A functional source separation algorithm to enhance error-related potentials monitoring in noninvasive brain-computer interface

(Article begins on next page)

20 February 2025

A functional source separation algorithm to enhance error-related potentials monitoring in noninvasive brain-computer interface

Francesco Ferracuti[a], Valentina Casadei[f], Ilaria Marcantoni[a], Sabrina Iarlori[a], Laura Burattini[a], Andrea Monteriù[a] and
Camillo Porcaro[b,a,c,d,e]

[a] Department of Information Engineering, Università Politecnica delle Marche, Ancona, Italy

[b] Institute of Cognitive Sciences and Technologies (ISTC) - National Research Council (CNR), Rome, Italy

[c] Research Center for Motor Control and Neuroplasticity, KU Leuven, Leuven, Belgium

[d] S. Anna Institute and Research in Advanced Neurorehabilitation (RAN) Crotone, Italy

[e] Centre for Human Brain Health, School of Psychology, University of Birmingham, Birmingham, United Kingdom

[f] Department of Electrical Engineering & Electronics, University of Liverpool, Liverpool, United Kingdom

f.ferracuti@univpm.it; valentina.casadei@liverpool.ac.uk; i.marcantoni@pm.univpm.it; s.iarlori@univpm.it; l.burattini@univpm.it;
a.monteriu@staff.univpm.it; camillo.porcaro@istc.cnr.it

Abstract

Background and Objectives: An Error related Potential (ErrP) can be noninvasively and directly measured from the scalp through electroencephalography (EEG), as response, when a person realizes they are making an error during a task (as a consequence of a cognitive error performed from the user). It has been shown that ErrPs can be automatically detected with time-discrete feedback tasks, which are widely applied in the Brain-Computer Interface (BCI) field for error correction or adaptation. In this work, a semi-supervised algorithm, namely the Functional Source Separation (FSS), is proposed to estimate a spatial filter for learning the ErrPs and to enhance the evoked potentials.

Methods: EEG data recorded on six subjects were used to evaluate the proposed method based on FFS algorithm in comparison with the xDAWN algorithm. FSS- and xDAWN-based methods were compared also to the Cz and FCz single channel. Single-trial classification was considered to evaluate the performances of the approaches. (Both the approaches were evaluated on single-trial classification of EEGs.)

Results: The results presented using the Bayesian Linear Discriminant Analysis (BLDA) classifier, show that FSS (accuracy 0.92, sensitivity 0.95, specificity 0.81, F1-score 0.95) overcomes the other methods (Cz - accuracy 0.72, sensitivity 0.74, specificity 0.63, F1-score 0.74; FCz - accuracy 0.72, sensitivity 0.75, specificity 0.61, F1-score 0.75; xDAWN - accuracy 0.75, sensitivity 0.79, specificity 0.61, F1-score 0.79) in terms of single-trial classification.

Conclusions: The proposed FSS-based method increases the single-trial detection accuracy of ErrPs with respect to both single channel (Cz, FCz) and xDAWN spatial filter.

Keywords: Brain Computer Interface (BCI), Electroencephalography (EEG), Error-related Potential (ErrP), Functional Source Separation (FSS), P300, Spatial filter

*Corresponding author:

Prof. Camillo Porcaro, PhD

Address: Via Palestro 32, 00185 Rome Italy

E-mail address: camillo.porcaro@istc.cnr.it

Phone: +39 06 4436 2370 int 0

AUC: Area Under Curve; BLDA: Bayesian Linear Discriminant Analysis; BCI: Brain-Computer Interface; C: Correct trials; CSP: Common Spatial Pattern; EEG: electroencephalography; ErrP: Error-Related Potential; EA: Evoked Activity; FDR: False Discovery Rate; FSS: Functional Source Separation; ICA: Independent Component Analysis; NC: Non-Correct trials; ROC: Receiver Operating Characteristic; SNR: Signal-to-Noise Ratio; SSNR: Signal-to-Signal plus Noise Ratio.

1. Introduction

Brain-Computer Interface (BCI) is a noninvasive technology that enables communication between the user's brain and a digital device (e.g. smart wheelchairs, computers, or prosthesis), usually named agent. BCI allows the recognition of the user's intention by decoding the neural activity through electroencephalography (EEG) in order to control the agent and improve its performances. Reaching this goal implies high cognitive attention and effort, since the user is asked a considerable concentration on the stimuli provided while operating a BCI.

In literature, many works analyzed the capability of the BCI to recognize erroneous behavior of agents directly from the user's brain signals [1–7]. Errors in the recognition of the user's intention elicit potentials called evoked Error-related EEG Potentials (ErrPs). ErrPs were analyzed for the first time in 1990, in a study about choice-reaction tasks [8]; in the same work, the typical ErrPs waveform was defined. In this paradigm, the user monitors the agent's actions providing a feedback that can be used to improve the overall performance of the agent. ErrP is physiologically defined as a two components brain signal, consisting in negative and positive peaks, associated with the response monitoring and error detection processes. Both peaks originate in the anterior cingulate cortex, a frontal brain structure involved in the cognitive and affective brain processes [9]. Typically, the signal is generated within 500 ms from the erroneous agent decision, where the first component is a negative peak at almost 50-100 ms. After the negative peak, a positive peak is generated, further divided into fronto-central and centroparietal components [10].

The works [1,2] provide and explain some examples where ErrP signals are generated when a user monitors the performance of an agent, without performing a direct control. Unlike traditional BCI systems, the user does not provide continuous commands, but only monitors the agent's performance, thus making possible to tailor the agent's behavior to the user's needs and preferences [3]. Furthermore, in the experimental protocol proposed in [1], the user tried to move a cursor towards a target location (either using a keyboard or mental commands). Moreover, it showed the possibility to recognize and correct an erroneous decision of the agent exploiting the EEG signals.

The application of ErrPs in BCI technology has increased during the last years, especially for the correction of the system behavior through what is called reinforcement learning. Precisely, the most common application of ErrP was done in BCI spellers, where during the spelling of a word, a character can be discarded if wrong [4]. ErrPs can be involved as a suitable alternative or a complementary signal for BCI systems, especially as supervision or feedback signal during the execution of the task [7]. ErrPs have turned out to be used for fixing this kind of problems, as demonstrated in an experiment carried in [5], where a biofeedback based on ErrPs is applied to a closed-loop system for the behavioral correction of a robot. The feasibility in using the ErrPs combined to BCI signals for correcting the erroneous commands was investigated in [6]; in addition, the authors in [11] found out that ErrPs are also elicited as a misinterpretation of user's intent. Another research [12] proposed the classification of error related potentials from EEG during a real-world driving task. While the subject was driving, a directional cue was shown before reaching an intersection and the proposed system infers whether the cued direction coincided with the subject's intention. Other works investigated the co-adaptation of human-agent using ErrPs and decoding of ErrPs in tasks with continuous feedback [13,14].

The most important challenge of non-invasive BCI applications relates to the performances, since the control cannot offer a constant level of assistance due to the weakness and noisy of EEG signals [15–17]. Consequently, the application of spatial filters to improve the Signal-to-Noise Ratio (SNR) and the single-trial classification is worthy of investigation. Spatial filters are proposed in the literature with the aim to increase the SNR by using a weighted sum of all electrodes rather than relying on a single, or a small sub-set, of EEG channels. Some examples of spatial filters are the so-called xDAWN and the Common Spatial Pattern (CSP) [15,18]. Variants and extensions of CSP are proposed in [18–20], trying to overcome the drawbacks of CSP and improve the classification of single-trial EEG. In [21,22], the authors proposed adaptive spatial filters, the former based on ensembles of CSP patches whereas the latter based on a combination of blind source separation and regression analysis.

