

Towards Classifier Visualisation in 3D Source Space

Laurens R. Krol
Biological Psychology and
Neuroergonomics
TU Berlin
Berlin, Germany
Email: lrkrol@gmail.com

Mahta Mousavi
Electrical and
Computer Engineering
UC San Diego
La Jolla, CA, USA
Email: mahta@ucsd.edu

Virginia R. de Sa
Cognitive Science
UC San Diego
La Jolla, CA, USA
Email: desa@ucsd.edu

Thorsten O. Zander
Zander Laboratories
Amsterdam, the Netherlands
Email:
thorsten@zanderlaboratories.com

Abstract—In the context of brain-computer interfacing, it is important to investigate what regions of the brain a classifier focuses on. For one, this will clarify to what extent the classifier relies on brain activity, as opposed to undesirable non-cortical signals. More generally, the practice is informative as it allows conclusions to be drawn about the cortical regions—and thus, cortical functions—that contribute to the effect under investigation. In this study, we start to investigate different methods to visualise the regions of interest of classifiers based on windowed means and on common spatial patterns. Specifically, we take individually reconstructed source spaces and transform the classifier filter weights into relevance weights indicating the relative contribution of each source to the classifier. This is visualised across participants in an average brain. By decomposing the classifier weights into separate sources and localising these in the brain, this method provides a tool to evaluate classifiers and test hypotheses.

I. INTRODUCTION

A brain-computer interface (BCI) allows an output channel to be established from a user’s brain to a computer—an output channel “that is neither neuromuscular nor hormonal” [1]. Such a channel can be used in various ways. For example, it allows paralysed or locked-in patients to communicate with the outside world using mental spellers [2] and brain-actuated prostheses [3]. BCI-based systems enable people to control such and other devices using only their brain activity.

A *passive* brain-computer interface (pBCI) [4] is a BCI system that uses similar hard- and software in order to interpret ongoing, “natural” brain activity [5] that is not meant to control a device. Instead, such brain activity reflects the human user’s cognitive or affective state. Passive BCI-based quantifications of mental states are used as *implicit input* to support ongoing human-computer interaction [6].

With recent trends in neuroadaptive technology and pBCI itself [7], [8], [9], [10], as well as advances in signal acquisition hardware [11], such real-world pBCI applications have become increasingly close at hand. In particular, electroencephalography (EEG) has become increasingly mobile [12]. An important issue with EEG however, is that many different sources of activity combine to form the final signal measured at the scalp. This includes not just numerous cortical sources, but all electromagnetic activity present in the body as well as in the environment. Notably, eye and muscle artefacts can contaminate the data.

During EEG experiments in realistic contexts, a great amount of eye and muscle activity is to be expected. Furthermore, this activity can be highly correlated to the experimental conditions. A classifier trained to distinguish between data reflecting different conditions can thus be highly influenced by such non-brain activity. Because of this, it is important to verify to what extent an advertised *brain*-computer interface system is indeed based on *brain* activity. For example, when a system is intended to measure negative affect from brain activity, it may inadvertently use activity from the corrugator muscle of the face [13]. Such a system may not be as reliable in different contexts (e.g. where sunshine leads to excessive squinting), and may be more accurately categorised as a form of physiological computing [14].

While detection of non-brain influences is one specific case where extended feature analysis can provide important answers, understanding the relative contributions of different brain sources is another. A classifier is trained to distinguish between different conditions using any all information available. This makes it a powerful method to distinguish between brain activity in different conditions [15], but also means that the classifier can be influenced by a variety of cognitive processes. For example, even when targeting a specific cognitive process such as motor imagery, activity from other processes may interfere with the recordings [16] and affect the classifier. Knowing what information carries what weight can help to validate the classifier and the experimental paradigm. If the cortical areas that a classifier focuses on can be identified, this may provide insights into the different cognitive processes underlying the investigated effects. When a classifier’s primary source of distinguishing brain activity comes from the visual cortex, for example, this may hint that low-level sensory aspects of the stimuli were more important than further cognitive interpretation of those stimuli. Such insights can also be used to improve experimental paradigms and BCI training methods [17].

In this paper we explore, with simulated data, aspects of a method used earlier [18] to visualise in a three-dimensional head volume the areas of the brain a BCI classifier focuses on. The original method applied only to windowed-means classifiers of event-related potentials [19]; here we describe an adaptation of that method, as well as an extension that applies to common spatial patterns [20].

