction

Since Hans Berger recorded the first electroencephalographs (EEGs) almost a century ago (1924) [1], the research for the optimization of the clinical procedures and the interpretation of the acquired signals has been growing constantly and over the last decade gained increased attention due to its implementation in a variety of scientific fields and applications [2,3]. Each of the aforementioned signals do not represent directly the activity of the specific brain area below each electrode, but conversely a mixture of source signals that are generated from different neuronal sub-areas inside the brain. Some of those areas are connected with each other functionally or structurally and are activated concurrently [4]. More specifically, the inhibitory and excitatory potentials that the cortical nerve cells generate in order to communicate with each other, summate in the cortex and can be recorded in the scalp surface as an EEG signal [5]. The simultaneous activity of different sources creates a linear mixture of electrical signals that each scalp electrode receives in different volumes.

A major contribution to the aforementioned signal comes from biological artifacts like muscle activity (EMG noise) [6], eye blinks (EOG noise) [7] or heartbeats (ECG noise) [8] which corrupt the EEG signal even more. The ocular activity in particular, poses as the major obstacle of good quality EEG recordings [9], due its unpredicted occurrence and because of the subject's natural need to blink. Such artifacts give rise to a number of problems in EEG related studies, like for example corrupting the Event Related Potential (ERP) analysis, or tricking the scientists in epilepsy studies as the ocular artifacts may be misinterpreted as epileptogenic spikes [10,11], leading to incorrect conclusions. Instructing the subject to restrict his/her eye movements and blinks may be an effective way around this problem, but often comes with the cost of affecting its cognitive processes [12,13], resulting in misleading results also. Additionally, in numerous studies the option of repeating the experiment if the subject blinks or moves his/her eyes is undesired, as for example in studies including children with Attention Deficit Hyperactivity Disorder (ADHD) or in habituation studies [14]. Considering these, the clinicians and the researchers are called to find an efficient

* Corresponding author.
E-mail address: mklados@york.citycollege.eu (M.A. Klados).

https://doi.org/10.1016/j.bspc.2021.103292
Received 10 August 2021; Received in revised form 4 October 2021; Accepted 23 October 2021
Available online 8 November 2021

and accurate method to extract the underlying neural activations, free of artifactual activity.

For this reason, a variety of artifact rejection techniques have been developed over the last decades, which can be mainly divided into two groups [9]. One group contains the regression – based techniques [15,16], where the algorithms try to find the portion of an artifactual referral signal in the EEG recordings, while the second group contains the Blind Source Separation (BSS) – based methods [17,18], where a BSS algorithm is used to decompose the signal into independent components (ICs) and then the artifactual sources are removed manually or automatically [9,19–20]. Independent Component Analysis (ICA) [21], is a well-established family of BSS algorithms, which is used to decompose the EEG signals into statistically independent components (ICs) [22]. This family contains different implementations of ICA, like Informax [21], FastICA [21], Adaptive Mixture ICA (AMICA) [23] etc., where their main difference is spotted on the way they conclude to statistically independent sources [21]. Additionally, to the ICA family, there are some other BSS algorithms that are based on second-order statistics, like Second Order Blind Identification (SOBI) [24], Blind Identification with Robust Orthogonalization (SOBIRO) [25] and Algorithm of Multiple Unknown Signal Extraction (AMUSE) [26].

The variety of the existing BSS methods raises the issue of which algorithm is the optimal choice in terms of achieved separation between the generated components. The importance of this matter lies in the fact that after the algorithmic decomposition of the filtered EEG signal, the generated components that represent the ocular artifacts must contain as little neural information as possible, so by removing those components and reconstructing the signal, the distortion of the original brain information is minimized. Thus, a good separation between the desired brain components which carry the neural information and the artifactual components that will eventually be removed from the data is desirable.

Previous studies [27,28,29] in this area have covered a wide range of different assessment criteria and set the path for the current research. Initial studies [28] have tested the more standard algorithms, like Infomax, FastICA and Joint Approximation Diagonalization of Eigenmatrices (JADE) [30], using MATLAB flops and the Signal to Noise Ratio (SNR) of the separated outputs. The results highlighted the higher computational cost of Infomax, as its performance is highly dependent on the gaussianity of the instantaneous mixtures to be separated. Other studies [29] focusing on similar algorithms (Principal Component Analysis, AMUSE, JADE, SOBI and fastICA) used simultaneous electrocorticography in order to evaluate the components that each algorithm generated, as this method captures most of the original neural sources. So, by computing the correlation coefficient between the outcome of BSS and the ECoG recordings, measurable results were obtained about the performance of each algorithm. Another category of comparison criteria, has to do with the performance of the different algorithms in terms of hardware compatibility and achievement. The assessment of five commonly applied algorithms like Infomax, JADE, fastICA and SOBI, in the basis of better running time, less allocated memory and optimal accuracy and scalability, indicated the superiority of SOBI against the rest of the algorithms as it presented the best results in terms of stability in the performed tests [28]. Similar studies but with different types of evaluation criteria, like the dipolarity of the components and the achieved mutual information reduction (MIR) have been conducted [27] in order to assess a wider range of algorithms and investigate the correlation between the aforementioned metrics. The results showed that the algorithms that used the natural gradient descent as the method for the determination of the maximally independent components proved to be the most well-performing in terms of separative ability and component definition, with AMICA ranking first.

The current study comes to examine the performance of five of the most commonly used BSS algorithms from a mathematical and statistical point of view. The comparison features were drawn both from previous works [27], such as the MIR, but also from the desire to investigate some of the measures that a large number of ICA methods do use in their methodology, like the different entropy features. In addition, the Spearman correlation coefficient was also considered as an assessment tool, although differentiating from the above features as it requires the knowledge of the original brain sources. An important breakthrough in the herein proposed methodology is the introduction of a supplementary EEG dataset which was artificially contaminated with ocular artifacts [31]. For the assessment of the algorithms in this dataset the Euclidean Distance and the correlation between the generated components and the initial Electrooculographic (EOG) signals were deployed. The results from both datasets were then compared and the best performing algorithm was nominated.

The rest of the paper is organized as follows. In Section 2 the datasets, the features and the algorithms used in this study are described. Section 3 presents the results for the 7 in total features used herein, while at Sections 4 and 5 we provide a discussion of our findings in light of the current literature and a conclusion of our work. Finally, a future work plan is proposed in Section 6.

2. Materials and methods

2.1. Data and preprocessing steps

2.1.1. Dataset 1 – Real data

Real continuous EEG recordings obtained from 26 healthy subjects in resting state, with their eyes closed. The duration of each dataset varied, with the mean duration for the 26 datasets being 12,84 s ± 2,99 s. Left (Right) channels were referred to the left (right) mastoid, while all the central electrodes were referred to the half of the sum of the left and right mastoid. The sampling rate was 250 Hz. A total of 126 scalp electrodes were placed according to the Geodesic Sensor Net (EGI, Inc.). In order to access and process our data, EEGLAB [32] was used. The continuous EEG data were first band pass filtered between 0.5 and 100 Hz using a FIR filter and then notch filtered at 50 Hz.

In order to examine the effectiveness of the BSS algorithms against the number of the provided channels, the total number of electrodes was gradually increased from 19 to 99 by one resulting in 81 different channel sets. For the 26 data sets and the 5 algorithms the total number of ICA decompositions reached to 26x5x81 = 10.530.

2.1.2. Dataset 2 - Semi-Simulated data

Pre-contaminated data

Real continuous EEG recordings were obtained from 27 healthy subjects with normal vision, 14 males (mean age: 28,2 ± 7,5 years) and 13 females (mean age: 27,1 ± 5,2 years), in an eyes closed session. The duration of each dataset was 30 s and in total 54 datasets were obtained (two for each subject). Nineteen scalp EEG electrodes (FP1, FP2, F3, F4, C4, P3, P4, O1, O2, F7, F8, T3, T4, T5, T6, Fz, Cz, Pz) were placed according to the 10–20 International system. Odd indices referred to the left and even indices to the right mastoid respectively, while the three central electrodes (Fz, Cz, Pz) were referenced to the half of the sum of the left and right mastoids. The sampling frequency was 200 Hz and the signals were then band pass filtered at 0,5–40 Hz and Notch filtered at 50 Hz. Separate, Electrooculography (EOG) signals were obtained during an eyes-opened condition, with four electrodes placed above and below the left eye and on either side of the canthi. As a result, two bipolar signals emerged, with the one accounting for the vertical eye movements/eye blinks (equal to the upper minus lower EOG electrode recordings) and the other referring to the horizontal eye movements (equal to the left minus right EOG electrode recordings). Those signals were band-pass filtered at 0.5–5 Hz and were used for the contamination of the eyes closed datasets, in order to generate the artificially contaminated datasets. For more detailed information about this dataset, you can refer to [9,31].

