Proceedings of the
7th Graz Brain-Computer Interface Conference 2017 DOI: 10.3217/978-3-85125-533-1-25

BAYESIAN REGRESSION FOR ARTIFACT CORRECTION IN ELECTROENCEPHALOGRAPHY

Karl-Heinz Fiebig[1], Vinay Jayaram[3], Thomas Hesse[2], Alexander Blank[1], Jan Peters[1,3], Moritz Grosse-Wentrup[3]

[1]Autonomous Systems Labs, Technische Universität Darmstadt, Darmstadt, Germany
[2]Institute for Mechatronic Systems in Mechanical Engineering, Technische Universität Darmstadt, Darmstadt, Germany
[3]Department Empirical Inference, Max Planck Institute for Intelligent Systems, Tübingen, Germany

E-mail: karl-heinz.fiebig@stud.tu-darmstadt.de

ABSTRACT: Many brain-computer interfaces (BCIs) measure brain activity using electroencephalography (EEG). Unfortunately, EEG is highly sensitive to artifacts originating from non-neural sources, requiring procedures to remove the artifactual contamination from the signal. This work presents a probabilistic interpretation for artifact correction that unifies session transfer of linear models and calibration to upcoming sessions. A linear artifact correction model is derived within a Bayesian multi-task learning (MTL) framework, which captures influences of artifact sources on EEG channels across different sessions to correct for artifacts in new sessions or calibrate with session-specific data. The new model was evaluated with a cross-correlation analysis on a real world EEG data set. We show that the new model matches state-of-the-art correlation reduction abilities, but ultimately converges to a simple group mean model for the experimental data set. This observation leaves the proposed MTL approach open for a more detailed investigations of artifact tasks.

INTRODUCTION

As opposed to artifact-computer interfaces, a brain-computer interface (BCI) relies on decoding signals of neural origin. Unfortunately, electroencephalography (EEG) based BCIs are very prone to contamination with non-neural noise sources. On the one hand, such artifacts may deteriorate the signal-to-noise ratio and decrease BCI performance. On the other hand, the performance may misleadingly increase due to exploitation of artifact patterns in the learning process. Hence, reducing the effect of artifacts is a key requirement for BCIs in order to reliably decode brain activity from EEG signals.

Many successful techniques to enhance the signal-to-noise ratio in EEG signals have been proposed for BCIs over the last decades. However, advanced methods like beamforming [1], Independent Component Analysis [8] or Common Spatial Patterns [2] often require manual selection of components by an expert for optimal performance. While hybrid approaches have been reported to work well [9], computing filters in the first place is also prone to noise and may miss signals of interest in favor of artifacts. It may therefore be advantageous to correct the signal from known artifact sources before applying further techniques. A popular method is the correction of electrooculographic (EOG) artifacts caused by eye movements and blinks using linear regression. This technique aims to learn the influence of EOG electrodes on EEG and subtracts EOG artifacts from the EEG signals. Note that EOG electrodes may accidentally capture brain activity from the frontal area, which is then unwantedly removed in the process. Influence coefficients are usually determined by regressing the observed EEG signal [10, 11, 12]. It has been found that averaging over coefficients from different signal segments, trials or subjects may increase performance of the artifact correction [13, 14]. However, approaches based on calibration data from the upcoming session or time based re-calculation of the coefficients were also suggested [3, 15]. These findings motivate a common framework in order to exploit stability of transfer models while still retaining the ability to calibrate regression models with new data.

In this work, a theoretical framework is presented that unifies the combination of influence coefficients from different sessions and the adaptation with new artifacts. The artifact correction problem is put into a probabilistic interpretation and approached within a Bayesian multi-task learning (MTL) framework that is already used to decode brain activity [4, 5, 16]. The presented algorithm is able to learn a matrix Gaussian prior distribution over artifact influences from different signal segments, sessions and subjects. The trained prior can be either used to directly correct artifacts in new sessions or be calibrated with new artifacts.