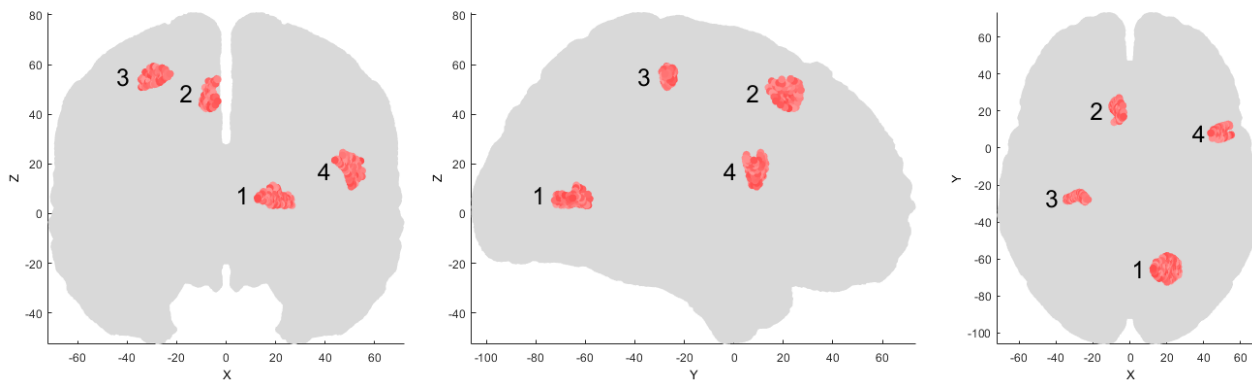


Fig. 1. Ground-truth locations in the brain of the simulated class-dependent signals, covering the 1.5 cm maximum deviation. In data set 1, source numbers 1 and 2 generated ERPs in one class, but not in the other. In data set 2, 1 and 2 generated alpha activity in one class, while 3 and 4 did so in the other.

II. METHODS

A. Data Simulation

In order to have a known ground truth to evaluate the method, we used simulated EEG data. This was generated using SEREEGA¹, an open-source toolbox dedicated to simulating event-related EEG activity. Using the New York Head model and lead field [21], two sets of 64-channel data were simulated. Each data set consisted of two conditions (classes) for the classifier to distinguish.

a) *Data set 1:* This data set contained systematic class differences in the temporal domain. One class consisted entirely of brown noise emanating from 64 sources spaced throughout the brain. The second class consisted of equally arranged noise, along with two event-related potentials (ERPs) emanating from two selected sources, respectively (numbered 1 and 2 in Figure 1). The second source’s ERP had an average amplitude four times greater than the other; however, this larger-amplitude ERP only occurred randomly in 25% of epochs of that class. Single epoch amplitudes varied by 20%. Both peaked at $300 \text{ ms} \pm 50$ with a width of 200 ± 40 .

b) *Data set 2:* This data set contained systematic class differences in the spectral domain. Both classes contained brown noise emanating from 64 sources spaced throughout the brain. Furthermore, each class contained additional uniform white noise filtered between 8 and 14 Hz, emanating from two sources each. In both classes, one source’s signal amplitude was twice that of the other. (From Figure 1, source numbers 1 and 2 were active in one class; 3 and 4 in the other.)

Note that the sources selected here do not reflect any deliberately chosen functional region. For the purposes of this paper, the selected regions’ neuroscientific significance is not relevant: we merely wish to reconstruct their locations.

Each data set mentioned above consisted of 10 simulated ‘participants’. The non-noise brain source locations were relatively consistent across these participants, differing up to 1.5 cm across participants. (Note that a small shift in position can lead to a significant change in projection as, due to the

cortical folds, the source will be oriented differently relative to the scalp.) The other sources were randomly distributed for each participant. A total of 100 epochs of 800 ms each were simulated for each class.

The signal-to-noise ratio was controlled in such a way that a cross-validated estimate of the classifier accuracy would roughly average 75–80%.

B. Classifier Visualisation in Putative Source Space

1) *Separation of the Data into Sources:* We wish to visualise a classifier in putative source space. To that end, independent of any BCI classifier, the data must first be separated into putative sources. Different methods exist for this purpose. In this paper, we use independent component analysis (ICA) [22] as an example. ICA is a blind source separation method that decomposes the data x into statistically maximally independent sources s by finding a transformation or ‘unmixing’ matrix A such that $x = As$. A is a filter matrix weighting the individual channel activations in sensor space to obtain the identified independent component activations. Inversely, A^{-1} contains the forward model of these components, i.e., their projections onto the scalp. Other source separation methods will produce different contents of A , but their application in this method is essentially the same.

