

EFFECTIVE SELECTION OF ELECTRODE SUBSETS IN BCI EXPERIMENTS

Andrey Eliseyev^{1,2,3}, Cecile Moro², Jean Faber^{1,2}, Alexander Wyss², Napoleon Torres²
Corinne Mestais², Tetiana Aksenova^{1,2} and Alim-Louis Benabid^{2,3}

¹Foundation Nanosciences, Grenoble, France

²CEA/LETI/CLIMATEC, MINATEC Campus, Grenoble, France

³Joseph Fourier University of Grenoble, Grenoble, France

Keywords: Tensor Factorization, Partial Least Squares, L1-Regularization, Multi-Way Analysis, Brain-Computer Interface, Self-Paced.

Abstract: Recently N-way Partial Least Squares (NPLS) were reported as an effective tool for neuronal signal decoding and BCI system calibration. This method simultaneously analyses data in several domains. It is based on the projection of a data tensor to a low dimensional space using all variables to create a final model. In the present paper the L1-Penalized NPLS is proposed for sparse BCI system calibration allowing to combine the projection technique with an effective selection of subset of features. The L1-Penalized NPLS was applied for binary self-paced BCI system calibration providing a subset of electrodes selection. Our BCI system is designed for animal research in particular for research in non-human primates.

1 INTRODUCTION

Based on neuronal activity recordings from the brain, Brain Computer Interface (BCI) aims to provide an alternative non-muscular communication pathway to send commands to the external world. Over the last decades several approaches and methods have been developed to improve neuronal signal decoding. Amongst others, recently multi-way analysis was reported as an effective tool for neuronal signal processing (Eliseyev et al., 2011; Fatourehchi et al., 2008; Müller-Putz et al., 2010; Bashashati et al., 2007). Data from several domains are treated simultaneously (e.g. space, frequency and time modalities). In particular, the multi-way analysis was applied in a binary self-paced BCI designed to function in animals (rats) (Eliseyev et al., 2011). In the above mentioned study rats were trained to push a pedal to activate a food dispenser without any cue or external stimulus. Neuronal activity was monitored and intentional control patterns were recognized by the BCI system. To map the neuronal recordings to the spatial-temporal-frequency space, continuous wavelet transform (CWT) was applied to form a tensor of observation. To identify the predictive model N-way Partial Least

Squares (NPLS) (Bro, 1996) was applied. It projects the feature tensor into a low dimensional feature space of latent variables. In parallel, a regression model predicting the intentional control was created. As opposed to other tensor-based methods which recently have been applied in BCI studies (Nazarpour et al., 2006; Zhao et al., 2009; Mørup et al., 2008) the N-way PLS involves class information to perform the tensor decomposition which significantly increases the efficiency of the model. As the NPLS works without any prior knowledge, it can efficiently be applied to automatically generate a model predicting BCI events from recordings of the neuronal brain activity. That is why this method has been chosen as a basic approach in the present study. Note that the NPLS is a projection based method. It involves all variables generating the final model. Throughout BCI experiments neuronal signals of the brain are processed in real-time. Thus computational efficiency of the BCI system is of crucial importance. Selecting an effective subset of features optimizes the computational efficiency and improves the quality of control. In the present article we propose ℓ_1 -Penalized NPLS to directly include feature selection in the modelling process. While generic NPLS lead to a linear combination of all

features the ℓ_1 -Penalized NPLS provide a sparse solution in different directions of analyses (e.g. space, frequencies, or time modalities). In the present study the ℓ_1 -Penalized NPLS was applied in binary self-paced BCI system calibration providing at the same time, a subset of electrodes selection. Corresponding BCI experiments were done in nonhuman primates.

2 METHODS

2.1 Generic NPLS

The N-way PLS algorithm is based on the data projection to a low dimensional feature space (the space of latent variables), with further construction of a linear regression. This method was introduced by Bro, 1996 as a generalization of the ordinary Partial Least Squares (PLS) (Geladi and Kowalski, 1986) to multi-way data sets (tensors). The PLS regression models a linear relationship between a vector of output variables and a vector of input variables on the basis of observation matrices \mathbf{X} and \mathbf{Y} : $\mathbf{Y} = \mathbf{X}\mathbf{C} + \mathbf{V}$, where \mathbf{V} and \mathbf{C} are noise and coefficient matrices. To build the model, the observations are projected into the low dimensional spaces in such a way that the maximum variances of \mathbf{X} and \mathbf{Y} are explained simultaneously. The PLS approach is an iterative procedure. First, the matrices \mathbf{X} and \mathbf{Y} are represented as