The remainder of this paper is organized as follows. The method section introduces a probabilistic interpretation for artifact correction and restates the problem with an equivalent formulation in order to derive a closed form solution. It is then shown how the new model can be immediately used for artifact correction and later on adapted with calibration data. Afterwards, the experimental setup

 DOI: 10.3217/978-3-85125-533-1-25

and the evaluation of the new model against current artifact regression models is described. The results section shows that the new model operates on lower correlation levels in comparison with other models. However, the MTL algorithm is found to train a simple group mean over influence coefficients for the artifacts in the data set. Finally, this paper elaborates on the results with a discussion and concludes with a short summary and future work on the proposed framework.

METHODS

In this paper, scalars are denoted with lowercase, vectors with bold lowercase, matrices with upper case and sets with calligraphic uppercase letters.

Probabilistic Artifact Regression: In accordance with the literature, we assume that an EEG measurement sample $\boldsymbol{y} \in \mathbb{R}^k$ from k channels can be modeled with $\boldsymbol{y} = \boldsymbol{s} + W\boldsymbol{n}$, where $\boldsymbol{s} \in \mathbb{R}^k$ are the EEG signals, $\boldsymbol{n} \in \mathbb{R}^m$ are the artifact sources as measured by m artifact channels and $W \in \mathbb{R}^{k\times m}$ is the weighting matrix. W explains the influence of the m artifact sources on each of the k EEG channels. However, the signal that is observed at the recording sites is additionally contaminated by noise contributions arising from other sources that we do not keep track of. We therefore extend the model to $\boldsymbol{y} = \boldsymbol{s} + W\boldsymbol{n} + \boldsymbol{\varepsilon}$, where $\boldsymbol{\varepsilon} \in \mathbb{R}^k$ represents the signal contribution from other noise. This model can be put into a probabilistic relation by assuming that the noise is distributed according to a zero-mean Gaussian $\boldsymbol{\varepsilon} \sim \mathcal{N}\left(\mathbf{0}, \sigma^2 I_k\right)$ with some variance σ^2, where $I_k \in \mathbb{R}^{k\times k}$ denotes the identity matrix in k dimensions. An observed EEG sample $\boldsymbol{y}$ is then drawn from a Gaussian distribution

$$\boldsymbol{y} \sim \mathcal{N}\left(\boldsymbol{s} + W\boldsymbol{n}, \sigma^2 I_k\right), \tag{1}$$

centered at the linear model output and deviating according to some noise encoded in σ^2.

Multi-task Learning with Artifacts: The weight matrix W is usually determined by linear regression on an artifactual data set in order to find the influences of artifact sources on EEG channels. However, these influences may vary across subjects, sessions and trials. We therefore regard the regression problem for an artifact as an individual task and denote the gathered data set of q tasks with $\mathcal{T} = \left\{\mathcal{D}^{(t)}\right\}_{t=1}^{q}$. Each task data set $\mathcal{D}^{(t)} \in \mathcal{T}$ takes the form

$$\mathcal{D}^{(t)} = \left\{\left(\boldsymbol{n}_i^{(t)}, \boldsymbol{y}_i^{(t)}\right)\right\}_{i=1}^{n_t} \subset \mathbb{R}^m \times \mathbb{R}^k \tag{2}$$

consisting of a single artifact contaminated segment with n_t EEG samples $\boldsymbol{y}_i^{(t)}$ measured at k channels and artifact samples $\boldsymbol{n}_i^{(t)}$ recorded at m channels. Each data set $\mathcal{D}^{(t)} \in \mathcal{T}$ is associated with a linear regression model defined by its weight matrix $W^{(t)} \in \mathbb{R}^{k\times m}$. We denote the set of weight matrices with $\mathcal{W} = \left\{W^{(t)}\right\}_{t=1}^{q}$. Following the Bayesian MTL framework presented in [16], we can state a data *likelihood* from the probabilistic interpretation in (1) and introduce a *prior* distribution over the weight matrices. The prior aims to capture commonalities in the influence of artifact sources to EEG channels across artifacts. In particular, we assume a matrix Gaussian distribution $p(W) = \mathcal{MN}\left(W \mid M_W, \Sigma_{r;W}, \Sigma_{c;W}\right)$ as prior model, where $M_W \in \mathbb{R}^{k\times m}$ is the mean weight matrix, $\Sigma_{r;W} \in \mathbb{R}^{k\times k}$ is the row covariance matrix that captures correlations in the influence between the EEG channels and $\Sigma_{c;W}$ is the column covariance capturing correlations between the artifact channels. Unfortunately, pulling everything together to state a *posterior* objective does not yield a closed form solution. However, the relation between matrix and multi-variate Gaussians is exploited in the next section in order to obtain an analytic solution for MTL artifact regression.