Independent EEG sources are dipolar [23]. We can thus fit an equivalent dipole model to the forward ICA decomposition, e.g. using the EEGLAB toolbox DIPFIT 2.3 [24]. The dipole model provides a 3D localisation for each independent component that minimises the residual variance between the dipole and the component projections.

For the results in this paper, since we used simulated data, we did not calculate A on the data. Instead, we obtained the ground-truth mixing matrix directly from the simulation. The equivalent dipole model however was calculated separately using DIPFIT.

2) *Obtaining a Classifier:* The method presented here applies to LDA-based spatio-temporal classifiers. In particular, we present results for common spatial patterns and the windowed-means approach.

¹Krol et al., in submission. See <https://github.com/lrkrol/SEREEGA>

In the windowed-means approach (WM) [19], the mean amplitudes of the scalp activations at each electrode and in each time window are extracted as features, on which a shrinkage linear discriminant analysis (LDA) is trained to separate the classes. This results in LDA filter weights $w_{\text{WM}} = \Sigma_{\text{WM}}^{-1}(\mu_1 - \mu_2)$ where μ_1 and μ_2 are the mean of the features for classes 1 and 2 respectively and Σ_{WM} is the common class covariance.

Common spatial patterns (CSP) [20] are used to extract features in the frequency domain. CSP finds the optimal set of filter weights that maximise the variance of the filtered signal for one class while simultaneously minimising it for the other. Usually the top $m < \frac{C}{2}$ filters for each class are selected, where C is the number of channels. The data in each epoch is spatially filtered to this new pseudo-channel space and the log of the signal's variance on each pseudo-channel is selected as the set of features for the classifier. We then apply a shrinkage LDA, resulting in LDA filter weights w_{CSP} .

3) *Obtaining a Classifier's Forward Model:* In case of the WM classifier, the LDA filter weights cannot be neuro-physiologically interpreted. As Haufe et al. explain, "classifier weights can exhibit small amplitudes for measurement channels containing the signal-of-interest, but also large amplitudes at channels not containing this signal" [25]. Therefore, we must first transform the LDA filter weights into *patterns*. For this paper, we did this by multiplying the LDA filter weights of the backward model by the regularised LDA's common covariance matrix (as opposed to the non-regularised version [25]). Thus, the patterns $p_{\text{WM}} = \Sigma_{\text{WM}} w_{\text{WM}} = \mu_1 - \mu_2$ show the differential scalp activity between classes.

For the CSP classifier, let CSP filters be columns of the matrix $W \in \mathbb{R}^{C \times C}$. CSP patterns are then defined to be columns of $P = (W^{-1})^T \in \mathbb{R}^{C \times C}$. Essentially, P and W are the forward and backward models, respectively. However, not each of the selected CSP filters contributes equally to classification. Their respective contributions depend on the LDA weights that were trained on their features. Here, the same LDA filter weight issues apply that were mentioned above. We thus transform these into interpretable weightings or forward weights indicating their relative contributions: $\tilde{w}_{\text{CSP}} = \Sigma_{\text{CSP}} w_{\text{CSP}}$ where $\Sigma_{\text{CSP}} \in \mathbb{R}^{2m \times 2m}$ is the common covariance of the CSP features from classes 1 and 2. We use these forward weights to scale the CSP patterns for visualisation: let $P_{\text{sel}} \in \mathbb{R}^{C \times 2m}$ be the CSP patterns corresponding to the $2m$ selected CSP filters, then the i -th column in p_{CSP} is the i -th column in P_{sel} scaled by the i -th element in \tilde{w}_{CSP} .

We now have patterns representing the forward model for these classifiers in the form of weights for each electrode. For WM, we have one pattern for each time window. For CSP, we have m weighted patterns for each class.

4) *Projecting the Classifier Patterns into Source Space:* The ICA's unmixing matrix A^{-1} transforms sensor-level scalp activations into source activations. Similarly, it can transform sensor-level weights into corresponding weights in source space. When patterns contain class-relevant scalp projections, projecting these into source space gives us the *relevance*

weights $w_r = |A^{-1}p|$ that indicate how the different source activations linearly combine to create these patterns. In other words, the relevance weights reflect to what extent the sources contribute to the patterns. By extension, the relevance weights thus reflect to what extent the sources contribute to the class differences. For our purpose, the sign of the weights is not relevant at this point, hence we take the absolute value.

In case of WM, we now have one set of relevance weights per time window per participant. In case of CSP, we have $2m$ sets per participant. Each set contains as many weights as we have independent components for that participant.