$$\mathbf{X} = \mathbf{t}_1 \mathbf{p}_1^T + \mathbf{E}_1, \quad \mathbf{Y} = \mathbf{u}_1 \mathbf{q}_1^T + \mathbf{F}_1,$$

where \mathbf{t}_1 and \mathbf{u}_1 are the latent variables (score vectors), whereas \mathbf{p}_1 and \mathbf{q}_1 are the loading vectors. \mathbf{E}_1 and \mathbf{F}_1 are the matrices of residuals. The score vectors are calculated to maximize the covariance between \mathbf{t}_1 and \mathbf{u}_1 (Geladi and Kowalski, 1986). The coefficient b_1 of a regression $\mathbf{u}_1 = b_1 \mathbf{t}_1 + \mathbf{r}_1$ is calculated to minimize the norm of the residuals \mathbf{r}_1 . The procedure is iteratively applied to the residual matrices.

Similar to the PLS, the NPLS projects the tensor of data into the space of latent variables. Tensors (multi-way arrays) are a higher-order generalization of vectors and matrices. Elements of a tensor $\underline{\mathbf{X}} \in R^{I_1 \times I_2 \times \dots \times I_N}$ are denoted as $x_{i_1 i_2 \dots i_N}$. Here, N is the order of the tensor, i.e., the number of dimensions (ways or modes). The number of the variables I_i in the mode i shows the dimensionality of this mode (Kolda and Bader, 2007). Let us

consider the case of a fourth-order tensor of observations $\underline{\mathbf{X}} \in R^{n \times I_1 \times I_2 \times I_3}$ which contains n samples $\underline{\mathbf{x}}_i \in R^{I_1 \times I_2 \times I_3}, i=1, \dots, n$. Each samples $\underline{\mathbf{x}}_i \in R^{I_1 \times I_2 \times I_3}$ is the tensor of the third-order (cube). This case corresponds to simultaneous analysis of neuronal activity in three domains (e.g. space, frequency and time). As an output a vector $\mathbf{y} \in R^n$ of n observations of scalar variables is considered. The particular case of binary $\mathbf{y} \in R^n$ corresponds to binary self-paced BCI experiments.

The NPLS method decomposes the tensor $\underline{\mathbf{X}}$ as:

$$\underline{\mathbf{X}} = \mathbf{t}_1 \circ \mathbf{w}_1^1 \circ \mathbf{w}_1^2 \circ \mathbf{w}_1^3 + \underline{\mathbf{E}}_1, \quad (1)$$

where the operation “ \circ ” is called the outer product (see Kolda and Bader, 2007). The latent variable $\mathbf{t}_1 \in R^n$ is extracted from the first mode of the tensor $\underline{\mathbf{X}}$ providing maximum of covariance between \mathbf{t}_1 and \mathbf{y} . In parallel, the algorithm forms the factor, i.e. the set of projectors $\{\mathbf{w}_1^1 \in R^{I_1}, \mathbf{w}_1^2 \in R^{I_2}, \mathbf{w}_1^3 \in R^{I_3}\}, \|\mathbf{w}_1^i\|=1, i=1,2,3$ related to the second, the third, and the fourth modes of $\underline{\mathbf{X}}$, respectively, in such a way that the projection of the tensor $\underline{\mathbf{X}}$ on these vectors results in \mathbf{t}_1 . The projectors correspond to each modality of analyses (e.g. space, frequency and time). To build the projectors, a tensor of correlation $\underline{\mathbf{Z}} = \underline{\mathbf{X}} \times_1 \mathbf{y}$ is calculated (\times_1 is the first-mode vector product of the tensor $\underline{\mathbf{X}}$ and the vector \mathbf{y}). Then the vectors $\mathbf{w}^1, \mathbf{w}^2, \mathbf{w}^3$ are estimated by the tensor $\underline{\mathbf{Z}}$ decomposition: $\underline{\mathbf{Z}} = \mathbf{w}^1 \circ \mathbf{w}^2 \circ \mathbf{w}^3 + \underline{\mathbf{E}}$, $\|\underline{\mathbf{Z}} - \mathbf{w}^1 \circ \mathbf{w}^2 \circ \mathbf{w}^3\|_F \rightarrow \min$, where $\|\cdot\|_F$ is the Frobenius norm, which is the generalization of the Euclidean norm for tensors (Kolda and Bader, 2007). To solve the optimization problem the Alternating Least Squares (ALS) (Yates, 1933) algorithm can be applied. It fixes all the projectors except one, which is estimated in a least square sense. The procedure is repeated for all projectors until convergence. A coefficient b_1 of a regression $\mathbf{y} = b_1 \mathbf{t}_1 + \mathbf{f}_1$ is calculated with the Minimal Least Squares (MLS). Next, factors are calculated decomposing the residuals. After the stop of iterations all the particular regressions $\hat{\mathbf{y}}_f = \mathbf{T}_f \mathbf{b}_f, f = \overline{1, F}$ are summarized into a final model $\hat{\mathbf{y}} = \sum_{f=1}^F \mathbf{T}_f \mathbf{b}_f = \mathbf{T} \mathbf{b}$. Vector \mathbf{b} summarized the

