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

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-SF) 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 movements (palmar grasp). MRCPs have been identified from the outputs of the two filters by matching with a template built by averaging responses to movement intentions in the training set. NL-SF had a median accuracy on the overall dataset of 84.6%, which is significantly better than that of OSF (i.e., 76.9%). Being a filter and feasible for self-paced applications, it could be of interest in online BCI system design.

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

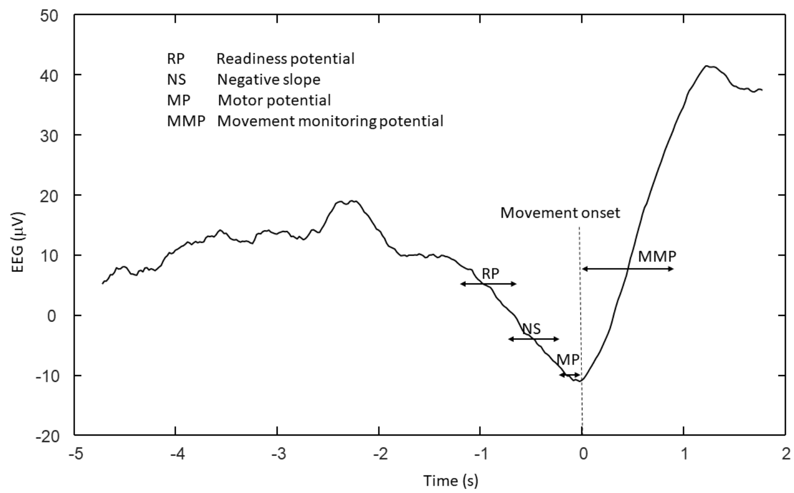


Figure 1: MRCPs of a healthy subject (participant number 1) in the case of motor execution. The wave was obtained by an average of 50 large Laplacian spatial filtered EEG trials.

a common issue in BCIs, i.e., the signal to noise ratio (SNR), which is very 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 waveforms of interest and maps them on a prototype, thus imposing similar filter outputs during each movement intention. We expect that using a non-linear approach (instead of the usual linear one) and imposing similar output waveforms could result 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.

2.1. Experimental data

2.1.1. Subjects

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

2.1.2. Experimental setup

The subjects were placed in a chair in front of the computer with a hand force transducer (Noraxon USA, Scottsdale, AZ) in the right hand. They performed maximum voluntary contraction (MVC) three times and the highest value was retained. Then, grasp trials were executed. A feedback was given to the participants to perform the grasp at 60% MVC force level during this motor execution task. The force data was sampled at 2000 Hz. All participants performed 50 trials of both motor execution and motor imagination of palmer grasp. Each movement type was performed 2×25 times with a 2-3-minute break after the 25^{th} movement. The movements were performed in blocks; the order was randomized. The subjects were visually cued (see Figure 2) by a custom-made program (Aalborg University). The force produced during motor execution tasks was recorded and used as input, so the subjects had continuous visual feedback. During motor imagination tasks, a cursor moved over the template (shown in Figure 2) to cue the subjects. Each trial had the following phases: 1) rest from -5 s to -3 s, 2) preparation for the task from -3 s to 0 s, 3) reach 60% MVC target level from 0 s to 0.5 s, 4) hold the contraction from 0.5 s to 1 s, 5) rest from 1 s to 3–5 s (thus, in total, the rest period varied between 5-7 seconds). For the tasks where the movements were executed, the force was used to determine the movement onset. This was defined as the instant where all values in a 200-ms wide moving time window were above the baseline. The baseline was calculated from the recordings during the rest phase. All onsets were visually inspected.

