Proceedings of the
7th Graz Brain-Computer Interface Conference 2017 DOI: 10.3217/978-3-85125-533-1-57

TIKHONOV REGULARIZATION ENHANCES EEG-BASED SPATIAL FILTERING FOR SINGLE-TRIAL REGRESSION

A. Meinel[1], F. Lotte[2], M. Tangermann[1]

[1]Brain State Decoding Lab, Cluster of Excellence BrainLinks-BrainTools,
Dept. of Computer Science, Albert-Ludwigs-University, Freiburg, Germany,
[2]Inria Bordeaux Sud-Ouest / LaBRI, Talence, France
E-mail: andreas.meinel@blbt.uni-freiburg.de

ABSTRACT: Robust methods for continuous brain state decoding are of great interest for applications in the field of Brain-Computer Interfaces (BCI). When capturing brain activity by an electroencephalogram (EEG), the Source Power Comodulation (SPoC) algorithm enables to compute spatial filters for the decoding of a continuous variable. However, as high-dimensional EEG data generally suffer from low signal-to-noise ratio, the method reveals instabilities for small data sets and is prone to overfitting. We introduce a framework for applying Tikhonov regularization to SPoC by restricting the solution space of filters. Our findings show that additional trace normalization of covariance matrices is a necessary prerequisite to tune the sensitivity of the resulting algorithm. In an offline analysis on data of $N = 18$ subjects, the introduced trace normalized and Tikhonov regularized SPoC variant (NTR-SPoC) outperforms standard SPoC for the majority of individuals. With this proof-of-concept study, a generalizable regularization framework for SPoC has been established which allows for implementing different regularization strategies in the future.

INTRODUCTION

Designing electroencephalography (EEG)-based Brain-Computer Interfaces (BCIs) require translating EEG signals into messages or commands for an application, e.g. by converting EEG activity recorded during imagined hand movements into cursor movements [1]. In most current BCIs for communication and control, this is typically achieved using machine learning and classification algorithms [2]. During online use, EEG signals are then assigned to a discrete set of classes (e.g. left or right hand imagined movements).

A widely used component for effective classification of EEG signals is spatial filtering [3]. Addressing volume conduction effects, spatial filter methods estimate sources, whose signals are more different between classes than signals obtained at the sensor level. The most popular spatial filter algorithm to classify oscillatory EEG activity is the Common Spatial Pattern (CSP) algorithm [3], [4]. It aims at finding filters such that the spatially filtered signals have a variance (and thus a band power in narrow-band filtered signals) that is maximally different between classes. While the CSP algorithm proved very efficient and has become a gold standard in BCI, it is sensitive to noise, non-stationarity and limited data. To address these limitations, various regularized variants of CSP have been proposed making this algorithm effectively more robust [5], [6], [7]. Typically, these approaches inject prior knowledge into the CSP objective function, e.g. in the form of regularization terms. The regularization approaches try to guide the optimization process towards good solutions, despite noise and non-stationarities.

However, not all BCIs are based on classification methods. Several brain signal decoding problems require regression techniques to estimate continuous rather than discrete mental states. For instance, BCIs can be used to estimate continuous workload levels [8] or reaction time from oscillatory activity [9]. As for classification techniques, regression models can also significantly benefit from the use of spatial filters. Thus, Dähne et al. proposed the Source Power Comodulation (SPoC) algorithm, which can be seen as an extension of CSP to regression problems [10]. Indeed, SPoC aims at finding spatial filters such that the power of the filtered EEG signals maximally covaries with a continuous target variable.

Due to similar mathematical formulations, CSP and SPoC share a number of pros and cons. Both algorithms can deliver informative oscillatory signal features but are prone to noise, non-stationarity and limited data. However, while robust variants of CSP have been proposed based on regularization approaches [5], there are no such robust variants for SPoC. Hence, this leaves SPoC with sub-optimal performances when used on noisy data such as those encountered outside laboratories for practical BCI use. In this paper, we aim at addressing this limitation. In particular, we present a novel method to apply Tikhonov regularization to the existing SPoC algorithm. We show how this regularization approach combined with appropriate normalization can indeed outperform the basic SPoC approach. We also illustrate the impact of various regularization parameters on the resulting oscillatory components.