regression coefficients for whole set of latent variables $\mathbf{T} = [\mathbf{t}_1 | \dots | \mathbf{t}_F]$. Latent variables \mathbf{t}_f , $f = \overline{1, F}$ correspond to projectors $(\mathbf{w}^1, \mathbf{w}^2, \mathbf{w}^3)_f$.

2.2 L1-Penalized N-PLS Algorithm

The NPLS can be generalized to include additional opportunities of feature selection. For this purpose, the Alternating Least Squares algorithm can be substitute for its penalized version decomposing tensor $\underline{\mathbf{Z}} = \underline{\mathbf{X}} \times_1 \mathbf{y}$. In this case the optimization problem has the form:

$$\{\hat{\mathbf{z}}_1, \hat{\mathbf{z}}_2, \hat{\mathbf{z}}_3\} = \arg \min_{\mathbf{z}_1, \mathbf{z}_2, \mathbf{z}_3} (\|\underline{\mathbf{Z}} - \mathbf{z}_1 \circ \mathbf{z}_2 \circ \mathbf{z}_3\|_F^2 + \lambda P(\mathbf{z}_1, \mathbf{z}_2, \mathbf{z}_3)), \quad (2)$$

where, $P(\cdot)$ is a penalization term, λ is a nonnegative parameter of penalization. Depending on the penalization operator $P(\cdot)$, several optimization tasks can be considered: The Least Absolute Shrinkage Selection Operator (LASSO), $P(\mathbf{A}) = \|\mathbf{A}\|_1$, (Tibshirani, 1996); the Fusion Lasso $P(\mathbf{A}) = \|D\mathbf{A}\|_1$, where D is a difference operator, (Land and Friedman, 1996); the Elastic Net (Enet) (Zou and Hastie, 2005) which includes weighted ℓ_1 -norm and ℓ_2 -norm penalisations etc.

To obtain a sparse solution, the ℓ_1 -norm penalty (LASSO) is often used. The LASSO can be implemented easily providing a sufficient level of selectivity. In the present study, to solve the problem the ℓ_1 -penalty was integrated into the ALS algorithm. At each step of the algorithm all the projectors are fixed except one leading to the optimization:

$$\hat{\mathbf{z}}_i = \arg \min_{\mathbf{z}_i} (\|\underline{\mathbf{Z}} - \mathbf{z}_1 \circ \mathbf{z}_2 \circ \mathbf{z}_3\|_F^2 + \lambda \|\mathbf{z}_i\|_1), \quad (3)$$

$i = 1, 2, 3.$

Considering the particular case $i = 1$:

$$\hat{\mathbf{z}}_1 = \arg \min_{\mathbf{z}_1} (\|\underline{\mathbf{Z}} - \mathbf{z}_1 \circ \mathbf{z}_2 \circ \mathbf{z}_3\|_F^2 + \lambda \|\mathbf{z}_1\|_1). \quad (4)$$

The optimization problem (4) can be rewritten as matrix:

$$\hat{\mathbf{z}}_1 = \arg \min_{\mathbf{z}_1} (\|\mathbf{Z} - \mathbf{z}_1 \mathbf{z}_{2,3}^T\|_F^2 + \lambda \|\mathbf{z}_1\|_1), \quad (5)$$

where $\mathbf{Z} = \underline{\mathbf{Z}}_{(1)}$ is unfolding of the tensor $\underline{\mathbf{Z}}$, and $\mathbf{z}_{2,3} = \text{vect}(\mathbf{z}_2 \circ \mathbf{z}_3)$.