2.2. BSS and ICA algorithms

Several different algorithms were compared in this study. The coveted aim in ICA is the determination of the unmixing matrix that maximizes the independence between the component time courses. Although, the strategy behind the estimation of this independence differs for each ICA algorithm. The following algorithms were implemented:

2.2.1. Infomax-ICA

Infomax-ICA [27] main aim is to maximize independence, by maximizing negentropy: $J(y) = H(y_{gauss}) - H(y)$. Negentropy J(y) is a normalized version of differential entropy H(y). Its use is preferred over H(y), as it provides us with a non-negative metric for measuring non-gaussianity. Negentropy is also invariant of invertible linear transformations and in some sense is the optimal estimator of non-gaussianity, although computationally expensive. By definition maximizing negentropy comes to the minimization of the differential entropy: $H(x) = -\int p_y(\eta)\cdot log p_y(\eta)\cdot d\eta$. It presents a gradient method for the estimation of the unmixing matrix W and the computation of each maximally non-gaussian direction (each independent component). As the name of this algorithm implies, the maximization of entropy results in the maximization of the mutual information I(x,y) between the input and the output vectors, as the mutual information of two random variables y is defined as:

$$I(X;Y) = H(X) - H(X/Y)$$

2.2.2. Adaptive mixture ICA (AMICA)

AMICA [23] differentiates from the standard ICA mixture methods, as it adapts the source density models of each computed component by using a mixture of generalized Gaussian density models. As a result, a very good match is achieved between the density model and the actual density of the source under study. More specifically AMICA's mixing procedure is based on a three-layered model. In the first two layers the standard ICA methodology is being employed and the active ICA mixture model is being computed. The third layer approximates the probability density function (PDF) of each component as a mixture of the Generalized Gaussians, which thereupon is parametrized. This final step makes AMICA special, as the other Infomax related algorithms assume pre-defined sub-Gaussian and super-Gaussian PDFs for each of the sources [33].

2.2.3. Second order statistics-based approaches

This group of algorithms are based on the joint diagonalization of an arbitrary set of time delayed covariance matrices, computed from $x(t)$ [34]. The following expression describes the above assumption: $R_x(t) = \{x(t)x(t+\tau)^T\} = AR_s(\tau)A(t), \forall\tau$, where R_s and R_x are the correlation matrix of the source signals and the time delayed matrix respectively.

2.2.3.1. Algorithm for Multiple Unknown signals extraction (AMUSE):. AMUSE [26] is one of the most popular second order statistics BSS approaches. The decorrelation of $x(t)$ here, takes place at two temporal lags, which is also its weakness, as it challenges its robustness [24]. This algorithm simultaneously diagonalizes two symmetric matrices and obtains the longed-for estimated sources. First of all, the estimation of the covariance matrix $\widehat{R}_x(0)$ at time lag $\tau = 0$ takes place, followed by the computation of the eigenvalue decomposition (EVD) or the singular value decomposition (SVD) of $\widehat{R}_x(0)$. Then the algorithm performs the pre-whitening transformation $x\_(k) = Qx(k)$. It also estimates the covariance matrix $\widehat{R}_x(\tau)$ for a specific time lag other than $\tau = 0$ and performs SVD on the aforementioned matrix. Finally, it estimates the covariance mixing matrix as $\widehat{A} = Q^{+}U_{x\_}$ and the source signals as $y(k) = U_{x\_}^{T}x\_(k)$. Another great feature of this algorithm is its ability to use the spectral content of the extracted components in order to arrange them by descending linear predictability [35].

2.2.3.2. Second order blind identification (SOBI):. SOBI (Belouchrani & Cichocki, 2000) shares a very similar methodology with AMUSE. The big difference between the two algorithms is that instead of the simultaneous diagonalization with EVD and SVD techniques which AMUSE adapts, SOBI utilizes joint diagonalization in order to find the orthogonal matrix U that is responsible for the diagonalization of the set of covariance matrices [24]. For this purpose, a large number of different algorithms can be implemented, with most of them including Jacobi techniques, parallel factor analysis or alternating least squares etc. Other than that, SOBI also uses a pre-whitening transformation and after the determination of the covariance matrices at specific time lags, the final estimation of the sources follows the exact same procedure as in AMUSE.

2.2.3.3. Second order blind identification - robust (SOBIRO):. SOBIRO introduces a robust orthogonalization step, in order to adapt for noisy signals [25]. The orthogonalization stands by the following expression:

$$x\_ = Qx(t)$$

for the mixed signals. As a result, the delayed covariance matrices are defined as:

$$R_{x\_}(p_i) = \left(\frac{1}{N}\right)\sum_{k=1}^{N} x\_(k)x\_(k-p_i) = QR_{x\_}(p_i)Q^T$$

where $(p_1, p_2, p_3, \ldots p_k)$ are the time delays. Applying Joint Approximation Diagonalization (JAD), with respect to the orthogonalized mixing matrix $A\_ = QA$, results in:

$$R_{x\_}(p_i) = QR_{x\_}(p_i)Q^T = AR_s(p_i)A^T = UD_iU^T$$

The estimation of the orthogonal mixing matrix becomes $\widehat{A} = Q\widehat{H} = U$, where H is the original mixing matrix. Finally, the longed-for estimated source signals are given as: $\widehat{S}_k = U^T Qx(k)$ and the mixing matrix $\widehat{H} = Q + U$ [25].

2.3. Comparison criteria

In this study the normalized entropy, the Spearman correlation coefficient and the amount of mutual information reduction (MIR) will be used as the features in order to compare the separation performance of the five different BSS algorithms (Infomax-ICA, SOBI, SOBIRO, AMUSE, AMICA). Entropy is a measure that has not been yet considered as an assessment criterion before, while MIR and Spearman CC has been used in similar previous studies [27,36]. We are also going to conduct statistical analysis and investigate the performance of each algorithm as the number of EEG channels (electrodes) increases. It is shown that although entropy and MIR share a similar statistical foundation, (more specifically MIR is based on entropy measures in its mathematical formula) they generated results that are not entirely matching. Moreover, we show that comparing the algorithms using the Spearman correlation coefficient produced results that greatly agreed with the MIR based algorithm assessment.

2.3.1. Entropy as a criterion of Non-gaussianity

Entropy is considered a novel criterion on measuring non-gaussianity and it is generally preferred over kurtosis, due to the fact that in many cases kurtosis's robustness is challenged [21]. In information theory, entropy is one of the basic concepts, as it measures the amount of information that is contaminated in a variable. As the randomness of a variable increases, the entropy also increases. In this paper Shannon entropy is going to be used, which is one of the most common types of entropy. More specifically we are going to examine the normalized Shannon entropy for comparative reasons.

2.3.1.1. Computing the normalized Shannon entropy. As it was previously

mentioned the normalized Shannon entropy will be used as the first criterion for the evaluation of the performance of the different algorithms. It is known that as the distribution of a component diverges from the normal (Gaussian) one, it decreases its entropy value [21]. This theoretical observation derives from the fundamental result of information theory that a gaussian variable has the largest entropy. Thus, according to the central limit theorem, the optimal separation between the brain sources is achieved when the components are maximally non-gaussian and present the minimal values in terms of entropy, as in that case each of them is defined by a combination of less original brain sources.

In this study, three different types of normalized Shannon entropy of the component activations were computed:

i) The first one is based on a formula proposed by Kumar and Kapur [37] who ended up in the equation: $H\_(P) = -\frac{1}{ln(n)}\sum_{i=1}^{n} p_i ln(p_i)$, where p_i is the probability distribution.
ii) The second one, utilizes the signal itself instead of the probability distribution. The corresponding MATLAB function (myEntropy. m) follows the equation: $H\_(x) = -\frac{1}{n}\sum_{i=1}^{n} x^2 \cdot ln(x^2)$.
iii) The third mathematical computation for the entropy was based on the wavelet entropy. The function of MATLAB (wentropy.m) for this type of entropy is expressed as: $H\_(x) = -\sum_{i=1}^{n} x_i{}^2 log(x_i{}^2)$.

An important fact that needs to be mentioned, is that MATLAB uses the amplitude of the signal as the input for the computation of the entropy function and not the probability p(n), usually generating negative values. To overcome this problem, we normalized the signal by its total energy, by using $\frac{x(n)}{sum(x(n)^2)}$ instead of $x(n)$.