The reminder of this paper first presents in detail the original SPoC algorithm and the regularized variant we

 DOI: 10.3217/978-3-85125-533-1-57

propose. Then it presents an evaluation of these two methods on real EEG data sets for motor performance prediction, before discussing the results.

MATERIALS AND METHODS

1) Source Power Comodulation (SPoC): Supervised spatial filtering algorithms are widely used in EEG-BCI applications. Those filters represent a linear transformation to project the multi-variate EEG data to a lower dimensional subspace. This work focuses on the Source Power Comodulation algorithm (SPoC; [10]) which optimizes a spatial filter by solving a linear regression problem.

In the following, $\mathbf{x}(t) \in \mathbb{R}^{N_c}$ describes the time course of the multivariate bandpass-filtered EEG data acquired from N_c sensors. In accordance with the generative model of the EEG [11], a spatial filter $\mathbf{w} \in \mathbb{R}^{N_c}$ describes the linear projection of the sensor space data $\mathbf{x}(t)$ to a one-dimensional source component $\hat{s}(t) = \mathbf{w}^\top \mathbf{x}(t)$. Translating $\mathbf{x}(t)$ into segments of N_e single epochs $\mathbf{x}(e) \in \mathbb{R}^{N_c \times N_s}$ with N_s sample points per epoch, then SPoC learns a spatial filter $\mathbf{w}$ such that the band-power $\Phi(e) = \mathrm{Var}[\hat{s}(e)]$ of $\hat{s}$ has maximal covariance with a given epoch-wise univariate target variable $z(e)$. Formally, this translates to maximizing the objective function

$$J_1(\mathbf{w}) = \mathrm{Cov}[\Phi(e), z(e)] = \mathbf{w}^\top \Sigma_z \mathbf{w} \qquad (1)$$

by defining a z-weighted averaged covariance matrix $\Sigma_z := \langle \Sigma(e)\, z(e) \rangle$ based on the trial-wise spatial covariance $\Sigma(e) = (N_s - 1)^{-1} \mathbf{x}(e)^\top \mathbf{x}(e)$. $\langle . \rangle$ defines the average across N_e epochs. Furthermore, a norm constraint on $\mathbf{w}$ is applied by setting $J_2(\mathbf{w}) = \mathrm{Var}[\hat{s}(e)] = \mathbf{w}^\top \Sigma_{avg} \mathbf{w} \stackrel{!}{=} 1$, where $\Sigma_{avg} = \langle \Sigma(e) \rangle$ describes the averaged covariance matrix. Overall, this translates to the Rayleigh quotient of the original SPoC$_\lambda$ formulation [10]:

$$J(\mathbf{w}) = \frac{J_1}{J_2} = \frac{\mathbf{w}^\top \Sigma_z \mathbf{w}}{\mathbf{w}^\top \Sigma_{avg} \mathbf{w}} \qquad (2)$$

Technically, maximizing $J(\mathbf{w})$ can be solved as a generalized eigenvalue problem and returns a set $\{\mathbf{w}^{(j)}\}_{j=1,..,N_c}$ of N_c spatial filters with j indexing the rank which is determined in descending order of the eigenvalues and thereby according to the covariance.

Spatial filters allow for a visual interpretation by estimating the corresponding activity pattern $\mathbf{a} = \Sigma_{avg} \mathbf{w}$ as highlighted by Haufe et al. [12].