In this work, Functional Source Separation (FSS) has been considered to estimate a spatial filter for learning the ErrPs in BCI context. In order to enhance evoked ErrPs, FSS algorithm is designed by considering the ErrPs as a functional constraints [23–26]. A direct comparison between the FSS [23] and the xDAWN algorithm [15,27] is presented to show the capability of the spatial filters to enhance the evoked ErrPs. Moreover, a single-trial classification was reported to assess the performances of FSS with respect to xDAWN. Moreover, the FSS and xDAWN based methods are also compared with the single channels Cz and FCz, usually selected to monitor ErrPs [1], in terms of single-trial classification.

Section 1 of this paper presents the experimental protocol, the spatial filters and the Bayesian Linear Discriminant Analysis (BLDA) classification algorithm used in the study. Section 2 exposes the idea of using the FSS as spatial filter for learning the ErrPs in BCI. Experimental results are presented in Section 3 including a qualitative evaluation of the spatial filters and the single-trial classification. In Section 4, we discuss the overall results and draw conclusion. Finally, in Section 5, we highlight the future works.

2. Materials and Methods

In this section the experimental details are presented and the FSS and xDAWN algorithms, together with the BLDA classification algorithm, are described. The Fig. 1 shows the proposed procedure which consist of a training and a testing phase. A detailed description of each part is reported in the following sections.
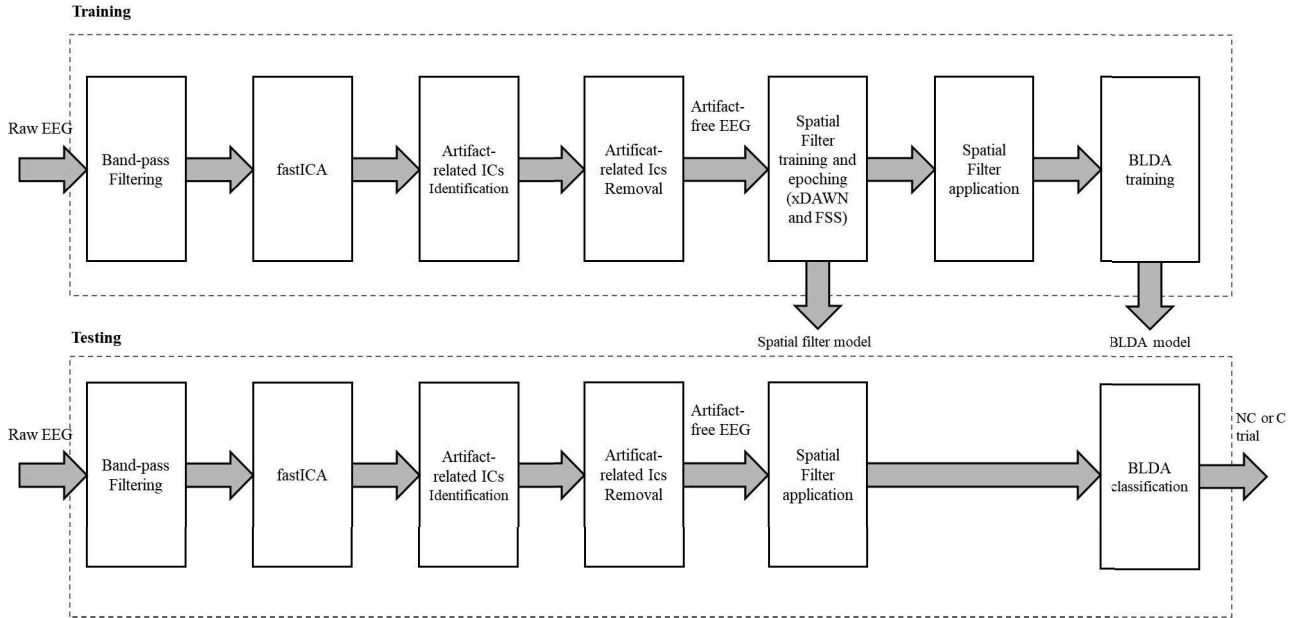

Fig. 1 Flowchart of the proposed procedure.

2.1 Experimental protocol and dataset description

The considered dataset, whose signals were acquired in the experimental protocols proposed by [1], served as an experiment on EEG ErrPs. Six subjects (mean age 27.83 ± 2.23 years standard deviation) performed two recording sessions (session 1 and session 2) separated by several weeks. Both session 1 and session 2 consisted of 10 blocks: each block was composed of approximately 50 trials and each trial was about 2000 ms long. In each trial, the user, without sending any command to the agent, only assessed whether an autonomous agent performed the task properly. In particular, the task consisted in a cursor reaching a target on a computer screen. Specifically, at the beginning of each trial the user was asked to focus on the center of the screen, while during the trial was asked to follow the movement of the cursor, knowing the goal of the task. Thus, ErrPs were elicited by monitoring the behavior of the agent.

Practically, each user was seated in front of a monitor where the visual task was displayed, defined by a green and square moving cursor and by a blue (or red) square target. The target could be placed on the left or on the right of the cursor: in the first case it was blue, while in the second red. The working area consisted of 20 horizontal locations along the center of the monitor. At each time step (i.e., at each trial) the cursor moved horizontally of discretized steps toward the target location, either the left or right.

Once the target was reached, the cursor remains fixed until a new target location is drawn. The new location will be at no more than three positions away from the current one. If the new target location is outside of the working area, it was relocated at the center of the screen. The probability for the cursor to move in the wrong direction (i.e., opposite to the target location) was set to 0.20. In the paper, the Correct trials (C) refer to a successful reaching of the final target, whereas the Non-Correct trials (NC) refer to those trials where the cursor moves in the wrong direction and then fails to reach the target. The experimental protocol is shown in Fig. 2.
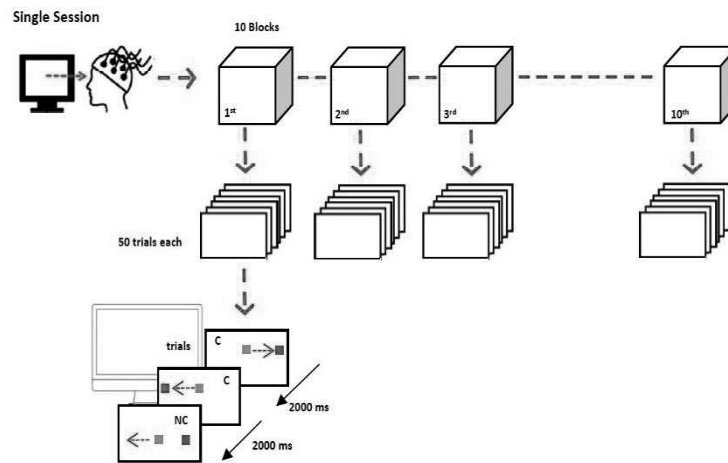

Fig. 2 Experimental Protocol - The figure shows the timing and division into sessions, blocks and trials in which the experiment is articulated. Moreover, in the scheme a representation of the experiment scenario is also integrated, where the user is in front of the screen where the different trials are projected.