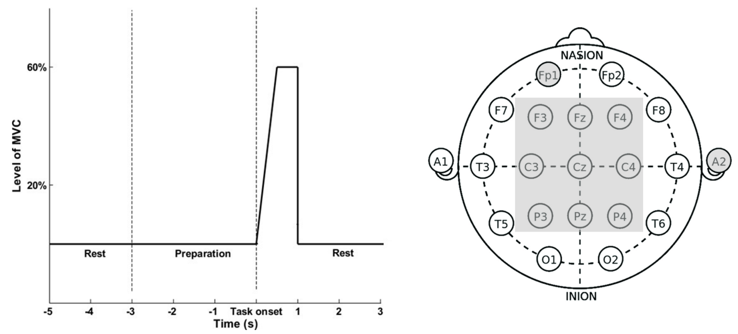


Figure 2: Left – visual cue provided to the participants. The entire template was visible from the beginning to the end of the trial. Subjects were asked to follow the template with the help of a moving cursor. For the Motor execution task, the output of the force transducer was used as a moving cursor, while, for motor imagination, the cursor moved over the template to cue the subjects. Right – EEG electrodes used in the study.

2.1.3. EEG Recording

Continuous 9 channel monopolar (Ag/AgCl ring electrodes) EEG (EEG Amplifiers, Nuamps Express, Neuroscan) was recorded from the following channels (according to the International 10-20 system): F3, Fz, F4, C3, Cz, C4, P3, Pz and P4. The signals were referenced to the right ear lobe and grounded at nasion. Electrooculography (EOG) was recorded from FP1. The EEG and EOG were sampled at 500 Hz and converted with 32-bit precision. The impedance of all electrodes was below 5 kΩ. During the recordings, the subjects were asked to minimize eye blinks and facial and body movements. A digital trigger was sent from the visual cueing program to the EEG amplifier at the beginning of each trial.

2.2. Signal processing

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

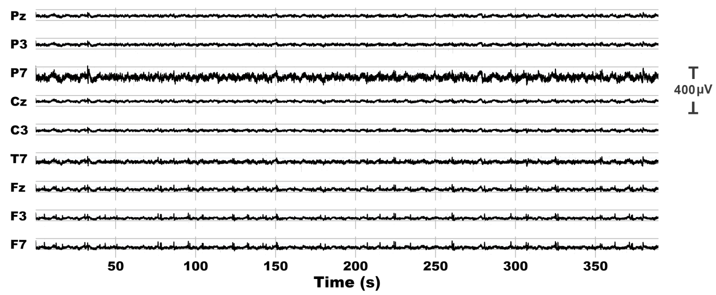
The data were divided as follows:

- All measurements from one participant (subject number 16) during motor execution (50 trials) were devoted for hyper-parameter optimization;
- Every remaining session was divided in 2 parts, 70% for training and 30% for testing.

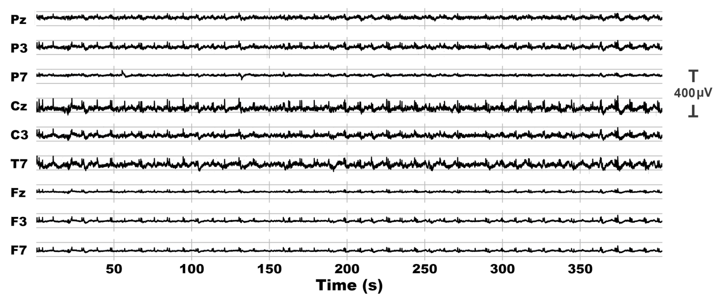
Every test set was consequent in time to the corresponding training set, as to simulate a realistic calibration procedure. Some tests were also performed considering a limited number of channels and a reduced training set.

The signals were high-pass filtered at 0.04 Hz, to remove low frequency drifts, reflecting a measurement artifact (Butterworth filter with 40 dB per decade of attenuation outside of the pass band) [2]. Blink artifacts exhibit a power significantly higher than the rest of the signal, rendering filtering ineffective as the low frequency components overlaying the MRCP are non-negligible [22]. Second Order Blind-source Identification (SOBI) [23] algorithm was shown to be capable of reliably identifying and isolating blink artifacts [24]. Specifically, the artifact was identified as included in the component (among those provided by SOBI algorithm) with lowest fractal dimension (computed by the Sevcik's method [25]). Such a component was removed before reconstructing the signal. Some examples of filtered data to which blink artifacts have been removed are shown in Figure 3.

