Neural component analysis: source localisation for motor imagery classification

Ian Daly[1] and Milan Rybář[1]

***Abstract*— The electroencephalogram (EEG) records a summed mixture of multiple sources of neural activity distributed throughout the brain. Source separation methods aim to un-mix the EEG in order to recover activity generated by the original sources. However, most current state-of-the-art source separation methods do not take into account the physical locations of sources of EEG activity.**

We present a new source separation method which uses an accurate model of the head to un-mix the EEG into individual sources based on their physical locations.

We apply our method to an EEG dataset recorded during motor imagery and show that it is able to identify sources that are located in distinct physical regions of the brain. We compare our method to independent component analysis and show that our sources have higher spatial specificity and, furthermore, allow higher classification accuracies (a mean improvement in accuracy of 8.6 % was achieved $p = 0.039$).

I. INTRODUCTION

The electroencephalogram (EEG) records electrophysiological activity originating from cortical neurons [1]. A single EEG electrode records activity that arises from multiple distinct sources of activity distributed throughout the brain. These sources can include both neural activity and other physiological sources, such as muscle and eye movements, and non-physiological sources, such as environmental noise [2]. Un-mixing theses sources from the EEG presents an ill-posed problem [3].

Nonetheless, if additional information about the mixing process is available, or appropriate assumptions are made, source separation is very useful for EEG analysis. If the EEG can be reliably separated into sources (either neural sources or other sources from outside of the brain) this can aid analysis of underlying neural processes.

Source separation is particularly useful in brain-computer interfacing (BCI). BCIs are devices that translate brain activity into control commands for an external device, such as a communications interface or a mobility aid [4]. Source separation methods that increase the signal to noise ratio of sources related to attempted BCI control have the potential to greatly improve the performance of many BCI systems.

One of the most widely used source separation methods for EEG analysis is, arguably, independent component analysis (ICA), a blind source separation method [3]. Blind source separation (BSS) methods assume that the observed EEG arises from a linear mixture of sources. This can be modelled via $x = AS$, where x denotes the observed EEG, A denotes a linear mixing matrix, and S denotes the original neural sources of activity [3].

Source separation is than encapsulated in the process of identifying the inverse mixing matrix A^{-1}, from which the original sources S may be recovered via $S = A^{-1}x$.

BSS methods, therefore, comprise a set of algorithms that attempt to estimate the de-mixing matrix A via some reasonable assumptions.

ICA attempts to estimate A^{-1} by making the assumption that the sources are statistically independent in the time domain. Multiple variants of the matrix A^{-1} are then explored and the variant of A^{-1} that maximises the statistical independence of the estimated sources is assumed to be the optimal de-mixing matrix [3].

ICA is widely used in EEG analysis, including analysis of neural correlates of cognitive events, such as event-related potentials (ERPs) [5] and EEG artefact removal [6]. ICA is also widely used in BCI systems to separate task-related EEG activity from background noise [7].

However, ICA is limited by its assumption of statistical independence of sources in the time domain. In reality many neural sources form integrated complex networks of inter-related brain regions that act in communication with one another to achieve cognition [8]. Thus, different brain regions are highly unlikely to exhibit activity that is truly statistically independent from one another. Furthermore, by imposing a statistical independence criteria, ICA is at risk of excluding physiologically meaningful sources of neural activity simply because they exhibit activity that is statistically related to other, less meaningful, sources or artefacts.

As an example, movement control typically involves a functional network of brain regions, including, but not limited to, M1 and the supplementary motor area (SMA) [9]. These brain regions are spatially separate, but, during movement control, exhibit neural activity that is statistically related. Thus, ICA may describe the activity in this multi-region network via a single source.

When ICA is applied to separate EEG into sources related to specific cognitive processes - for example, for the purpose of aiding classification in BCI - this limitation can lead to identification of sub-optimal components.

Thus, there is a need for a source separation technique that is able to un-mix EEG into sources that originate from physically separate brain regions, even in the case of high levels of correlation in the time domain between the regions.

Our proposed method, neural component analysis, seeks to do this. We propose a BSS method that seeks a de-

*This work was not supported by any organization

[1]Ian Daly and Milan Rybář are both with the Brain-Computer Interfacing and Neural Engineering Laboratory, Department of Computer Science and Electronic Engineering, University of Essex, Wivenhoe Park, Colchester, Essex CO4 3SQ, UK {i.daly, milan.rybar} @essex.ac.uk

mixing matrix that separates the EEG into components that are maximally independent in the spatial domain.

