

Non-linear optimized spatial filter for single-trial identification of movement related cortical potential

Original

Non-linear optimized spatial filter for single-trial identification of movement related cortical potential / Mascolini, A.; Niazi, I. K.; Mesin, L.. - In: BIOCYBERNETICS AND BIOMEDICAL ENGINEERING. - ISSN 0208-5216. - 42:1(2022), pp. 426-436. [10.1016/j.bbe.2022.02.013]

Availability:

This version is available at: 11583/2959492 since: 2022-03-28T15:01:28Z

Publisher:

Elsevier B.V.

Published

DOI:10.1016/j.bbe.2022.02.013

Terms of use:

This article is made available under terms and conditions as specified in the corresponding bibliographic description in the repository

Publisher copyright

Elsevier preprint/submitted version

Preprint (submitted version) of an article published in BIOCYBERNETICS AND BIOMEDICAL ENGINEERING © 2022, <http://doi.org/10.1016/j.bbe.2022.02.013>

(Article begins on next page)

Non-linear Optimized Spatial Filter for Single-Trial Identification of Movement Related Cortical Potential

Abstract

To investigate the optimal filter settings for pre-processing of Movement Related Cortical Potentials (MRCP) for the detection through EEG in single trial, we have proposed a novel Non-Linear Optimized Spatial Filter (NL-OSF) and compared it to the Optimized Spatial Filtering (OSF) used in literature. MRCPs from EEG recordings are emphasized, calculating the optimal non-linear combination of channels which isolates the signal of interest. The method is applied to EEG data recorded from 16 healthy patients either executing or imagining 50 self-paced upper limb movement (palmar grasp). NL-OSF had average true positive rates of about $92\pm1\%$ and $82\pm4\%$ (mean \pm std) in motor execution and imagination, respectively, which are significantly better than those of OSF applied to the same dataset. The proposed method can be potentially used for online BCI system design for neuro-rehabilitation purposes.

Keywords: Surface EEG, Brain computer interface, Spatial filters

1. Introduction

The Movement Related Cortical Potential (MRCP) is a low frequency negative shift in the EEG signal appearing around 2 seconds before a planned or executed voluntary movement [1][2]. Its detection can be instrumental in the development of Brain Computer Interfaces (BCI) which allow communication of patients who are otherwise unable, as well as in the neurorehabilitation of people with motor impairments [3]. An improvement in accuracy of the detectors could lead to a significant advancement in the field of neuroprosthetics [4].

BCIs are a relatively recent subject of research, with the first paper on the topic

published in 1973 [5]. The term BCI encompasses multiple types of techniques to allow machine-brain communication, which are helpful for patients with conditions which do not allow them to communicate with the external world, such as locked-in syndrome [6], amyotrophic lateral sclerosis [7] and cerebral palsy [8]. This kind of assistive technology gives these patients the ability to communicate, providing a significant improvement of their quality of life [9].

Nonetheless current BCIs have many challenges, such as providing precise biofeedback to the user: lack of touch, pressure, muscle lengthening and proprioception render the feedback poorly effective [10]. Indeed, the subject can usually only use sight to understand the difference between the desired action and the actual BCI output. Another important issue is latency: if the delay between the action and its feedback is too long, the ability of the patient to learn and improve the effective control of the BCI can be severely affected [11].

Different approaches have been explored in the literature of BCI systems, e.g., event-related potentials like P300 [12], steady-state visual evoked potentials (SSVEP, [13]), low frequency asynchronous switch design [14]. Here we focus on the detector performance of MRCP [1][2] (see an example in Figure 1). This EEG potential can be seen before a planned voluntary movement, both when it is executed and when it is simply imagined [1]. Moreover, the MRCP is found even if the patient is not physically capable of performing the movement, rendering its detection a good candidate for a BCI application [15].

MRCPs have been studied for decades [1]. Research in the field has shown that their size and delay are adjusted according to the participants' mental state and characteristics of the executed movement, such as speed, accuracy and frequency. Moreover, these potentials contain important information, including the intended limb, grasp force, speed and direction of the movement [16].

Efforts have been devoted to developing systems for single trial MRCP detection for application in BCIs [17]. These attempts have been hindered by what is a common issue in BCIs, i.e., the signal to noise ratio (SNR), which is very

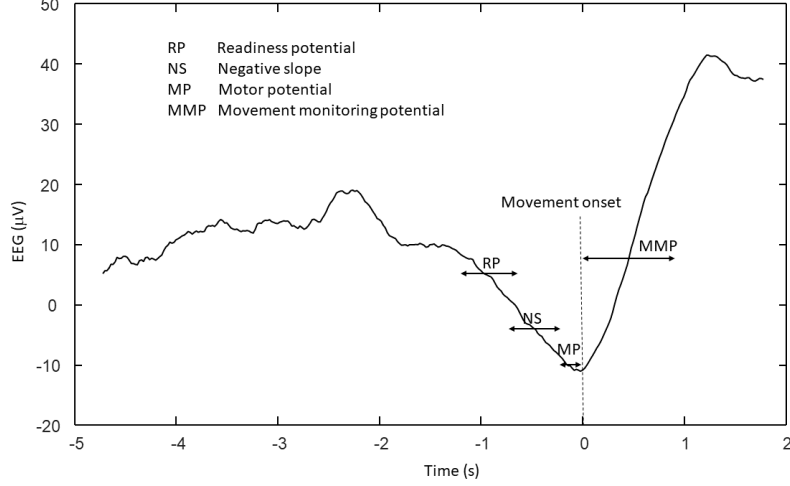


Figure 1: MRCPs of a healthy subject (participant number 1) in the case of motor execution. The wave was obtained by an average of 50 large Laplacian spatial filtered EEG trials.

low (like most endogenous brain potentials recorded through the EEG). This reduces the accuracy of detection methods [18]. However, clinical studies have shown that participants can learn how to control and amplify MRCPs through training [19][20]. Individually calibrating endogenous BCIs has been postulated to be the solution to these problems [21]. Another important issue is the need of performing MRCP identification in order to give the user the impression to control the BCI in real-time [22].

In this paper, an innovative technique is proposed to identify the MRCP. It is based on the estimation of an optimal non-linear combination of channels which isolates the waveform of interest, resulting in better performance for the MRCP based detector compared to previously proposed methods.

2. Materials and Methods

In the following sections, the data collection will be outlined as well as the analysis used in the current study.

53 2.1. Experimental data

54 2.1.1. Subjects

55 Sixteen healthy subjects aged 28 ± 12 years, 4 men and 12 women, with no
56 history of neurological diseases, participated in the experiment. All subjects
57 gave their written informed consent. All procedures were approved by the local
58 ethical committee (number 20130081).