During the acquisition, 64 electrodes were placed according to the standard ("International 10/20 system") and EEG data were recorded by using a Biosemi ActiveTwo system at a sampling rate of 512 Hz. The data were re-referenced to common average reference. Off-line bandpass forward–backward filtering between 1 and 10 Hz (Butterworth second order filter) was applied [1]. Since lateral eye movements could be induced during the experiments, a semiautomatic Independent Component Analysis (ICA)-based procedure was applied [28,29]. In particular, ICA-based methods, as other Blind Source Separation (BSS) methods, decompose the EEG data in as many components (sources) as the number of EEG channels but with the property that those components are statistically independent [30–33]. Following ICA application, in the context of functional Magnetic Resonance Imaging (fMRI) [33,34] and fetal magnetoencephalography [31], we applied a modified version of [28] procedure to the filtered EEG data to classify artifactual components such as eye movements, cardiac artifacts, environmental and channel noise, without rejecting the contaminated epochs. This approach is based on spectral and statistical properties of the Independent Components (ICs). The pipeline consists of three main steps: (1) ICA application to the filtered EEG data (in this specific case we used fastICA [35]; (2) automatic Power Spectrum Density (PSD) correlation between FP1/FP2 channel (instead of the electrooculogram (EOG) that was missing in those recordings) and all the ICs estimated by fastICA, ($p < 0.01$) and kurtosis together with entropy properties to detect cardiac artifact and environmental/channel noise, respectively. More specifically, for the latter indexes (kurtosis and entropy) were applied to segmented (2 s long centered on the stimulus) data. For both indexes, segment distributions were normalized with respect to all ICs (mean 0 and standard deviation 1). Thus, standard deviation thresholds respect to the mean were applied (standard deviation threshold set at ±1.64) and, whether a percentage of segments (in our case the 33%) exceeded the threshold those ICs were marked as rejected. (3) A control cycle by the user was implemented as a third step based on the 'discrepancy', i.e., the difference between the original EEG data and the reconstructed data using only the ICs marked as non-artifactual after automatic artifact detection (step 2). The goal of the control cycle was to have quick visual feedback on the automatic artifact identification quality. The control cycle is based on PSD and (Event-Related Potential) ERP visualization of the EEG raw data, cleaned data (i.e., data reconstructed with the non-artifactual components after step 2) and the discrepancy. The feedback was positive when discrepancy contained only artifacts and noise compared to EEG raw data and cleaned data. In case of negative feedback (i.e., presence of evoked activity in the ERP and/or brain rhythms in the PSD of the discrepancy), the entropy and kurtosis percentage thresholds were reduced and step (2) repeated using the new thresholds.

In this work, the session 1 of each subject is used to train the considered spatial filters and the BLDA classifier, while session 2 is used to test the algorithms.

2.2 Functional Source Separation Algorithm

The FSS [23,32,33,36–40] algorithm is part of semi-Blind Source Separation (sBSS) methods [41], which works with *a priori* information from the data (in this specific case EEG signals). In this scenario, the goal of FSS is to boost the separation of the signals of interest by using some characteristics of the original signal waveform. FSS, similarly to ICA, represents the EEG data X as a linear combination of s sources by a mixing matrix A. The cost function for the FSS (with respect to standard ICA) is defined as: $F = J + \lambda R$ where J is the ICA statistical constraint, while R represents the *a priori* information to maximize. The λ weight (between these two parameters) should be chosen to maximize the functional constraint and, simultaneously, to minimize the computational time [32]. Furthermore, the FSS cost function F is optimized through the simulated annealing [42].

In the experimental condition proposed in [1], the *a priori* information to be verified is the ERP around 300 ms activated during the target identification. Thus, we noticed that the functional source underlying the P300 processes that maximize the P300 response, is named FS_{P300}. The functional constraints were defined as follows:

$$R_{FS_{P300}} = \frac{1}{(\varDelta_1 t_k + \varDelta_2 t_k)} \sum_{t_k - \varDelta_2 t_k}^{t_k + \varDelta_2 t_k} |EA(t)| - \frac{1}{500} \sum_{t=-500}^{0} |EA(t)| \quad (1)$$

where EA represents the evoked activity obtained by averaging the EEG signal on the stimulus ($t = 0$); t_k is the time instant where the power of the EEG signal is maximum around 300 ms post-stimulus; $\varDelta_1 t_k$ ($\varDelta_2 t_k$) represents the window(s) where the signal amplitude is 50% of the maximum value before (after) t_k; the baseline correction was chosen in the range from -500 to 0 ms.

2.3 xDAWN

The xDAWN spatial filter enhances the SNR through an unsupervised estimation of the evoked subspace, so that the evoked potentials are emphasized by projecting [15,27].
Let $X \in \mathbb{R}^{N_t \times N_s}$ be the matrix of the recorded EEG signals, where N_t is the number of samples and N_s the number of channels; the target stimuli elicited by evoked potential lead to the following model:

$$X = DW + N \quad (2)$$

where N is the noise term, D is the Toeplitz matrix and W represents the synchronous response with target stimuli. The spatial filter U is designed in such a way that maximizes the Signal-to-Signal plus Noise Ratio (SSNR) of X. This concept is formulated by the generalized Rayleigh quotient:

$$\hat{U} = \underset{U}{\arg\max} \frac{trace(U^T \hat{W}^T D^T D \hat{W} U)}{trace(U^T X^T X U)} \quad (3)$$

where $\hat{W} = (D^T D)^{-1} D^T X$ and the optimization problem in (3) can be solved by a QR factorization with a singular value decomposition. In order to improve the performance of the xDAWN algorithm, we considered the use of a regularization term during the learning stage [43].
The regularization operation allows overcoming the issues related to high dimensionality, where the applicability of the spatial filter is a demanding task since, in this situation, direct high dimension matrix operations are required [43]. In order to overcome these issues, the empirical, polled covariance matrices considered in the generalized Rayleigh quotient were replaced with:

$$S = (1 - \gamma)S + \gamma \cdot \frac{trace(S)}{d} I \quad (4)$$

where γ is a hyperparameter to be set and the matrix S is both $X^T X$ and $\hat{W}^T D^T D \hat{W}$. The experimental results described in the paper refer to the xDAWN algorithm with the addition of the regularization term (γ).

The detection of the evoked ErrPs is performed by the BLDA classifier, which was used in [15] to detect the evoked related potentials spatially filtered by the xDAWN. In this paper, BLDA was adopted also to evaluate FSS-based spatial filter and the single-trial classification by using the channels FCz and Cz as well.

Among the proposed classifiers for BCIs, BLDA [44,45] was chosen since it was efficient and fully automatic (i.e., no hyperparameters to adjust).

BLDA aims to fit data $\boldsymbol{x}$ using a linear function of the form:

$$y(\boldsymbol{x},\boldsymbol{w}) = \sum_{j=1}^{M} w_j \phi_j(x) = \boldsymbol{w}^T \boldsymbol{\phi}(\boldsymbol{x}) \qquad (5)$$

where $\boldsymbol{\phi}(\boldsymbol{x})$ is the feature vector, assuming that the target variable is equal to $t = y(\boldsymbol{x},\boldsymbol{w}) + \epsilon$, where ϵ is Gaussian noise. The objective of BLDA is to minimize the function:

$$J(\boldsymbol{w}) = \frac{\beta}{2} \| t - \boldsymbol{w}^T \boldsymbol{\phi}(x) \|_2^2 + \frac{\alpha}{2} \boldsymbol{w}^T \boldsymbol{w} \qquad (6)$$

where α and β are automatically inferred from data by using a Bayesian framework.

2.5 Error-Related Potentials behavior

Once the ErrPs were identified from Cz and FCz electrodes, ERP analysis at the average and single-trial level was performed to characterize and validate the quality of the results obtained for both sessions (session 1 and session 2) and C and NC conditions using the xDAWN and FSS algorithms. In particular, ERP analysis was performed on session 1 and session 2 to evaluate the stability of ERP at different time periods and between NC and C. All the trials were windowed (-1000 to 1000 ms) and were baseline-corrected for the interval from -500 to 0 ms. Pointwise statistical analysis was performed on the ERP by two-sample permutations t-test using 5000 permutations. False Discovery Rate (FDR) was also used for multiple comparisons corrections.

