Linear spatial integration for single trial detection in encephalography

Lucas Parra*, Chris Alvino*, Akaysha Tang†, Barak Pearlmutter†, Nick Yeung‡, Allen Osman§ and Paul Sajda‖

Vision Technologies Laboratory, Sarnoff Corporation, Princeton, New Jersey 08540, USA

† Department of Psychology and Department of Computer Science, University of New Mexico, Albuquerque, New Mexico 87131, USA

‡   Department of Psychology, Princeton University, Princeton, New Jersey 08544, USA

§Department of Psychology, University of Pennsylvania, Philadelphia, Pennsylvania 19104, USA

‖Department of Biomedical Engineering, Columbia University, New York, New York 10027, USA

Running title: Single trial detection in encephalography

Correspondence and requests for materials should be addressed to: Prof. Paul Sajda, Department of Biomedical Engineering, Columbia University, 351 Engineering Terrace: MC8904, New York, NY 10027. phone: (212) 854-5279, fax: (212) 854-8725, e-mail: ps629@columbia.edu.

Abstract

Conventional analysis of electroencephalography (EEG) and magnetoencephalography (MEG) often relies on averaging over multiple trials to extract statistically relevant differences between two or more experimental conditions. In this paper we demonstrate single trial detection by linearly integrating information over multiple spatially distributed sensors within a predefined time window. We report an average, single trial discrimination performance of $A_z \approx 0.80$ and fraction correct between 0.70-0.80, across three distinct encephalographic datasets. We restrict our approach to linear integration as it allows the computation of a spatial distribution of the discriminating component activity. In the present set of experiments the resulting component activity distributions are shown to correspond to the functional neuroanatomy consistent with the task (e.g. contralateral sensory-motor cortex and anterior cingulate). Our work demonstrates how a purely data-driven method for learning an optimal spatial weighting of encephalographic activity can be validated against the functional neuroanatomy.

Introduction

Trial averaging is often used in braining imaging to mitigate low signal-to-interference (SIR) ratios. For example, it is the basis for analysis of event-related potentials (ERPs) (Coles and Rugg, 1995). However for some encephalographic applications, such as seizure prediction, trial averaging is problematic. One application where the problem of single-trial averaging is immediately apparent is the brain-computer interface (BCI), i.e. interpreting brain activity for real-time communication. In the simplest case, where one wishes to communicate a binary decision, averaging corresponds to asking the same question over multiple trials and averaging the subject's binary responses. In order to obtain high-bandwidth communication, it is desirable to do as little averaging over time or across trials as possible.

More generally, single trial analysis of brain activity is important in order to uncover the origin of response variability, for instance in analysis of error related negativity (ERN). The ERN is a negative deflection in the EEG following perceived incorrect responses (Gehring et al., 1993, Falkenstein et al., 2000) or expected losses (Gehring and Willoughby, 2002) in a forced-choice task. Single trial detection of the ERN has been proposed as a means of correcting communication errors in a BCI system (Schalk et al., 2000). With the ability to analyze the precise timing and amplitude of the ERN, on individual trials, one can begin to study parameters that cannot be controlled across trial, such as reaction time or error perception. Such an approach opens up new possibilities for studying the behavioral relevance and neurological origin of the ERN.

With the large number of sensors in high density EEG and MEG, an alternative approach to trial averaging is to integrate information over space rather than across trials. A number of methods along these lines have been proposed. Blind source

separation analyzes the multivariate statistics of the sensor data to identify spatial linear combinations that are statistically independent over time (Makeig et.al 1996, Vigario et al., 2000, Tang et al., 2002). Separating independent signals and removing noise sources and artifacts, increases SIR. However, blind source separation does not exploit the timing information of external events that is often available. In most current experimental paradigms subjects are prompted with external stimuli to which they are asked to respond. The timing of the stimuli, as well as the timing of overt responses, are therefore available and should be exploited by the analysis method.

In the context of a BCI system, many methods have applied linear and non-linear classification to a set of features extracted from the EEG. For example, adaptive autoregressive models have been used to extract features across a limited number of electrodes, with features combined using either linear or non-linear classifiers to identify the activity from the time course of individual sensors (Pfurtscheller and Neuper, 2001). Others have proposed to combine sensors in space by computing maximum and minimum eigenvalues of the sensor covariance matrices. The eigenvalues, which capture the power variations of synchronization and desynchronization, are then combined non-linearly to obtain binary classification (Ramoser et al., 2000). Spatial filtering has also been used to improve the SNR of oscillatory activity. However, there has been no systematic effort to choose optimal spatial filters. In the context of the ERN, Gehring et. al (1993) use linear discrimination to identify characteristic time courses in individual electrodes, but do not exploit spatial information. Though many of these aforementioned methods obtain promising performance in terms of classifying covert (purely mental) processes, their neurological interpretation remains obscured.