The results of each entropy were stored in corresponding matrices. By the end of the computation part of all entropies, the 3D matrix *dataset* × *algorithm* × *channelset*, was constructed for all three of them. In order to ensure that the entropy values that filled this matrix were computed only for components that represented original brain sources and not artifacts, the MARA [38] classifier was used for all 10.530 decompositions. From the output of this classifier, we only considered as normal (brain sources) the components that throughout the whole range of the different algorithms and datasets, where identified as brain sources with 75% confidence or more.

2.3.2. Spearman correlation coefficient

Another criterion for the assessment of the BSS and ICA algorithms is the level of correlation between their produced components and the actual sources inside the brain [36]. In order for such an algorithm to be operating successfully there must be a resemblance between the components and the brain sources that generate the original data at the time of the EEG recording. For this purpose, a correlation-based index had to be drought. The Spearman correlation coefficient was chosen, for two major reasons:

1. It depends only on the relative shape of the signal and not on its absolute amplitude.
2. It poses as a non-parametric measure, in contrast with the Pearson correlation coefficient (CC) which performs well only for Gaussian sources, which is not the case in our study as in ICA the "maximal" non-gaussianity is desired [36].

Additionally, the Spearman CC magnitude ranges between 0 and 1, with the sign not being of importance here.

A very important detail in this methodology is the knowledge of the original sources for comparison reasons. In order to deal with this problem, the decomposition results from the AMICA algorithm were considered as the reference and the remaining four algorithms were compared. This choice was carefully made as the AMICA algorithm presented the best results out of the five, both in terms of the entropy of its components but also regarding the MIR values and the Euclidean Distance.

The principal formula for the computation of the Spearman correlation coefficient is:

$$r = 1 - 6\sum \frac{d^2}{N(N^2-1)}$$

where d represents the difference in statistical rank between the two compared variables and N is the total number of pairs of values.

In Dataset 1 and for every BSS algorithm under study, the CC in every channel set is computed as follows:

a) The correlation between every independent component and every "brain source" (AMICA component) is calculated. The higher CC is then chosen for this component, as it is more likely for this component to represent the corresponding source signal of the brain.
b) The mean value of the above CCs is then calculated and assigned for the specific channel set and algorithm.

By the end of the computation of all the Spearman Correlation Coefficient values for all the decompositions, a 26x4x81 matrix containing the corresponding results was constructed.

2.4. Mutual information reduction

2.4.1. Mutual information

The level of independence between a set of random variables can also be measured by the amount of the system's mutual information [27]. Mutual information is a non-negative metric that is highly associated with the entropy of the system. More specifically, for two random variables A and B, the mutual information between these two variables is defined as: $I(A;B) = h(A) + h(B) - h(A,B)$

where, $h(A)$ and $h(B)$are the individual entropies of A and A, and $h(A,B)$represents the joint entropy.

So, for the elements of a signal vector $y = [y_1, ..., y_n]$, the mutual information is described as:

$$I(y_1; ....; y_n) = h(y_1) + ... + h(y_n) - h(y)$$

The higher the values of mutual information, the smaller the divergence between the uncertainties (entropies) of the joint distribution against the individual entropies of the components. Therefore, in a system described by small values of mutual information, the constituting components present a higher level of independence as they make the vector y when described as a hole, to have relatively similar uncertainty $[h(y)]$ as if it were considered as a number of each component separately $[h(y_1) + ... + h(y_n)]$ [27].

As has been discussed above, in BSS the principal goal is the computation of the unmixing matrix W, so that the source signals $x = A \bullet$ s can be decomposed into independent components, $y = W \bullet$x, where A is the mixing matrix and s the original source signals. The entropy of such a component (source estimate) signal can be described by:

$$h(y) = log|det(W)| + h(x)$$

So, the corresponding mutual information of the transformed data will be:

$$I(y) = h(y_1) + ... + h(y_n) - log|det(W)| - h(x)$$

2.4.2. Mutual information reduction

Every BSS algorithm achieves a different amount of reduction of the mutual information during its implementation. The total volume of mutual information removed from a set of channels can be described by:

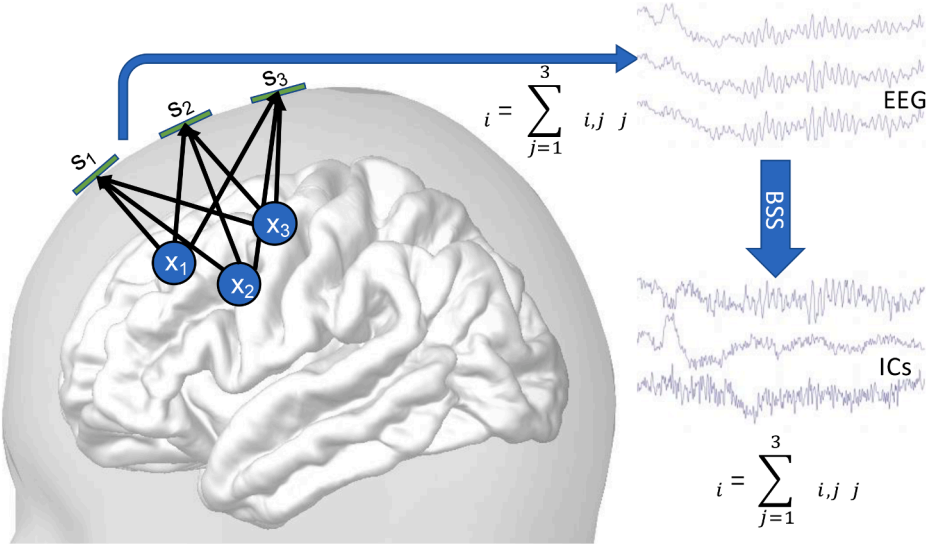


Fig. 1. The scalp electrodes (S1, S2, S3) receive the linear mix of the electrical activity generated from different neuronal sub-areas inside the brain (X1, X2, X3) in different volumes. Then the signal is amplified and illustrated on a computer screen. By using a BSS method, the raw EEG signal from the different electrodes is decomposed to the same number of Independent Components.

$$MIR = I(x) - I(y)$$
$$= [h(x_1) + \ldots + h(x_n)] - [h(y_1) + \ldots. + h(y_n)] - h(x) + log|detW| + h(x)$$
$$= log|detW| + [h(x_1) + \ldots + h(x_n)] - [h(y_1) + \ldots. + h(y_n)] \text{ [36]}$$

As can be seen, MIR depends only on the determinant of matrix W and the difference between the individual entropies of the original signals and the component time courses. The higher the value of MIR is, the more independent the final components are and so better the implementation of BSS methodology. Completing the computation of the MIR values for all the five decompositions and the sum 26 datasets, a 26x5x81 aggregate matrix was obtained containing the corresponding results. What also needs to be mentioned is that the software used for the computation of the MIR was taken from the Swartz Center of Computational Neuroscience (SCCN) of the University of San Diego (https://sccn.ucsd.edu/wiki/BSSComparison).

2.5. Euclidean distance

The Euclidean distance between two points is defined as the minimum distance between those two points in the Euclidean space and it is denoted as the length of a line segment between those two points. So, for two points **k** and **m** which are given by Cartesian coordinates in the **n**-dimensional Euclidean space, the corresponding mathematical formula is: $d(k,m) = \sqrt{(k_1 - m_1)^2 + (k_2 - m_2)^2 + \ldots. + (k_n - m_n)^2}$.

Now, the case with the ICs that are generated after every BSS decomposition is that they can be expressed as two-dimensional vectors. In addition, the VEOG and HEOG signals that were separately collected from an eyes-open session in Dataset 2 can also be expressed as vectors. In order to investigate the level of performance of the algorithms under study we can use the Euclidean distance between the ICs and the VEOG and HEOG signals. The logic behind this line of thought is the following: The optimal decomposition of an EEG signal would account for the minimum amount of brain information present into the artifactual components that needs to be removed. That means that after a good decomposition the artifactual components should only represent the biological sources that are unrelated to the crucial cortical activity. In the case of the semi-simulated data of Dataset 2, the EOG signals are known. As a result, it is possible to measure the relationship between the artifactual ICs that are generated by every decomposition with the VEOG and HEOG signals, through the Euclidean distance of their vectors. The smaller the distance, the stronger the relationship of the artifactual Independent Component with the EOG signal is, and therefore the less the amount of brain information that it contains. Eventually, the smaller the distance the better and more accurate the decomposition. The above procedure provides us with a measurable metric of the performance of the algorithm under study, by focusing on the artifactual ICs that the corresponding decomposition generates.