59 2.1.2. Experimental setup

60 The subjects were placed in a chair in front of the computer with a hand force
61 transducer (Noraxon USA, Scottsdale, AZ) in the right hand. They performed
62 maximum voluntary contraction (MVC) three times and the highest value was
63 retained. Then, grasp trials were executed. A feedback was given to the partici-
64 pants to perform the grasp at 60% MVC force level during this motor execution
65 task. The force data was sampled at 2000 Hz. All participants performed 50
66 trials of both motor execution and motor imagination of palmer grasp. Each
67 movement type was performed 2×25 times with a 2-3-minute break after the
68 25th movement. The movements were performed in blocks; the order was ran-
69 domized. The subjects were visually cued (see Figure 2) by a custom-made
70 program (Aalborg University), and the produced force was recorded and used
71 as input, so the subjects had continuous visual feedback. For the tasks where
72 the movements were executed, the force was used to determine the movement
73 onset. This was defined as the instant where all values in a 200-ms wide moving
74 time window were above the baseline. The baseline was calculated from the
75 recordings during the rest phase. All onsets were visually inspected.

76 2.1.3. EEG Recording

77 Continuous 9 channel monopolar (Ag/AgCl ring electrodes) EEG (EEG Am-
78 plifiers, Nuamps Express, Neuroscan) was recorded from the following channels
79 (according to the International 10-20 system): F3, Fz, F4, C3, Cz, C4, P3, Pz
80 and P4. The signals were referenced to the right ear lobe and grounded at na-
81 sion. Electrooculography (EOG) was recorded from FP1. The EEG and EOG

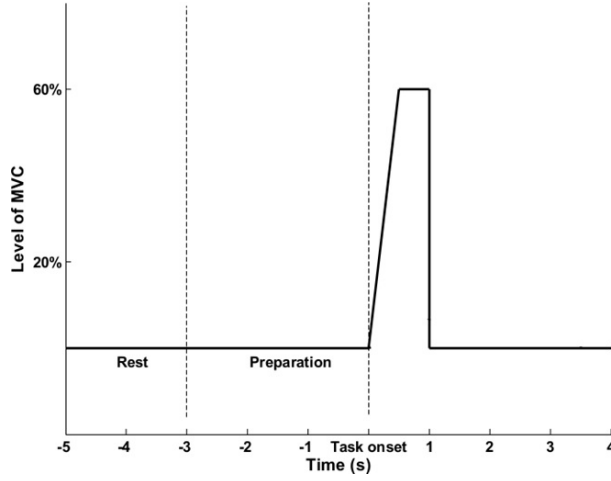


Figure 2: Visual cue provided to the participants.

82 were sampled at 500 Hz and converted with 32-bit precision. The impedance of
 83 all electrodes was below 5 k Ω . During the recordings, the subjects were asked
 84 to minimize eye blinks and facial and body movements. A digital trigger was
 85 sent from the visual cueing program to the EEG amplifier at the beginning of
 86 each trial.

87 2.2. Signal processing

88 We have developed an innovative filter to improve the SNR of EEG recordings
 89 containing MRCPs. The new method is compared to a state-of-the-art filter
 90 proposed in the literature [2].

91 The data were divided as follows:

- 92 • The measurement from one participant during motor execution was de-
 93 voted entirely for hyper-parameter optimization;
- 94 • Every remaining session was divided in 2 parts, 70% for training and 30%
 95 for testing.

96 Every test set was consequent in time to the corresponding training set, as
 97 to simulate a realistic calibration procedure. Some tests were also performed

98 considering a limited number of channels and a reduced training set.

99 The signals were high-pass filtered at 0.04 Hz, to remove low frequency drifts,
100 reflecting a measurement artifact (Butterworth filter with 40 dB per decade of
101 attenuation outside of the pass band) [2]. Some examples of filtered data are
102 shown in Figure 3.

103 Blink artifacts exhibit a power significantly higher than the rest of the signal,
104 rendering filtering ineffective as the small frequency components overlaying the
105 MRCP are non-negligible [22]. Second Order Blind-source Identification (SOBI)
106 [23] algorithm was shown to be capable of reliably identifying and isolating blink
107 artifacts [24]. Specifically, the artifact was identified as included in the compo-
108 nent (among those provided by SOBI algorithm) with lowest fractal dimension
109 (computed by the Sevcik’s method [25]). Such a component was removed before
110 reconstructing the signal. The same data considered in Figure 3 are shown after
111 removal of blink artifacts in Figure 4.

112 ECG lays outside the frequency band of MRCPs and can be removed by a low-
113 pass filter. Specifically, a low-pass filter with cut-off 20 Hz was used (Butter-
114 worth filter with roll-off 40 dB/decade). Moreover, the data were down-sampled
115 by a factor of 10, bringing the sampling frequency to 50 Hz.

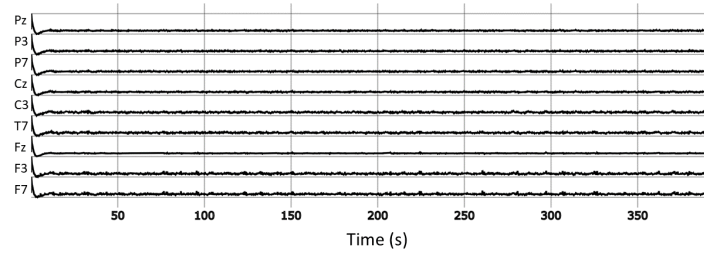
116 2.3. Non-linear optimal spatial filter

117 2.3.1. Linear approach

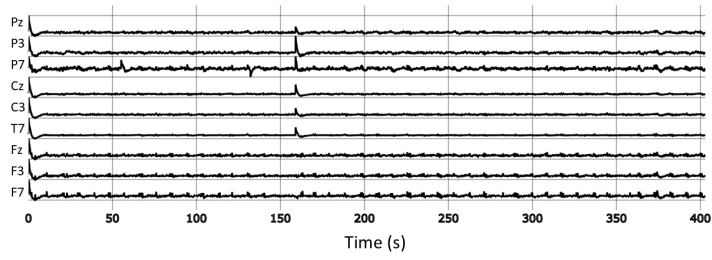
To introduce the problem, we discuss here the design of a linear filter, which
is an approximation of the non-linear technique detailed in the following. The
method strives to find the best weight vector W which, when multiplied by the
multivariate EEG collected in the rows of matrix S , gives the best approximation
of the MRCP component A of the signal

$$S \cdot W = A + \tau \quad (1)$$

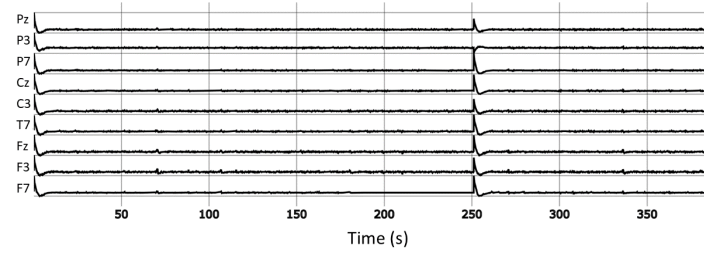
118 where τ is a residual error. This linear model can be considered only as an ap-
119 proximation of the real situation. Indeed, our ill-posed source separation prob-