In this paper we report on using conventional linear discrimination to compute the optimal spatial integration of a large array of sensors. We exploit timing information by discriminating and averaging within a short time window relative to a given external event. We restrict ourselves to a linear integration as it permits the computation of spatial distributions of the discriminating component[1] activity, which in turn can be compared to functional neuroanatomy.

We demonstrate the utility of the proposed method for three distinct datasets acquired via high-spatial density encephalography :

- Predicting explicit (overt) motor response using MEG (122 sensors).

- Classifying imagined (covert) motor activity using EEG (59 sensors).

- Detecting decision errors for a binary discrimination task using EEG (64 sensors).

Materials and Methods

Linear Discrimination

Denoting $\mathbf{x}(t)$ as the M sensor values sampled at time instance t, we compute the spatial weighting coefficients $\mathbf{v}$ such that

---

[1] Note that we use the term "component" instead of "source" so as to avoid confusion with an implied physiological source.

$$y(t) = \mathbf{v}^\mathrm{T}\mathbf{x}(t) \tag{1}$$

is maximally discriminating between the times t, corresponding to two different experimental conditions. For example, in the prediction of explicit motor response experiments (described below) the times correspond to a number of samples prior to an explicit button push. The samples corresponding to a left button push are to be discriminated from samples of a right button push. For each of N trials there are T samples, totaling NT training examples. We use conventional logistic regression (Duda et al., 2001) to find $\mathbf{v}$. After finding the optimal $\mathbf{v}$ we average over the T dependent samples of the kth trial to obtain a more robust result, $\bar{y}_k = \dfrac{1}{T} \sum_{t \in T_k} y(t)$, where T_k denotes the set of sample times corresponding to trial k. We evaluate performance using receiver operating characteristic (ROC) analysis (Swets, 1979) on the single-trial short-time averaged discrimination activities ($\bar{y}_k$). For visualization purposes, it is also useful to compute the trial averaged discrimination activities

$$\bar{y}_e(t) = \frac{1}{N} \sum_{k \in N_e} y_k(t) \tag{2}$$

where N_e denotes the set of samples for event e (e.g. left or right button push) with time measured relative to some common reference across trials. The separation of the means together with their corresponding variances provides an indication of whether single-trial discrimination is plausible within the analysis window.

Localization of Discriminating Components

In order to provide a functional neuroanatomical interpretation of the resultant spatial weighting, we treat $y(t)$ as a component which is maximally discriminating given the linear model and task. A simple way of visualizing the origin of a component's activity

is to display the coupling coefficients of the component with the sensors. The strength of the coupling roughly indicates the proximity and orientation of the component elative to the sensor. The coupling $\mathbf{a}$ is defined as the coefficients that multiply the component $y(t)$ to give its additive contribution $\mathbf{x}_y(t)$ to the sensor readings, $\mathbf{x}_y(t)=\mathbf{a}y(t)$. However $\mathbf{x}_y(t)$ is not observable in isolation and instead we observe, $\mathbf{x}(t) = \mathbf{x}_y(t) + \mathbf{x}_{y'}(t)$, where $\mathbf{x}_{y'}(t)$ represents the activity that is not due to the discriminating component. If the contributions of the other components, $\mathbf{x}_{y'}(t)$, are uncorrelated with $y(t)$ we obtain the coupling coefficients by the least-squares solution (Haykin, 1996). Arranging the samples $\mathbf{x}(t)$ for different t as columns in the a matrix $\mathbf{X}$, and $y(t)$ as a column vector $\mathbf{y}$ the solution is given by

$$\mathbf{a} = \frac{\mathbf{X}\mathbf{y}}{\mathbf{y}^T\mathbf{y}} \tag{3}$$

In the appendix we present a derivation of $\mathbf{a}$ and describe some of the underlying assumptions in its interpretation. In general other components are not guaranteed to be uncorrelated with the discriminating component. Therefore $\mathbf{a}$ represents the coupling of all component activities that are correlated with the discriminating component $y(t)$. We refer to $\mathbf{a}$ as a "sensor projection" as it measures the activity in the sensors that correlate with a given component. Our approach relies on the linearity of $y(t)$ and the fact that different components in EEG and MEG add linearly (Baillet et al., 2001).