Bayesian Kronecker Regression: While the MTL approach using a matrix Gaussian prior does not derive in closed form, the problem can be restated into a form that yields an analytic solution. First, note that the matrix Gaussian $\mathcal{MN}\left(W \mid M_W, \Sigma_{r;W}, \Sigma_{c;W}\right)$ is equivalent to a multi-variate Gaussian of the form $\mathcal{N}\left(\mathrm{vec}\left(W\right) \mid \mathrm{vec}\left(M_W\right), \Sigma_{c;W} \otimes \Sigma_{r;W}\right)$, where $\mathrm{vec} : \mathbb{R}^{k\times m} \to \mathbb{R}^{km}$ is the vectorization of a matrix stacking the columns into a column vector and $\otimes : \mathbb{R}^{k\times m} \times \mathbb{R}^{k\times m} \to \mathbb{R}^{km\times km}$ is the Kronecker product of two matrices [6]. Hence, instead of targeting the weight matrix W itself, the vectorized version $\mathrm{vec}\left(W\right)$ of the weights can be optimized. In fact, the model stated in (1) is equivalent to a Kronecker formulation of the form

$$\boldsymbol{y} \sim \mathcal{N}\left(\boldsymbol{s} + \left(\boldsymbol{n}^T \otimes I_k\right)\mathrm{vec}\left(W\right), \sigma^2 I_k\right). \tag{3}$$

It turns out that by assuming that the source $\boldsymbol{s}$ and noise signal $\boldsymbol{n}$ are independent, the artifact regression problem in this formulation is directly solvable by the MTL algorithm from [16]. Hence, the maximum a-posteriori (MAP) estimate of a task weight matrix $W \in \mathcal{W}$ for task t is given by

$$\mathrm{vec}\left(W\right) = \left(\Sigma_{\otimes} \sum_{i=1}^{n_t} {N_i^{(t)}}^T N_i^{(t)} + \sigma^2 I_{km}\right)^{-1} \left(\Sigma_{\otimes} \sum_{i=1}^{n_t} {N_i^{(t)}}^T \boldsymbol{y}_i^{(t)} + \sigma^2 \mathrm{vec}\left(M_W\right)\right), \tag{4}$$

where $N_i^{(t)} = {\boldsymbol{n}_i^{(t)}}^T \otimes I_k$ and $\Sigma_{\otimes} := \Sigma_{c;W} \otimes \Sigma_{r;W}$. Applying the iterative learning algorithm from [16] trains the parameters of the multi-variate Gaussian prior, i.e. the vectorized mean $\mathrm{vec}\left(M_W\right)$ and Kronecker covariance matrix $\Sigma_{\otimes}$. The original weight matrix W can be easily restored from the Kronecker model by reshaping the vectorized weights accordingly $W = \mathrm{unvec}\left(\mathrm{vec}\left(W\right)\right)$. The original row and column covariance matrices $\Sigma_{r;W}$ and $\Sigma_{c;W}$, respectively, may be obtained from the unvectorized weights using expectation maximization algorithms [17]. An algorithm to obtain the matrix Gaussian parameters is outlined in Algorithm 1.

 DOI: 10.3217/978-3-85125-533-1-25

Calibration-Free Correction: A prior model $\mathcal{MN}(M_W, \Sigma_{r;W}, \Sigma_{c;W})$ encodes shared characteristics across artifact regression models. In fact, the mean weight matrix M_W can be used to parameterize a matrix regression model $\boldsymbol{y} = \boldsymbol{s} + M_W\boldsymbol{n}$. A sample $(\boldsymbol{y}, \boldsymbol{n}) \in \mathbb{R}^k \times \mathbb{R}^m$ from a new session can be therefore immediately corrected by computing

$$\boldsymbol{s} = \boldsymbol{y} - M_W\boldsymbol{n} \tag{5}$$

and using $\boldsymbol{s}$ as the artifact-free sample. This correction is based on the commonalities found across artifact influences in different sessions or subjects. Hence, in contrast to the group mean approach, the MTL algorithm additionally updates the prior mean in dependency to the covariance relations.