(a) participant 1

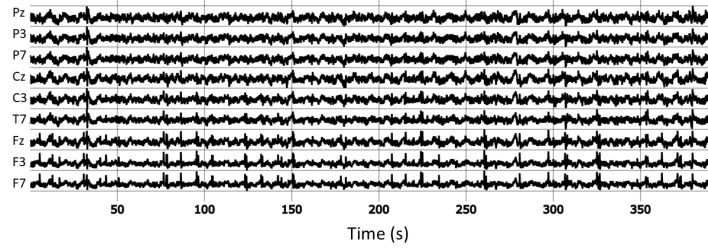


(b) participant 2

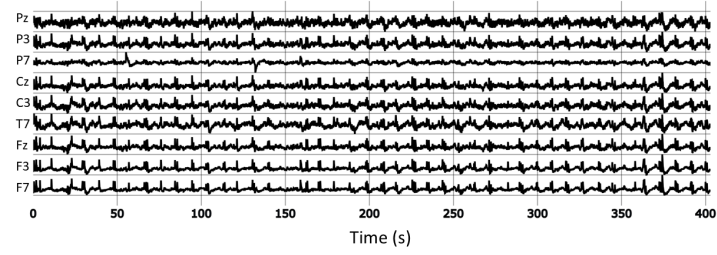


(c) participant 3

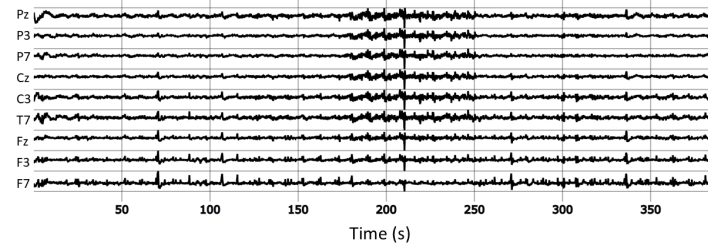
Figure 3: Representative examples of EEGs from different participants, bandpass filtered between 0.4 and 20 Hz.



(a) participant 1



(b) participant 2



(c) participant 3

Figure 4: Representative examples of EEGs from different participants, bandpass filtered and cleaned from artifacts.

120 lem could likely benefit from a non-linear model. This observation suggested us
121 to implement an algorithm able to learn non-linear mappings (described in the
122 following). This is despite most authors managed the extraction of MRCPs from
123 EEG recordings with fully linear models, which are surely simpler to manage
124 than non-linear ones.

125 Equation (1) is linear, in the canonical form of an Ordinary Least Squares
126 (OLS) problem. The OLS method allows to get the vector W which minimizes
127 the residual τ , under the following mathematical assumptions:

- 128 • residuals have zero conditional mean;
- 129 • predictors are linearly independent;
- 130 • residuals are spherical.

131 The validity of the above assumptions will be analyzed in the Appendix.

132 A calibration process is used to estimate A , i.e., the MRCP component included
133 in the signal. Assuming the vector W to be constant in time (which is reasonable
134 considering that the dipole sources generating the MRCP are primarily affected
135 by the physical properties of the skull and of the measuring system which are
136 supposed to be time-invariant), we can estimate A by knowing the instant in
137 which a movement was imagined or executed during a training session. A refer-
138 ence signal is then generated, by placing a prototype waveform in relation to the
139 movement onsets. Specifically, the prototype is a 1 s long wave starting from
140 0 and linearly decreasing until the instant of a moving onset; then, it instantly
141 reaches 0 in the following time sample (notice that different prototypes with du-
142 rations in the range 0.5 - 2 seconds have been tested, obtaining similar results).
143 Then, the vector W is calculated by solving the model on the training set and
144 the MRCP over time (i.e., A) is computed for new unseen EEG recordings based

145 on the estimated W :

$$\begin{aligned} W &= S^{-1} \cdot A \quad (\text{training set}) \\ A &= S \cdot W \quad (\text{testing set}) \end{aligned} \tag{2}$$

146 Notice that S is not square, so that it cannot be inverted. It was pseudo-
 147 inverted (Moore-Penrose inverse [26]). Replacing S^{-1} with its pseudo-inverse
 148 allows to minimize the square norm of the residual τ , obtaining the solution
 149 with minimum squared error. Notice that this solution is unlikely to feature a
 150 residual $\tau = 0$, but still represents the best linear combination of channels to
 151 map the MRCPs to our prototype (in the least mean squared sense).

152 2.3.2. Whitening Transformation

Applying a transformation to S (matrix collecting the EEG channels in its columns) that makes it spherical, i.e., with covariance equal to the identity matrix, can ensure that the model satisfies the last two assumptions of OLS method (i.e., orthogonality of predictors and sphericity of residuals), improving the reliability of the results.

Thus, whitening was employed, by using singular value decomposition (SVD). Consider the factorization of the matrix S written as

$$S = U\Sigma V^T$$

where U and V are orthonormal. The matrix Σ is square diagonal, so that its inversion is immediate and can be used to whiten matrix S

$$S^w = UV^T$$

$$S^{w+} = VU^T$$

where S^w and S^{w+} are the whitened matrix and its pseudo-inverse, respectively.

In summary, the linear model now works as follows. The optimal vector is obtained processing the training signal:

$$W = S^{w+} \cdot A \quad (\text{training set})$$

This vector is used to define the filter to be applied:

$$s_{est}(t) = S_{test}^w \cdot W \quad (testing\ set)$$

153 where $s_{est}(t)$ is the filtered signal obtained by processing the test data S_{test} ,
154 which ideally should be equal to the prototype waveform during an MRCP and
155 zero otherwise.

156 Notice that this method not only emphasizes the signal in the epochs containing
157 the movements intention, while reducing the amplitude out of those epochs, but
158 it also forces the MRCPs to be all similar, which could be useful to identify
159 them.

160 2.3.3. Non-linear method

161 Up until now, the method we devised is only able to infer linear mappings be-
162 tween the EEG signal and the MRCPs. In the field of machine learning, a
163 common strategy to allow separation of non-linear data (e.g., in the field of
164 support vector machines, SVM) is known as the kernel trick [27]. The method
165 is based on the assumption that non linearly separable data can be linearly sep-
166 arated when mapped in a different, usually higher dimensional, feature space
167 [28].

168 The idea of extending the dimensionality of the dataset by a non-linear trans-
169 formation was also applied here. The data, after being extended by a non-linear
170 function, were linearly classified, following the same method detailed in Section
171 2.3.1. The Radial Basis Function (RBF, which is a common kernel) was used
172 to transform our EEG data. It maps the data in an infinite dimensional space
173 and allows a linear classifier to learn any smooth non-linear function [27][28]. In
174 order to reduce the computational cost and memory storage, we approximated
175 the kernel in a finite dimensional feature space [29]. Specifically, the Fourier
176 transform of a RBF $p(\omega)$ is a Gaussian function, which is positive and real
177 (this property holds also for other common shift invariant kernels, by Bochner's
178 theorem [29]). Thus, after normalization, we can consider it as a probability
179 distribution function (i.e., a positive function with integral equal to 1). Hence,

180 writing the RBF as the inverse transform of $p(\omega)$, we can interpret it as the
181 mean value of the complex exponential, or of the cosine function, as both the
182 kernel and its transform are real. The RBF was then estimated using a set of
183 cosine functions with random frequencies with distribution $p(\omega)$ and uniformly
184 distributed phases (see [29] for details). The new kernel has finite dimensional-
185 ity and can be simply reconstructed from the sampled points, so we can use it
186 to explicitly map the EEG data to a high dimensional space before feeding it
187 to the linear algorithm fitting the MRCPs.