3. Results

In this section, a qualitative and quantitative comparison among single channel (FCz and Cz), xDAWN and FSS spatial filter is presented. Afterwards, classification results among single channel (FCz and Cz), xDAWN and FSS are reported as well.

3.1 Error-Related Potentials on Cz and FCz

Consistently with previous studies [1,23], EEG error-related activity appears in fronto-central areas, as illustrated by the scalp topographical maps in Fig. 3. Fig. 3 also shows the grand average ERP for the C and NC conditions for Cz and FCz electrodes in both sessions (sessions 1: thick blue line and session 2: red dashed line). In particular, NC condition refers to the trials that elicited ErrPs (i.e., the cursor moves in the wrong direction) and C condition refers to the trials that did not elicit ErrPs (i.e., the cursor moves in the right direction).
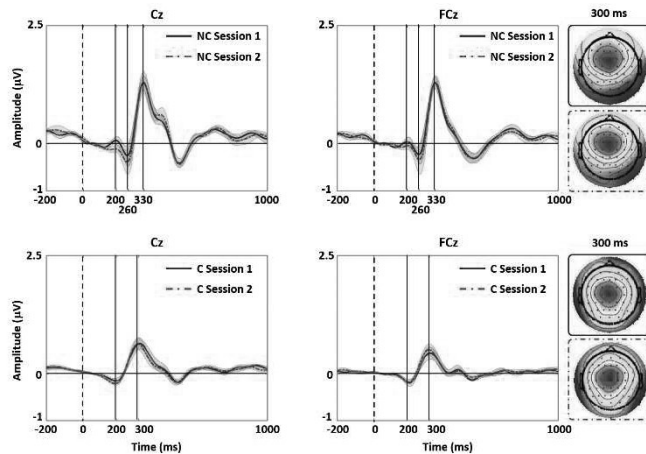
Fig. 3 ERPs and topography map session 1 vs. session 2 - Grand average event related potentials between session 1 (blue thick line) and session 2 (red dashed line) for Non-Correct trials (NC: first row) and Correct trials (C: second row). First column shows the ERPs on the Cz electrode for the two recording sessions, and second column shows the ERPs on the FCz electrode for the two recording sessions as well. Last column shows scalp topographical maps at 300 ms for session 1 (blue square thick line) and session 2 (red square dashed line). Shaded area of the same color highlights standard error. No sessions differences were observed between NC session 1 vs. NC session 2 and C session 1 vs. C session 2 (point-by-point permutation t-test at p < 0.05).

In all cases shown in Fig. 3, the waveforms are characterized by a small positive peak near 200 ms after delivery of feedback for the NC and a negative one for the C, followed by a positive peak around 330 and 300 ms for the NC and C, respectively. It should be noted that for the NC a negative deflection around 260 ms is also shown. The stability observed between session 1 and session 2 of these signals is a key issue for their use especially for BCI applications. In fact, comparison of the ERPs for the two different recording days (session 1 and session 2) shows that the signal remains stable over several weeks (see Fig. 3). In particular, the first three ERP components for the NC (i.e., negative peak at 260 ms and two positive peaks at 200 ms and 330 ms) and the first two peaks for the C (i.e., negative peak at 200 ms and the positive one at 300 ms) are stable between the two recording sessions. No significant difference was found between sessions (point-by-point permutation t-test at p < 0.05).
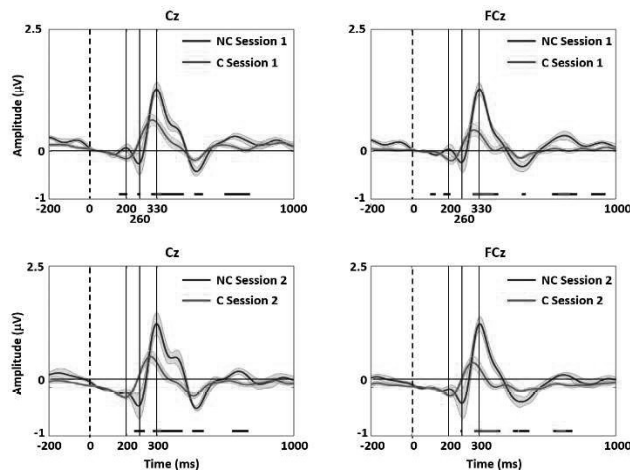

Fig. 4 ERPs Non-Correct trials (NC) vs. Correct trials (C) - Grand average event related potentials between Non-Correct trials (NC: blue line) and Correct (C: red line) for session 1 (first row) and session 2 (second row). First column shows the ERPs on the Cz electrode for NC and C, and second column shows the ERPs on the FCz electrode for the NC and C as well. Shaded area of the same color highlights standard error. Horizontal black and magenta lines, on the bottom of the figure, indicate a significant group difference between NC vs. C for session 1 (first row) and session 2 (second row). Permutation t-test at p < 0.05 (black line); pFDR < 0.05 (magenta line).

On the other hand, statistically significant differences (point-by-point permutation t-test at p<0.05) between NC and C ERPs were found in both sessions (Fig. 4).

3.2 Event-Related Potentials extracted by xDAWN and FSS

xDAWN and FSS (Fig. 5) as well as Cz and FCz electrodes show the stability between the two sessions (no significant difference was found, point-by-point permutation t-test at p < 0.05). Instead, statistical difference was observed comparing NC vs. C in both xDAWN and FSS algorithms. In particular, we can observe that only the FSS survives after the multiple comparison correction (pFDR<0.05), emphasizing that FSS at the average level is more robust in discriminating NC vs. C conditions (Fig. 6).
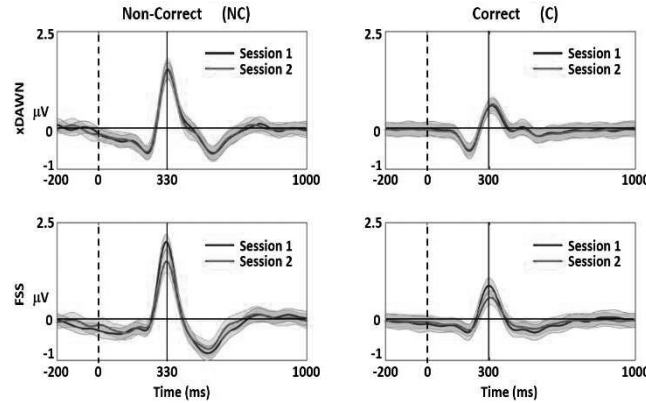

Fig. 5 ERPs session 1 vs. session 2 - Grand average event related potentials between session 1 (blue line) and session 2 (red line) for Non-Correct trials (NC: first column) and Correct trials (C: second column). First row shows the ERPs for the xDAWN for the two recording sessions and second row shows the ERPs for the FSS for the two recording sessions as well. Shaded area of the same color highlights standard error. No sessions differences were observed between NC session 1 vs. NC session 2 and C session 1 vs. C session 2 (point-by-point permutation t-test at p < 0.05).
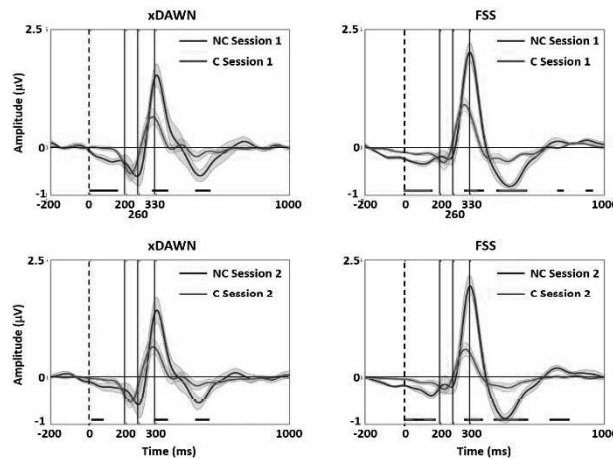

Fig. 6 ERPs Non-Correct trials (NC) vs. Correct trials (C) - Grand average event related potentials between Non-Correct trials (NC: blue line) and Correct trials (C: red line) for session 1 (first row) and session 2 (second row). First column shows the ERPs for the xDAWN for NC and C and second column shows the ERPs for the FSS for the NC and C as well. Shaded area of the same color highlights standard error. Horizontal black and magenta lines, on the bottom of the figure, indicate a significant group difference between NC vs. C for session 1 (first row) and session 2 (second row). Permutation t-test at p < 0.05 (black line); pFDR < 0.05 (magenta line).