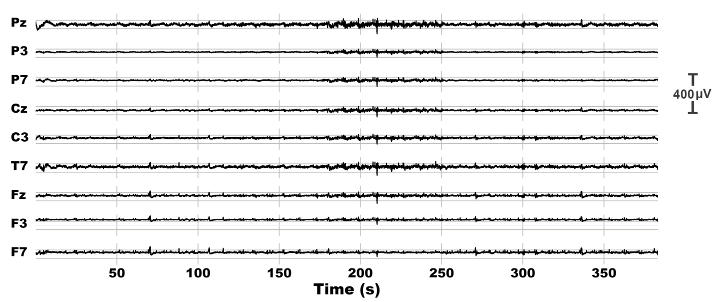
Some high frequency noise (e.g., due to EMG or electrical noise) was removed by a low-pass filter. Specifically, a low-pass filter with cut-off 20 Hz was used (Butterworth filter with roll-off 40 dB/decade). Moreover, the data were down-sampled by a factor of 10, bringing the sampling frequency to 50 Hz.



(a) participant 1



(b) participant 2



(c) participant 3

Figure 3: Representative examples of portions of EEG data, after bandpass filtering and artifacts removal.

2.3. Non-linear optimal spatial filter

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 EEG collected in the rows of matrix S, gives the best approximation of the MRCP component A of the signal

$$S \cdot W = A + \tau \tag{1}$$

where τ is a residual error, W is an unknown vector to be estimated (with nc elements, where nc is the number of EEG channels), A is a time series including a prototype of the MRCP (detailed below, with N_T samples, where N_T is the number of time samples considered) and S is a matrix (of dimension $nc \times N_T$) having in each row the EEG time series from a specific channel. This linear model can be considered only as an approximation of the real situation. Indeed, our ill-posed source separation problem could likely benefit from a non-linear model. This observation suggested us to implement an algorithm able to learn non-linear mappings (described in the following). This is despite most authors managed the extraction of MRCPs from EEG recordings with fully linear models, which are surely simpler to manage than non-linear ones.

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

- residuals have zero conditional mean;
- predictors are linearly independent;
- residuals are spherical.

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

A calibration process is used to estimate A, i.e., the MRCP component included in the signal. Assuming the vector W to be constant in time (which is reasonable considering that the dipole sources generating the MRCP are primarily affected by the physical properties of the skull and of the measuring system which are supposed to be time-invariant), we can estimate A by knowing the instant in which a movement was imagined or executed during a training session. A reference signal is then generated, by placing a prototype waveform in relation to the movement onsets. Specifically, the prototype is a 1 s long wave starting from 0 and linearly decreasing until the instant of a moving onset; then, it instantly reaches 0 in the following time sample (notice that different prototypes with durations in the range 0.5 - 2 seconds have been tested, obtaining similar results). Examples of time series A including prototypes corresponding to movement onsets are superimposed to filtered data in Figures 5. Then, the vector W is calculated by solving the model on the training set and the MRCP over time (i.e, A) is computed for new unseen EEG recordings based on the estimated W:

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

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

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:

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

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

$$s_{est}(t) = S^w_{test} \cdot W \quad (\textit{testing set})$$

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

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

2.3.3. Non-linear method

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

The idea of extending the dimensionality of the dataset by a non-linear transformation was also applied here. The data, after being extended by a non-linear function, were linearly classified, following the same method detailed in Section 2.3.1. The Radial Basis Function (RBF, which is a common kernel) was used to transform our EEG data. It maps the data in an infinite dimensional space and allows a linear classifier to learn any smooth non-linear function [27][28]. In order to reduce the computational cost and memory storage, we approximated the kernel in a finite dimensional feature space [29]. Specifically, the Fourier transform of a RBF $p(\omega)$ is a Gaussian function, which is positive and real (this property holds also for other common shift invariant kernels, by Bochner's theorem [29]). Thus, after normalization, we can consider it as a probability density function (i.e., a positive function with integral equal to 1). Hence, writing the RBF as the inverse transform of $p(\omega)$, we can interpret it as the mean value of the complex exponential, or of the cosine function, as both the kernel and its transform are real. The RBF was then estimated using a set of cosine functions with random frequencies with distribution $p(\omega)$ and uniformly distributed phases (see [29] for details). The new kernel has finite dimensionality and can be simply reconstructed from the sampled points, so we can use it to explicitly map the EEG data to a high dimensional space before feeding it to the linear algorithm fitting the MRCPs.