188 As shown by a fine tuning on preliminary tests, a dimension of 200 is enough
189 to provide a significant performance boost to the algorithm without overfitting.
190 The steps of this innovative non-linear filter are shown in Figure 5.

191 The output of the filter was lowpass filtered with an exponential filter of order
192 2. Then, a single shallow, CART-based binary decision tree with a maximum of
193 10 nodes computed the thresholds at which the signal is to be considered either
194 a MRCP or noise based on the univariate filter output.

195 *2.4. Comparison with a state-of-the-art method*

196 We have reproduced for comparison the Optimised Spatial Filter (OSF) with
197 quasi-Newton BFGS optimizer and likelihood ratio based detector [2].

The method calculates a virtual channel as a zero-mean linear combination of the EEG channels such as to emphasize the energy of the MRCPs with respect to the noise:

$$\begin{aligned} \text{maximize : } & 10 \cdot \log_{10} \left[\frac{P(\sum_{k=1}^{nc} x_k S_k(t))}{P(\sum_{k=1}^{nc} x_k N_k(t))} \right] \\ \text{subject to : } & \sum_{k=1}^{nc} x_k = 0 \end{aligned}$$

where $P(\cdot)$ indicates power, nc is the number of EEG channels, S the concatenation of signal epochs (in which MRCPs were present) and N the noise (concatenation of epochs in which the MRCP was absent). The windows in which a MRCP is present and absent are taken in the training data set. Starting from

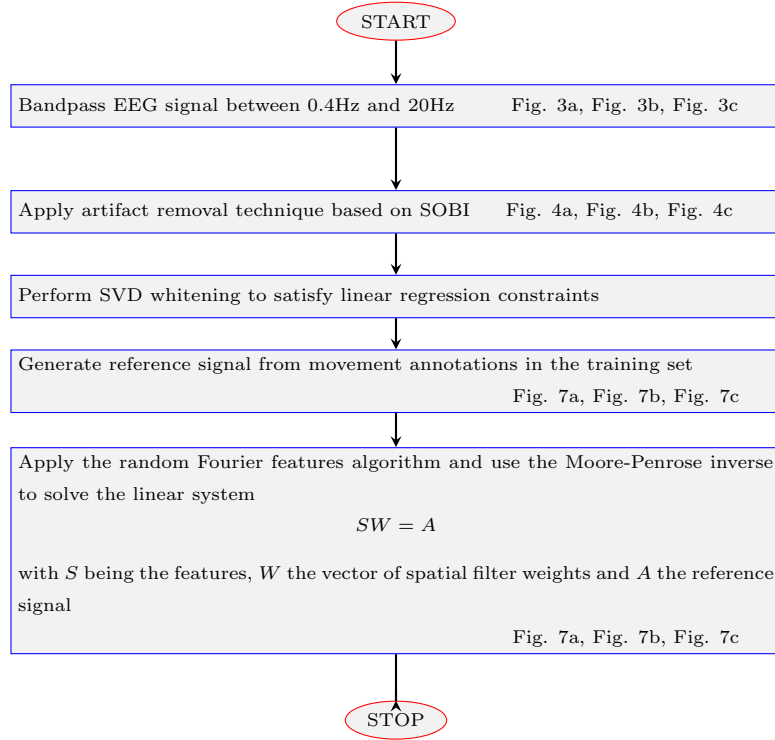


Figure 5: Overview of the NL-OSF algorithm in the training phase (during the test, the data are whitened, processed by the random Fourier features algorithm and applied to the vector of spatial filter weights W to estimate the surrogate signal).

the coefficients of the Laplacian spatial filter

$$x_k = \begin{cases} 1, & k = 1 \\ -\frac{1}{nc-1} & k \neq 1 \end{cases} \quad (3)$$

where $k = 1$ for the electrode Cz, the coefficients are updated using the BFGS algorithm in order to maximize the SNR of such a linear combination.

Here, the constraint that the filter coefficients have zero sum was implicitly parametrized inside the loss function by using a penalty term

$$\text{minimize : } \left(10 \cdot \log_{10} \left[\frac{P(\sum_{k=1}^{nc} x_k S_k(t))}{P(\sum_{k=1}^{nc} x_k N_k(t))} \right] \right)^{-1} + \left(\sum_{k=1}^{nc} x_k \right)^2$$

To smooth the output of the OSF, we used a lowpass exponential filter of order 2. The obtained surrogate signal was classified based on the likelihood ratio [30]. Thus, it was necessary to calculate a reference signal to use in the classification process. This reference was computed as the average of all the MRCPs in the training data, as 2 s windows ending in the negative peak of the potential. The optimal threshold has been calculated using cross-validation on the training data and the Receiver Operating Characteristics (ROC) curve.

2.5. Metrics

Training is performed on continuous traces, while the results are computed on 2 s segmented windows of EEG data taken from the testing set. For every movement of the user, a single window is taken containing the 2 s before the motion execution and a second window is taken from 4 to 6 s before the movement in an interval in which there are no MRCPs. The algorithms are then asked to solve a balanced classification problem.

The metrics chosen for the evaluation of the performances are the Accuracy, the True Positive Rate (TPR) and the False Positive Rate (FPR). They have been reported per-participant alongside the global mean and standard deviation. Performances of the different methods were compared using one-way Kruskal–Wallis ANOVA test by ranks, followed by post-hoc Wilcoxon signed rank test, if significant differences were obtained.

Some tests have also been made by changing some parameters from the default conditions. Specifically, the effect of reducing the number of detection channels was tested, by measuring classification performances when using a lower number of channels: the electrodes F3, P4 and Fz have been removed. Moreover, the effect of reducing the training data was investigated: instead of using the training set including the 70% of the data, performances were also computed reducing the training to the 40% of the MRCPs. The Wilcoxon signed rank test was applied to make specific paired comparisons of the performances of the methods when either the number of recording channels or the training set were reduced.

3. Results

The output of the two filters OSF and NL-OSF is shown in Figures 6 and 7, respectively, for a few representative data (i.e., from the first 3 participants, during the motor execution task). Notice that NL-OSF shows waveforms corresponding to movement onsets which are more similar among them, with respect to those obtained by the OSF. The mean and standard error of MRCPs (aligned and averaged on the basis of the instants of movement onsets) are shown in Figures 8 and 9, for the two filters, respectively, considering the same data of the previous figures. Notice that the average MRCPs obtained by the NL-OSF show smaller oscillations (with an almost monotonic decrease) than those provided by the OSF.

The performances of the two methods on every participant are reported in Tables 1 and 2, considering TPR and FPR (respectively), either in motor execution or imagination.