Model Adaptation: The Bayesian setting of the MTL model proposed in this work allows for a natural adaptation to calibration data. If a calibration data set $\mathcal{D}^{(*)}$ becomes available, an adapted weight matrix $W^{(*)}$ can be inferred using the MAP estimate from (4) on the data set $\mathcal{D}^{(*)}$. The prior then acts as a regularizer towards the shared structure that is controlled by the variance factor σ^2. The correction procedure then follows (5) where we replace the prior mean M_W with the adapted weights $W^{(*)}$. The adaptation is also eligible for calibration in which new artifacts are obtained in an online setting, e.g. by thresholding techniques [18] or the Riemannian Potato [7].

Experimental Setup: We performed an evaluation of the model on real EEG signals recorded from five subjects with four to five sessions. Each subject sat on a comfortable chair in front of a screen and began the session with a five minute resting state recording (eyes open, fixating a cross on the screen). The subject then performed five to 13 runs of mental imagery with nine trials per run. As we are only interested in the artifacts of the recordings, the exact design of the imagery experiment is of no further interest in this work (however, details on the data set can be requested from the authors).

Brain activity during the experiment was recorded using EEG with 128 electrodes positioned according to the extended 10-20 system (referenced at TPP10h). The signals were sampled at 500Hz using actiCHamp amplifiers [1] and active electrodes.

The recorded EEG signals were divided into training and test segments. The five minute resting state recording was used to extract EOG blinking artifacts for model training. Each training segment consisted of a one second window containing an EOG blink artifact in the center that was automatically extracted by variance thresholding. In absence of explicitly placed EOG channels in the recordings, the artifacts were measured at two frontal electrodes (Fp1 and Fp2) designated to act as sources for EOG artifacts that measure eye blinks. The data from the experimental runs were then used as test signals for a cross-correlation evaluation between artifact sources and EEG

[1] BrainProducts GmbH, Gilching, Germany

Algorithm 1: Multi-task Kronecker Regression

Data: Training sets $\mathcal{T}$ as described in (2)
Result: Matrix-variate Gaussian prior $\mathcal{MN}(M_W, \Sigma_{r;W}, \Sigma_{c;W})$

Initialize $M_W = \mathbf{0}$, $\Sigma_{r;W} = I_k$, $\Sigma_{c;W} = I_m$;
Initialize $\mathcal{W} = \left\{W^{(t)}\right\}_{t=1}^{q}$ with $W^{(t)} = \mathbf{0}$;
Set $\Sigma_{\otimes} = \Sigma_{c;W} \otimes \Sigma_{r;W}$;
for $\mathcal{D}^{(t)} \in \mathcal{T}$ **do**
    for $\left(\boldsymbol{n}_i^{(t)}, \boldsymbol{y}^{(t)}\right) \in \mathcal{D}^{(t)}$ **do**
        Compute $N_i^{(t)} = {\boldsymbol{n}_i^{(t)}}^T \otimes I_k$
while M_W *and* $\Sigma_{\otimes}$ *not converged* **do**
    for $W^{(t)} \in \mathcal{W}$ **do**
        Train $\mathrm{vec}\left(W^{(t)}\right)$ using (4) ;
        Restore $W^{(t)} := \mathrm{unvec}\left(\mathrm{vec}\left(W^{(t)}\right)\right)$;
    Update M_W with the sample mean of the weights in $\mathcal{W}$;
    Update $\Sigma_{\otimes}$ with the sample covariance of the vectorized weights in $\mathcal{W}$;
Estimate $\Sigma_{r;W}$ and $\Sigma_{c;W}$ from $\mathcal{W}$ [17];

channels. Each test segment consisted of a four second window where the subject was either in an imagery phase (with rare eye blinks and little noise) or a pause phase (with more frequent eye blinks and more noise). In summary, the models were only trained from contaminated EEG samples, while the test segments consisted of artifact-free as well as contaminated samples. All signals were preprocessed with a common average reference and band-passed in 1-40Hz (Butterworth, order 4).