2) Tikhonov Regularization: The SPoC objective function directly builds upon sample covariance matrices Σ_z and Σ_{avg} as stated by Eq. 2. Their estimation is very sensitive to noisy data, with small training data sets and a high dimensionality aggravating the problem. Finally, poorly estimated covariance matrices will not describe the intended neural processes well and thus mislead the spatial filter optimization. To overcome this, regularization by adding a penalty term $P(\mathbf{w})$ to the objective function's denominator is a common mitigation strategy [13], [14]. The penalty is expressed by a prior that restricts the possible solution space. In this paper, we assign a quadratic penalty term $P(\mathbf{w}) = \mathbf{w}^\top \mathbb{1} \mathbf{w} = \|\mathbf{w}\|^2$ with $\mathbb{1} \in \mathbb{R}^{N_c \times N_c}$ stating the identity matrix. As this penalty scales with the spatial filter norm, solutions with small weights are preferred. Overall, the penalty term is added to the denominator with a regularization parameter α leading to the following maximization problem:

$$J_P(\mathbf{w}) = \frac{\mathbf{w}^\top \Sigma_z \mathbf{w}}{\mathbf{w}^\top [(1-\alpha)\Sigma_{avg} + \alpha \mathbb{1}] \mathbf{w}} \qquad (3)$$

This formulation is known as Tikhonov regularization (TR, [15]) and has similarly been established for the CSP algorithm [5]. Directly solving Eq. 3 refers to SPoC with *Tikhonov regularization* (**TR-SPoC**) in this paper. In case of an extreme regularization expressed by $\alpha = 1$, the Rayleigh quotient in Eq. 3 collapses to the one of the Principal Component Analysis (PCA, [16]).

The Rayleigh quotient in Eq. 2 and 3 relates two sample covariance matrices. In order to control for their relative scaling, a normalization by the trace might be a suitable strategy [4], [17], e.g. $\Sigma'(e) = \Sigma(e)/tr(\Sigma(e))$. Trace normalization will be applied to $\Sigma(e)$ and Σ_{avg} entering Eq. 3, but not upon Σ_z as the z-weighting shall be maintained. This version of the algorithm will be stated as *normalized Tikhonov regularization* of SPoC (**NTR-SPoC**). Applying the same scheme of trace normalization to the standard SPoC algorithm (Eq. 2), will be referred to as *trace-normalized* SPoC (**TN-SPoC**).

3) Data Set for Offline Evaluation: To evaluate the introduced regularization algorithms, data of 18 subjects performing a visuomotor hand force task was used. The paradigm allowed to derive a trial-wise motor performance metric [18], [9]. Each subject completed one session with 400 trials. Within each trial, a "get-ready" interval preceded a "motor execution" phase which was initiated by a clear go-cue. EEG signals were used to predict the trial-wise reaction time (RT) of the motor task based on the time interval [-800, -50] ms prior to the go-cue. EEG activity was acquired by multichannel EEG amplifiers (BrainAmp DC, Brain Products) with a sampling rate of 1 kHz from 63 passive Ag/AgCl electrodes (EasyCap) placed according to the extended 10-20 system. After preprocessing and outlier rejection following the methods described in [9], we restricted our analysis to oscillatory features within the alpha-band frequency range of $[8, 13]$ Hz. The bandpass was realized applying a zero-phase butterworth filter of 6th order. The number of data points N_e remaining after outlier removal varied across the 18 subjects, ranging from 142 to 352 trials. In summary, the following evaluation is based upon RT as a trial-wise continuous target variable z_{true}, which we aim to predict utilizing the individual oscillatory bandpower features of the pre-go EEG activity.