From the 19 channels of each dataset, the EEG decomposition generated 19 different independent components. The distance was calculated between every IC vector and the VEOG and HEOG signals of the corresponding dataset and then the lowest value of Distance was taken into account. This value represented the maximum correlation between the main artifactual component that was obtained from the decomposition and the original vertical and horizontal eye artifacts. This procedure was implemented for all 54 artificially contaminated datasets and all the five algorithms. Finally, the average value of the Euclidean Distance was calculated across all datasets and this metric was used as the representative value of the performance of the algorithm under study.

The data and the code used in this paper is available in Github (https://github.com/ramsys28/BSSCompPaper), for replication purposes.

3. Results

3.1. Results regarding Dataset 1

For all five comparison features (three different entropies, Spearman CC value and MIR) that were used in Dataset 1, each algorithm's performance was plotted in order to carefully examine its behavior as the number of the total channels increased. From a statistical point of view a one-way ANOVA was chosen as the most appropriate analysis as the aim of the study was to compare the different BSS algorithms in terms of separating ability and performance, assessing the five aforementioned features.

3.1.1. Entropy based comparisons

Considering the lower entropy values as the indicator for the optimal separation performance, this section presents the results of our study based on the aforementioned criterion. Fig. 1 present the behavior of the average entropy value of the brain source components, of each

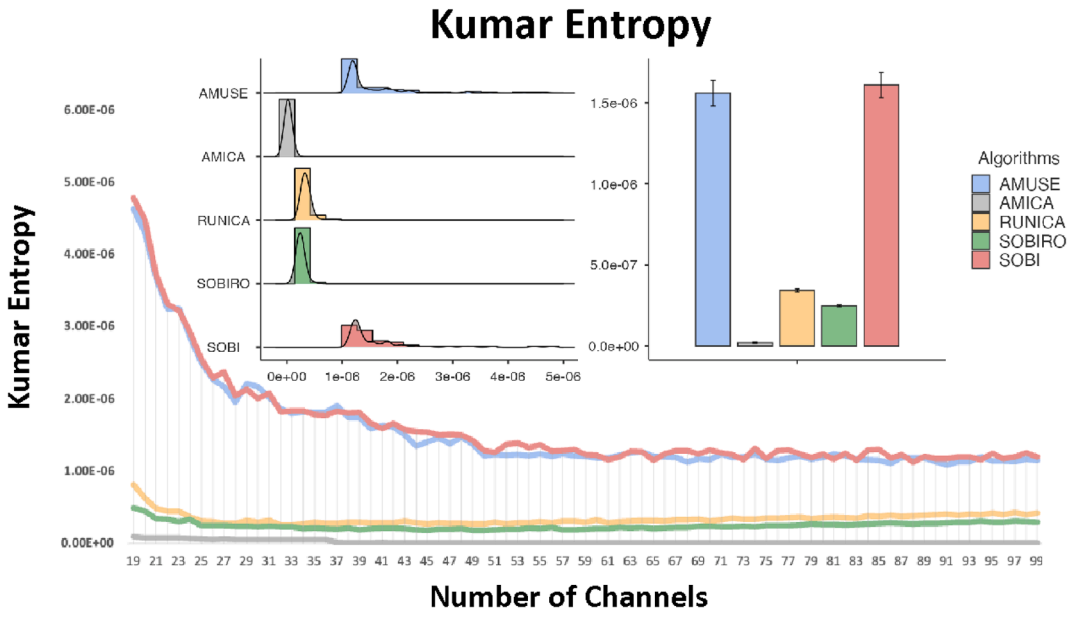


Fig. 2. Entropy values based on the entropy formula of Kumar and Kapur against the number of the data channels(electrodes) that were taken under consideration. The AMICA algorithm is obviously out-performing the other four algorithms in terms of lower entropy and consequently in better separation between the independent components. There are also two clustering zones, one including RUNICA and SOBIRO and the other AMUSE and SOBI.

decomposition, for the five algorithms under evaluation.

3.1.1.1. Kumar based entropy. There was a significant difference between the mean value of the entropies for the five algorithms [F (4, 125) = 36.75, $p < .001$]. As can be detected, AMICA is the best performing algorithm as it presents the lower values of entropy across the whole range of channel sets. Post hoc comparisons using the Tukey-Kramer test justified this observation as they indicated that the mean value of entropy using the AMICA algorithm ($M = 2.24{\cdot}10^{-8} \pm 2{,}43{\cdot}10^{-8}$) differed significantly from the corresponding values regarding AMUSE and SOBI ($p < .001$ in both) and non-significantly from RUNICA ($p = .34$) and SOBIRO ($p = .57$) algorithms. Two major clustering zones divided the remaining four algorithms. More specifically, SOBIRO which is the next best performing algorithm based on this type of entropy ($M = 2{,}49{\cdot}10^{-7} \pm 5.40{\cdot}10^{-8}$) behaves similarly with RUNICA ($M = 3.44{\cdot}10^{-7} \pm 8.09{\cdot}10^{-8}$), even though each method follows a different strategy for the extraction of the independent components. The second clustering zone belongs to AMUSE ($M = 1.56{\cdot}10^{-6} \pm 7.03{\cdot}10^{-7}$) and SOBI ($M = 1{,}61{\cdot}10^{-6} \pm 7.10{\cdot}10^{-7}$), the two less well-performing algorithms, which both are based on Second Order Statistics. Finally, from Fig. 2 it can be observed that the value of the entropy seems to be rapidly decreasing as we move from the first channel set (19 electrodes) to the tenth channel set (29 electrodes) and then presents a steadier and almost linear trend until we increase to the last channel set (99 electrodes). This generally descending character of the entropy lies in the fact that the more the channels under consideration, the less the number of the original brain sources that define a certain independent component. And according to the central limit theorem, this increases the non-gaussianity of this component and subsequently it decreases its entropy value.

3.1.1.2. Normalized Shannon entropy. Considering the myEntropy.m sample function of MATLAB as the comparison feature, similar results were obtained as with the Kumar entropy-based comparisons. Due to this similarity, further graphs considering this type of entropy are omitted. ANOVA results indicated that the difference between the mean values of entropy for each algorithm were significant [$F(4, 125) = 34.15$, $p < .001$]. AMICA was also the best performing algorithm ($M = 3.49{\cdot}10^{-11} \pm 2.57{\cdot}10^{-11}$), with SOBIRO ($M = 6.26{\cdot}10^{-10} \pm 1.36{\cdot}10^{-10}$) and RUNICA ($M = 8.55{\cdot}10^{-10} \pm 2.04{\cdot}10^{-10}$) following and creating the first clustering zone. As was observed using Kumar entropy, AMUSE ($M = 3.98{\cdot}10^{-9} \pm 1.76{\cdot}10^{-9}$) and SOBI ($M = 4.08{\cdot}10^{-9} \pm 1.78{\cdot}10^{-9}$), two of the second order statistics (SOS) algorithms presented the less good results regarding their separation performance. The decreasing rate of the entropy value in all the algorithms is also present in this case.

3.1.1.3. Wavelet entropy. The form of our results remained unchanged for this final type of entropy. The corresponding graph containing the entropy values for each algorithmic decomposition is almost identical to Fig. 2 and thus is omitted. The difference between the mean values for every algorithm remained significant [F (4, 125) = 37.80 , $p < .001$]. The best performing algorithm is AMICA ($M = 1.15{\cdot}10^{-7} \pm 8.42{\cdot}10^{-8}$), with SOBIRO ($M = 2.02{\cdot}10^{-6} \pm 4.38{\cdot}10^{-7}$) and RUNICA ($M = 2.80{\cdot}10^{-6} \pm 6.55{\cdot}10^{-7}$) following. As before, AMUSE ($M = 1.27{\cdot}10^{-5} \pm 5.68{\cdot}10^{-6}$) and SOBI ($M = 1.30{\cdot}10^{-5} \pm 5.75{\cdot}10^{-6}$) maintained higher entropy values compared to the other three algorithms.

There is an evident classification of the five algorithms under evaluation. The AMICA algorithm, which uses the three-layered mixing procedure, is the best performing algorithm in terms of component separation, as it generates components whose entropy values are significantly lower than in any other algorithm. Additionally, SOBIRO and RUNICA do present similar results even though they are based on different statistical strategies. The least well-performing algorithms, AMUSE and SOBI, cluster together and indicate that the more standard second order statistic approaches (those without the robust orthogonalization step like SOBIRO) do not meet the expectations in terms of separation between the generated components.