We based our evaluation on the assumption that EEG source signals are uncorrelated to artifact signals [13]. Accordingly, correction models with lower correlation between artifact and EEG were considered better. We compared the MTL regression model for artifact correction presented in this work against no correction, standard linear regression and a group mean of weight matrices. The MTL prior and group mean are transfer models and were trained from artifacts of the training segments of all subjects, but excluding the subject that was evaluated. The MTL prior model (MTL Reg (P)) was trained using Algorithm 1. The group mean model (Mean Reg) was constructed by averaging over the weight matrices trained from individual artifacts. The standard regression (Std Reg (A)) and adapted MTL regression (MTL Reg (A)) were calibrated models trained from the artifacts in the same session that was been evaluated.

The test segments were corrected with the models and the Pearson correlation coefficient between the time series of an artifact source and the cleaned EEG signals were computed. This procedure resulted in 125 normalized cross-correlation values for a total of 360 test segments (resulting in a total of 45000 correlations per artifact source). The correlation coefficients were pooled and compared according to their absolute total correlation, density

 DOI: 10.3217/978-3-85125-533-1-25

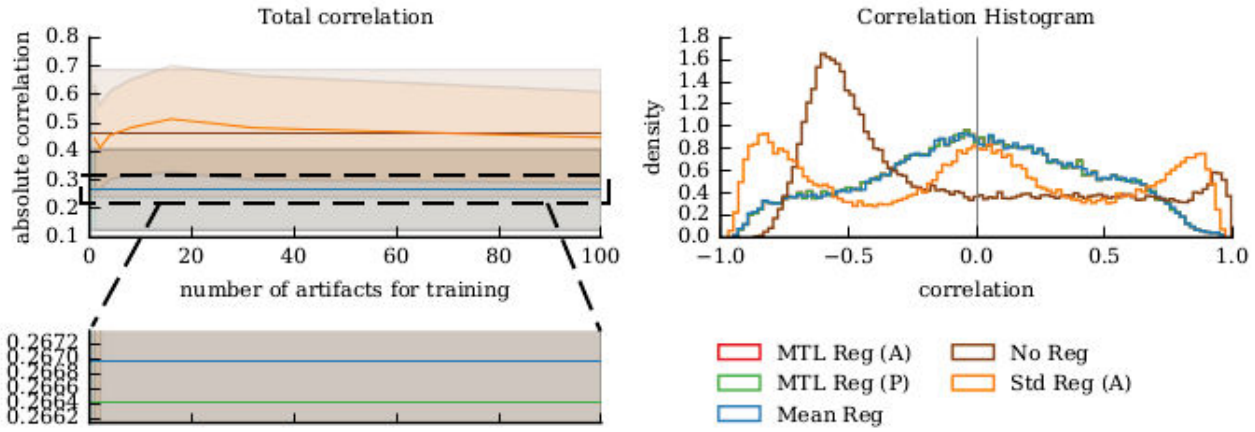

Figure 1: Correlation analysis between EOG electrode and the EEG channels. The Pearson correlation coefficient between EOG channel Fp1 and the 125 EEG channels were computed over all 360 test segments. The top and bottom left plot show the mean (solid lines) and standard deviation (shaded areas) of the total absolute correlation of each model. The bottom left plot rescaled the y-axis to highlight the differences between the transfer models. The histogram on the right shows the density distribution of the correlation coefficients for each model.

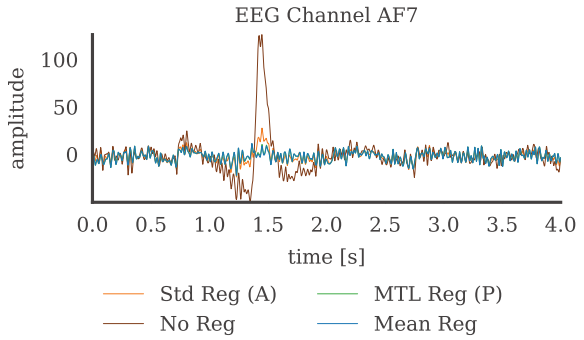

Figure 2: An exemplary comparison of the artifact correction models on a preprocessed EEG time series of four seconds recorded at channel AF7. The signal was bandpassed between 1 and 40Hz and contains an EOG blink artifact at 1.5 seconds. All artifact correction methods visibly attenuate the deflection and follow the original signal in the segments where no artifact occurs.

distribution and topographic relations. In the following section, we only present the results for EOG channel Fp1 and omit Fp2, as both channels showed identical behavior.