In this manuscript we first describe our proposed method. We then investigate how our method performs separating EEG recorded during a motor imagery task into sources that originate in physically distinct parts of the brain. We show how these identified sources have high explanatory power and, as an example application, compare the classification results to those achieved with sources identified via ICA.

II. METHODS

A. Proposed method

We start from the same initial assumptions of a linear mixing process used by BSS methods and model the observed EEG as a linear mixture of sources via

$$x = BS, \tag{1}$$

where S denotes the sources, B our mixing matrix, and x observed EEG. However, unlike BSS methods we modify the criterion to identify the mixing matrix B to include information about physical locations of the estimated sources.

To estimate the distribution of physical locations contributing to each source we use "Exact low resolution brain electromagnetic tomography" (eLORETA) [10]. This allows us to estimate the contribution to each source from a set of discrete volume locations in a model of the brain. Specifically, for each estimated mixing matrix B we estimate the scalp projection onto the electrodes $[1, ..., N]$ for each of the corresponding estimated sources $\hat{S}^i_T$ via

$$P^i = B\hat{S}^i \tag{2}$$

where i denotes the index of the source $i \in [1, ...N]$, P^i denotes the projection of the i'th source onto the N EEG channels (defined as $P^i_{N \times T}$, where T denotes the length of the EEG recording), and $\hat{S}^i$ denotes the i'th estimated source and is also of dimensions $N \times T$. Specifically, the i'th row of the $N \times T$ matrix $\hat{S}^i$ contains the time series of the estimated sources, while all other rows are set to zero. We then use eLORETA to estimate the contribution to this projection from every candidate source location in a model of the head [10].

The eLORETA method allows us to find the inverse solution to the source identification problem. Specifically, eLORETA uses a weighted minimum norm inverse solution in which weights are unique and enable localisation for any point source in the brain. The solution provided by eLORETA is exact but low resolution (we use voxel sizes of 1.5 cm), which is appropriate for low density EEG, in which the volume conduction problem limits spatial resolution [10].

Our eLORETA method uses a state-of-the-art anatomical model of the head, its electrical conductivity properties, and locations of EEG channels to produce a lead-field matrix. This allows us to solve the source localisation method.

Our head model is built by segmenting a standard T1 weighted MRI anatomical scan into five layers comprising grey matter, white matter, cerebral spinal fluid, the skull, and the scalp. A mesh is then generated for each segmented region from 2000 vertices per region before these meshes are used to create a boundary element model [11] of conductivity changes between different anatomical regions in the head.

EEG channel locations are then automatically warped and transformed to fit locations of anatomical landmarks on the head, before manual adjustment. Finally, a mesh is overlaid on the head model with a grid size of 1.5 cm per voxel.

Voxel dimensions were chosen as a trade-off between convergence speed and accuracy. The Fieldtrip toolbox (v. 20190724) was used to generate this model [12].

Our method seeks a de-mixing matrix B^{-1} that results in a set of estimated sources that produces (via eLORETA) estimated activity distributions at a set of source locations in the head model that are maximally spatially independent.

Spatial independence is evaluated by measuring mutual information between spatial distributions of voxel source projections. For each voxel in our head model eLORETA estimates the contribution of that voxel to the scalp projected component P^i. Specifically, for each projection we estimate the contribution to that projection from every voxel in our head model V^i_M, where M denotes the number of voxels in our head model. We then calculate the median mutual information between all pairs of voxel sets.

A global optimisation technique [13] is used to identify the mixing matrix B that minimises the mutual information between the estimated voxel contributions to all pairs of components. The algorithm is partly inspired by a similar approach proposed by Abbas et.al. to optimise the classic ICA de-mixing matrix via particle swarm optimisation (we use a swarm of 100 particles with an inertia range of 0.1-100 and a stopping criterion of a maximum of 100 iterations) [14]. The algorithm is described by the pseudo-code.