One of the approaches to solve an optimization problem with the ℓ_1 -penalization is the Gauss-Seidel algorithm (Shevade and Keerthi, 2003; Schmidt, 2005). The advantages of this algorithm are its simplicity and low iteration cost, as well as low memory consumption. We have applied this approach to solve the optimization task (5). Namely, the anti-gradient of $RSS = \|\mathbf{Z} - \mathbf{z}_1 \mathbf{z}_{2,3}^T\|_F^2 + \lambda \|\mathbf{z}_1\|_1$ was considered: $-G(\mathbf{z}_1) = 2\mathbf{z}_{2,3}^T (\mathbf{Z}^T - \mathbf{z}_{2,3} \mathbf{z}_1^T) - \lambda \text{sign}(\mathbf{z}_1)$. For the first iteration, \mathbf{z}_1 is set equal to zero consequently the anti-gradient $-G(+\mathbf{0}) = 2\mathbf{z}_{2,3}^T \mathbf{Z}^T - \lambda \mathbf{1}$. Then, the elements of \mathbf{z}_1 with the largest magnitude of the anti-gradient are added to a set of 'free' variables. These 'free' variables are optimized in a 'one at a time' way. For details see Shevade and Keerthi, 2003. Note, that if $\lambda \geq \lambda_{\max} = \max(2\mathbf{z}_{2,3}^T \mathbf{Z}^T)$, the method returns as a solution $\hat{\mathbf{z}}_1 = \mathbf{0}$.

Penalized decomposition of tensor $\underline{\mathbf{Z}} = \underline{\mathbf{X}} \times_1 \mathbf{y}$ results in factor $\{\mathbf{w}^1, \mathbf{w}^2, \mathbf{w}^3\} = \{\hat{\mathbf{z}}_1, \hat{\mathbf{z}}_2, \hat{\mathbf{z}}_3\}$.

To automatically select the optimal value of λ different approaches can be used: Cross-validation (Devijver and Kittler, 1982), generalized cross-validation (Golub et al., 1979), Akaike's Information Criterion (Akaike, 1974), or Schwartz's Bayesian Information Criterion (Schwartz, 1978). The L1-Penalized NPLS algorithm combines computational simplicity and moderate memory consumption with sufficient selectivity. This method was applied for binary self-paced BCI system calibration and for a subset of electrodes selection in the context of BCI experiments in non-human primates.

2.3 Influence Analysis

The elements of the input data have an implicit impact on the NPLS regression model through the latent variables. The Modality Influence (MI) analysis (Cook and Weisberg, 1982) allows estimating the relative importance of the elements of each mode for the final model. In case of tensor input and scalar output variables, the MI procedure is as follows. Latent variables are normalized $\mathbf{t}_f^* = \mathbf{t}_f / \|\mathbf{t}_f\|$ and the regression model takes the form: $\hat{\mathbf{y}} = \mathbf{T}^* \mathbf{b}^*$, $b_f^* = b_f \|\mathbf{t}_f\|$, $f = \overline{1, F}$. Then for chosen modality $i = 1, 2, 3$ coefficients \mathbf{b}^* and components of all factors related to this modality $\{\mathbf{w}_f^i\}_{f=1}^F$ form the matrix $\mathbf{A}^i = [b_1^* \mathbf{w}_1^i | \dots | b_F^* \mathbf{w}_F^i]$.

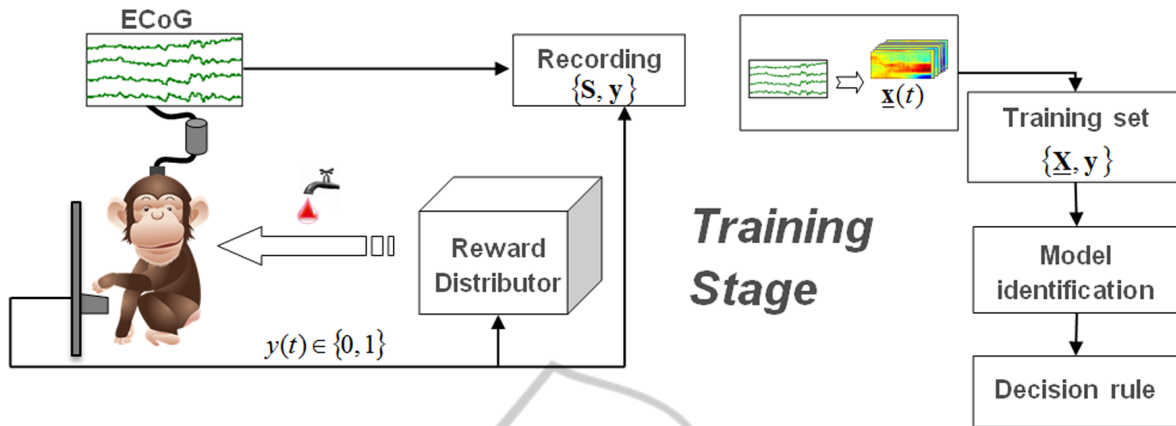


Figure 1: BCI system calibration results in a decision rule; $y(t)$ characterizes the position of the pedal at the moment t ; contains the signal from the brain, recorded during the experiment; \underline{X} is used for internal representation of the signal in the system.