3.1.2. Mutual information reduction based comparisons

Comparing the amount of mutual information that each algorithm managed to remove from the raw data, we extracted the following results. The dependence between these two metrics is described by an almost linear trend line. As the number of channels increased, the additional amount of input data resulted in a gradual increase of the MIR values.

A one-way ANOVA was conducted between the algorithms for the investigation of the potential differences between the mean values of MIR after the implementation of every algorithm. The result of this statistical analysis indicated that this difference was not as significant [F

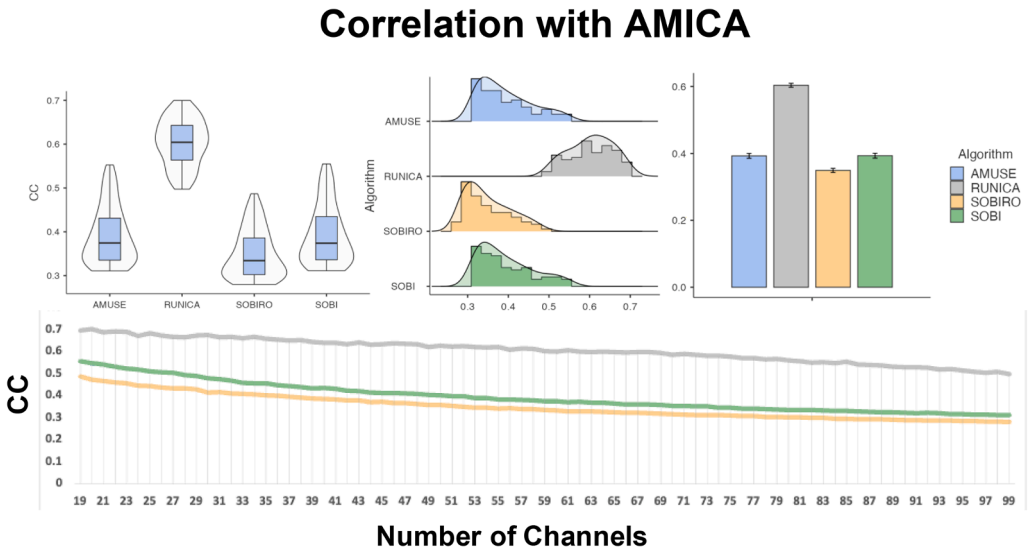


Fig. 3. Spearman CC values against the number of the data channels(electrodes). AMICA is taken as the exemplar algorithm and the remaining four algorithms are compared. RUNICA is the best performing algorithm in terms of higher correlation with the ground truth sources. AMUSE and SOBI do cluster together and present lower levels of correlation, while SOBIRO is the least performing algorithm in terms of correlation.

$(4, 125) = 3.58$, $p = .15$]. This finding occurs as the divergence between the mean MIR values of the algorithms in small channel sets (<26) was far less significant ($p > .12$) than the corresponding value in the input data set where a large number of channels was taken into account ($p < 0.001$ for 45 channels and above).

Similarly, to the entropy-based comparison, implementing the AMICA algorithm in order to decompose the raw EEG data, presented the optimal results in terms of component separation ($M = 118{,}57 \pm 47.40$ kbits/sec). A very close competitor of AMICA was the RUNICA algorithm ($p = 0.98$) which was the next best performing ($M = 116{,}55 \pm 46.82$ kbits/sec). A second clustering zone was created by the less well-performing algorithms, with AMUSE ($M = 108.68 \pm 42.38$ kbits/sec) and SOBI ($M = 108.66 \pm 42.35$ kbits/sec) having almost identical values throughout the whole range of the channel sets and SOBIRO ($M = 105.61 \pm 40.98$ kbits/sec) achieving the less amount of mutual information reduction.

Again, the three second order statistics (SOS) based algorithms were out-performed by the gradient method of RUNICA and the multi-layered approach of AMICA.

3.1.3. *Correlation based comparisons*

Setting the exemplary AMICA algorithm as the reference, the four

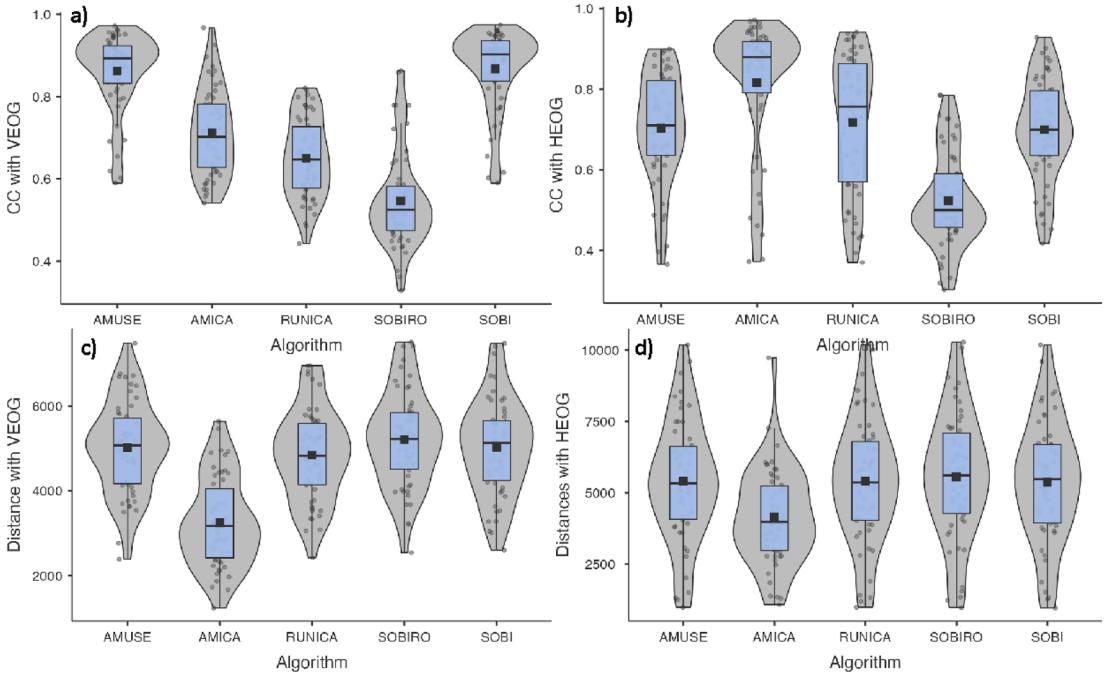


Fig. 4. Violin plots presenting the Spearman CC and the Euclidean Distance between the main artifactual IC and the original VEOG and HEOG signals. a) AMUSE and SOBI present the highest mean correlation values between the main artifactulal component and the original VEOG signal. b) AMICA is the best performing algorithm, as it generated artifactual components with the highest correlation with the original HEOG signal. c) AMICA outperforms the other four algorithms in terms of smaller Euclidean distance between the artifactual IC and the VEOG signal. d) Again, AMICA achieves the smallest distance with the horizontal eye-movement (HEOG) signal.

remaining algorithms were compared on the basis of their components' correlation with the corresponding components ("brain sources") produced by AMICA. The divergence between the mean Spearman CC value, of the components of each algorithm's decomposition, for the 26 subjects and across every channel set, was significant [F (3, 100) = 502.40 , $p < .001$]. Interestingly, the level of this divergence was much more intense ($\sim 10^{-32}$) than the equivalent that emerged both from the entropy and the MIR features. Fig. 3 illustrates the mean Spearman CC value of the components, against the number of channels that were taken into account.

RUNICA algorithm ($M = 0.60 \pm 0.05$) outperformed the other three as it presented components with higher correlation values regarding the "original brain sources". Of course, this may be related to the fact that AMICA's components were chosen as the ground truth and the two algorithms share a similar methodology, as both are Infomax related. So, a higher correlation between them is not a surprising result. Additionally, there is a clear clustering between the SOS algorithms. AMUSE ($M = 0{,}39 \pm 0.07$) and SOBI ($M = 0{,}39 \pm 0.06$) did not perform as well as RUNICA, but achieved slightly better results than SOBIRO ($M = 0{,}35 \pm 0.06$) where the correlation with the original brain sources was smaller compared with all the other algorithms.

3.2. Results regarding Dataset 2

In this section the results of the comparison features regarding the semi-simulated data of Dataset 2 are provided. These results concern the computation of the Spearman Correlation Coefficient (CC) and the Euclidean Distance (D) between the independent components that are generated and the initial vertical-EOG (VEOG) and horizontal-EOG (HEOG) signals that were acquired separately in an eyes-open EEG session. The one-way ANOVA was chosen as the statistical analysis tool while the p-values were calculated for comparative reasons.