RESULTS

We first compared the performance of each model in terms of total correlation. Therefore, the mean correlation values over all 360 test segments was taken at each EEG electrode. Then, the mean and standard deviation of the absolute correlations were computed across all channels. The results for the different models are shown in Figure 1 (top left and bottom left). The calibrated models (suffixed by (A)) were trained on an increasing number of artifacts from the calibration session. The standard regression approach (orange) shows similar total correlation as opposed to not performing any regression at all (brown). The MTL and mean regression models show equal performance by decreasing the total correlation and variance compared to both, no and standard regression. The bottom left plot scales towards the MTL models and the mean regression. Here, the MTL prior (green) and adapted model (red) have minimally lower correlation than the mean model (blue). The MTL adaptation performance is equal to the MTL prior and does not change with more session-specific artifacts to train on.

The top right plot of Figure 1 shows the density distribution of the correlations in a histogram. Performing no regression at all (brown) exhibits a clear peak at negative correlations and a smaller at the positive tail. Standard regression (orange) also has modes at the positive and negative tails, but induces another peak around zero. The MTL models (red and green) and mean regression (blue) are again indistinguishable and centered with most correlation values around zero. Hence, the transfer models are able to keep more correlation values closer to zero than standard or no regression.

The topographic relations of the cross-correlations are depicted in Figure 3. The topographies show the correlation difference on each electrode between the two models pairs in a row and column. Red areas are positive differences indicating that the row model has lower correlation than the column model. Likewise, blue areas are negative differences and indicate that the row model has higher correlation than the column model. Performing regression lowers correlations mainly at the very frontal and occipital region. The transfer models (Mean Reg, MTL Reg (P) and MTL Reg (A)) yield reduced occipital and parietal correlations compared to standard regression (Std Reg (A)). While topographic differences between the MTL models compared to mean regression are present, there are no clear brain regions where one model outper-

 DOI: 10.3217/978-3-85125-533-1-25

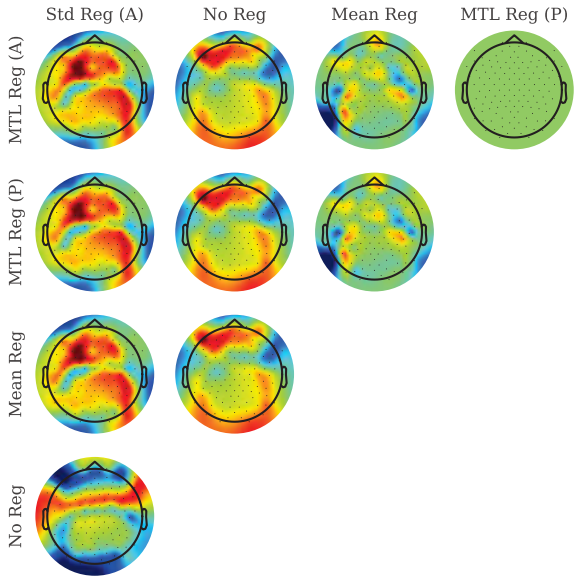

Figure 3: Topographies of the correlation difference between two models for artifact source Fp1 and the EEG channels. Red regions are positive differences and state the the row-model induces a lower correlation at that region than the column-model. Likewise, blue regions are negative differences and imply a lower correlation of the column-model compared to the row-model. Notable differences emerge for EEG channels near the frontal artifact sources (associated with No Reg) and for the frontal and occipital areas (associated with regression).

forms the other. The MTL prior model (MTL Reg (P)) and its calibrated version (MTL Reg (A)) show no differences at all.
Finally, Figure 2 shows an example of an EEG times series from a four second test segment at the frontal electrode AF7. An eye blink occurred at 1.5 seconds and after the preprocessing it is still clearly visible without EOG regression (green). Standard regression (blue) manages to reduce the amplitude, but the deflection is still visible. The transfer models (red and purple) manage to even further reduce the amplitude and seem to have visually corrected the artifact well. The corrected signals follow the original signal before and after the artifact occurred.