 DOI: 10.3217/978-3-85125-533-1-57

4) Evaluation Scheme: In an offline analysis, we evaluated the proposed algorithms NTR-SPoC and TR-SPoC w.r.t. the regularization parameter α by varying its value in the range $\{0; [10^{-8}, 10^{0}]\}$. Overall, 40 discrete, logarithmically spaced evaluation points were chosen. At each α-value the following $K = 10$-fold chronological cross-validation (CV) scheme was applied:
First, the spatial filter set $\{\mathbf{w}^{(j)}\}_{j=1,..,N_c}$ was gained on training data $\mathbf{x}_{tr}$ and the first $N_f = 4$ highest ranked components selected. Second, a linear regression model with coefficients $\{\beta_j\}_{j=0,..,N_f}$ was trained upon the bandpower features $\Phi_{j,tr} = \mathrm{Var}[\mathbf{w}_{tr}^{(j)}\mathbf{x}_{tr}]$. Finally, those coefficients were used to predict the trial-wise target variable $z_{est}(e)$ using the bandpower features $\Phi_{j,te}(e) = \mathrm{Var}[\mathbf{w}_{tr}^{(j)}\mathbf{x}_{te}](e)$ of unseen test data $\mathbf{x}_{te}$:

$$z_{est}(e) = \beta_0 + \sum_{j=1}^{N_f} \beta_j \Phi_{j,te}(e) \qquad (4)$$

In order to compare the estimated z_{est} with the true motor performance z_{true}, different metrics can be calculated [9]. For simplification, we focused on a single evaluation metric only. Per data set, the evaluation was carried out by transferring the continuous labels z_{true} into a two-class scenario according to the 50th percentile of z_{true}. This enabled the utilization of the receiver operating characteristics (ROC) curve which is calculated upon the estimated target variable z_{est} given the true two-class labels [19]. As ROC performance can be reduced to a scalar value by calculating the area under the ROC curve (AUC), we will name this metric z-AUC as it characterizes the separability of the estimated target variable z_{est}.

5) Selection of Regularization Parameter: To estimate the future performance of the proposed algorithms, we compared two alternative selection strategies in order to determine a suitable regularization parameter α_{opt} for each subject.
First, a leave-one-subject-out (LOSO) cross-validation was applied by determining the regularization parameter α_{opt} based on the grand average z-AUC across $N_{sub}-1$ subjects, calculated for the 40 evaluation points of α.
Second, α_{opt} was selected by a subject-wise nested $K = 10$-fold chronological CV. The inner CV served for estimating the individually optimal regularization parameter α_{opt} among 10 logarithmically scaled values ranging from $\alpha \in [10^{-6}, 10^{-2}]$. In comparison with the LOSO scheme, we chose fewer α values for computational reasons. The value maximizing the z-AUC metric was selected and applied to the outer CV in order to train the respective spatial filtering algorithm and the linear regression model.

RESULTS

Sensitivity to Regularization Parameter: The sensitivity of the introduced approaches TR-SPoC and NTR-SPoC in terms of the regularization parameter α is displayed in Fig. 1. For each α, the z-AUC is reported as a grand average (GA) across all 18 subjects (solid lines). For an individual subject, the impact of the regularization may depend upon the initial performances obtained with basic SPoC. Hence, we decided to separately report the performance for five good subjects (dashed lines). They were selected as the subjects with the best SPoC performance values z-AUC(SPoC) > 0.55. However, as the absolute best subject with z-AUC(SPoC) $= 0.77$ represents a very strong (positive) outlier compared to the full set of 18 subjects, it was not included into this group. (compare with Fig. 2)

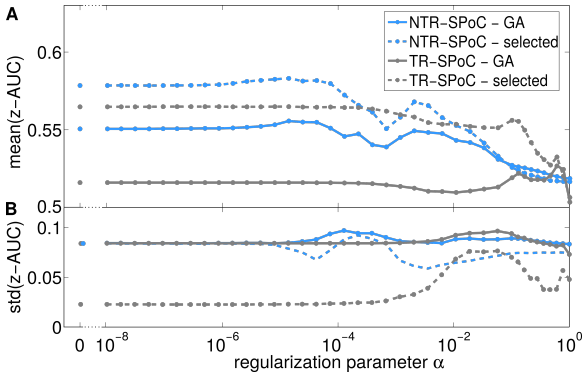

Figure 1. Sensitivity analysis wrt. the regularization strength for TR-SPoC and NTR-SPoC. (A) Averaged z-AUC performance is reported as grand average (GA) across all 18 subjects (solid line) and for five good subjects (dashed line). (B) Standard deviations of the corresponding averages are depicted. For comparison, the displayed performance at $\alpha = 0$ corresponds to the non-regularized SPoC versions.