5) *Alternative Source Weights:* Using different methods can give different perspectives on the data. We have explained how sensor-level patterns can be transformed to obtain source-level relevance weights. Alternatively, source-level weights can be calculated in source space directly, independent of the classifier. For example, we can use Pearson correlation to obtain a set of weights for the WM case.

To that end, we first repeat the feature extraction in source space, i.e., we obtain the mean of each source activation in each time window. We then calculate the correlation coefficient between these source features F_s and the vector of true class labels L to obtain a set of weights $w_{\text{WMcorr}} = |\text{corr}(F_s, L)|$. The sign is irrelevant for our purpose.

6) *Visualising Relevance Weights in Source Space:* In case of WM, we generate one visualisation per time window, each illustrating the sources contributing to the class differences at that time. In case of CSP, we generate one per class, illustrating the sources distinctive for that class.

For each time window or class, respectively, the obtained relevance weights are distributed to the dipole locations of the corresponding sources for each participant. We then generate a weighted 3D kernel density plot containing these weights for all participants. For this, we use the EEGLAB plug-in `dipoleDensity v0.36` [26] which aligns the output to slices of the mean MNI brain. For the figures in this paper, we used a smoothing kernel of 12 mm.

III. RESULTS

Figure 4 shows the output of the presented method for one case. The top left shows the sorted distribution of the relevance weights. The other panels show different slices of the mean MNI brain along with the colour-coded weighted dipole density, or 'relevance density'. The sparse distribution indicates that a relatively small number of sources receive a relatively high percentage of weights. In other words, the method shows that the differences between the classes can be traced back to a relatively specific area in the brain.

Figure 4 visualises class 1 of the CSP classifier calibrated on data set 2. We see that the most dense, i.e. the most relevant area is near source number 1 in figure 1. A second, weaker area of relevance is near source number 2. This accurately reflects the simulation of data set 2, where class 1 was defined by activity in these two sources, with the signal amplitude of source number 1 twice that of source number 2.

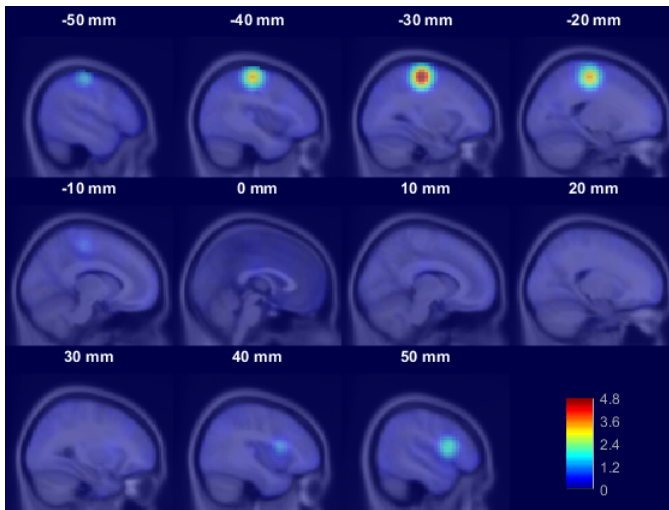


Fig. 2. Partial visualisation (sagittal section) of the CSP classifier: class 1.

To preserve space, the other figures only show the sagittal density slices. Figure 2 visualises the second class of data set 2, as identified by the CSP classifier. This accurately reflects the correct locations of source numbers 3 and 4.

Figure 5 visualises the regions of interest of the LDA classifier calibrated on data set 1 in a selected time window. This, too, accurately reflects the ground truth that sources in the right-occipital and left-frontal lobes (i.e. numbers 1 and 2 in figure 1) were generators of the class differences. Notably, we see that the sources are weighted roughly equally, with source 2 weighted slightly more than source 1. This is in line with the signals that were generated by these two sources. The amplitude of source 2 was four times that of source 1. However, the probability of a signal occurring at all in source 2 was only 25%, whereas source 1 was active in each simulated trial. The mean amplitude of these sources over all trials was thus roughly the same, but their predictive value was not.

Figure 6 visualises the alternative weights for data set 1, described in section II-B5. We see a high correlation density for source number 1 and a significantly lower density for source number 2 reflecting the difference in predictive value between sources 1 and 2.

IV. SUMMARY

We simulated two data sets of 10 simulated participants each. Each participant’s data was simulated with unique locations for the generator sources, with some consistency maintained for the sources that generated the class differences. In one data set, these class differences were caused by the presence of event-related potentials from two sources. In the other, differences were caused by the presence of alpha-band activity in different sources. An ICA solution was available for the simulated data.