3.3 Single-Trial Classification

The use of ErrPs in practical BCI applications requires their accurate recognition on a single-trial basis. Following previous studies, we classify the signals using a classifier based on the BLDA, as described in [15,27].

Classification analysis was performed on Cz and FCz electrodes and using advanced source extraction algorithms such as FSS [23,41] and xDAWN. In this work, for xDAWN training, we considered $\gamma=0.8$ since with this value we achieved the best classification accuracy in the training dataset. We assess single-trial classification of ErrPs using the first dataset to train the spatial filters and the BLDA classifier and the second dataset for testing. So the evaluation of classifier results, obtained on data recorded several weeks before, allows us to evaluate the feasibility of recognizing ErrP signals.

Fig. 7 shows the Receiver Operating Characteristic (ROC) curves for classification using the training data (i.e., session 1). The methods FSS, xDAWN and the single channels Cz, FCz are compared. Fig. 7 shows the classification results for each subject, and Fig. 8 shows the ROC curves for classification using the testing data (i.e., session 2). It is clear that the FSS algorithm is able to detect the elicited ErrPs better than the other methods, whereas the xDAWN achieves comparable results to the single channels FCz and Cz.

Table 1 shows in detail the classification accuracy and the Area Under Curve (AUC) for each subject and method using the training data (i.e., session 1). The results, shown in the Table 1, reveal how the FSS outperforms the other methods in terms of classification accuracy both in the case of NC and of C. Table 2 shows the classification accuracy for each subject and method using the testing data (i.e., session 2). The results of FSS, shown in Table 2, reveal that successful single-trial classification is achieved for both classes with higher detection of C (mean classification accuracy of 95% and 81% for C and NC, respectively).

Best performances are observed for subjects 1, 2, 3, 4 and 5 for whom the recordings were about seven weeks apart for subjects 1, 2 and 3 and 200 days apart for subject 4 and finally 600 days apart for subject 5. In addition, it must be noticed that reasonably good performances are also achieved for subject 6, whose recordings were around 650 days apart. The comparison of the ERPs for the two different recording days shows that classification accuracy remains stable over a long time.

Finally, Table 3 shows the overall performances in terms of accuracy, sensitivity, specificity and F1-score, revealing that the FSS-based spatial filter overcomes the other methods by about 20% in all indexes.

It is worth noting that the average classification accuracy reached by the FSS algorithm outperforms the classification results obtained in [1].

Fig. 7 ROC for classification by BLDA, training data (i.e., session 1) and single channel. The single channels are FCz, Cz and the best channel of FSS and xDAWN. (a) ROC for classification of ErrPs elicited by subject 1, (b) ROC for classification of ErrPs elicited by subject 2, (c) ROC for classification of ErrPs elicited by subject 3, (d) ROC for classification of ErrPs elicited by subject 4, (e) ROC for classification of ErrPs elicited by subject 5, (f) ROC for classification of ErrPs elicited by subject 6.
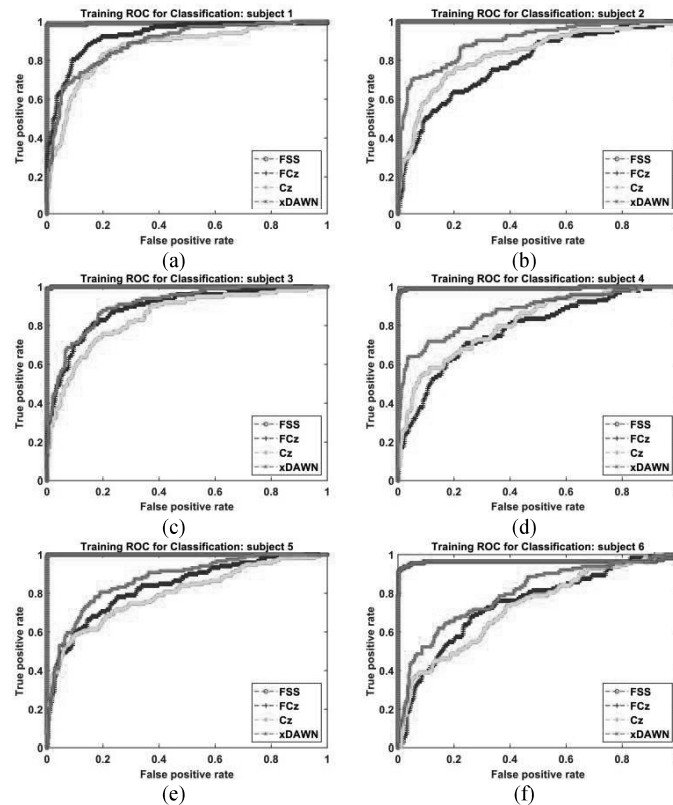
Fig. 8 ROC for classification by BLDA, testing data (i.e., session 2) and single channel. The single channels are FCz, Cz and the best channel of FSS and xDAWN. (a) ROC for classification of ErrPs elicited by subject 1, (b) ROC for classification of ErrPs elicited by subject 2, (c) ROC for classification of ErrPs elicited by subject 3, (d) ROC for classification of ErrPs elicited by subject 4, (e) ROC for classification of ErrPs elicited by subject 5, (f) ROC for classification of ErrPs elicited by subject 6.
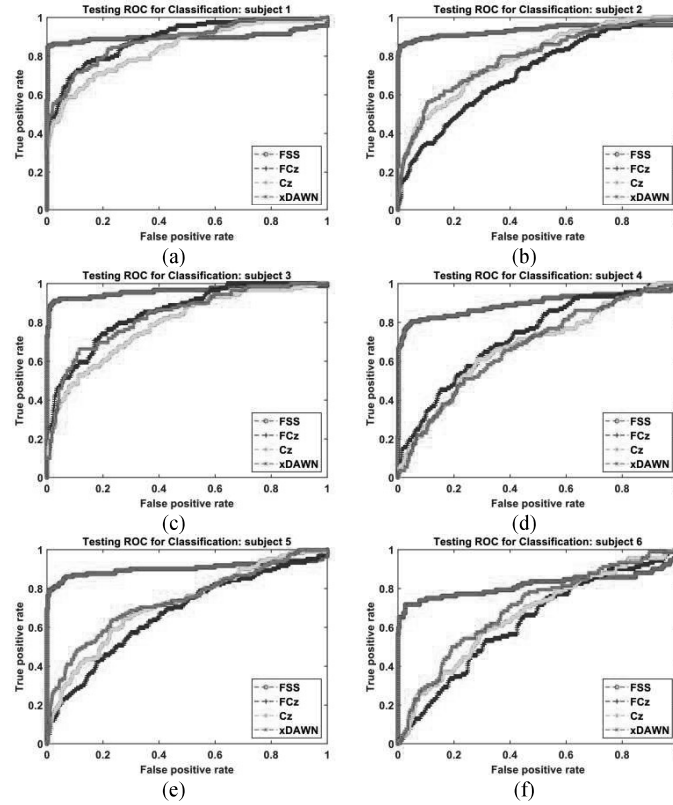
Table 1
Classification accuracy and area under curve of training data (session 1)