The vector of leverages $\mathbf{h}^i = \text{diag}(\mathbf{A}^i (\mathbf{A}^i)^T)$ shows the summarized influence of elements of this modality on the predicted output.

The MI analysis can be applied to estimate the importance of electrodes, frequency bands, and time intervals related to control events (Eliseyev, 2011).

3 RESULTS

3.1 Data Description

Data was collected from behavioral experiments in non-human primates based on a simple reward-oriented task. During the experiment the monkey is sitting in a custom made primate chair minimally restrained, its neck collar hooked to the chair. The monkey has to push a pedal which can be mounted in for different positions ('left', 'right', 'up', and 'down') on a vertical panel facing the monkey. Every correct push event activates a food dispenser. We used no cue or conditioning stimulus to tell the monkey when to push the pedal. A set of ECoG recordings was collected from 32 surface electrodes chronically implanted in the monkeys brain. Simultaneously, information about the state of the pedal was stored. One recording of each position was used to calibrate the BCI system. Training data sets included all event-related epochs and randomly selected 'non-event' epochs.

3.2 BCI System Calibration

To calibrate the BCI system the brain activity signal

of the training recording was mapped to the temporal–frequency–spatial space to form a tensor of observation. For each epoch j (determined by its final moment t), electrode c , frequency f and time shift τ , elements $x_{j,\tau,f,c}$ of the tensor \underline{X} were calculated as norm of CWT of ECoG signal (see Fig. 1). Frequency band $[10, 300]$ Hz with step $\delta f = 2$ Hz and sliding windows $[t - \Delta\tau, t]$, $\Delta\tau = 0.5$ s with step $\delta\tau = 0.01$ s were considered for all electrodes $c = \overline{1, 32}$. The resulting dimension of a point is $(146 \times 51 \times 32)$. Meyer wavelet was chosen as the mother wavelet taking into account its computational efficiency (Sherwood and Derakhshani, 2009). The binary dependent variable was set to one, $y_j = 1$, if the pedal was pressed at the moment t , and $y_j = 0$, otherwise.

The resulting tensor and the binary vector, indicating the pedal position, were used for calibration. Five factors (the number is defined by the cross-validation procedure) and the corresponding latent variables t_i , $i = \overline{1, 5}$ were extracted by the NPLS algorithm for each pedal position. Due to computational restrictions, the L1-penalized version of the NPLS algorithm ($\lambda = 0.9\lambda_{\max}$) was applied to find a subset of electrodes impacting most the final model. The coefficients b_i^* , of the normalized predictive model $\hat{y} = \sum_{i=1}^5 t_i^* b_i^* + b_0$ correspond to weights of the related factors in the final decomposition:

'left': 0.346, 0.273, 0.232, 0.111, 0.038;
 'right': 0.346, 0.217, 0.195, 0.138, 0.104;

'up': 0.383, 0.263, 0.158, 0.151, 0.045;
 'down': 0.278, 0.210, 0.194, 0.182, 0.138.

Resulting predictive models are based on subsets of few electrodes: 6, 6, 7, and 9 for 'left', 'right', 'up', and 'down' positions of the pedal, respectively. MI analysis revealed the leverages of the elements of each modality (Fig. 2).

3.3 Comparison to Generic NPLS

To compare the L1-Penalized NPLS method with the generic NPLS recordings corresponding to one of positions of the pedal ('up' position) were used. The BCI system was calibrated with both algorithms. Resulted models were applied to the test recording. The computational experiment has demonstrated that the L1-Penalized NPLS outperformed the generic NPLS approach for all tested number of factors from 1 to 5 (Fig. 3).

4 DISCUSSION

Clinical application of BCI is one of the most challenging tasks in neuroengineering. Over the last decades, promising results were obtained both in animal (Chapin et al., 1999; Wessberg et al., 2000) and in human (Leuthardt et al., 2004; Wolpaw et al., 2002) studies. Nevertheless, an effective solution does not exist yet. Most BCI experiments were made in the context of cue-paced (synchronized) approaches where subjects wait for an external cue that drives the interaction (Wolpaw et al., 2002). As a consequence, users are supposed to generate commands only during specific periods. Only the last years an increasing number of laboratories started to apply self-paced BCI paradigms (Leeb et al., 2007; Scherer et al., 2008; Fatourechí et al., 2008; Müller-Putz et al., 2010; Qian et al., 2010). Users control a self-paced BCI at their own intention making these devices more suitable for real-life applications. However, the BCI performances reported by the authors are still not suited for practical application. Our study addresses the problem of neuronal signal decoding in self-paced BCI experiments.