In summary, the non-linear mapping is obtained by the following steps.

- Set a probability distribution $p(\omega)$ as the discrete Fourier transform of the kernel normalized to have unitary integral (notice that our kernel is Gaussian and its Fourier transform is too).

- Set $\Omega = \{\omega_i\}$ as N vectors of nc elements extracted from the distribution $p(\omega)$ (notice that N is a hyper-parameter to be selected).

- Set $B = \{b_i\}$ (random phases) as N samples taken from the uniform distribution between 0 and 2π.
- Define the i^{th} dimension of the non-linear map as

$$cos(\omega_i \cdot \mathbf{x} + b_i) \tag{3}$$

where $\mathbf{x}$ indicates the EEG data.

In this way, N time series are obtained, which are non-linear combinations of the EEG data. To those data, the linear method (described before) is applied to fit the prototype MRCP component A.

As shown by a fine tuning on preliminary tests using the measurements from subject 16 (as mentioned above), a dimension $N = 200$ is enough to provide a significant performance boost to the algorithm without overfitting. The steps of this innovative non-linear filter are shown in Figure 4.

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

We have reproduced for comparison the Optimized Spatial Filter (OSF), with quasi-Newton BFGS algorithm [2].

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

$$maximize: \quad 10 \cdot log_{10} \left[\frac{P(\sum_{k=1}^{nc} x_k S_k(t))}{P(\sum_{k=1}^{nc} x_k N_k(t))} \right]$$

$$subject\ to: \quad \sum_{k=1}^{nc} x_k = 0$$

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 the coefficients of the Laplacian spatial filter

$$x_k = \begin{cases} 1, & k = 1 \\ -\frac{1}{nc-1} & k \neq 1 \end{cases} \tag{4}$$

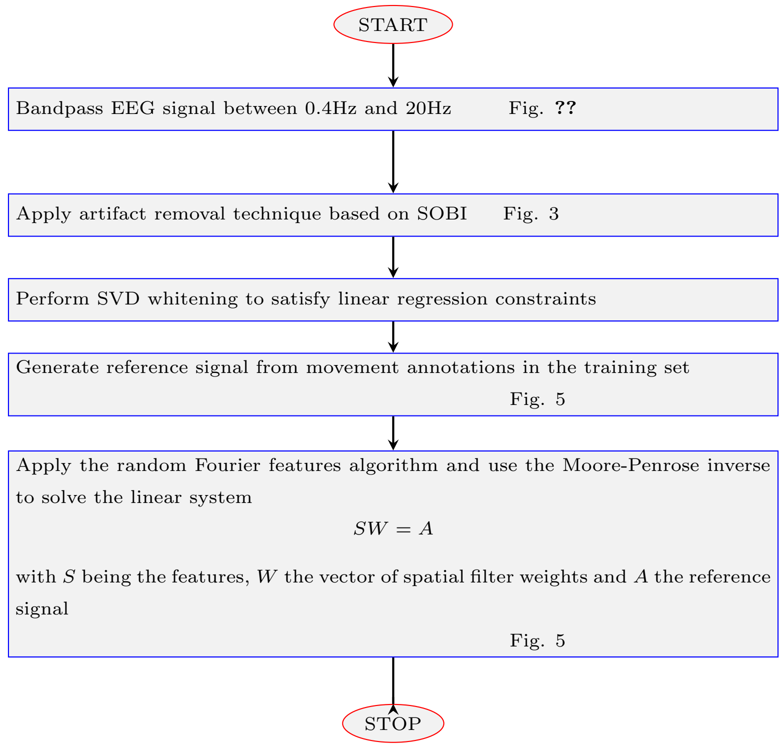


Figure 4: Overview of the NL-SF algorithm in the training phase (during the test, the data are whitened, processed by the random Fourier features algorithm and applied to the vector of spatial filter weights W to estimate the filtered signal).