3.2.1. Spearman CC between ICs and VEOG/HEOG signals

3.2.1.1. Vertical-EOG (VEOG) signals. The difference between the mean Spearman CC values across the 54 datasets for the five different algorithmic decompositions was significant [F (4, 265) = 98.80 , $p < .001$] in the case of the correlation with the VEOG signals/eye blinks.

Fig. 4 (a) illustrates the mean CC values, the interquartile range and the distribution for every algorithm. The best performing algorithms concerning this metric proved to be SOBI ($M = 0.87 \pm 0.1$) and AMUSE ($M = 0.86 \pm 0.09$) as the main eye artifactual IC that they generated introduced the highest correlation with the initial VEOG signal that was used for the artificial contamination of Dataset 2. The two higher order statistic algorithms, AMICA ($M = 0.71 \pm 0.1$) and RUNICA ($M = 0.65 \pm 0.09$), achieved slightly lower levels of performance, while the remaining second order statistics algorithm SOBIRO ($M = 0.55 \pm 0.12$) produced ICs with the smallest amount of correlation with the VEOG signal out of all.

3.2.1.2. Horizontal-EOG (HEOG) signals. The results concerning the relation of the ICs with the bipolar signal that represented the horizontal eye movements showed a different picture compared to the VEOG one. Fig. 4 (b) presents the aforementioned results.

Again, in this case, there was a significant difference between the mean correlation for the five different algorithms [F (4, 265) = 29.24 , $p < .001$]. Here, the AMICA algorithm ($M = 0.82 \pm 0.16$) outperformed the other four, with RUNICA ($M = 0.73 \pm 0.17$) following and leading AMUSE ($M = 0.70 \pm 0.13$) and SOBI ($M = 0.70 \pm 0.12$) which clustered together and presented almost identical results. Finally, SOBIRO ($M = 0.52 \pm 0.11$) emerged again as the algorithm that did not perform as well as the other four as it presented the smallest amount of correlation. The above ranking seems to greatly agree with the results that we obtained from the MIR comparison in Dataset 1. It also shares many characteristics with the corresponding ranking of the Entropy features, with the difference that in the latter case, the SOBIRO algorithm clustered with RUNICA and outperformed the other two Second Order Statistics algorithms (SOBI and AMUSE).

Table 1
Computational time (sec) needed for each BSS algorithm.

Number of Channels	RUNICA	SOBI	SOBIRO	AMUSE	AMICA
19	13.87	1.79	1.79	1.66	28.22
32	16.51	1.93	1.93	1.67	38.82
64	27.82	2.69	2.99	1.68	65.26
99	41.97	4.23	4.67	1.75	98.21
Mean Time:	25.01	2.66	2.85	1.70	57.60

3.3. Euclidean distance between ICs and VEOG/HEOG signals

3.3.1. Vertical-EOG (VEOG) signals

In this section the results concerning the mean minimum Euclidean Distance between the main artifactual IC and the initial VEOG signal that represents the eye blinks are presented. The main results are illustrated in Fig. 4 (c), while a more detailed statistical analysis showed that the algorithms were characterized by an important difference between the mean distance values [F (4, 265) = 29.36 , $p < .001$].

One more time, AMICA ($M = 3.25{\cdot}10^3 \pm 1.04{\cdot}10^3$) posed as the optimal algorithm regarding its performance as it generated artifactual components whose distance with the VEOG signal was significantly smaller than the remaining four algorithms. More specifically, the other higher order statistics algorithm RUNICA ($M = 4.84{\cdot}10^3 \pm 1.07{\cdot}10^3$) joined AMUSE ($M = 5.01{\cdot}10^3 \pm 1.07{\cdot}10^3$), SOBI ($M = 5.02{\cdot}10^3 \pm 1.13{\cdot}10^3$) and SOBIRO ($M = 5.20{\cdot}10^3 \pm 1.12{\cdot}10^3$).

3.3.2. Horizontal-EOG (HEOG) signals

In the same manner as above, the study of the distance of the artifactual components with the horizontal-EOG signals this time, resulted in similar observations, as can be seen in Fig. 4 (d). The level of divergence between the mean distance values across the different algorithms remained significant [F (4, 265) = 4.21 , $p = .002$], but the deviation between the algorithms was smoother than the one concerning the VEOG signals ($p = .014$ between AMICA and RUNICA compared to $p < .001$ *in VEOG).*

AMICA's main artifactual components introduce the highest similarity with the HEOG signal ($M = 4.16{\cdot}10^3 \pm 1.65{\cdot}10^3$), while SOBI ($M = 5.37{\cdot}10^3 \pm 2.12{\cdot}10^3$), RUNICA ($M = 5.40{\cdot}10^3 \pm 2.16{\cdot}10^3$) and AMUSE ($M = 5.41{\cdot}10^3 \pm 2.13{\cdot}10^3$) cluster together creating a zone of intermediate performance. The algorithm whose main eye artifactual component diverges the most from the original horizontal-EOG bipolar signal is SOBIRO ($M = 5.55{\cdot}10^3 \pm 2.19{\cdot}10^3$), with the aforementioned distance being the longest of the above.

3.3.3. Execution time

Although our study was mainly focused around the level of separation that each algorithm achieved, we also proceeded to the computation of the overall time needed for the algorithm to run and generate the decomposed components. The computational cost of each algorithm is presented in Table 1. According to this table, the three Second Orden Statistics (SOS) algorithms decompose the raw EEG signal considerably faster than the Infomax based algorithms. More specifically, AMUSE (mean time = 1.70 sec) is the fastest algorithm with SOBI (mean time = 2.66 sec) and SOBIRO (mean time = 2.85 sec) following close by. On the other hand, the Infomax (RUNICA) algorithm (mean time = 25.01 sec) requires significantly more computational time and finally AMICA (mean time = 57.60 sec) emerges as the slowest algorithm of all, soaring the mean computational time at almost 1 min per decomposition.

In order to summarize all the herein presented results regarding the quality of the separation of all the methods under study, Table 1 was

Table 2
Ranking table of all the BSS algorithms.

Algorithms	Dataset 1				Dataset 2				Mean Ranking	Final Ranking
	Entropy	MIR	CC	Time	CC VEOG	CC HEOG	ED VEOG	ED HEOG		
AMICA	1	1	–	5	3	1	1	1	1.62	1
AMUSE	4	3	2	1	2	3	3	4	2.75	3
RUNICA	3	2	1	4	4	2	2	3	2.62	2
SOBI	5	4	2	2	1	4	4	2	3	4
SOBIRO	2	5	3	3	5	5	5	5	4.12	5

formed. In this table the rank of each algorithm and for each criterion is noted and then the mean ranking was computed by averaging all the ranks for each algorithm separately. According to the mean rank the final ranking was concluded. Table 2 indicates AMICA as the best performing algorithm (mean ranking = 1.62), RUNICA comes second (mean ranking = 2.62), then AMUSE (mean ranking = 2.75) and SOBI (mean ranking = 3) follow, while SOBIRO (mean ranking = 4.12) seems to be the less good performing BSS algorithm.

4. Discussion and conclusion

The implementation of BSS separation methods in the pre-processing stage of electroencephalographic data, has been practiced for over 20 years now. A very large number of studies [27,28,29] have been conducted, trying to understand the underlying validity of the available algorithms and categorize them in terms of achieved performance and plausibility. A good decomposition implies that the independent components are optimally separated with each other, meaning that they approach accurately the original brain sources that generated the EEG signal. Consequently, the amount of the neural information included in the artifactual components is minimized and thus the loss of information is prevented when those components are removed. The comparison between the variety of the different algorithms becomes an ever-increasing need, as the effective pre-processing of the raw EEG signals is of vital importance for the next stages of any EEG related study.

In this piece of work, we deployed a range of mathematical and statistical metrics in order to appoint the algorithm with the optimal separation capabilities between the generated Independent Components. We tried to spherically approach the matter by using two different datasets. The first dataset was from 26 real human EEG recordings in an eyes closed session, while the second dataset contained 27 human recordings that were artificially contaminated with electrooculographic artifacts. The assessment of the algorithms was made on a software basis, meaning that only the characteristics of specific statistical and mathematical metrics were taken into consideration and not the computational load and time for the decomposition of the filtered EEG signals. Some of the comparison features were drawn both from previous studies like the Mutual Information Reduction (MIR) [27] and the Spearman Correlation Coefficient (CC) [36], while others like the different entropy features were used for the first time to the best of our knowledge. The consideration of both real EEG data (Dataset 1) and artificially contaminated datasets (Dataset 2) come as a solution to the problem of not knowing the ground sources that generated the EEG signals, which is always present in studies that investigate the performance of the Blind Source Separation algorithms. As it is presented below, the overall agreement between the results from the two datasets reinforces the validity of our study and justifies our choice for the specific metrics as the comparison features.