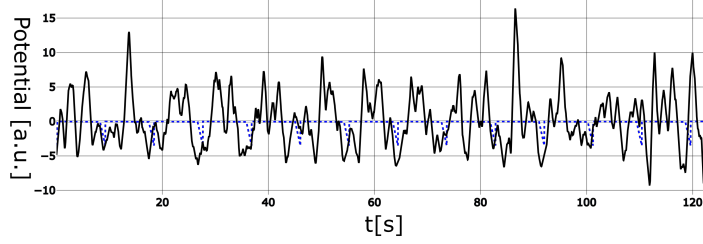
With three one-way ANOVA tests, we see that all performance indexes show some statistically significant variation among different methods. The post-hoc test shows that in Motor Execution the accuracies of NL-OSF is better ($p < 0.001$), its true positive rate is larger ($p = 0.016$) and the false positive rate

True Positive Rate				
Participant ID	Motor Execution		Motor Imagination	
	OSF	NL-OSF	OSF	NL-OSF
1	0.86	0.93	0.92	0.85
2	0.74	0.93	0.62	0.92
3	0.84	0.69	0.62	0.77
4	0.63	0.85	0.07	0.71
5	0.50	0.86	0.79	0.79
6	0.86	0.93	0.69	0.69
7	0.77	1.00	0.69	0.69
8	0.71	0.50	0.57	0.79
9	1.00	0.92	0.62	0.69
10	0.77	0.87	0.29	0.79
11	0.57	0.86	0.43	0.93
12	1.00	1.00	0.07	0.86
13	0.50	0.86	0.92	0.85
14	0.59	0.93	1.00	0.93
15	0.38	1.00	0.86	0.93
16			0.29	1.00

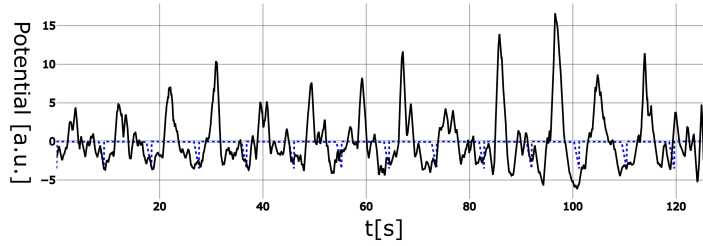
Table 1: True Positive Rate of methods based on optimal spatial filter (OSF) and non-linear optimal spatial filter (NL-OSF) applied to EEG data acquired during either Motor Execution or Imagination.

False Positive Rate				
Participant ID	Motor Execution		Motor Imagination	
	OSF	NL-OSF	OSF	NL-OSF
1	0.27	0.07	0.31	0.08
2	0.43	0.07	0.54	0.08
3	0.15	0.15	0.62	0.46
4	0.38	0.23	0.07	0.36
5	0.07	0.14	0.50	0.21
6	0.50	0.00	0.69	0.08
7	0.15	0.00	0.69	0.15
8	0.58	0.36	0.50	0.29
9	0.52	0.38	0.46	0.23
10	0.33	0.13	0.21	0.29
11	0.40	0.21	0.57	0.07
12	0.43	0.21	0.07	0.07
13	0.57	0.21	1.00	0.15
14	0.43	0.07	1.00	0.29
15	0.15	0.23	1.00	0.00
16			0.50	0.00

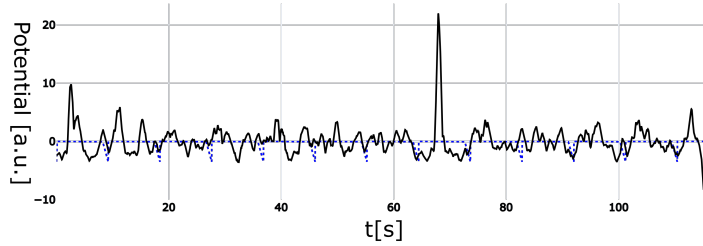
Table 2: False Positive Rate of methods based on optimal spatial filter (OSF) and non-linear optimal spatial filter (NL-OSF) applied to EEG data acquired during either Motor Execution or Imagination.



(a) Result of the OSF Algorithm - Testing Set - Reference in dashed blue - participant 1



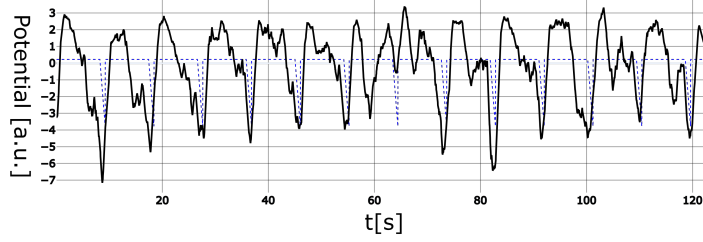
(b) Result of the OSF Algorithm - Testing Set - Reference in dashed blue - participant 2



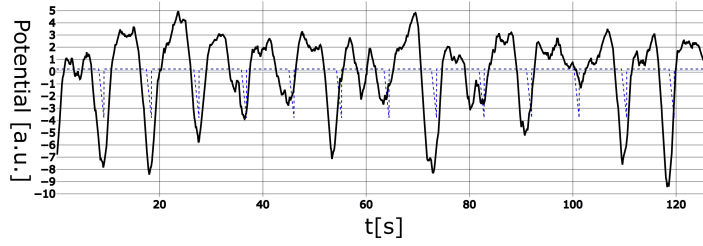
(c) Result of the OSF Algorithm - Testing Set - Reference in dashed blue - participant 3

Figure 6: Representative surrogate data obtained by the OSF Algorithm, during motor execution.

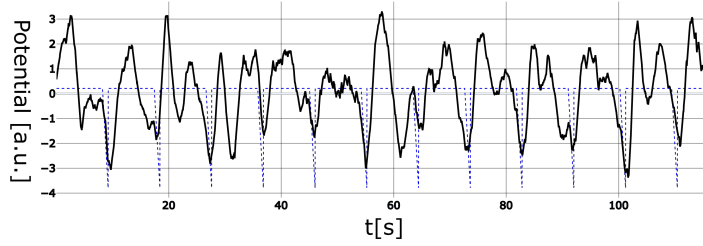
248 is lower ($p=0.001$) than for the OSF. Considering Motor Imagination, the NL-
 249 OSF is superior than OSF in terms of accuracy ($p<0.001$), true positive rate
 250 ($p=0.009$) and false positive rate ($p=0.001$).



(a) Result of the NL-OSF Algorithm - Testing Set - Reference in dashed blue - participant 1



(b) Result of the NL-OSF Algorithm - Testing Set - Reference in dashed blue - participant 2



(c) Result of the NL-OSF Algorithm - Testing Set - Reference in dashed blue - participant 3

Figure 7: Representative surrogate data obtained by the NL-OSF Algorithm in different participants, during motor execution.