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

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

2.5. Identification of the MRCP and performance metrics

The outputs of the two filters were post-processed in the same way to perform a classification task, i.e., discriminating between windows that include an MRCP or not. Specifically, a lowpass exponential filter of order 2 was first used. Then, a template was obtained by averaging aligned MRCPs extracted from the training set. This template was used to build a match filter, with optimal threshold computed based on the Receiver Operating Characteristics curve applied on the training data. Then, it was applied on test data and the results were computed on 2 s segmented windows of EEG. For every movement intention of the user (thus, either executed or imagined movement), a single window is taken containing the 2 s before it and a second window is taken from 4 to 6 s before it, 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 two methods were compared using Wilcoxon signed rank test.

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-SF is shown in Figure 5 for a few representative data (i.e., from the first 3 participants, during the motor execution task). Notice that NL-SF 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 Figure 6 for the two filters, considering the same data of the previous figures. Notice that the average MRCPs obtained by the NL-SF show smaller oscillations (with an almost monotonic decrease) than those provided by the OSF. The variabilities of the MRCPs identified by the two filters were quantified as the average over time of the standard error (across epochs) divided by the mean of the absolute value of the average MRCP: their mean±std were 0.35±0.15 and 0.16±0.03, for the OSF and the NL-SF, respectively; Wilcoxon signed rank test indicated a statistically significant difference, with p=0.002.

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.

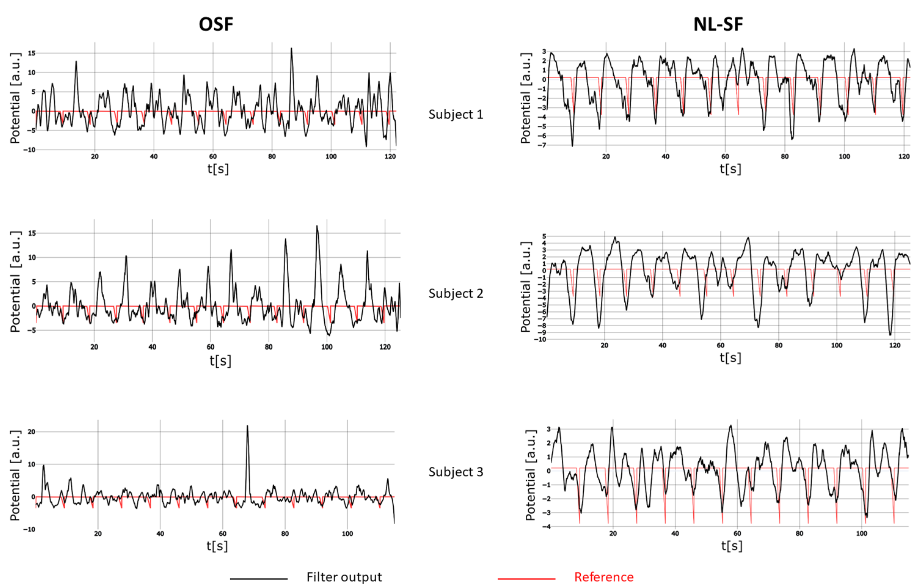


Figure 5: Representative filtered data obtained by the OSF (left) and the NL-SF (right) superimposed to the prototype (indicating movement onset at the peak of the triangular waveforms), during motor execution.

Performance indexes show some statistically significant variation among the two

True Positive Rate				
Participant ID	**Motor Execution**		**Motor Imagination**	
	OSF	**NL-SF**	**OSF**	**NL-SF**
1	0.54	0.92	1.00	0.83
2	1.00	0.77	0.92	0.92
3	0.67	0.67	0.83	0.67
4	0.67	0.83	0.38	0.62
5	0.77	0.85	0.69	0.77
6	0.77	0.85	0.75	0.92
7	1.00	1.00	0.58	0.83
8	0.85	0.38	0.38	0.69
9	0.92	0.83	0.50	0.83
10	0.57	0.71	0.92	0.69
11	0.62	0.69	0.62	0.85
12	0.69	0.85	0.23	0.85
13	0.54	0.92	0.92	0.92
14	0.92	0.92	0.69	0.85
15	0.75	0.75	0.69	0.92
16			1.00	0.92

Table 1: True Positive Rate of methods based on optimal spatial filter (OSF) and non-linear optimal spatial filter (NL-SF) applied to test EEG data (last 30% of our recordings) acquired during either Motor Execution or Imagination (subject number 16 during motor execution is excluded, as the data were used for hyper-parameter optimization).