For TR-SPoC, an increased regularization strength parameter does not affect performance within a wide range of $10^{-8} < \alpha < 10^{-3}$. While even stronger regularization with values of $10^{-1} < \alpha < 10^{0}$ sightly improves the grand average performance, it comes at the cost of a higher standard deviation (see Fig. 1(B)).
The situation looks different for NTR-SPoC as three major effects can be observed when increasing regularization strength controlled by α. Within the range $10^{-8} < \alpha < 10^{-6}$, the performance remains on a stable level. Enlarging the regularization strength to $10^{-6} < \alpha < 10^{-2}$, the performance increases (two local maxima) with best average performance obtained at $\alpha = 1.4 \cdot 10^{-5}$. Extreme regularization expressed by $\alpha > 10^{-2}$ results in a drastic drop of performance. The standard deviation of the reported performance means constantly remain on a high level, indicating that the sensitivity wrt. the regularization parameter α varies strongly across subjects.
In Fig. 1, the effect of trace normalization upon performance can be evaluated for $\alpha = 0$, which describes the absence of any regularization. Thus, the performance reported for TR-SPoC at $\alpha = 0$ corresponds to that of standard SPoC, while $\alpha = 0$ for NTR-SPoC directly

 DOI: 10.3217/978-3-85125-533-1-57

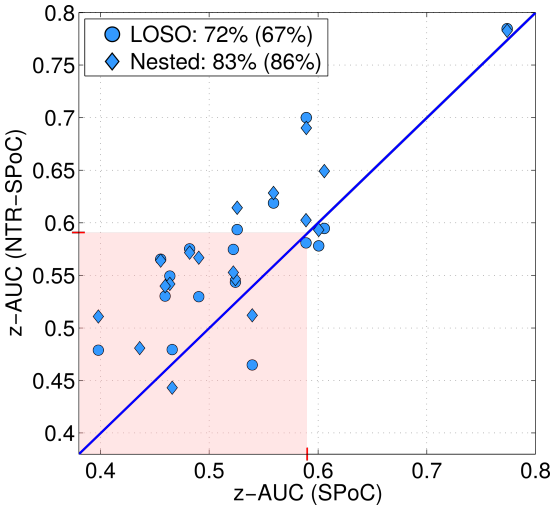

Figure 2. Performance comparison of NTR-SPoC with standard SPoC, using both LOSO and individual nested cross-validation. Each blue marker corresponds to one of the 18 subjects. For both of the selection strategies, the first number reports the percentage of subjects for which NTR-SPoC outperforms SPoC. Percentage values given in brackets are restricted to those data points located outside the red shaded area. It encloses all subjects which do not reach a threshold criterion on z-AUC for meaningful predictions.

maps to TN-SPoC. On the grand average evaluation, the performance of TN-SPoC is increased by 5 %, compared to SPoC while the corresponding standard deviations of both methods are on comparable levels.

As these intermediate results show, that NTR-SPoC effectively outperforms the non trace-normalized alternative TR-SPoC, we will restrict the evaluations in the next paragraphs to NTR-SPoC.

Individual Selection of Regularization Strength: In Fig. 2, the subject-wise performance comparison of NTR-SPoC to standard SPoC is reported for the two regularization parameter selection strategies LOSO and subject-wise nested cross-validation.