Two different classifiers were trained on the two data sets respectively: A CSP classifier, focusing on the spectral differences between the classes in data set 2, and a WM classifier, focusing on temporal differences in data set 1.

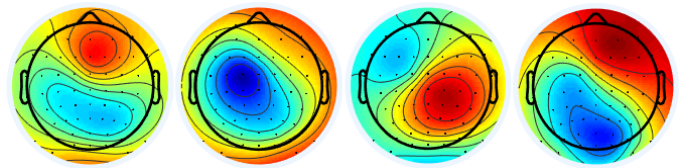


Fig. 3. Combined projection patterns of the two class-dependent sources in data set 1, for four of the simulated participants.

For the WM classifier, we transformed the LDA filter weights into patterns representing the forward model. For the CSP classifier, we separated the produced CSP patterns and weighted them by computed forward weights.

We then transformed these patterns into relevance weights connected to independent components using the unmixing matrix of each participant’s data’s ICA solution. Since a 3D position in the brain was calculated for each of these independent components, we were able to generate a ‘relevance density plot’ indicating the classifiers’ regions of interest. As we used simulated EEG data, we compared the obtained results to the known ground truth and could verify that the generated plots corresponded to the original sources.

For the WM classifier, we also presented an alternative method to calculate relevance weights directly in source space. This method can provide a different perspective on the brain dynamics underlying the class differences.

V. DISCUSSION

In our earlier work [18] we presented a combination of the two WM-based methods discussed here, where the classifier’s regions of interest were additionally weighted by those regions’ class-correlation. In this paper we simulated data that highlights how these two methods can provide different perspectives when used separately.

The use of simulated data enabled us to compare the method’s results to a known ground truth. We see that under these circumstances, the method accurately recovers the correct sources in the brain. Simulated data also allowed us to use a ground-truth ICA decomposition to initially control for the varying results that different ICA methods provide. In future work, we will apply the method to simulated data with ICA decompositions of varying quality.

Since estimating the covariance matrix is a fundamental step in producing the patterns, we will further investigate the influence of different covariance estimation methods on the patterns and their projections into source space. The previous iteration of this method has already been shown to produce results in line with hypotheses on real EEG data [18], and we will continue this work by validating the current method on other real EEG recordings. We will also extend the method to filter-bank CSP [27], and make all code available for free.

The method presented here visualises the areas in the brain that the classifier focuses on, for two popular classification methods. The patterns that can be obtained from these classifiers can be neurophysiologically interpreted on their own

[25]. However, the current method provides two additional advantages. First of all, the patterns are decomposed into individual sources. A single pattern can consist of the combined projections of any number of different sources, and different source projections can interfere with each other to the point of making interpretation difficult. For example, figure 3 shows the combined projection patterns of the two class-dependent sources in data set 1 for four of the simulated participants. As we can see, minor changes in source location and orientation produce large differences in the projection pattern. From these patterns, it is not obvious that they are produced by two distinct sources, let alone their location. With accurate ICA models, we can untangle this relation and show individual source contributions to the classifier. Secondly, this method uses an equivalent dipole model to visualise the sources in 3D space. Projection patterns coming from roughly the same cortical area can vary markedly between participants due to anatomical differences. A 3D visualisation of the cortical areas corrects for these differences in a way that e.g. a mean projection pattern cannot. We visualise the combined relevance weights of all participants to highlight the most consistently relevant areas.

It is important to be able to perform an inspection of a classifier's regions of interest, and compare the results to our hypotheses and other perspectives on the data. When we design a BCI application, we hypothesise what functions (and thus, what regions) of the brain will be targeted. Visualisation methods such as this one enable us to compare a classifier's actual regions of interest to these hypotheses, and validate our assumptions—and to gather new insights about the cortical processes underlying the observed effects.

ACKNOWLEDGEMENTS

This work was supported by the Deutsche Forschungsgemeinschaft (ZA 821/3-1), a Short-Term Research Grant awarded to MM by the German Academic Exchange Service, and the National Science Foundation (IIS 1528214).

REFERENCES

[1] J. R. Wolpaw and E. W. Wolpaw, "Brain-computer interfaces: something new under the sun," in *Brain-Computer Interfaces: Principles and practice*, J. R. Wolpaw and E. W. Wolpaw, Eds. Oxford, UK: Oxford University Press, 2012, pp. 3–12.