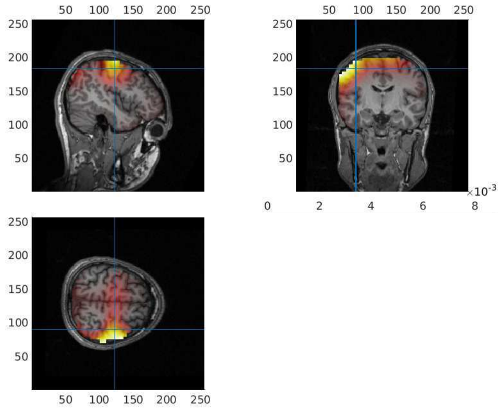
1) Estimate a candidate mixing matrix $\bar{B}$.
2) Estimate a candidate de-mixing matrix $\bar{B}^{-1}$ and from this estimate a set of sources $\hat{S}$.
3) For each source ($\hat{S}^i$), use eLORETA to estimate the contribution from each voxel in the head model V^i_M.
4) Measure the mutual information of voxel projection strengths between all pairs of estimated sources.
5) Use a global optimisation technique to minimise the median mutual information between all pairs of sources. Run until a stopping criteria is reached.

B. Data

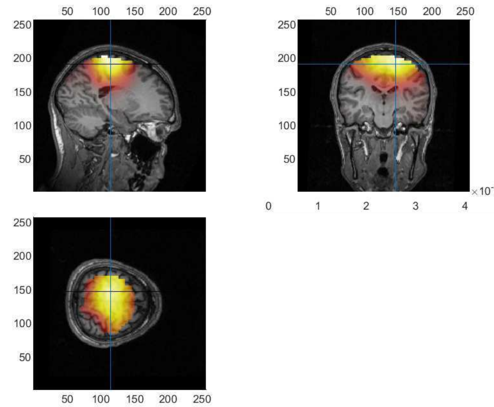
We use an EEG dataset to evaluate our method. The dataset contains event-related de-synchronisations (ERDs) evoked by kinaesthetic motor imagery.

The dataset was obtained from the BNCI horizon 2020 open EEG collection (http://bnci-horizon-2020.eu/database/data-sets, downloaded on 11/12/2018). Dataset 13, 'Individual imagery' was used. This dataset contains EEG recorded from 9 BCI users with disabilities of movement (spinal cord injury or stroke). EEG was recorded at a sample rate of 256 Hz from 30 electrodes.

Participants were asked to perform one of 5 tasks on cue: word association, mental subtraction, mental spatial rotation,



(a) Proposed method.



(b) ICA.

Fig. 1. Examples of source localisation of the components identified by our proposed method and ICA with activity closest to the locus of the ERD.

hand movement motor imagery, and feet movement motor imagery. Participants were given 6.5 s to perform each task, with cues appearing in random order. An inter stimulus interval of 2.5-3.5 s was used, with a pre-cue fixation cross on screen for 3.5 s. Further details may be found in [15].

C. Analysis

We evaluate our methods ability to characterise and classify the ERD, a relative reduction in strength of the EEG within the alpha (8-13 Hz) and beta (13-25 Hz) bands in the motor cortex during motor planning and execution. Identification of the ERD within a continuous EEG recording may be used as a control signal for BCI systems [16]. Thus, a filter that is able to accurately characterise the ERD has numerous potential applications.

The brain regions known to be most strongly involved in the ERD during hand motor control include the left and right hand representation areas of the sensorimotor cortex. Joint EEG-fMRI analysis of motor control has shown the locus of activity to vary slightly, depending on whether motor execution or motor imagery is used, but to be focussed at MNI coordinates (8, -6, 56) for left hand motor imagery and (-8, -26, 52) for right hand motor imagery [17].

To evaluate our method, we identify the source with an estimated centre of mass closest to the left hand representation area of the motor cortex to characterise EEG related to left hand motor imagery. We then use the time series of this component to characterise the ERD. We compare the ERD characterised via our method to ERD characterised from the best neural source identified by ICA.

Specifically, to compare our method to ICA we use the fastICA implementation of ICA in EEGlab [18]. We then use eLORETA to estimate the contribution each voxel in our head model makes to each IC and select the IC with a centre of mass closest to the left hand motor area.

We attempt to perform single trial classification of the components containing ERD characterised by our method and ICA. Specifically, we low-pass filter the EEG at 45 Hz (6'th order Butterworth). We extract features from this filtered signal by first calculating the magnitude of the EEG bandpower in the alpha frequency band (8-13 Hz) and then splitting the time series into a set of trials of length 0-6.5 s from the presentation time of the movement cue with a baseline correction of -2 to 0 s relative to stimulus presentation. The features given to the classifier are the mean amplitudes of the EEG band-powers from this time window.