Subjects	FCz	Cz	xDAWN	FSS
S1 NC	0.88	0.83	0.78	0.98
S1 C	0.85	0.80	0.83	1.00
S1 AUC	0.93	0.86	0.89	0.99
S2 NC	0.64	0.76	0.87	1.00
S2 C	0.80	0.79	0.77	1.00
S2 AUC	0.78	0.83	0.90	1.00
S3 NC	0.81	0.76	0.86	0.99
S3 C	0.84	0.80	0.81	1.00
S3 AUC	0.89	0.85	0.91	1.00
S4 NC	0.71	0.72	0.72	0.98
S4 C	0.76	0.73	0.89	0.99
S4 AUC	0.78	0.81	0.87	0.99
S5 NC	0.78	0.71	0.81	0.99
S5 C	0.75	0.75	0.81	1.00
S5 AUC	0.83	0.80	0.87	1.00
S6 NC	0.68	0.73	0.67	0.96
S6 C	0.74	0.61	0.80	0.95
S6 AUC	0.75	0.73	0.80	0.96
AVG NC	0.75±0.09	0.75±0.04	0.78±0.08	**0.98±0.02**
AVG C	0.79±0.05	0.75±0.07	0.82±0.04	**0.99±0.02**

Table 2
Classification accuracy and area under curve of testing data (session 2)

Subjects	FCz	Cz	xDAWN	FSS
S1 NC	0.77	0.65	0.79	0.87
S1 C	0.85	0.86	0.82	0.90
S1 AUC	0.89	0.84	0.88	0.90
S2 NC	0.53	0.61	0.67	0.81
S2 C	0.77	0.78	0.76	1.00
S2 AUC	0.71	0.79	0.79	0.93
S3 NC	0.81	0.69	0.75	0.84
S3 C	0.71	0.72	0.73	0.99
S3 AUC	0.85	0.79	0.84	0.96
S4 NC	0.63	0.61	0.51	0.80
S4 C	0.70	0.71	0.75	0.95
S4 AUC	0.73	0.68	0.67	0.89
S5 NC	0.57	0.61	0.52	0.79
S5 C	0.69	0.73	0.85	0.99
S5 AUC	0.67	0.72	0.74	0.90
S6 NC	0.37	0.61	0.41	0.76
S6 C	0.76	0.63	0.84	0.84
S6 AUC	0.63	0.66	0.70	0.82
AVG NC	0.61±0.16	0.63±0.03	0.61±0.15	**0.81±0.04**
AVG C	0.75±0.06	0.74±0.08	0.79±0.05	**0.95±0.06**

Table 3
Overall performances

		Accuracy	Sensitivity	Specificity	F1-score
Training	FCz	0.78	0.79	0.75	0.79
	Cz	0.75	0.75	0.75	0.75
	xDAWN	0.81	0.82	0.78	0.82
	FSS	**0.99**	**0.99**	**0.98**	**0.99**
Testing	FCz	0.72	0.75	0.61	0.75
	Cz	0.72	0.74	0.63	0.74
	xDAWN	0.75	0.79	0.61	0.79
	FSS	**0.92**	**0.95**	**0.81**	**0.95**

4. Discussion

In this study, a semi-supervised algorithm named FSS was proposed to improve the SNR at the single-trial level with the aim to better classify ErrPs, as response, when a person realizes they are making an error during a task. Accurate classification of those responses become crucially important in a BCI framework. The data used in this study, were presented in [1] and the results obtained by FSS were compared with those obtained by xDAWN [15,27] and single channels Cz and FCz [1,23], which are usually selected for monitoring ErrPs. Using BLDA, we found that the higher ErrPs classification was obtained by FSS with 92% of accuracy respect to the other methods that obtained lower accuracy (75% for xDAWN and 72% for Cz/FCz).

The ErrPs morphology achieved by FSS was consistent with the results obtained from single channels Cz and FCz [1,23] and other spatial filter approach such as xDAWN [15,27] consisting of a small positive peak near 200 ms after delivering the feedback, followed by a negative deflection around 260 ms and another positive peak around 300 ms (Fig. 3 and Fig. 5). A strong ErrPs stability between session 1 and session 2 for all the methods tested was observed as well. Stability of the ErrPs among different days is essential for BCI applications [1,23]. In particular, comparison of the ErrPs for two different sessions (session 1 and session 2) shows that the signal remains stable over several weeks (Fig. 3 and Fig. 5). Differences, instead, were observed between Non-Correct trials vs. Correct trials in both xDAWN and FSS algorithms and channels Cz and FCz. Stronger difference between C and NC was observed for channel FCz and FSS algorithm with C vs NC difference surviving multiple comparison correction (pFDR<0.05) emphasizing that FCz channel and FSS, at the average level, were more robust in discriminating C vs. NC conditions respect to xDAWN and CZ electrode (Fig. 4 and Fig. 6).

However, the use of ErrPs in practical BCI applications requires their accurate recognition and classification at the single-trial level [15,27]. Classification results on the tested data (session 2) for each subjects, using BLDA and session

1 as training data, were reported in Table 2, while Table 3 reported the results as an overall performance. In both cases FSS-based spatial filter provides a better single-trial.

Despite the obtained results are without a doubt in favor of FSS approach, those results have been still tested off-line. On the other hand, the results achieved by FSS are promising since were robust respect to the single subject performance. Instead, other studies have already highlighted how individual differences may affect the overall performance of the ErrPs mechanism [46–48]. An important point for EEG-based BCI application is the number of artifacts that are mixed in the EEG recordings. This aspect is not trivial especially for the algorithms that use all the EEG information for building the spatial filter, which might be corrupted by those artifacts. Fortunately, nowadays, different online algorithms for BCI artifact removal are available (e.g., [22,49,50]). The availability of these kind of algorithms for BCI, strongly increases the efficacy of the spatial filter algorithms and importantly decreases the computational time to estimate it.

Using single channel approach might be, perhaps, a solution to avoid computational time. However, single channel solution has problems in all the cases the task under investigation involves a long range and distributed brain network [3,23,24]. In fact, although the noninvasive nature of electrophysiological techniques, such as EEG, provides the ability to directly measure large-scale neuronal activity, much still needs to be done in depicting it. Specifically, when an electrical potential is generated by a neuronal group, its activity is not only recorded from the electrode closest to this source but also from distant ones, due to the electric field propagation phenomenon. Consequently, each channel on the scalp derives its signal from more than one source [23,41,51,52]. This problem worsens with the increasing of the number of sources activated at the same time. Since the P300 definitely derives its activities from a broadly distributed network [53–56], picking a single channel or an average of channels based on a topographic map could be misleading, especially if the aim is to describe the whole neuronal communication system. In this regard, methods able to extract the under investigated neural source by combining the activity from all the electrodes (such as FSS and xDAWN), are suitable for overcoming possible misleading results by avoiding the choice of a single electrode or averaging group of electrodes.

In this respect using ErrPs EEG dataset proposed by [1], as a benchmark, we tested channels approach (Cz, FCz) and spatial filter approach such as xDAWN vs. FSS. In conclusion, the results obtained by FSS respect to the other methods makes FSS a cutting-age tool for BCI.

5. Future Works

As future works it could be possible to put the algorithm online for the testing phase after an offline training phase, removing BCI signal artifacts and allowing the possibility to online detect the ErrPs. In this way the BCI system can be employed as an interface in an online task, during which a person realizes they are making an error as a consequence of a cognitive mistake.

Conflict of interest statement

All authors have no financial and personal relationships with other people or organizations that could inappropriately bias the work.