False Positive Rate				
Participant ID	**Motor Execution**		**Motor Imagination**	
	OSF	**NL-SF**	**OSF**	**NL-SF**
1	0.15	0.38	0.00	0.00
2	0.00	0.07	0.08	0.00
3	0.25	0.17	0.42	0.33
4	0.33	0.08	0.46	0.46
5	0.00	0.23	0.15	0.23
6	0.00	0.00	0.17	0.17
7	0.00	0.00	0.25	0.25
8	0.23	0.08	0.15	0.23
9	0.25	0.25	0.00	0.25
10	0.00	0.00	0.08	0.31
11	0.31	0.31	0.15	0.08
12	0.54	0.15	0.15	0.00
13	0.00	0.23	0.08	0.17
14	0.15	0.00	0.31	0.15
15	0.17	0.08	0.38	0.23
16			0.00	0.00

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

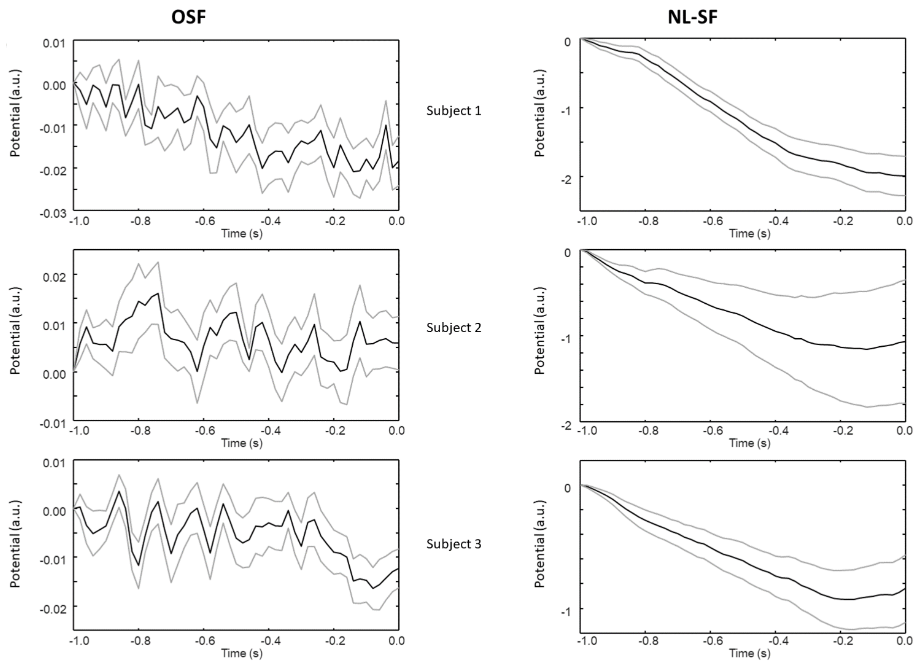


Figure 6: Mean and Standard Error of MRCPs in the testing set of 3 subjects, obtained by the OSF (left) and the NL-SF (right).

methods. The median accuracy was always larger for the NL-SF than for the OSF (see Figure 7), but the sample number was small and statistical significance was not disclosed in the motor execution and motor imagination groups. The post-hoc test shows that in motor imagination the true positive rate of NL-SF is better ($p<0.05$) than for the OSF. Pooling together motor execution and imagination (thus increasing the sample number), the NL-SF is superior than OSF in terms of accuracy ($p=0.04$, median accuracy of 76.9% and 84.6%, for the OSF and NL-SF, respectively) and true positive rate ($p=0.04$, median TPR of 69.2% and 83.3%, for the OSF and NL-SF, respectively).