DISCUSSION

The results suggest that the group mean and MTL models outperform standard regression in terms of reducing correlation between artifact sources and EEG channels. In fact, standard correction seems to perform worse at some EEG channels with a varying total correlation, similar to the uncorrected signal and high negative and positive correlation modes. The differences seem to also occur at relevant brain regions within the frontal and parietal areas. A possible explanation may be that the standard regression is able to regress out artifacts well, but corrupts the signal at some channels when there is no artifact present. The transfer methods on the other hand account for variability in the artifacts, resulting in more stable regression models. These results agree with findings that group means may outperform individual regression models [13, 14].

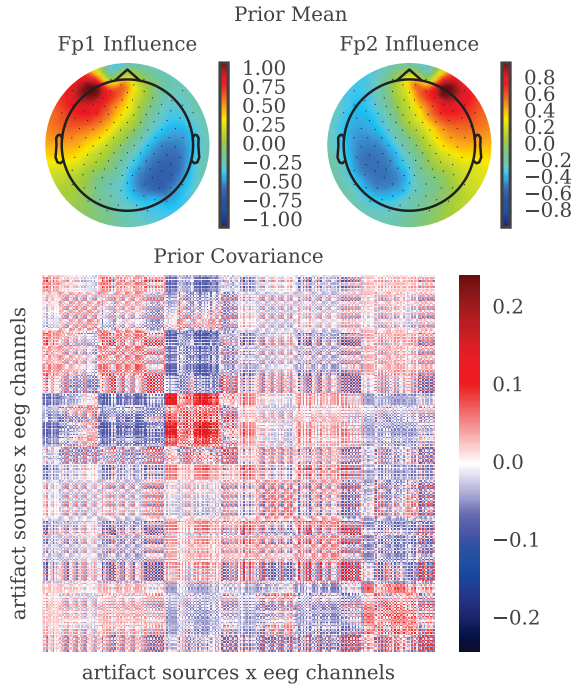

Figure 4: Visualization of the matrix Gaussian prior parameters trained by the MTL algorithm. The top two plots show topographies of the learned weights associated with the corresponding artifact source Fp1 or Fp2 on each channel. In both cases, the influence gradually decreases with increasing distance to the artifact source. The bottom plot shows a heatmap of the prior covariance matrix in Kronecker form that was estimated from the vectorized form with the standard sample covariance. The covariance matrix shows spatial block structures captured across artifacts from different subjects and sessions.

Notice that the group mean is a special case of the MTL learning algorithm where only a single prior update is performed. This equivalence led to our expectation that the MTL prior will perform at least as good as the group mean. Moreover, as the prior was used to regularize adaptation with session-specific calibration data, we expected to further increase performance as opposed to using the plain prior or group mean. Unfortunately, the MTL prior and adaptation have trained the same weights and there are only minimal differences between the MTL and group mean model. The neglectable difference in total correlation and density distribution indicate that MTL and group mean essentially trained the same weights. This conclusion is supported by the lack of correlation differences at clear brain regions shown in the topographic maps. We analyzed the MTL training process and found that the prior quickly converges within a few iterations. A possible explanation for the quick convergence may arise from the rather low dimensionality of the feature

 DOI: 10.3217/978-3-85125-533-1-25

space compared to the large amount of data points for training. Hence, the resulting MAP estimates are mainly based on the data likelihood and do not need to rely much on the prior regularization. It is further worth noticing that the MTL adaptation process to calibrate with session specific artifacts does not increase performance over the MTL prior. The covariance prior was also analyzed and showed structure captured across artifacts (see Figure 4), which implies general spatial feature directions for training. However, the captured structure did not seem to be of relevance in case of EOG artifacts, as the final prior ultimately converge to the group mean and could not be improved through calibration. A solution to this problem may have been to not only consider eye blinks, but further horizontal and vertical saccades.

CONCLUSION

This work presented a probabilistic interpretation of artifact correction that unifies inter-subject linear models and session-specific calibration. The introduced method combines influence distributions of artifact sources on EEG channels within a Bayesian MTL framework in which individual artifacts across sessions and subjects constitute the tasks. However, the final model ultimately converges to a group mean of the weight matrices, implicating that there was no additional session-specific structure across EOG artifacts that further improved performance of the model. The MTL framework has already proven to work well in the case where tasks have few data points compared to the feature dimensionality. In this sense, promising follow up work is the evaluation of this approach for other tasks, artifacts and data sets that may contain exploitable structure across artifact tasks. Further promising future work may investigate and interpret the influence structure captured by the covariance matrix. Such an analysis is likely to give further insights into the behavior of artifacts across sessions or subjects and may aid the development of new models for artifact correction or regression techniques on EEG signals.