[2] N. Birbaumer, N. Ghanayim, T. Hinterberger, I. Iversen, B. Kotchoubey, A. Kübler, J. Perelmouter, E. Taub, and H. Flor, "A spelling device for the paralysed," *Nature*, vol. 398, no. 6725, pp. 297–298, 1999.

[3] G. R. Müller-Putz and G. Pfurtscheller, "Control of an electrical prosthesis with an SSVEP-based BCI," *IEEE Transactions on Biomedical Engineering*, vol. 55, no. 1, pp. 361–364, 2008.

[4] T. O. Zander and C. A. Kothe, "Towards passive brain-computer interfaces: applying brain-computer interface technology to human-machine systems in general," *Journal of Neural Engineering*, vol. 8, no. 2, p. 025005, 2011.

[5] L. R. Krol, L. M. Andreessen, and T. O. Zander, "Passive Brain-Computer Interfaces: A Perspective on Increased Interactivity," in *Brain-Computer Interfaces Handbook: Technological and Theoretical Advances*, C. S. Nam, A. Nijholt, and F. Lotte, Eds. Boca Raton, FL, USA: CRC Press, 2018, pp. 69–86.

[6] T. O. Zander, J. Brönstrup, R. Lorenz, and L. R. Krol, "Towards BCI-based Implicit Control in Human-Computer Interaction," in *Advances in Physiological Computing*, S. H. Fairclough and K. Gilleade, Eds. Berlin, Germany: Springer, 2014, pp. 67–90.

[7] L. R. Krol and T. O. Zander, "Passive BCI-based neuroadaptive systems," in *Proceedings of the 7th Graz Brain-Computer Interface Conference 2017*, 2017, pp. 248–253.

[8] T. O. Zander, K. Shetty, R. Lorenz, D. R. Leff, L. R. Krol, A. W. Darzi, K. Gramann, and G.-Z. Yang, "Automated task load detection with electroencephalography: Towards passive brain-computer interfacing in robotic surgery," *Journal of Medical Robotics Research*, vol. 2, no. 1, p. 1750003, 2017.

[9] L. R. Krol, S.-C. Freytag, M. Fleck, K. Gramann, and T. O. Zander, "A task-independent workload classifier for neuroadaptive technology: Preliminary data," in *2016 IEEE International Conference on Systems, Man, and Cybernetics (SMC)*, 2016, pp. 003 171–003 174.

[10] T. O. Zander and L. R. Krol, "Team PhyPA: Brain-computer interfacing for everyday human-computer interaction," *Periodica Polytechnica Electrical Engineering and Computer Science*, vol. 61, no. 2, pp. 209–216, 2017.

[11] T. O. Zander, L. M. Andreessen, A. Berg, M. Bleuel, J. Pawlitzki, L. Zawallich, L. R. Krol, and K. Gramann, "Evaluation of a dry EEG system for application of passive brain-computer interfaces in autonomous driving," *Frontiers in Human Neuroscience*, vol. 11, p. 78, 2017.

[12] T. R. Mullen, C. A. E. Kothe, Y. M. Chi, A. Ojeda, T. Kerth, S. Makeig, T.-P. Jung, and G. Cauwenberghs, "Real-time neuroimaging and cognitive monitoring using wearable dry EEG," *IEEE Transactions on Biomedical Engineering*, vol. 62, no. 11, pp. 2553–2567, 2015.

[13] J. T. Larsen, C. J. Norris, and J. T. Cacioppo, "Effects of positive and negative affect on electromyographic activity over zygomaticus major and corrugator supercilii," *Psychophysiology*, vol. 40, no. 5, pp. 776–785, 2003.

[14] S. H. Fairclough, "Fundamentals of physiological computing," *Interacting with computers*, vol. 21, no. 1, pp. 133–145, 2009.

[15] E. Noh and V. de Sa, "Discriminative dimensionality reduction for analyzing EEG data," in *Proceedings of the 36th Annual Meeting of the Cognitive Science Society*, 2014, pp. 1090–1095.

[16] M. Mousavi, A. S. Koerner, Q. Zhang, E. Noh, and V. R. de Sa, "Improving motor imagery BCI with user response to feedback," *Brain-Computer Interfaces*, vol. 4, no. 1–2, pp. 74–86, 2017.

[17] M. Mousavi and V. R. de Sa, "Towards elaborated feedback for training motor imagery brain computer interfaces," in *Proceedings of the 7th Graz Brain-Computer Interface Conference 2017*, 2017, pp. 332–337.