The effect of a reduction of either the number of EEG channels or the size of the training set is shown in Figure 7. Considering the overall dataset (including motor execution and imagination), the accuracy obtained using the NL-SF is statistically greater than OSF when decreasing the number of channels (median accuracy 75.0% for OSF, 84.6% for NL-SF, p=0.035). Moreover, possible differences in performances when considering motor execution or imagination are tested (paired test, removing from the motor imagination the participant whose data during motor execution were used for hyper-parameter optimization). Notice that performances decrease only in a few conditions, showing that the methods are quite stable to problems or to a reduction of information in the data (either due to motor imagination instead of execution or to a reduction of channels or training examples).

4. Discussion

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

In the tests, the performances of our method overcome those of another filter, i.e., the OSF [2]. The main focus of the OSF is in increasing the energy of the potential in the epochs in which the MRCP is present and decreasing it when it is absent. However, the filter responses during different MRCPs are not constrained to be similar. On the other hand, our filter forces both that the output is large only when the MRCP is present and that it is similar for different MRCPs. The result is that the output of our filter is much more consistent during motor intention of the participants than that of the OSF (Figures 5).

It is worth noticing that the OSF presented here was coupled with pre-processing techniques which are adapted to our data and to the need of assessing the performance in realistic online conditions (in which subjective removal of perturbed epochs cannot be applied). Thus, the pre-processing was different from that

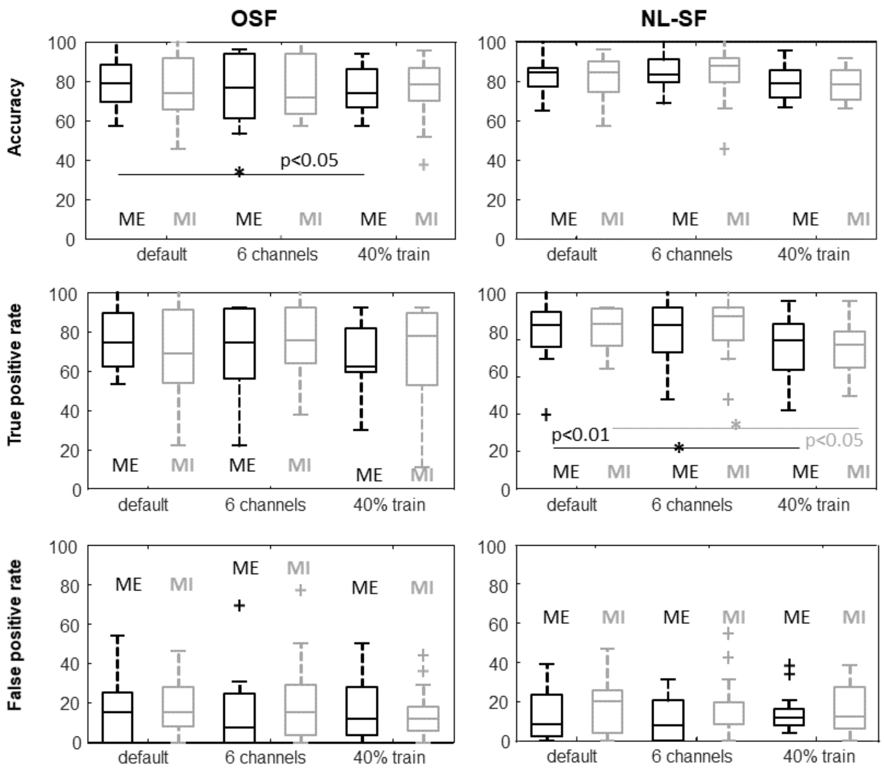


Figure 7: Performances of the filters on the testing set as a function of the experimental modality (either motor execution or imagination) and the reduction of either the number of channels (6 instead of the 9 channels of the default model) or the size of the training set (40% of the MRCPs instead of the 70% of the default model). Box and whiskers plots are shown, indicating median, quartiles, range and outliers (using + markers). Statistical differences in paired comparisons ($p<0.01$) are shown with marker * and a segment joining the two tested distributions.

used in the original paper in which it was proposed, where the blink was not attenuated automatically by a filter, but epochs with a clear blink were removed [2].