251 The effect of a reduction of either the number of EEG channels or the size of
 252 the training set is shown in Figure 10. Moreover, possible differences in per-
 253 formances when considering motor execution or imagination are tested (paired

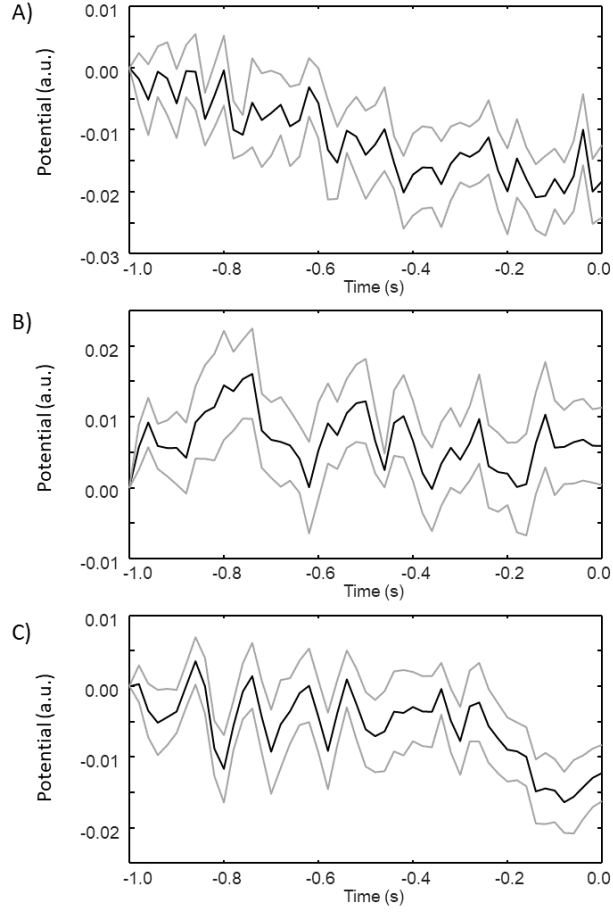


Figure 8: Mean and Standard Error of MRCPs in the testing set. A) participant 1, B) participant 2 and C) participant 3.

254 test, removing from the motor imagination the participant whose data during
 255 motor execution were used for hyper-parameter optimization). Notice that per-
 256 formances decrease only in a few conditions, showing that the methods are quite
 257 stable to problems or to a reduction of information in the data (either due to
 258 motor imagination instead of execution or to a reduction of channels or training
 259 examples).

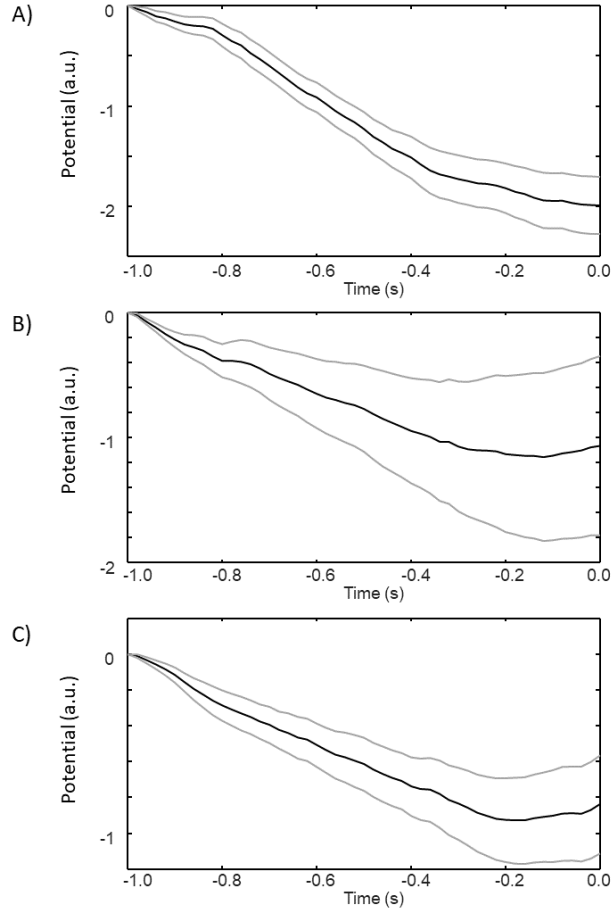


Figure 9: Mean and standard error of MRCPs in the testing set after NL-OSF. A) participant 1, B) participant 2 and C) participant 3.

260 4. Discussion

261 A method for extracting the MRCP component from EEG recordings has been
 262 developed and tested on 15 recordings from different healthy subjects performing
 263 self-paced hand movements and 16 recordings of the same subjects imagining
 264 to perform such hand movements. Our approach is based on a non-linear filter,
 265 mapping multi-channel EEG into a surrogate signal. This signal should be
 266 ideally zero except when the user either performs or imagines a movement, in
 267 which case a prototype similar to an MRCP emerges.

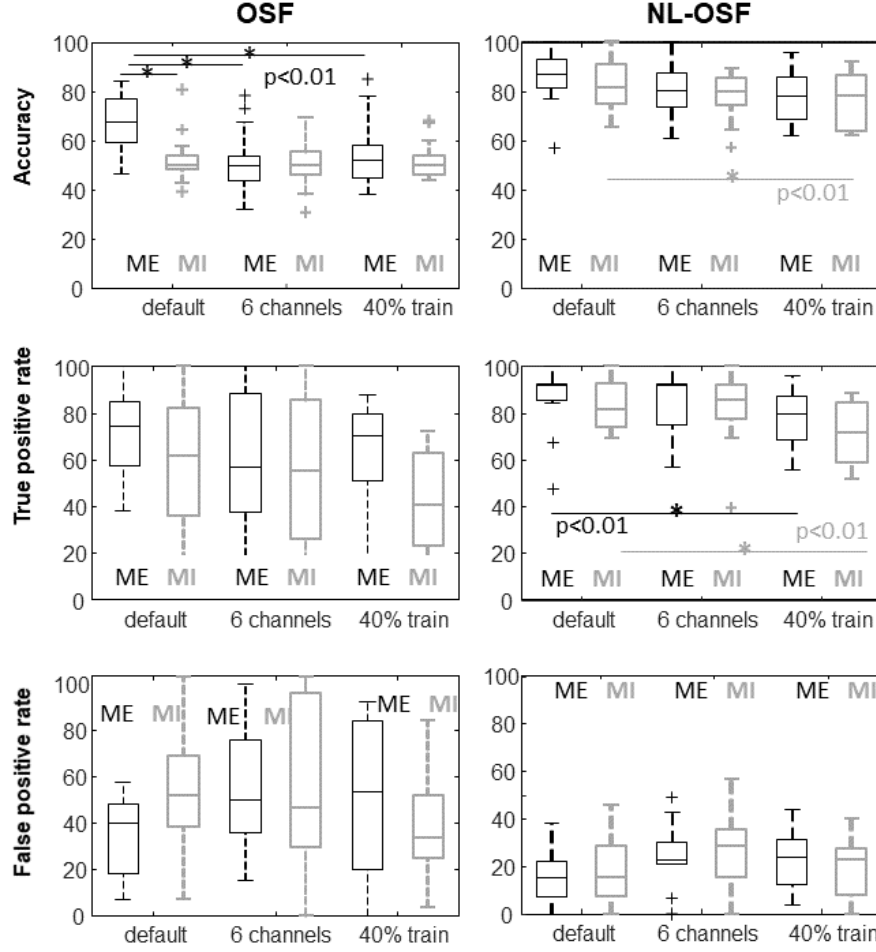


Figure 10: Performances of the filters on the testing set as a function of the experimental modality (either motor execution or imagination) and the reduction of either the number of channels (6 instead of the 9 channels of the default model) or the size of the training set (40% of the MRCs instead of the 70% of the default model).