[18] T. O. Zander, L. R. Krol, N. P. Birbaumer, and K. Gramann, "Neuroadaptive technology enables implicit cursor control based on medial prefrontal cortex activity," *Proceedings of the National Academy of Sciences*, vol. 113, no. 52, pp. 14 898–14 903, 2016.

[19] B. Blankertz, S. Lemm, M. S. Treder, S. Haufe, and K.-R. Müller, "Single-trial analysis and classification of ERP components – a tutorial," *NeuroImage*, vol. 56, no. 2, pp. 814–825, 2011.

[20] H. Ramoser, J. Müller-Gerking, and G. Pfurtscheller, "Optimal spatial filtering of single trial EEG during imagined hand movement," *IEEE Transactions on Rehabilitation Engineering*, vol. 8, no. 4, pp. 441–446, 2000.

[21] Y. Huang, L. C. Parra, and S. Haufe, "The New York Head—a precise standardized volume conductor model for EEG source localization and tES targeting," *NeuroImage*, vol. 140, pp. 150–162, 2016.

[22] A. J. Bell and T. J. Sejnowski, "An information-maximization approach to blind separation and blind deconvolution," *Neural Computation*, vol. 7, no. 6, pp. 1129–1159, 1995.

[23] A. Delorme, J. Palmer, J. A. Onton, R. Oostenveld, and S. Makeig, "Independent EEG sources are dipolar," *PLoS ONE*, vol. 7, no. 2, p. e30135, 2 2012.

[24] R. Oostenveld and A. Delorme. (2003) DIPFIT 2.x. [Online]. Available: <http://sccn.ucsd.edu/eeglab/plugins>

[25] S. Haufe, F. Meinecke, K. Görgen, S. Dähne, J.-D. Haynes, B. Blankertz, and F. Bießmann, "On the interpretation of weight vectors of linear models in multivariate neuroimaging," *NeuroImage*, vol. 87, no. 0, pp. 96–110, 2014.

[26] M. Miyakoshi. (2003) dipoleDensity. [Online]. Available: http://sccn.ucsd.edu/wiki/Plugin_list_process

[27] K. K. Ang, Z. Y. Chin, H. Zhang, and C. Guan, "Filter bank common spatial pattern (FBCSP) in brain-computer interface," in *IEEE International Joint Conference on Neural Networks (IJCNN), 2008*, June 2008, pp. 2390–2397.