Consider also that the techniques we employed to pre-process the signal could be not optimal in other applications or they could have poor generalization. Indeed, the literature in the field of EEG processing and multivariate signal analysis presents many interesting techniques which could be tested as preliminary step (e.g., to remove artifacts or reduce noise) to select the optimal combination of pre-processing for the specific application.

In summary, our technique is based on a filter that provides better performances than OSF in identifying the MRCP (when they are post-processed by the same method, based on template matching). Even being aware of the still limited accuracy and inter-subject variability, we expect that there is room for improving performances of our method as the patient learns and adapts to the BCI during self-paced sessions [11]. Results hold up with a low number of channels as well as in the case of a reduced training set, as shown in Figure 7.

5. Conclusions

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

Appendix - OLS Assumptions

The main OLS assumptions are discussed.

The residuals should have zero conditional mean. This is also known as the exogeneity constraint. The main causes of failure of exogeneity are the following [30]:

- Measurement error;
- Reverse causality;
- Omitted variables;
- Omitted sample selection;
- Lagged dependent variables.

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

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

The residuals should be spherical. This implies that the variance of the residual is diagonal and not dependent on time. If we assume that the MRCPs are small 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:10.1088/1741-2560/8/2/025004.

[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. Nasseroleslami, 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`.

[20] K. Dremstrup, I. Niazi, M. Jochumsen, N. Jiang, D. Mrachacz-Kersting, N. Farina, Rehabilitation using a brain computer interface based on movement related cortical potentials. a review, XIII Mediterranean Conference on Medical and Biological Engineering and Computing 2013. Ed. by Laura M. Roa Romero. Cham: Springer International Publishing (2014) 1659–1662.

[21] R. Scherer, J. Faller, P. Sajda, C. Vidaurre, Eeg-based endogenous online co-adaptive brain-computer interfaces: Strategy for success?, 10th Computer Science and Electronic Engineering (CEEC), Colchester, United Kingdom (2018) 299–304`doi:10.1109/CEEC.2018.8674198`.

[22] F. Karimi, J. Kofman, N. Mrachacz-Kersting, D. Farina, N. Jiang, Detection of movement related cortical potentials from eeg using constrained ica for brain-computer interface applications, Front. Neurosci. 11 (2017) 356. `doi:10.3389/fnins.2017.00356`.

[23] L. De Lathauwer, J. Castaing, Second-order blind identification of underdetermined mixtures. independent component analysis and blind signal separation, Ed. by Justinian Rosca, Deniz Erdogmus, José C. Príncipe, Simon Haykin. Berlin, Heidelberg: Springer Berlin Heidelberg (2006) 40–47.

[24] G. Gomez-Herrero, W. Clercq, H. Anwar, K. Kara, O. Egiazarian, S. Huffel, W. Paesschen, Automatic removal of ocular artifacts in the eeg without an eog reference channel., Proceedings of the 7th Nordic Signal Processing Symposium (2006) 130–133.

[25] C. Sevcik, A procedure to estimate the fractal dimension of waveforms, Complexity International 5 (2018).

[26] C. Rao, S. Mitra, Generalized inverse of matrices and its applications, New York: John Wiley & Sons. (1971).

[27] T. Hofmann, B. Scholkopf, A. Smola, Kernel methods in machine learning, The Annals of Statistics 36 (3) (2008) 1171–1220. doi:10.1214009053607000000677.

[28] S. Theodoridis, K. Koutroumbas, Pattern recognition, Academic Press (2008).

[29] A. Rahimi, B. Recht, Random features for large-scale kernel machines, Advances in Neural Information Processing Systems 20. Ed. by J. C. Platt, Daphne Koller, Yoram Singer, Sam T. Roweis, Curran Associates, Inc. (2008) 1177–1184.

[30] N. Uttam Singh, K. Das, A. Roy, How to test endogeneity or exogeneity: an e-learning hands on sas.