We attempt 2-class classification to differentiate word association and left hand motor imagery trials. A linear discriminant analysis classifier is applied within a 5×5 cross-fold train and validation scheme on a per-participant basis.

The performance of each of these methods is evaluated by visualing comparing the components they identify. We inspect the estimated sources and compare the resulting recovered ERD signal components between the methods.

III. RESULTS

The components identified by our proposed method each contain neural activity from distinct spatial regions. An example of this may be seen in Figure 1(a), which illustrates the eLORETA source localisation for the component identified by our method that is closest to the locus of the ERD. This may be compared with Figure 1(b), which shows the source localisation for the source identified by ICA that is closest to the locus of the ERD.

It may be noted that the source identified by our method is considerably more spatially focussed over the right motor cortex, while the source identified by ICA is a mixture of locations in the left and right motor cortices.

We used the components identified by our method and identified via ICA to attempt to classify the EEG. The results for each participant are illustrated in Figure 2.

It should be noted that ICA alone is typically not sufficient as a feature extraction method when attempting to classify ERD. This is reflected in our results, which show that ICA-extracted features cannot be classified at a statistically significant level of accuracy for the majority of participants. By contrast, our proposed method affords statistically significant accuracies for the majority of participants.

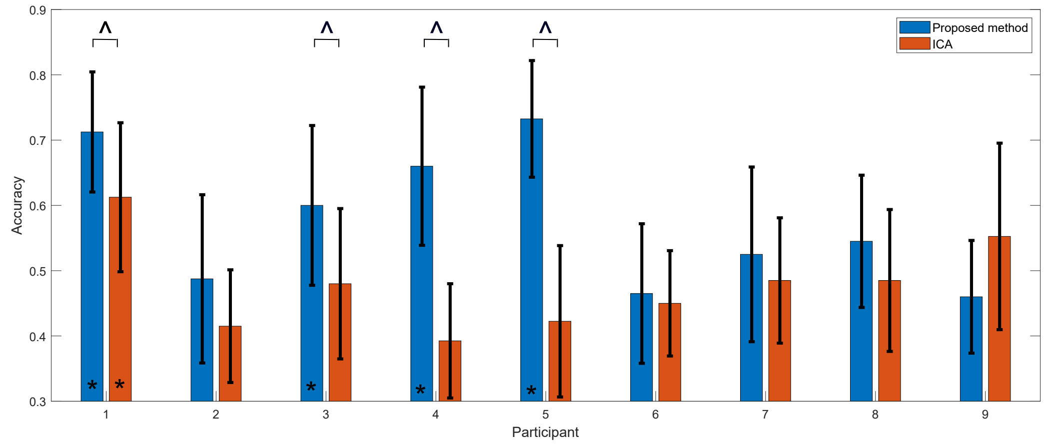


Fig. 2. ERD classification results achieved using components identified via our proposed method and via ICA. The asterisks indicate individual participant results that are significantly higher than random chance ($p < 0.05$, as assessed against the binomial distribution). The hats indicate pairs of results for a participant that are significantly different (paired t-test, $p < 0.05$).

Indeed, our proposed method produces accuracies that are statistically significant for 4 out of 9 participants (as assessed against a binomial distribution and indicated in Figure 2, $p < 0.05$). Furthermore, the mean classification results over all participants achieved using components identified by our proposed method are significantly higher than the accuracies produced by ICA ($p = 0.039$, Wilcoxon signed rank test).

IV. Conclusion

There is a need, in EEG analysis, for a method that is able to seperate EEG into components that are independent in the spatial domain. Our proposed method has advantages over ICA in several specific application areas. In this paper we illustrate an example of one such area, analysis and classification of ERD activity during motor imagery.

Our results demonstrate that our proposed method is better able to isolate brain activity sources to specific physical locations in the head. In addition to the demonstrated application to motor imagery classification this method has numerous potential applications which we will explore further in future.

References

[1] L. Schomer and F. Lopes de Silva, eds., *Niedermeyer's electroencephalography: Basic principles, clinical applications, and related fields*. Lippincott Williams & Wilkins, 6 ed., 2011.

[2] E. Niedermeyer and F. Lopes da Silva, *Electroencephalography: Basic principles, clinical applications and related fields*. Williams and Wilkins, 1999.