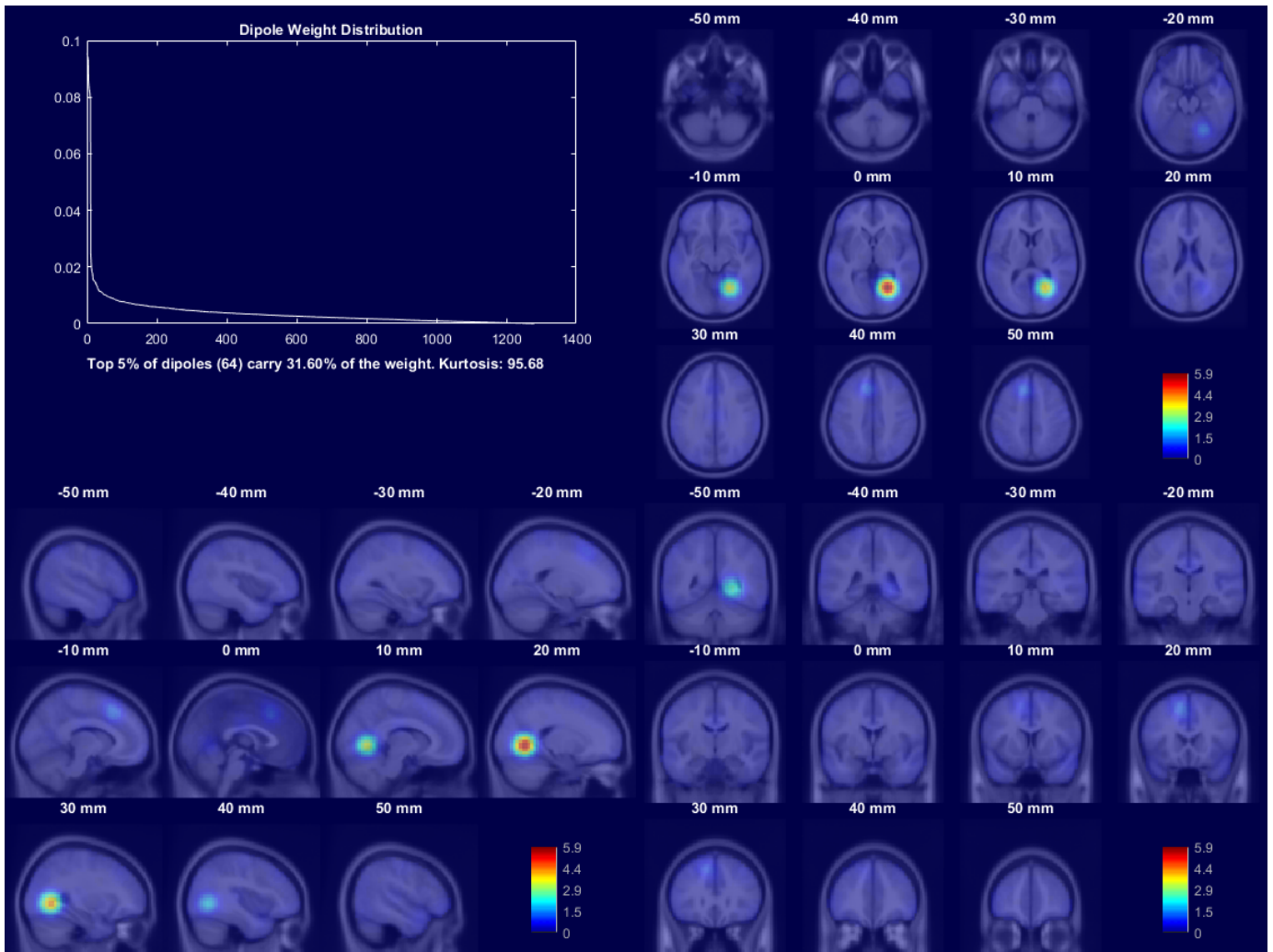


Fig. 4. Visualisation of the CSP classifier: class 2. Slices are labelled with their corresponding MNI coordinates. Top left: sorted dipole weight distribution.

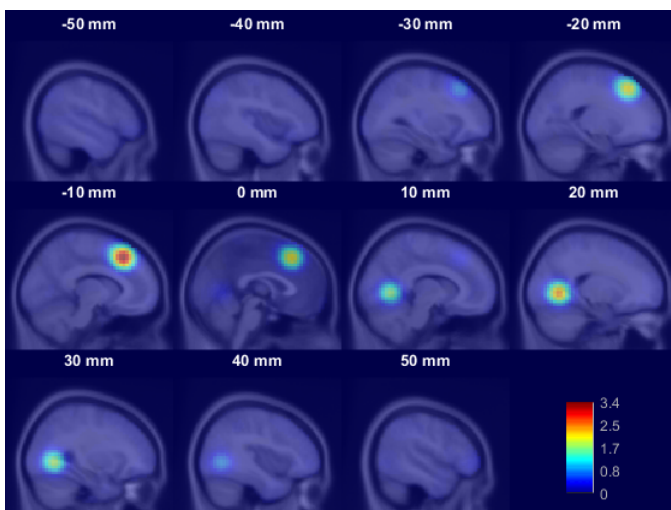


Fig. 5. Partial visualisation (sagittal section) of the WM classifier: projected LDA weights.

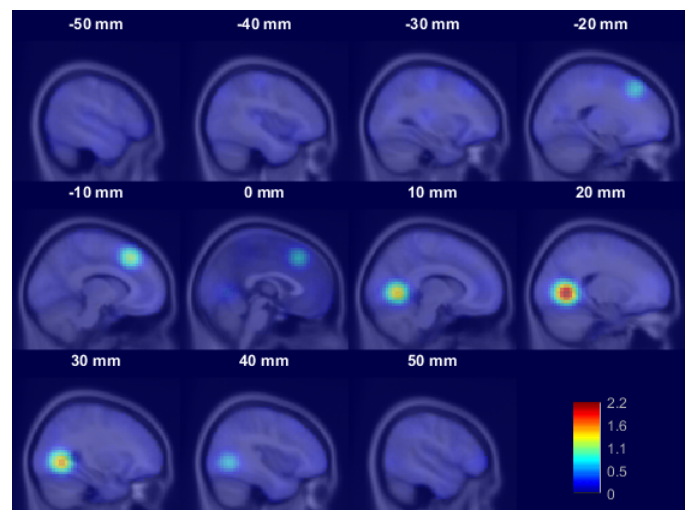


Fig. 6. Partial visualisation (sagittal section) of source feature class correlation in data set 1: source correlation weights.