A common approach in brain signal processing intended for event detection/prediction consists in extraction of event related features from neuronal activity. Information from spatial (Rakotomamonjy et al., 2005), frequency (Schlögl et al., 2005), and temporal (Vidaurre et al., 2009) domains is analysed. Note that standard methods are designed for vector input variables which generally represent only one domain (modality) of analysis. However,

using only one domain often does not provide satisfactory results. In most cases two or three ways of analyses are applied sequentially. From the other hand, a tensor-based approach allows simultaneous treatment of several domains. Recently this approach was reported as a prospective tool for neuronal signal processing.

However simultaneous signal processing in several domains increases the dimension of feature space. Reported methods of the multimodal analysis are based on tensor factorization and projection of the data into the low dimensional feature space. They keep all the variables in a final model. Sparse solutions, excluding non-informative electrodes and/or frequency bands will provide better computational efficiency and quality of control. This was the particular objective of the present study.

To do so, we have applied the Penalized NPLS which combines the advantages of the projection technique, the variable selection as well as the advantage of the integrated regression model. The penalized version NPLS was applied to real data collected during BCI experiments in non-human primates to calibrate the self-paced BCI system. Penalization was applied to the spatial modality only. BCI system calibration resulted in predictive models based on subsets of few electrodes (6 - 9 electrodes among 32) in all experimental protocols. The Modality Influence analysis indicates that the electrode #22 located in the primary motor cortex has the highest impact on the decision rule (84%, 97%, 89%, and 75% of extracted information for 'left', 'right', 'up', and 'down' positions of the pedal, respectively). High frequencies (≥ 100 Hz) significantly contribute to the decision in the frequency modality, however, the influence of the lower frequencies (< 100 Hz) is also considerable, especially for the 'left' position of the pedal. In the time domain the interval $[-0.2, 0]$ s before the event is the most significant for all positions of the pedal.

Comparison of the L1-Penalized NPLS with the generic NPLS algorithm demonstrated that the proposed method outperformed the generic approach. This advantage can be explained by the overfitting effect suppression. Additional computational experiments including different tasks will allow better comparison of methods.

Application of sparse predictive models in on-line real-time experiments will be the next step of this study.