4.1. Ranking

The most efficient algorithm proved to be AMICA. In reference to Dataset 1, this was true both in terms of entropy (Fig. 1) and MIR. In the artificially contaminated signals of Dataset 2, AMICA was also the dominant algorithm when the Euclidean distance was computed for both the vertical and horizontal EOGs. The ascendance of this algorithm was challenged though, when the Spearman Correlation Coefficient (CC) was calculated between the artifactual components of each decomposition and the original vertical EOGs, something that was not the case for the corresponding value regarding the horizontal eye movement artifacts. The exceptional performance of the AMICA algorithm, compared with RUNICA, the other natural gradient descent-based algorithm, probably lies in the fact that AMICA adapts the source density models by implementing a mixture of Generalized Gaussian density models [33]. Thus, a better separation between the components was achieved in contrast with the more standard infomax algorithm that assumes one fixed parametric template for the PDF of each independent component. The difference in performance between those two algorithms was far more significant when the normalized entropies were assessed [KUMAR: $p = 0.34$, Normal: $p = 0.36$, Wavelet: $p = 0.32$] compared with the corresponding performances regarding the MIR [$p = 0.98$]. Regarding the comparison features in the semi-simulated Dataset 2, the aforementioned algorithms had a largest divergence between their achieved performances, [$p = 0.014$ for the Euclidean Distance] and [$p = .008$ for the Spearman CC].

From the Second Order Statistics based algorithms SOBIRO did cluster with Infomax (RUNICA) (a higher order statistics-based algorithm), when the algorithms were assessed by their lower entropy values [$p = 0.96$ *in all*] in Dataset 1, despite no visible connection between the methodologies that they follow for the extraction of the independent components. This was not the case for the MIR, neither for the correlation-oriented comparisons, as SOBIRO was the least well-performing algorithm and clustered partially with SOBI and AMUSE [$p = 0.94$ *in both]* in the case of the MIR and on the other hand differed significantly [$p < .001$ *in both]* for the Spearman CC. Those two last algorithms presented a similar decomposition character, throughout the whole range of comparison features, something that is not surprising as they share a very similar methodology. The difficulty of SOBIRO for optimal separation was also observed when the artificially contaminated signals of Dataset 2 were decomposed, as it achieved the smallest amount of correlation and at the same time the highest Euclidean Distance with the original artifactual EOG signals. What derives from the above, is that the existence of the robust orthogonalization step that SOBIRO presents improved the ability of the algorithm to separate the components, compared with the other SOS based algorithms, only in terms of lower entropy values of the component's time courses and not for the amount of mutual information that the decomposition reduces, the final correlation that the independent components present compared with the original sources and the Euclidean Distance between the generated artifactual components and the original EOG signals.

The results of our study confirmed that even though each BSS algorithm managed to decompose the original EEG data into maximally independent brain components, there is an evident classification between those algorithms as they achieve different levels of performance. This finding comes to supplement existing studies, that also punctuate the importance of ranking amongst the different ICA algorithms. This need originates from the fact that in real time situations, a computationally fast and accurate decomposition is of vital importance during the

Table 3
Pros and cons of the BSS algorithms under study.

Algorithms	Pros	Cons
AMICA	Accurate decompositions (Ranked 1st), Scales well to high dimensions	Computational Time (Impractical for real life applications)
AMUSE	Computationally fast (Ranked 1st)	Average performance regarding separation
RUNICA	Accurate decompositions (Ranked 2nd)	Computational Time, Assumes non-gaussian sources
SOBI	Computational Time	Average performance regarding separation
SOBIRO	Computational Time	Least well performing algorithm regarding separation (Ranked Last), Malfunctions in decompositions with large number of electrodes (after 116 in our experience)

clinical procedure, as it composes the pre-processing stage of the raw EEG signal. So, the idea behind choosing the finer BSS algorithm was that the more precise the decomposition is, the more efficient the artifact removal procedure becomes and the loss of neural information is limited. We found that AMICA presented the optimal results. Throughout both datasets and the entire range of channel sets under study, it generated independent components whose time courses were characterized by the lower normalized entropy values compared with the other four algorithms, something that dictates superior statistical independence between them according to the central limit theorem. The superiority of AMICA extended also in a more likelihood-oriented metric, that of MIR. Another interesting result was that by setting AMICA's produced components as the ground truth brain sources and assessing the performance of the remaining algorithms in terms of correlation of their components with them, a very similar ranking with the other features was extracted, as RUNICA outperformed the SOS based algorithms. Thus, it is safe to say that when an accurate and clean decomposition of a raw EEG signal is desired, the higher order statistics algorithms and especially AMICA pose as the current optimal solution compared with other algorithms that follow the Second Order Statistics methodologies.

It is suggested that in order to be able to safely appoint an algorithm (or a group of algorithms) as optimally performing, a more thorough research must be conducted. This realization lies in the fact that in addition to the performance of an algorithm in terms of achieved separation of the computed "maximally independent" components, the computational complexity and demand should also be assessed in a more detailed manner. For example, in our study, where the hardware used was pretty standard (8 GB RAM, Intel i3 processor), SOBIRO malfunctioned after the input channels reached 116, with the required computational time exceeding the admissible. So, from that perspective, we encourage the line of thought of feature studies to be the simultaneous assessment of the computational load and the separation performance of different and newer algorithms in a variety of experimental setups. Achieving this, will get us a step closer to the optimization of the real time pre-processing approaches and the establishment of EEG as a 3D, real time brain imaging technique.

Table 3 summarizes all the aforementioned results and presents an overview of the algorithms under comparison, by noting both the pros and the key features but also some of the drawbacks during their implementation. The connection between the quality of the separation achieved and the computational cost that is demanded for the use of each algorithm is also visible here, enabling us to say that there is a clear tradeoff between performance and execution time.

CRediT authorship contribution statement

Christos Stergiadis: Methodology, Software, Formal analysis, Writing – original draft, Writing – review & editing. **Vasiliki-Despoina Kostaridou:** Investigation, Resources, Data curation, Writing – original draft, Writing – review & editing. **Manousos A. Klados:** Conceptualization, Methodology, Visualization, Supervision, Project administration, Writing – review & editing.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

References

[1] Haas, L. F. (2003). Hans Berger (1873-1941), Richard Caton (1842-1926), and electroencephalography. *Journal of Neurology, Neurosurgery & Psychiatry*, *74*(1). https://doi.org/10.1136/jnnp.74.1.9.

[2] Pham, T., Ma, W., Tran, D., Nguyen, P., & Phung, D. (2013). *EEG-Based User Authentication in Multilevel Security Systems*. https://doi.org/10.1007/978-3-642-53917-6_46.

[3] M.V. Ruiz-Blondet, Z. Jin, S. Laszlo, CEREBRE: A Novel Method for Very High Accuracy Event-Related Potential Biometric Identification, IEEE Trans. Inf. Forensics Secur. 11 (7) (2016) 1618–1629, https://doi.org/10.1109/TIFS.2016.2543524.

[4] L.Q. Uddin, Complex relationships between structural and functional brain connectivity, Trends in Cognitive Sciences 17 (12) (2013) 600–602, https://doi.org/10.1016/j.tics.2013.09.011.

[5] D.P. Subha, P.K. Joseph, R. Acharya U, C.M. Lim, EEG Signal Analysis: A Survey, J. Med. Syst. 34 (2) (2010) 195–212, https://doi.org/10.1007/s10916-008-9231-z.

[6] I.I. Goncharova, D.J. McFarland, T.M. Vaughan, J.R. Wolpaw, EMG contamination of EEG: spectral and topographical characteristics, Clin. Neurophysiol. 114 (9) (2003) 1580–1593, https://doi.org/10.1016/S1388-2457(03)00093-2.

[7] P. Senthil Kumar, R. Arumuganathan, K. Sivakumar, C. Vimal, Removal of Ocular Artifacts in the EEG through Wavelet Transform without using an EOG Reference Channel, In *Int. J. Open Problems Compt. Math* 1 (2008) Issue 3).

[8] J.-A. Jiang, C.-F. Chao, M.-J. Chiu, R.-G. Lee, C.-L. Tseng, R. Lin, An automatic analysis method for detecting and eliminating ECG artifacts in EEG, Comput. Biol. Med. 37 (11) (2007) 1660–1671, https://doi.org/10.1016/j.compbiomed.2007.03.007.