References

[1] R. Chavarriaga, J.D.R. Millán, Learning from EEG error-related potentials in noninvasive brain-computer interfaces, IEEE Trans. Neural Syst. Rehabil. Eng. 18 (2010) 381–388. doi:10.1109/TNSRE.2010.2053387.

[2] P.W. Ferrez, J.D.R. Millán, You are wrong! - Automatic detection of interaction errors from brain waves, in: IJCAI Int. Jt. Conf. Artif. Intell., 2005: pp. 1413–1418.

[3] R. Chavarriaga, A. Sobolewski, J. del R. Millán, Errare machinale est: The use of error-related potentials in brain-machine interfaces, Front. Neurosci. (2014). doi:10.3389/fnins.2014.00208.

[4] B. Dal Seno, M. Matteucci, L. Mainardi, Online detection of P300 and error potentials in a BCI speller, Comput. Intell. Neurosci. 2010 (2010). doi:10.1155/2010/307254.

[5] A.F. Salazar-Gomez, J. Delpreto, S. Gil, F.H. Guenther, D. Rus, Correcting robot mistakes in real time using EEG signals, in: Proc. - IEEE Int. Conf. Robot. Autom., 2017: pp. 6570–6577. doi:10.1109/ICRA.2017.7989777.

[6] X. Artusi, I.K. Niazi, M.F. Lucas, D. Farina, Performance of a simulated adaptive BCI based on experimental classification of movement-related and error potentials, IEEE J. Emerg. Sel. Top. Circuits Syst. 1 (2011) 480–488. doi:10.1109/JETCAS.2011.2177920.

[7] I. Iturrate, L. Montesano, J. Minguez, Shared-control brain-computer interface for a two dimensional reaching task using EEG error-related potentials, in: Proc. Annu. Int. Conf. IEEE Eng. Med. Biol. Soc. EMBS, 2013: pp. 5258–5262. doi:10.1109/EMBC.2013.6610735.

[8] M. Falkenstein, J. Hohnsbein, J. Hoormann, L. Blanke, Effects of errors in choice reactiontasks on the ERP under focused and divided attention, Brunia, C.H.M., Gaillard, A.W.K., Kok,A. (Eds.), Psychophysiological Brain Res. Tilbg. Univesity

Press. Tilbg. (1990) 192 − 195.

[9] G. Hajcak, N. McDonald, R.F. Simons, To err is autonomic: Error-related brain potentials, ANS activity, and post-error compensatory behavior, in: Psychophysiology, 2003. doi:10.1111/1469-8986.00107.

[10] M. Spüler, C. Niethammer, Error-related potentials during continuous feedback: using EEG to detect errors of different type and severity, Front. Hum. Neurosci. 9 (2015) 1–10. doi:10.3389/fnhum.2015.00155.

[11] P.W. Ferrez, J. Del R. Millán, Error-related EEG potentials generated during simulated brain-computer interaction, IEEE Trans. Biomed. Eng. 55 (2008) 923–929. doi:10.1109/TBME.2007.908083.

[12] R. Chavarriaga, Z. Khaliliardali, L. Gheorghe, I. Iturrate, J.D.R. Millán, EEG-based decoding of error-related brain activity in a real-world driving task, J. Neural Eng. 12 (2015). doi:10.1088/1741-2560/12/6/066028.

[13] S.K. Ehrlich, G. Cheng, Human-agent co-adaptation using error-related potentials, J. Neural Eng. 15 (2018). doi:10.1088/1741-2552/aae069.

[14] C. Lopes Dias, A.I. Sburlea, G.R. Müller-Putz, Masked and unmasked error-related potentials during continuous control and feedback, J. Neural Eng. 15 (2018). doi:10.1088/1741-2552/aab806.

[15] B. Rivet, A. Souloumiac, V. Attina, G. Gibert, xDAWN Algorithm to Enhance Evoked Potentials: Application to Brain-Computer Interface, IEEE Trans. Biomed. Eng. 56 (2009) 2035–2043. doi:10.1109/TBME.2009.2012869.

[16] G. Di Pino, C. Porcaro, M. Tombini, G. Assenza, G. Pellegrino, F. Tecchio, P.M. Rossini, A neurally-interfaced hand prosthesis tuned inter-hemispheric communication, Restor. Neurol. Neurosci. 30 (2012) 407–418. doi:10.3233/RNN-2012-120224.

[17] M. Tombini, J. Rigosa, F. Zappasodi, C. Porcaro, L. Citi, J. Carpaneto, P.M. Rossini, S. Micera, Combined analysis of cortical (EEG) and nerve stump signals improves robotic hand control, Neurorehabil. Neural Repair. 26 (2012) 275–281. doi:10.1177/1545968311408919.

[18] O. Falzon, K.P. Camilleri, J. Muscat, The analytic common spatial patterns method for EEG-based BCI data, J. Neural Eng. 9 (2012). doi:10.1088/1741-2560/9/4/045009.

[19] K. Yu, K. Shen, S. Shao, W.C. Ng, X. Li, Bilinear common spatial pattern for single-trial ERP-based rapid serial visual presentation triage, J. Neural Eng. 9 (2012). doi:10.1088/1741-2560/9/4/046013.

[20] C. Sannelli, C. Vidaurre, K.-R. Müller, B. Blankertz, CSP patches: An ensemble of optimized spatial filters. An evaluation study, J. Neural Eng. 8 (2011). doi:10.1088/1741-2560/8/2/025012.

[21] C. Sannelli, C. Vidaurre, K.-R. Müller, B. Blankertz, Ensembles of adaptive spatial filters increase BCI performance: An online evaluation, J. Neural Eng. 13 (2016). doi:10.1088/1741-2560/13/4/046003.

[22] R. Guarnieri, M. Marino, F. Barban, M. Ganzetti, D. Mantini, Online EEG artifact removal for BCI applications by adaptive spatial filtering, J. Neural Eng. 15 (2018). doi:10.1088/1741-2552/aacfdf.

[23] C. Porcaro, J. Henk, D. Mantini, I.H. Robertson, N. Wenderoth, NeuroImage P3b amplitude as a signature of cognitive decline in the older population : An EEG study enhanced by Functional Source Separation, Neuroimage. 184 (2019) 535–546. doi:10.1016/j.neuroimage.2018.09.057.

[24] S. Migliore, G. Curcio, C. Porcaro, C. Cottone, I. Simonelli, G. D'aurizio, D. Landi, M.G. Palmieri, A. Ghazaryan, F. Squitieri, M.M. Filippi, F. Vernieri, Emotional processing in RRMS patients: Dissociation between behavioural and neurophysiological response, Mult. Scler. Relat. Disord. 27 (2019) 344–349. doi:10.1016/j.msard.2018.11.019.

[25] D. Gutiérrez, D.I. Escalona-Vargas, EEG data classification through signal spatial redistribution and optimized linear discriminants, Comput. Methods Programs Biomed. (2010). doi:10.1016/j.cmpb.2009.05.005.

[26] M. Mohseni, V. Shalchyan, M. Jochumsen, I.K. Niazi, Upper Limb Complex Movements Decoding From Pre-Movement EEG Signals Using Wavelet Common Spatial Patterns, Comput. Methods Programs Biomed. 183 (2020) 105076. doi:10.1016/j.cmpb.2019.105076.

[27] B. Rivet, H. Cecotti, A. Souloumiac, E. Maby, J. Mattout, Theoretical analysis of XDAWN algorithm: Application to an efficient sensor selection in a P300 BCI, in: Eur. Signal Process. Conf., 2011: pp. 1382–1386.