ACKNOWLEDGEMENTS

We thank the Athena-Minerva Team for providing the data set used in this work. We further thank Natalie Faber and Daniel Tanneberg for discussions on the topic and the reviewers for their valuable feedback.

REFERENCES

(Conference proceeding)
[1] Grosse-Wentrup M, Liefhold C, Gramann K, Buss M. Beamforming in noninvasive brain–computer interfaces, IEEE Transactions on Biomedical Engineering, 2009, 56(4):1209–1219.
[2] Lotte F, Guan C. Regularizing common spatial patterns to improve BCI designs: unified theory and new algorithms, IEEE Transactions on biomedical Engineering, 2011, 58(2):355–362.
[3] Van Vliet M. Effectiveness of Automatic EOG Regression, University of Twente, Netherlands, 2006.
[4] Alamgir M, Grosse-Wentrup M, Altun Y. Multitask Learning for Brain-Computer Interfaces, AISTA, 2010, 9(5):17–24.
[5] Fiebig KH, Jayaram V, Peters J, Grosse-Wentrup M. Multi-task logistic regression in brain-computer interfaces, Proceedings of the IEEE International Conference on Systems, Man, and Cybernetics (SMC 2016), 2016.
[6] Stegle O, Lippert C, Mooij JM, Lawrence ND, Borgwardt KM. Efficient inference in matrix-variate Gaussian models with\ iid observation noise, Advances in Neural Information Processing Systems 24, 2011, 630–638.
[7] Barachant A, Andreev A, Congedo M. The Riemannian Potato: an automatic and adaptive artifact detection method for online experiments using Riemannian geometry, TOBI Workshop lV, 2013, 19–20.
(Journal)
[8] Hyvärinen A, Oja E. Independent component analysis: algorithms and applications, Neural networks, 2000, 13(4):411–430.
[9] Klados MA, Papadelis C, Braun C, Bamidis PD. REG-ICA: a hybrid methodology combining blind source separation and regression techniques for the rejection of ocular artifacts, Biomedical Signal Processing and Control, 2011, 6(3):291–300.
[10] Wallstrom GL, Kass RE, Miller A, Cohn JF, Fox NA. Automatic correction of ocular artifacts in the EEG: a comparison of regression-based and component-based method, International journal of psychophysiology, 2004, 53(2):105–119.
[11] Croft RJ, Chandler JS, Barry RJ, Cooper NR, Clarke AR. EOG correction: a comparison of four methods, Psychophysiology, 2005, 42(1):16–24.
[12] Fatourechi M, Bashashati A, Ward RK, Birch GE. EMG and EOG artifacts in brain computer interface systems: A survey, Clinical neurophysiology, 2007, 118(3):480–494.
[13] Verleger R, Gasser T, Möcks J. Correction of EOG Artifacts in Event-Related Potentials of the EEG: Aspects of Reliability and Validity, Psychophysiology, 1982, 19(4):472–480.
[14] Gasser T, Sroka L, Möcks J. The correction of EOG artifacts by frequency dependent and frequency independent methods, Psychophysiology, 1986, 23(6):704–712.
[15] Schlögl A, Keinrath C, Zimmermann D, Scherer R, Leeb R, Pfurtscheller G. A fully automated correction method of EOG artifacts in EEG recordings, Clinical neurophysiology, 2007, 118(1):98–104.
[16] Jayaram V, Alamgir M, Altun Y, Schölkopf B, Grosse-Wentrup M. Transfer learning in brain-computer interfaces, IEEE Computational Intelligence Magazine, 2016, 11(1):20–31.
[17] Dutilleul P. The MLE algorithm for the matrix normal distribution, Journal of statistical computation and simulation, 1999, 64(2):105–123.
[18] Nolan H, Whelan R, Reilly RB. FASTER: fully automated statistical thresholding for EEG artifact rejection, 2010, 192(1):152–162.