[9] M.A. Klados, C. Papadelis, C. Braun, P.D. Bamidis, REG-ICA: A hybrid methodology combining Blind Source Separation and regression techniques for the rejection of ocular artifacts, Biomed. Signal Process. Control 6 (3) (2011) 291–300, https://doi.org/10.1016/j.bspc.2011.02.001.

[10] S.R. Benbadis, W.O. Tatum, Overintepretation of EEGs and Misdiagnosis of Epilepsy, J. Clin. Neurophysiol. 20 (1) (2003) 42–44, https://doi.org/10.1097/00004691-200302000-00005.

[11] W.O. Tatum, Artifact-related epilepsy, Neurology 80 (Issue 1, Supplement 1) (2013) S12–S25, https://doi.org/10.1212/WNL.0b013e3182797325.

[12] C.J. Ochoa, J. Polich, P300 and blink instructions, Clin. Neurophysiol. 111 (1) (2000) 93–98, https://doi.org/10.1016/S1388-2457(99)00209-6.

[13] R. Verleger, The instruction to refrain from blinking affects auditory P3 and N1 amplitudes, Electroencephalogr. Clin. Neurophysiol. 78 (3) (1991) 240–251, https://doi.org/10.1016/0013-4694(91)90039-7.

[14] J.J.M. Kierkels, J. Riani, J.W.M. Bergmans, G.J.M. van Boxtel, Using an Eye Tracker for Accurate Eye Movement Artifact Correction, IEEE Trans. Biomed. Eng. 54 (7) (2007), https://doi.org/10.1109/TBME.2006.889179.

[15] P. He, G. Wilson, C. Russell, Removal of ocular artifacts from electro-encephalogram by adaptive filtering, Med. Biol. Eng. Compu. 42 (3) (2004) 407–412, https://doi.org/10.1007/BF02344717.

[16] A. Schlögl, C. Keinrath, D. Zimmermann, R. Scherer, R. Leeb, G. Pfurtscheller, A fully automated correction method of EOG artifacts in EEG recordings, Clin. Neurophysiol. 118 (1) (2007) 98–104, https://doi.org/10.1016/j.clinph.2006.09.003.

[17] J. Escudero, R. Hornero, D. Abasolo, A. Fernandez, M. Lopez-Coronado, Artifact Removal in Magnetoencephalogram Background Activity With Independent Component Analysis, IEEE Trans. Biomed. Eng. 54 (11) (2007) 1965–1973, https://doi.org/10.1109/TBME.2007.894968.

[18] C.A. Joyce, I.F. Gorodnitsky, M. Kutas, Automatic removal of eye movement and blink artifacts from EEG data using blind component separation, Psychophysiology 41 (2) (2004) 313–325, https://doi.org/10.1111/psyp.2004.41.issue-210.1111/j.1469-8986.2003.00141.x.

[19] M.A. Klados, C. Bratsas, C. Frantzidis, C.L. Papadelis, P.D. Bamidis, A Kurtosis-Based Automatic System Using Naïve Bayesian Classifier to Identify ICA Components Contaminated by EOG or ECG Artifacts, XII Mediterranean Conference on Medical and Biological Engineering and Computing 49–52 (2010), https://doi.org/10.1007/978-3-642-13039-7_13.

[20] M.A. Klados, C.L. Papadelis, P.D. Bamidis, REG-ICA: A new hybrid method for EOG Artifact Rejection, in: 2009 9th International Conference on Information Technology and Applications in Biomedicine, 2009, pp. 1–4, https://doi.org/10.1109/ITAB.2009.5394295.

[21] A. Hyvärinen, J. Karhunen, E. Oja, S. Haykin (Eds.), Adaptive and Learning Systems for Signal Processing, Communications, and ControlIndependent Component Analysis, John Wiley & Sons, Inc., New York, USA, 2001.

[22] S. MAKINO Blind Source Separation of Convolutive Mixtures of Speech in Frequency Domain E88-A 7 2005 1640 1655 10.1093/ietfec/e88-a.7.1640.
[23] J.A. Palmer K. Kreutz-Delgado S. Makeig AMICA: An Adaptive Mixture of Independent Component Analyzers with Shared Components 2012.
[24] Antoni, J., & Chauhan, S. (2011). *Second Order Blind Source Separation techniques (SO-BSS) and their relation to Stochastic Subspace Identification (SSI) algorithm.*
[25] Turnip, A., Munandar, A., Redhyka, G. G., Sebleku, P., Firmanto, A. D., Saragi, T., & Tumbelaka, B. Y. (2014, August). Removing ocular artifact of EEG signal using SOBI-RO on motor imagery experiment. *2014 2nd International Conference on Technology, Informatics, Management, Engineering & Environment.* https://doi.org/10.1109/TIME-E.2014.7011636.
[26] L. Tong, V.C. Soon, Y.F. Huang, R. Liu, AMUSE: a new blind identification algorithm, IEEE International Symposium on Circuits and Systems (1990), https://doi.org/10.1109/ISCAS.1990.111981.
[27] A. Delorme, J. Palmer, J. Onton, R. Oostenveld, S. Makeig, L.M. Ward, Independent EEG Sources Are Dipolar, PLoS ONE 7 (2) (2012) e30135, https://doi.org/10.1371/journal.pone.0030135.
[28] Y. Li D. Martin W. Powers D. Powers J. Peach Comparison of blind source separation algorithms 2001 https://www.researchgate.net/publication/228605059.
[29] N. Oosugi, K. Kitajo, N. Hasegawa, Y. Nagasaka, K. Okanoya, N. Fujii, A new method for quantifying the performance of EEG blind source separation algorithms by referencing a simultaneously recorded ECoG signal, Neural Networks 93 (2017) 1–6, https://doi.org/10.1016/j.neunet.2017.01.005.
[30] J.-F. Cardoso, High-Order Contrasts for Independent Component Analysis, Neural Comput. 11 (1) (1999) 157–192, https://doi.org/10.1162/089976699300016863.
[31] M.A. Klados, P.D. Bamidis, A semi-simulated EEG/EOG dataset for the comparison of EOG artifact rejection techniques, Data in Brief 8 (2016) 1004–1006, https://doi.org/10.1016/j.dib.2016.06.032.
[32] A. Delorme, S. Makeig, EEGLAB: an open source toolbox for analysis of single-trial EEG dynamics including independent component analysis, J. Neurosci. Methods 134 (1) (2004) 9–21, https://doi.org/10.1016/j.jneumeth.2003.10.009.
[33] S.-H. Hsu, L. Pion-Tonachini, J. Palmer, M. Miyakoshi, S. Makeig, T.-P. Jung, Modeling brain dynamic state changes with adaptive mixture independent component analysis, NeuroImage 183 (2018) 47–61, https://doi.org/10.1016/j.neuroimage.2018.08.001.
[34] R. Romo-Vázquez, H. Velez-Perez, R. Ranta, V. Louis-Dorr, D. Maquin, L. Maillard, R. Romo Vázquez, H. Vélez-Pérez, R. Ranta, V. Louis Dorr, D. Maquin, L. Maillard, Blind source separation, wavelet denoising and discriminant analysis for EEG artefacts and noise cancelling, Biomed. Signal Process. Control 7 (4) (2012) 389–400, https://doi.org/10.1016/j.bspc.2011.06.005i.
[35] J. Escudero R. Hornero D. Abásolo A. Fernández Comparison of Blind Source Separation Preprocessings Applied to Magnetoencephalogram Recordings to Improve the Classification of Alzheimer's Disease Patients In International Journal of Bioelectromagnetism www.ijbem.org Vol. 12, Issue 4 2010 www.ijbem.org.
[36] M. Klemm, J. Haueisen, G. Ivanova, Independent component analysis: comparison of algorithms for the investigation of surface electrical brain activity, Med. Biol. Eng. Compu. 47 (4) (2009) 413–423, https://doi.org/10.1007/s11517-009-0452-1.
[37] U. KUMAR, V. KUMAR, J.N. KAPUR, NORMALIZED MEASURES OF ENTROPY, Int. J. Gen Syst 12 (1) (1986) 55–69, https://doi.org/10.1080/03081078608934927.
[38] I. Winkler, S. Haufe, M. Tangermann, Automatic Classification of Artifactual ICA-Components for Artifact Removal in EEG Signals, Behavioral and Brain Functions 7 (1) (2011) 30, https://doi.org/10.1186/1744-9081-7-30.