[3] M. Pal, R. Roy, J. Basu, and M. S. Bepari, "Blind source separation: A review and analysis," in *2013 International Conference Oriental COCOSDA held jointly with 2013 Conference on Asian Spoken Language Research and Evaluation (O-COCOSDA/CASLRE)*, pp. 1–5, IEEE, nov 2013.

[4] J. R. Wolpaw, N. Birbaumer, D. J. McFarland, G. Pfurtscheller, and T. M. Vaughan, "Brain-computer interfaces for communication and control.," *Clin Neurophysiol*, vol. 113, pp. 767–791, jun 2002.

[5] S. Makeig, T. P. Jung, A. J. Bell, D. Ghahremani, and T. J. Sejnowski, "Blind separation of auditory event-related brain responses into independent components.," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 94, pp. 10979–84, sep 1997.

[6] I. Daly, R. Scherer, M. Billinger, and G. Muller-Putz, "FORCe: Fully Online and automated artifact Removal for brain-Computer interfacing.," *IEEE transactions on neural systems and rehabilitation engineering*, vol. 23, aug 2014.

[7] G. Bauernfeind, S. C. Wriessnegger, I. Daly, and G. R. Müller-Putz, "Separating heart and brain: on the reduction of physiological noise from multichannel functional near-infrared spectroscopy (fNIRS) signals.," *Journal of neural engineering*, vol. 11, p. 056010, oct 2014.

[8] M. E. Raichle, "The restless brain.," *Brain connectivity*, vol. 1, pp. 3–12, jan 2011.

[9] I. Daly, S. J. Nasuto, and K. Warwick, "Brain computer interface control via functional connectivity dynamics," *Pattern Recognition*, vol. 45, pp. 2123–2136, jun 2012.

[10] R. D. Pascual-Marqui, D. Lehmann, M. Koukkou, K. Kochi, P. Anderer, B. Saletu, H. Tanaka, K. Hirata, E. R. John, L. Prichep, R. Biscay-Lirio, and T. Kinoshita, "Assessing interactions in the brain with exact low-resolution electromagnetic tomography," *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences*, vol. 369, pp. 3768–3784, oct 2011.

[11] H. Hallez, B. Vanrumste, R. Grech, J. Muscat, W. De Clercq, A. Vergult, Y. D'Asseler, K. P. Camilleri, S. G. Fabri, S. Van Huffel, and I. Lemahieu, "Review on solving the forward problem in EEG source analysis.," *Journal of neuroengineering and rehabilitation*, vol. 4, p. 46, nov 2007.

[12] R. Oostenveld, P. Fries, E. Maris, and J.-M. Schoffelen, "FieldTrip: Open source software for advanced analysis of MEG, EEG, and invasive electrophysiological data.," *Computational intelligence and neuroscience*, vol. 2011, p. 156869, dec 2011.

[13] R. Horst, *Introduction to Global Optimization*. Springer, 2000.

[14] N. A. Abbas and H. M. Salman, "Independent component analysis based on quantum particle swarm optimization," *Egyptian Informatics Journal*, vol. 19, pp. 101–105, jul 2018.

[15] R. Scherer, J. Faller, E. V. C. Friedrich, E. Opisso, U. Costa, A. Kübler, and G. R. Müller-Putz, "Individually Adapted Imagery Improves Brain-Computer Interface Performance in End-Users with Disability," *PLoS ONE*, vol. 10, no. 5, p. e0123727, 2015.

[16] C. Neuper, R. Scherer, S. Wriessnegger, and G. Pfurtscheller, "Motor imagery and action observation: modulation of sensorimotor brain rhythms during mental control of a brain-computer interface.," *Clinical neurophysiology : official journal of the International Federation of Clinical Neurophysiology*, vol. 120, pp. 239–247, feb 2009.

[17] C. Zich, S. Debener, C. Kranczioch, M. G. Bleichner, I. Gutberlet, and M. De Vos, "Real-time EEG feedback during simultaneous EEGfMRI identifies the cortical signature of motor imagery," *NeuroImage*, vol. 114, pp. 438–447, jul 2015.

[18] A. Hyvarinen, "Fast and robust fixed-point algorithms for independent component analysis," *IEEE Transactions on Neural Networks*, vol. 10, pp. 626–634, may 1999.