[28] G. Barbati, C. Porcaro, F. Zappasodi, P.M. Rossini, F. Tecchio, Optimization of an independent component analysis approach for artifact identification and removal in magnetoencephalographic signals, Clin. Neurophysiol. 115 (2004) 1220–1232. doi:10.1016/j.clinph.2003.12.015.

[29] C. Porcaro, M.T. Medaglia, A. Krott, Removing speech artifacts from electroencephalographic recordings during overt picture naming, Neuroimage. 105 (2015) 171–180. doi:10.1016/j.neuroimage.2014.10.049.

[30] M.T. Medaglia, F. Tecchio, S. Seri, G. Di Lorenzo, P.M. Rossini, C. Porcaro, Contradiction in universal and particular reasoning, Hum. Brain Mapp. 47 (2009) S111. doi:10.1002/hbm.20838.

[31] C. Porcaro, F. Zappasodi, G. Barbati, C. Salustri, V. Pizzella, P.M. Rossini, F. Tecchio, Fetal auditory responses to external sounds and mother's heart beat: Detection improved by Independent Component Analysis, Brain Res. 1101 (2006) 51–58. doi:10.1016/j.brainres.2006.04.134.

[32] C. Porcaro, G. Barbati, F. Zappasodi, P.M. Rossini, F. Tecchio, Hand sensory-motor cortical network assessed by functional source separation, Hum. Brain Mapp. 29 (2008) 70–81. doi:10.1002/hbm.20367.

[33] C. Porcaro, D. Ostwald, A.P. Bagshaw, Functional source separation improves the quality of single trial visual evoked potentials recorded during concurrent EEG-fMRI, Neuroimage. 50 (2010) 112–123. doi:10.1016/j.neuroimage.2009.12.002.

[34] C.F. Beckmann, S.M. Smith, Probabilistic Independent Component Analysis for Functional Magnetic Resonance Imaging, IEEE Trans. Med. Imaging. 23 (2004) 137–152. doi:10.1109/TMI.2003.822821.

[35] A. Hyvärinen, Fast and robust fixed-point algorithms for independent component analysis, IEEE Trans. Neural Networks. 10 (1999) 626–634. doi:10.1109/72.761722.

[36] G. Barbati, C. Porcaro, A. Hadjipapas, P. Adjamian, V. Pizzella, G.L. Romani, S. Seri, F. Tecchio, G.R. Barnes, Functional source separation applied to induced visual gamma activity, Hum. Brain Mapp. 29 (2008) 131–141. doi:10.1002/hbm.20375.

[37] G. Barbati, R. Sigismondi, F. Zappasodi, C. Porcaro, S. Graziadio, G. Valente, M. Balsi, P.M. Rossini, F. Tecchio, Functional source separation from magnetoencephalographic signals, Hum. Brain Mapp. 27 (2006) 925–934. doi:10.1002/hbm.20232.

[38] C. Porcaro, G. Coppola, G. Di Lorenzo, F. Zappasodi, A. Siracusano, F. Pierelli, P.M. Rossini, F. Tecchio, S. Seri, Hand

somatosensory subcortical and cortical sources assessed by functional source separation: An EEG study, Hum. Brain Mapp. 30 (2009) 660–674. doi:10.1002/hbm.20533.

[39] C. Porcaro, D. Ostwald, A. Hadjipapas, G.R. Barnes, A.P. Bagshaw, The relationship between the visual evoked potential and the gamma band investigated by blind and semi-blind methods, Neuroimage. 56 (2011) 1059–1071. doi:10.1016/j.neuroimage.2011.03.008.

[40] F. Tecchio, C. Porcaro, G. Barbati, F. Zappasodi, Functional source separation and hand cortical representation for a brain-computer interface feature extraction, J. Physiol. (2007) 703–721. doi:10.1113/jphysiol.2007.129163.

[41] C. Porcaro, F. Tecchio, Semi-blind Functional Source Separation Algorithm from Non-invasive Electrophysiology to Neuroimaging, in: G.R. Naik, W. Wang (Eds.), Blind Source Sep. Adv. Theory, Algorithms Appl., Springer Berlin Heidelberg, 2014: pp. 521–551. doi:10.1007/978-3-642-55016-4_19.

[42] S. Kirkpatrick, C.D. Gelatt, M.P. Vecchi, Optimization by simulated annealing, Science (80-.). 220 (1983) 671–680. doi:10.1126/science.220.4598.671.

[43] Y. Guo, T. Hastie, R. Tibshirani, Regularized linear discriminant analysis and its application in microarrays, Biostatistics. 8 (2007) 86–100. doi:10.1093/biostatistics/kxj035.

[44] U. Hoffmann, J.M. Vesin, T. Ebrahimi, K. Diserens, An efficient P300-based brain-computer interface for disabled subjects, J. Neurosci. Methods. 167 (2008) 115–125. doi:10.1016/j.jneumeth.2007.03.005.

[45] D.J.C. Mackay, N. Systems, Bayesian Interpolation, Neural Comput. 4 (1992) 415–447. doi:10.1162/neco.1992.4.3.415.

[46] P. Margaux, M. Emmanuel, D. Sébastien, B. Olivier, M. Jérémie, Objective and subjective evaluation of online error correction during P300-based spelling, Adv. Human-Computer Interact. 2012 (2012). doi:10.1155/2012/578295.

[47] N.M. Schmidt, B. Blankertz, M.S. Treder, Online detection of error-related potentials boosts the performance of mental typewriters, BMC Neurosci. 13 (2012). doi:10.1186/1471-2202-13-19.

[48] M. Spüler, M. Bensch, S. Kleih, W. Rosenstiel, M. Bogdan, A. Kübler, Online use of error-related potentials in healthy users and people with severe motor impairment increases performance of a P300-BCI, Clin. Neurophysiol. 123 (2012) 1328–1337. doi:10.1016/j.clinph.2011.11.082.

[49] W. Du Chang, J.H. Lim, C.H. Im, An unsupervised eye blink artifact detection method for real-time electroencephalogram processing, Physiol. Meas. 37 (2016) 401–417. doi:10.1088/0967-3334/37/3/401.

[50] Y. Zou, V. Nathan, R. Jafari, Automatic identification of artifact-Related independent components for artifact removal in EEG recordings, IEEE J. Biomed. Heal. Informatics. 20 (2016) 73–81. doi:10.1109/JBHI.2014.2370646.

[51] M. Rusiniak, M. Lewandowska, T. Wolak, A. Pluta, R. Milner, M. Ganc, A. Włodarczyk, A. Senderski, L. Śliwa, H. Skarżyński, A modified oddball paradigm for investigation of neural correlates of attention: A simultaneous ERP-fMRI study, Magn. Reson. Mater. Physics, Biol. Med. 26 (2013) 511–526. doi:10.1007/s10334-013-0374-7.

[52] M. Siegel, T.H. Donner, A.K. Engel, Spectral fingerprints of large-scale neuronal interactions., Nat. Rev. Neurosci. 13 (2012) 121–134. doi:10.1038/nrn3137.

[53] A.M. Fjell, K.B. Walhovd, B. Fischl, I. Reinvang, Cognitive function, P3a/P3b brain potentials, and cortical thickness in aging, Hum. Brain Mapp. 28 (2007) 1098–1116. doi:10.1002/hbm.20335.

[54] D.E.J. Linden, The P300: Where in the brain is it produced and what does it tell us?, Neuroscientist. (2005) 563–576. doi:10.1177/1073858405280524.

[55] F. Barcelo, Detection of change: event-related potential and fMRI findings, Clin. Neurophysiol. 115 (2004) 1712–1713. doi:10.1016/j.clinph.2004.02.002.

[56] J. Polich, Updating P300: An integrative theory of P3a and P3b, Clin. Neurophysiol. (2007) 2128–2148. doi:10.1016/j.clinph.2007.04.019.