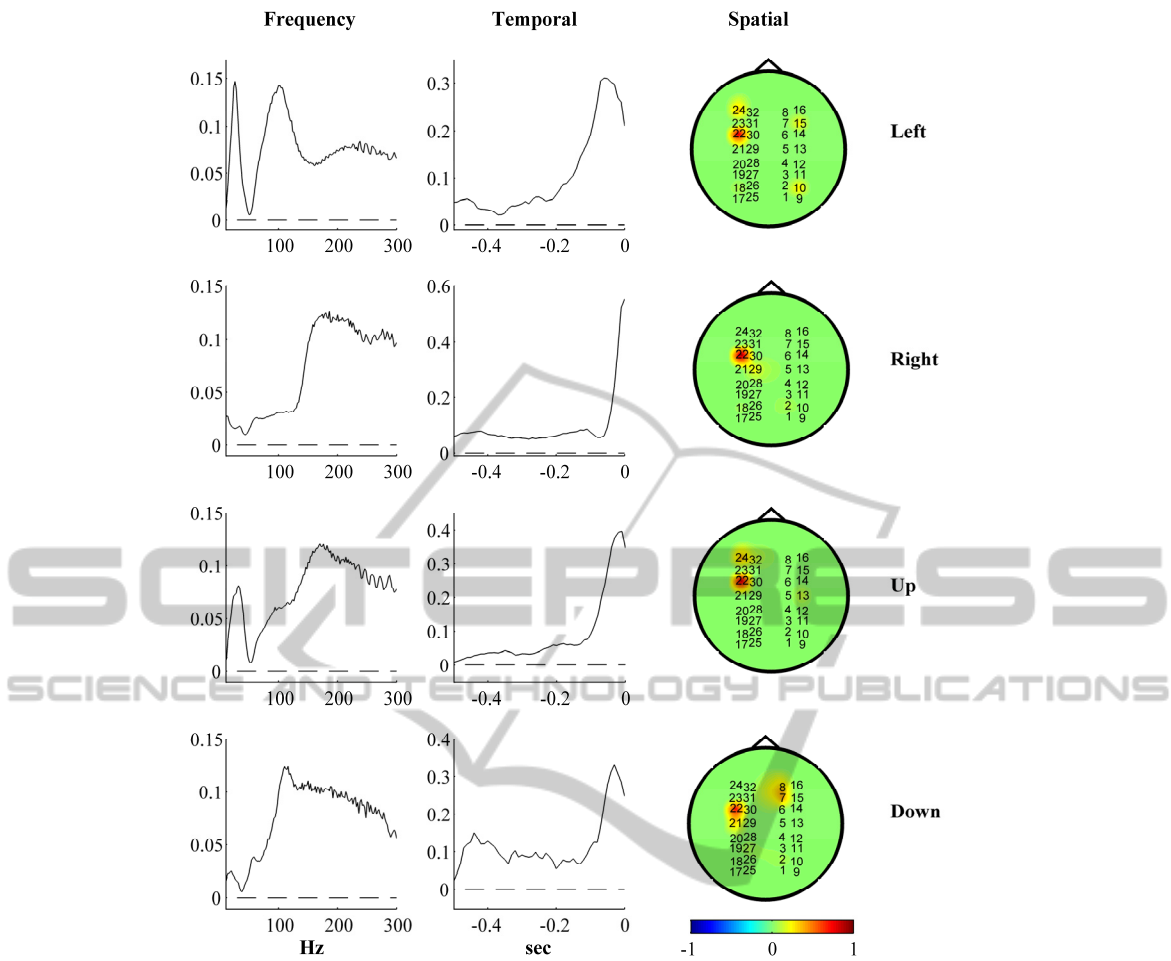


Figure 2: Impact on the predictive model of the components of different modalities (weights) according to the MI analysis for each pedal position; spatial modalities are represented by the graphs and the corresponding color map.

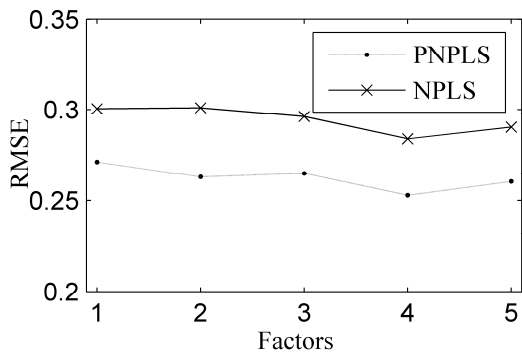


Figure 3: Comparison of prediction errors (root mean squared error, RMSE) for the NPLS and the PNPLS algorithms on the test set for different number of factors.

ACKNOWLEDGEMENTS

This work was partially supported by project CE ICoBI, Nanosciences Foundation RTRA; Edmond J. Safra Philanthropic Foundation; Fondation de l'Avenir, CEA, France.

REFERENCES

Akaike, H., 1974. *A new look at the statistical model identification*. IEEE Trans. Automat. Control 19, 716-723.

Bashashati, A., Ward, R. K., Birch, G.E., 2007. *Towards development of a 3-state self-paced brain computer interface*. J. Computational Intelligence & Neuroscience, 84386.

Bro, R., 1996. *Multway calibration. multilinear pls*. J. Chemom., 10, 47-61.