In Fig. 2, the circle-shaped markers report the individual performances obtained with LOSO. Building upon the sensitivity analysis, for 15 out of 18 subjects the regularization parameter is selected as $\alpha_{opt} = 7.9 \cdot 10^{-6}$, for the remaining three subjects it was chosen as $\alpha_{opt} = 2.4 \cdot 10^{-5}$. Those values are in good accordance with the global maximum of the grand average performance reported in Fig. 1. The diamond-shaped markers correspond to the validation by individual nested cross-validation. To compare the two selection strategies, we make use of two different group statistics. First, the overall ratio of subjects for which NTR-SPoC outperforms SPoC is provided (72 % for LOSO and 83 % for nested CV). In addition, the values in brackets consider only those individual performances which manage to cross a

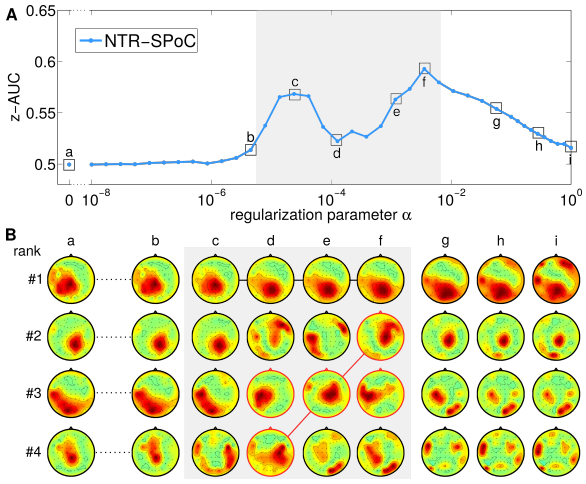

Figure 3. Effect of regularization parameter upon the performance and the spatial patterns for an exemplary subject VP9. (A) Performance of NTR-SPoC for choosing $\alpha \in [10^{-8}, 10^{0}]$. (B) Corresponding activity patterns along first four ranked components at the marked evaluation points (a)-(i).

threshold of minimum meaningful performance at z-AUC$_{th}$ = 0.59 (red shaded area). Details on how this threshold is determined have been reported in [9]. A two-sided Wilcoxon rank sum test on the full data set yields for both selection strategies that NTR-SPoC achieves statistically significant higher performances than SPoC. (The corresponding p-values obtained were $p_{\text{LOSO}} = 2.2 \cdot 10^{-2}$ and $p_{\text{Nested}} = 1.4 \cdot 10^{-3}$.)

Effect of Regularization Upon Oscillatory Components: For an exemplary subject, the effect of the regularization strength in NTR-SPoC upon the underlying first four ranked oscillatory components is depicted in Fig. 3. In (A) the z-AUC is reported for the different α values, in (B) the first four ranked patterns of the marked evaluation point (a)-(i) are shown. As the sign of a pattern $\mathbf{a}$ is arbitrary, they have been corrected to be consistent across the displayed patterns and scaled by their norm.

Consistent with the results presented in Fig. 1, three different α ranges can be identified for the selected subject. The ranges can be characterized according to two aspects, the performance (A) and the underlying spatial patterns (B).

For very small regularization values, represented by evaluation points (a) and (b), the performance as well as the spatial patterns are stable despite of increased α. In other words, NTR-SPoC is not sensitive for such small values of α. The gray shaded area in Fig. 3(A) encloses the evaluation points (c)–(f). This range is sensitive to changes of α which is revealed by performance improvements as well as rank switches among the spatial patterns (e.g. rank #4 from (d) tracked by a solid red line to (f)) or even novel patterns that appear among the top four ranks (e.g. rank #3 at position

 DOI: 10.3217/978-3-85125-533-1-57

(d)). In Fig. 3(B), novel patterns are marked by a red circumference. Regularization beyond $\alpha > 10^{-2}$ ((g)–(i)) leads to a slight drop in performance. This is accompanied by an increased number of components among the first ranks, which display higher spatial frequencies in their activation patterns. The latter observation was made for most subjects and usually affected patterns of ranks 2–4.