268 In the tests, the results of our method are compared to those of another filter
 269 (i.e., the OSF [2]) showing higher performances. Other methods have been pro-
 270 posed in the literature which have shown good performances, but they need to
 271 process epochs of EEG, making difficult the application in self-paced: the Lin-
 272 earity Preserving Projections (LPP) with Linear Discriminant Analysis (LDA)

273 [31]; the Adaptive Riemann Kernel (ARK) with SVM [32]. They have been also
274 compared to our approach (not shown results), achieving performances which
275 are not statistically different from those of our method.

276 The main focus of the OSF is in increasing the energy of the potential in the
277 epochs in which the MRCP is present and decreasing it when it is absent.
278 However, the filter responses during different MRCPs are not imposed to be
279 similar. On the other hand, our filter imposes both that the output is large
280 only when the MRCP is present and that it is similar for different MRCPs.
281 The result is that the output of our filter is much more consistent during motor
282 intention of the participants than that of the OSF (Figures 6-9).

283 It is worth noticing that the OSF presented here was coupled with pre-processing
284 techniques which are adapted to our data and to the need of assessing the perfor-
285 mance in realistic online conditions (in which subjective removal of perturbed
286 epochs cannot be applied). Thus, the pre-processing was different from that
287 used in the original paper in which it was proposed, where the blink was not
288 attenuated automatically by a filter, but epochs with a clear blink were removed
289 [2].

290 Consider also that the techniques we employed to pre-process the signal could
291 be not optimal in other applications or they could have poor generalization.
292 Indeed, the literature in the field of EEG processing and multivariate signal
293 analysis presents many interesting techniques (e.g., the constrained ICA [22])
294 which could be tested as preliminary step to select the optimal combination for
295 the specific application.

296 In summary, our technique is based on a filter providing better performances
297 than OSF. Furthermore, not shown results indicate that it has performances
298 comparable to those of window based techniques, but it allows self-paced appli-
299 cation. This is important, as it allows the patient to learn and adapt to the BCI
300 during self-paced sessions [11]. Results hold up with a lower number of channels
301 as well and in the case of a reduced training set, as shown in Figure 10.

302 5. Conclusions

303 An innovative non-linear EEG filter has been developed for identification of
304 MRCP during motor execution or imagination. The results are promising,
305 showing better performances than a previous state-of-the-art filter. Thus, our
306 algorithm could be of interest for application in self-paced BCI.

307 Appendix - OLS Assumptions

308 Here, we analyze whether the main OLS assumptions are verified.

309 *The residuals should have zero conditional mean.* This is also known as the
310 exogeneity constraint. The main causes of failure of exogeneity are the following
311 [33]:

- 312 • Measurement error;
- 313 • Reverse causality;
- 314 • Omitted variables;
- 315 • Omitted sample selection;
- 316 • Lagged dependent variables.

317 We can easily see that our predictor matrix S should not be affected by these
318 items (under proper measurement conditions and provided the assumption that
319 the process which maps the source of the MRCPs to each channel does not affect
320 its phase is verified).

321 *The predictors should be linearly independent.* There is no guarantee that this
322 assumption is verified. In fact, different channels could record the activity of the
323 same sources in the brain or of different sources which have correlated activity.
324 Whitening the data imposes this hypothesis to hold.

325 *The residuals should be spherical.* This implies that the variance of the residual
326 is diagonal and not dependent on time. If we assume that the MRCPs are small
327 compared to the matrix S and thus the EEG signal, we can ensure that this

assumption is close to be verified, by imposing the matrix S to be spherical
itself.

Acknowledgments

Competing interests: None declared

Funding: None

Ethical approval: All subjects participating to the recordings gave their written
informed consent and procedures were approved by the local ethical committee
(number 20130081).

References

- [1] A. Shakeel, M. Navid, M. Anwar, S. Mazhar, M. Jochumsen, I. Niazi, A review of techniques for detection of movement intention using movement-related cortical potentials, *Comput Math Methods Med.* 2015 (2015) 346217. doi:10.1155/2015/346217.
- [2] I. Niazi, N. Jiang, O. Tiberghien, J. Nielsen, K. Dremstrup, D. Farina, Detection of movement intention from single-trial movement-related cortical potentials, *J. Neural Eng.* 8 (6) (2011) 066009.
- [3] L. van Dokkum, T. Ward, I. Laffont, Brain computer interfaces for neurorehabilitation – its current status as a rehabilitation strategy post-stroke, *Annals of Physical and Rehabilitation Medicine* 58 (1) (2015) 3–8. doi:10.1016/j.rehab.2014.09.016.
- [4] E. Leuthardt, G. Schalk, J. Roland, A. Rouse, D. Moran, Evolution of brain-computer interfaces: going beyond classic motor physiology, *Neurosurg Focus*. 27 (1) (2009) E4. doi:10.3171/2009.4.FOCUS0979.
- [5] J. Vidal, Toward direct brain-computer communication, *Annual Review of Biophysics and Bioengineering* 2 (1) (1973) 157.

- [6] C. Guger, R. Spataro, B. Allison, A. Heilinger, R. Ortner, W. Cho, V. La Bella, Complete locked-in and locked-in patients: Command following assessment and communication with vibro-tactile p300 and motor imagery brain-computer interface tools, *Front. Neurosci.* 11 (2017) 251. doi:10.3389/fnins.2017.00251.
- [7] Y. Shahriari, T. Vaughan, L. McCane, B. Allison, J. Wolpaw, D. Krusienski, An exploration of bci performance variations in people with amyotrophic lateral sclerosis using longitudinal eeg data, *J. Neural Eng.* 16 (5) (2019) 056031. doi:10.1088/1741-2552/ab22ea.
- [8] R. Alcaide-Aguirre, S. Warschausky, A. Brown, D. Aref, J. Huggins, Asynchronous brain-computer interface for cognitive assessment in people with cerebral palsy, *J Neural Eng.* 14 (6) (2017) 066001. doi:10.1088/1741-2552/aa7fc4.
- [9] E. Holz, L. Botrel, T. Kaufmann, A. Kubler, Long-term independent brain-computer interface home use improves quality of life of a patient in the locked-in state: a case study, *Arch Phys Med Rehabil.* 96 ((3 Suppl)) (2015) S16–26. doi:10.1016/j.apmr.2014.03.035.
- [10] A. Ramos-Murguialday, M. Schurholz, V. Caggiano, M. Wildgruber, A. Caria, E. Hammer, S. Halder, N. Birbaumer, Proprioceptive feedback and brain computer interface (bci) based neuroprostheses, *PLoS One* 7 (10) (2012) e47048. doi:10.1371/journal.pone.0047048.
- [11] M. Grosse-Wentrup, D. Mattia, K. Oweiss, Using brain-computer interfaces to induce neural plasticity and restore function., *J. Neural Eng.* 8 (2011) 025004. doi:10.1088/1741-2556/8/2/025004.
- [12] M. L.M., S. Heckman, D. McFarland, G. Townsend, J. Mak, E. Sellers, D. Zeitlin, L. Tenteromano, J. Wolpaw, T. Vaughan, P300-based brain-computer interface (bci) event-related potentials (erps): People with amyotrophic lateral sclerosis (als) vs. age-matched controls., *Clin Neurophysiol.* 126 (11) (2015) 2124–2131. doi:10.1016/j.clinph.2015.01.013.