- Chapin, J. K., Moxon, K. A., Markowitz, R. S., Nicolelis, M.A., 1999. *Real-time control of a robot arm using simultaneously recorded neurons in the motor cortex*. Nature America Inc.
- Cook, R. D., Weisberg, S., 1982. *Residuals and Influence in Regression*. London: Chapman and Hall.
- Devijsver, P. A., Kittler, J., 1982. *Pattern Recognition: A Statistical Approach*. Prentice-Hall, London, 1982.
- Eliseyev, A., Moro, C., Costecalde, T., Torres, N., Gharbi, S., Mestais, C., Benabid, A.L., Aksenova, T., 2011. *Iterative N-way PLS for self-paced BCI in freely moving animals*. Journal of Neural Engineering, 8, 046012.
- Fatourechi, M., Ward, R. K., Birch, G. E., 2008. *A self-paced brain-computer interface system with a low false positive rate*. J Neural Eng 5:9-23. doi:10.1088/1741-2560/5/1/002.
- Geladi, P., Kowalski, B. R., 1986. *Partial least-squares regression: a tutorial*. Anal. Chim. Acta. 185, 1-17.
- Golub, G., Heath, M., Wahba, G., 1979. *Generalization cross-validation as a method for choosing a good ridge parameter*. Technometrics 21, 215-223.
- Kolda, T. G., Bader, B. W., 2007. *Tensor Decompositions and applications*. Sandia report, SAND2007-6702.
- Land, S., Friedman, J., 1996. *Variable fusion: a new method of adaptive signal regression*. Technical Report. Department of Statistics, Stanford University, Stanford.
- Leeb, R., Settgast, V., Fellner, D. W., Pfurtscheller, G., 2007. *Self-paced exploring of the Austrian National Library through thoughts*. International Journal of Bioelectromagnetism, 9, 237-244.
- Leuthardt E. C., Schalk G., Wolpaw J. R., Ojemann J.G., Moran D. W., 2004. *A brain-computer interface using electrocorticographic signals in humans*. J Neural Eng. 1(2): 63-71.
- Mörup, M., Hansen, L. K., Arnfred, S. M., 2008. *Algorithms for Sparse Nonnegative Tucker Decomposition*. Neural Computation, 20, 2112-2131.
- Müller-Putz, G. R., Kaiser, V., Solis-Escalante, T., Pfurtscheller, G., 2010. *Fast set-up asynchronous brain-switch based on detection of foot motor imagery in 1-channel EEG*. International Federation for Medical and Biological Engineering 2010, 10.1007/s11517-009-0572-7.
- Nazarpour, K., Sanei, S., Shoker, L., Chambers, J. A., 2006. *Parallel space-time-frequency decomposition of eeg signals for brain computer interfacing*. EUSIPCO 2006.
- Qian, K., Nikolov, P., Huang, D., Fei, D. Y., Chen, X., Bai, O., 2010. *A motor imagery-based online interactive brain-controlled switch: paradigm development and preliminary test*. Clinical Neurophysiology, 121, 1303-1313.
- Rakotomamonjy, A., Guigue, V., Mallet, G., Alvarado, V., 2005. *Ensemble of SVMs for improving brain computer interface P300 speller performances*. In International Conference on Artificial Neural Networks.
- Scherer, R., Lee, F., Schlögl, A., Leeb, R., Bischof, H., Pfurtscheller, G., 2008. *Toward self-paced brain-computer communication: navigation through virtual worlds*. IEEE Transactions on Biomedical Engineering, 55, 675-682.
- Schlögl, A., Lee, F., Bischof, H., Pfurtscheller, G., 2005. *Characterization of four-class motor imagery EEG data for the BCI-competition 2005*. Journal of Neural Engineering 2(4): L14-L22.
- Schmidt, M., 2005. *Least Squares Optimization with L1-Norm Regularization*. Cs542B Project Report, December 2005.
- Schwartz, G., 1978. *Estimating the dimension of a model*. Ann. Statist. 6, 461-464.
- Sherwood, J., Derakhshani, R., 2009. *On Classifiability of Wavelet Features for EEG-Based Brain-Computer Interfaces*. Proceedings of the 2009 international joint conference on Neural Networks, 2508-2515.
- Shevade, S. K., Keerthi, S.S., 2003. *A simple and efficient algorithm for gene selection using sparse logistic regression*. Bioinformatics, 19(17), 2246-2253.
- Tibshirani, R., 1996. *Regression shrinkage and variable selection via the lasso*. J. Roy. Statist. Soc. Ser. B 58, 267-288.
- Vidaurre, C., Krämer, N., Blankertz, B., Schlögl, A., 2009. *Time domain parameters as a feature for EEG-based brain computer interfaces*. Neural Networks, 22, 1313-1319.
- Wessberg J., Stambaugh C. R., Kralik J. D., Beck P. D., Laubach M., Chapin J. K., Kim J., Biggs S.J., Srinivasan M. A., Nicolelis M. A., 2000. *Real-time prediction of hand trajectory by ensembles of cortical neurons in primates*. Nature 408(6810), 361-365.
- Wolpaw, J. R., Birbaumer, N., McFarland, D. J., Pfurtscheller, G., Vaughan, T. M., 2002. *Brain-computer interfaces for communication and control*. Clin. Neurophysiol., 113, 767-791.
- Yates, F., 1933. *The analysis of replicated experiments when the field results are incomplete*. The empire journal of experimental agriculture, 1, 129.
- Zhao, Q., Caiafa, C.F., Cichocki, A., Zhang, L., Phan, A. H., 2009. *Slice Oriented Tensor Decomposition of EEG Data for Feature Extraction in Space, Frequency and Time Domains*. Lecture Notes in Computer Science, 5863, 221-228.
- Zou, H., Hastie, T., 2005. *Regularization and variable selection via the elastic net*. J. Roy. Statist. Soc. Ser. B 67, 301-320.