DISCUSSION

We introduced the concept of Tikhonov regularization for the SPoC algorithm. Its performance was evaluated by applying a regression model upon the $N_f = 4$ top-ranked components. Nonetheless, this number was not optimized and thus leaves room for improvement.
As reported on the grand average of 18 subjects, the pure Tikhonov regularization of SPoC (TR-SPoC) is not detrimental for a low to medium amount of regularization. Under strong regularization the approach becomes only slightly favorable in terms of performance. However, most strongly regularized TR-SPoC components seemed to show less plausible patterns in the subjective opinion of the authors (data not shown here). Based on even a full regularization parameter sweep, an operating range has neither been observed for a linear nor a logarithmic spacing of α-values along different orders of magnitude as the Rayleigh coefficient of TR-SPoC stated in Eq. 3 is mostly insensitive to the added penalty term.
We tackled this issue by incorporating the trace normalization to the Tikhonov regularization of SPoC (NTR-SPoC). As a result, the NTR-SPoC becomes sensitive to the penalty term over a range of small regularization parameters (see Fig.1 and 3). In addition, we have shown that similar regularization parameters mostly lead to comparable performance increases. As originally proposed by Ramoser et al. [4], the trace normalization of the covariance matrices involved in Eq. 3 balances the influence of the nominator and the denominator. We have shown that NTR-SPoC can improve the performance of individuals, but have also observed that optimal operating ranges varied across subjects. As a positive side effect, the approach brings up new meaningful components. This observation is in accordance with Tikhonov regularization of the CSP algorithm [13]. However, NTR-SPoC comes at a prize, as for some individuals we observed a performance decrease. In these cases, trace normalization may unfortunately remove meaningful information. Though we did not identify a predictor yet, that could indicate, if trace normalization is useful in a new subject or not. Thus, we propose to evaluate both approaches in parallel. In the future, we aim to run the algorithms also on simulation data as this enables e.g. to control for the underlying SNR of the data. As this proof-of-concept motivated the inclusion of the trace norm, applying this to other variants of regularization will help to characterize the effect behind it.

CONCLUSION

In summary, we have presented an approach to regularize the existing SPoC algorithm using a Tikhonov regularization which favors spatial filter solutions with small norms. Our findings show that by simply adding a penalty term to the original SPoC objective function, the filter solutions are insensitive to the regularization term and performance does not improve. We applied an additional trace normalization as a remedy and observed, that it enhances the algorithm's sensitivity for the regularization. This enabled us to define a subject-specific operating range of the regularization and thus improve the achievable performance for most subjects. This proof-of-concept study opens up future research upon different regularization strategies and allows for an in-depth characterization of data sets.

ACKNOWLEDGMENT

This work was fully supported by BrainLinks-BrainTools Cluster of Excellence funded by the German Research Foundation (DFG), grant number EXC 1086. For the data analysis, the authors acknowledge support by the state of Baden-Württemberg through bwHPC and the German Research Foundation (DFG) through grant no INST 39/963-1 FUGG. For some aspects of the data analysis, the BBCI Toolbox was utilized [20].

REFERENCES

[1] Millán J. d. R, Rupp R, Mueller-Putz G, Murray-Smith R, Giugliemma C, Tangermann M, Vidaurre C, Cincotti F, Kübler A, Leeb R, Neuper C, Müller K.-R, and Mattia D. Combining brain–computer interfaces and assistive technologies: State-of-the-art and challenges. *Frontiers in Neuroscience*, 4:161, 2010.

[2] Müller K.-R, Krauledat M, Dornhege G, Curio G, and Blankertz B. Machine learning techniques for brain–computer interfaces. *Biomedical Technologies*, 49:11–22, 2004.

[3] Blankertz B, Tomioka R, Lemm S, Kawanabe M, and Müller K.-R. Optimizing spatial filters for robust EEG single-trial analysis. *Signal Processing Magazine, IEEE*, 25(1):41–56, 2008.