- [13] S. Ajami, A. Mahnam, V. Abootalebi, Development of a practical high frequency brain-computer interface based on steady-state visual evoked potentials using a single channel of eeg., *Biocybernetics and Biomedical Engineering* 38 (1) (2018) 106–114. doi:10.1016/j.bbe.2017.10.004.
- [14] G. Birch, Z. Bozorgzadeh, S. Mason, Initial on-line evaluations of the lfasd brain-computer interface with able-bodied and spinal-cord subjects using imagined voluntary motor potentials., *IEEE Trans Neural Sys Rehab Eng.* 10 (4) (2002) 219–224. doi:10.1109/TNSRE.2002.806839.
- [15] R. Xu, N. Jiang, A. Vuckovic, M. Hasan, N. Mrachacz-Kersting, D. Allan, M. Fraser, B. Nasserroleslami, B. Conway, K. Dremstrup, D. Farina, Movement-related cortical potentials in paraplegic patients: abnormal patterns and considerations for bci-rehabilitation, *Front. Neuroeng.* 7 (2014) 35. doi:10.3389/fneng.2014.00035.
- [16] A. Schwarz, M. Holler, J. Pereira, P. Ofner, G. Muller-Putz, Decoding hand movements from human eeg to control a robotic arm in a simulation environment, *J Neural Eng.* 17 (3) (2020) 036010. doi:10.1088/1741-2552/ab882e.
- [17] Y. Gu, D. Farina, A. Ramos-Murguialday, K. Dremstrup, N. Birbaumer, Comparison of movement related cortical potential in healthy people and amyotrophic lateral sclerosis patients, *Front. Neurosci.* 7 (2013) 65.
- [18] N. Jiang, N. Mrachacz-Kersting, R. Xu, K. Dremstrup, D. Farina, An accurate, versatile, and robust brain switch for neurorehabilitation, Guger, Christoph & Vaughan, Theresa & Allison, Brendan. (2013). *Brain-Computer Interface Research: A State-of-the-Art Summary* 3 (2013) 58. doi:10.1007/978-3-319-09979-8.
- [19] C. Jeunet, E. Jahanpour, F. Lotte, Why standard brain-computer interface (bci) training protocols should be changed: an experimental study., *J. Neural Eng.* 13 (3) (2016) 036024. doi:10.1088/1741-2560/13/3/036024.

- 410 [20] K. Dremstrup, I. Niazi, M. Jochumsen, N. Jiang, D. Mrachacz-Kersting,
411 N. Farina, Rehabilitation using a brain computer interface based on move-
412 ment related cortical potentials. a review, XIII Mediterranean Conference
413 on Medical and Biological Engineering and Computing 2013. Ed. by Laura
414 M. Roa Romero. Cham: Springer International Publishing (2014) 1659–
415 1662.
- 416 [21] R. Scherer, J. Faller, P. Sajda, C. Vidaurre, Eeg-based endogenous on-
417 line co-adaptive brain-computer interfaces: Strategy for success?, 10th
418 Computer Science and Electronic Engineering (CEECE), Colchester, United
419 Kingdom (2018) 299–304doi:10.1109/CEECE.2018.8674198.
- 420 [22] F. Karimi, J. Kofman, N. Mrachacz-Kersting, D. Farina, N. Jiang, Detec-
421 tion of movement related cortical potentials from eeg using constrained ica
422 for brain-computer interface applications, Front. Neurosci. 11 (2017) 356.
423 doi:10.3389/fnins.2017.00356.
- 424 [23] L. De Lathauwer, J. Castaing, Second-order blind identification of underde-
425 termined mixtures. independent component analysis and blind signal sepa-
426 ration, Ed. by Justinian Rosca, Deniz Erdogmus, José C. Príncipe, Simon
427 Haykin. Berlin, Heidelberg: Springer Berlin Heidelberg (2006) 40–47.
- 428 [24] G. Gomez-Herrero, W. Clercq, H. Anwar, K. Kara, O. Egiazarian, S. Huffel,
429 W. Paesschen, Automatic removal of ocular artifacts in the eeg without an
430 eeg reference channel., Proceedings of the 7th Nordic Signal Processing
431 Symposium (2006) 130–133.
- 432 [25] C. Sevcik, A procedure to estimate the fractal dimension of waveforms,
433 Complexity International 5.
- 434 [26] C. Rao, S. Mitra, Generalized inverse of matrices and its applications, New
435 York: John Wiley & Sons.
- 436 [27] T. Hofmann, B. Scholkopf, A. Smola, Kernel methods in machine

- 437 learning, The Annals of Statistics 36 (3) (2008) 1171–1220. doi:10.
438 1214009053607000000677.
- 439 [28] S. Theodoridis, K. Koutroumbas, Pattern recognition, Academic Press.
- 440 [29] A. Rahimi, B. Recht, Random features for large-scale kernel machines,
441 Advances in Neural Information Processing Systems 20. Ed. by J. C. Platt,
442 Daphne Koller, Yoram Singer, Sam T. Roweis, Curran Associates, Inc.
443 (2008) 1177–1184.
- 444 [30] Science direct, likelihood ratio - an overview,
445 [https://www.sciencedirect.com/topics/earth-and-planetary-](https://www.sciencedirect.com/topics/earth-and-planetary-sciences/likelihood-ratio)
446 [sciences/likelihood-ratio](https://www.sciencedirect.com/topics/earth-and-planetary-sciences/likelihood-ratio).
- 447 [31] E. Lew, R. Chavarriaga, S. Silvoni, J. d. R. Millán, Detection of self-paced
448 reaching movement intention from eeg signals, Frontiers in Neuroengineering
449 5 (2012) 13.
- 450 [32] A. Barachant, S. Bonnet, M. Congedo, C. Jutten, Riemannian geom-
451 etry applied to bci classification, 9th International Conference Latent
452 Variable Analysis and Signal Separation (2010) 629–636doi:10.1007/
453 978-3-642-15995-4_78.
- 454 [33] N. Uttam Singh, K. Das, A. Roy, How to test endogeneity or exogeneity:
455 an e-learning hands on sas.