[4] Ramoser H, Müller-Gerking J, and Pfurtscheller G. Optimal spatial filtering of single trial EEG during imagined hand movement. *Rehabilitation Engineering, IEEE Transactions on*, 8(4):441–446, 2000.

[5] Lotte F and Guan C. Regularizing common spatial patterns to improve BCI designs: Unified theory and new algorithms. *IEEE Transactions on Biomedical Engineering*, 58(2):355–362, Feb 2011.

[6] Devlaminck D, Wyns B, Grosse-Wentrup M, Otte G, and Santens P. Multisubject learning for common spatial patterns in motor-imagery BCI. *Intell. Neuroscience*, 2011:8:8–8:8, January 2011.

 DOI: 10.3217/978-3-85125-533-1-57

[7] Samek W, Kawanabe M, and Müller K.-R. Divergence-based framework for common spatial patterns algorithms. *IEEE Reviews in Biomedical Engineering*, 2014.

[8] Frey J, Daniel M, Hachet M, Castet J, and Lotte F. Framework for electroencephalography-based evaluation of user experience. In *Procedings of CHI*, pages 2283–2294, 2016.

[9] Meinel A, Castaño-Candamil S, Reis J, and Tangermann M. Pre-trial EEG-based single-trial motor performance prediction to enhance neuroergonomics for a hand force task. *Frontiers in Human Neuroscience*, 10:170, 2016.

[10] Dähne S, Meinecke F. C, Haufe S, Höhne J, Tangermann M, Müller K.-R, and Nikulin V. V. SPoC: a novel framework for relating the amplitude of neuronal oscillations to behaviorally relevant parameters. *NeuroImage*, 86(0):111–122, 2014.

[11] Parra L. C, Spence C. D, Gerson A. D, and Sajda P. Recipes for the linear analysis of EEG. *NeuroImage*, 28(2):326–341, 2005.

[12] Haufe S, Meinecke F, Görgen K, Dähne S, Haynes J.-D, Blankertz B, and Bießmann F. On the interpretation of weight vectors of linear models in multivariate neuroimaging. *NeuroImage*, 87:96–110, 2014.

[13] Samek W, Vidaurre C, Müller K.-R, and Kawanabe M. Stationary common spatial patterns for brain–computer interfacing. *Journal of Neural Engineering*, 9(2):026013, 2012.

[14] Blankertz B, Kawanabe M, Tomioka R, Hohlefeld F. U, Nikulin V. V, and Müller K.-R. Invariant common spatial patterns: Alleviating nonstationarities in brain-computer interfacing. In *Advances in Neural Information Processing Systems*, pages 113–120, 2007.

[15] Tikhonov A. N. Regularization of incorrectly posed problems. *Soviet Mathematics Doklady*, 4:1624–1627, 1963.

[16] De Bie T, Cristianini N, and Rosipal R. Eigenproblems in pattern recognition. In *Handbook of Geometric Computing*, pages 129–167. Springer, 2005.

[17] Lu H, Eng H. L, Guan C, Plataniotis K. N, and Venetsanopoulos A. N. Regularized common spatial pattern with aggregation for EEG classification in small-sample setting. *IEEE Transactions on Biomedical Engineering*, 57(12):2936–2946, Dec 2010.

[18] Reis J, Schambra H. M, Cohen L. G, Buch E. R, Fritsch B, Zarahn E, Celnik P. A, and Krakauer J. W. Noninvasive cortical stimulation enhances motor skill acquisition over multiple days through an effect on consolidation. *Proceedings of the National Academy of Sciences*, 2009.

[19] Fawcett T. An introduction to ROC analysis. *Pattern Recognition Letters*, 27(8):861–874, 2006.

[20] Blankertz B, Acqualagna L, Dähne S, Haufe S, Schultze-Kraft M, Sturm I, Uumlic M, Wenzel M. A, Curio G, and Müller K.-R. The berlin brain-computer interface: Progress beyond communication and control. *Frontiers in Neuroscience*, 10:530, 2016.