Datasets for Analysis

*Predicting explicit (overt) motor response using MEG***:** This dataset was provided by AT and BP. Four subjects performed a visual-motor integration task. A "trump" experiment was defined whereby subjects were simultaneously presented with two visual stimuli on a CRT, one of which is the target and "trumps" (beats-out) the other.

Subjects were instructed to push a left hand or right hand button, depending on which side the target (trump stimulus) was present. The subject was to discover the target by trial-and-error using auditory feedback. Each trial began with visual stimulus onset, followed by button push, followed by auditory feedback, indicating if the subject responded correctly. The interval between the motor-response and the next stimulus presentation was 3.0 ± 0.5 sec. Each subject performed 90 trials, which took approximately 10 minutes. MEG data was recorded using 122 sensor at a sampling rate of 300 Hz and high-pass filtered to remove DC drifts.

Classifying imagined (covert) motor activity using EEG: This dataset was provided by AO. Nine subjects performed a visual stimulus driven finger (L/R) tapping task. Subjects were asked to synchronize an explicit or imagined tap by the left, right, or both index fingers to the presentation of a brief temporally predictable signal. Subjects were trained until their explicit taps occurred consistently within 100 ms of the synchronization signal. Subjects were presented visual stimuli indicating with which index finger to tap and if it should be an explicit or imagined tap. 1.25 s after the last instruction symbol a fixation point was replaced for 50 ms by the letter "X." This letter served as a signal to which the instructed tap (whether overt or imagined) was to be synchronized. Each trial lasted for 6 s. After training, each subject received 10 blocks of trials. Each 72-trial block consisted of nine replications of the eight trial types (Explicit vs. Imagined x Left vs. Right vs. Both vs. No Tap) presented in a random order. Trials with noise due to eye blinks were not considered in the EEG analysis. The electromyogram (EMG) was recorded to detect muscle activity during imagined movements. The 59 EEG channels were sampled at 100 Hz and high-pass filtered to remove DC components.

Detecting decision errors for a binary discrimination task using EEG: This dataset was provided by NY. Seven subjects performed a visual target detection amongst distractors task. On each trial, subjects were presented with a stimulus for 100 ms. There were four possible stimuli, each consisting of a row of five arrows. Subjects were told to respond by pressing a key on the side indicated by the center arrow. They were to ignore the four flanking arrows. On half of the trials, the flanking arrows pointed in the same direction as the target (e.g. <<<<<), on the other half the flankers pointed in the opposite direction (e.g. << ><<). Subjects were slower and made many more errors in the latter case. Following their response, there was an inter-trial interval of 1.5 seconds, after which a new stimulus was presented. Subjects performed 12 blocks of 68 trials each. The 100 ms interval prior to the response was used as the baseline period (separately for each trial and electrode). The sampling rate was 250 Hz. Following the baseline period, trials were manually edited to remove those with blinks, large eye movements, instrument artifacts and amplifier saturation.

Results and Discussion

Single trial discrimination results are shown for the three different data sets and include trial averaged discriminating component activity $\bar{y}_e(t)$, sensor projections **a**, and detection/prediction performance using single-trial, short-time averaged $\bar{y}_k$. Performance is reported using ROC analysis computed with a leave-one-out training and testing procedure (Duda et al., 2001). ROC analysis is a reasonable method for quantifying performance for these three datasets since it enables one to incorporate an independent cost for false positives and false negatives. For example, in an error

correction application using the ERN, it important to detect error events with high

confidence. The desired operating point of such a detector is therefore at a low false

positive rate (high specificity). In contrast, an application which looks to exploit motor

imagery for communicating a binary decision is best assessed at the operating point

were sensitivity equals specificity i.e. the error rates for the two possible outcomes are

equal. A metric that quantifies the overall performance of a detector for arbitrary

operating points is the area under the ROC curve (A_z). In the following sections we

report A_z , as well as the fraction of correct classifications, for all three tasks. A

summary of the results for the three datasets is given in the following table.

	ROC area (A_z) Mean ± std.	Fraction correct Mean ± std.	N	N_e	Sensors	Detection time window
Explicit L/R button push prediction	0.82 ± 0.06	0.79 ± 0.09	4	45/45	122 MEG	100ms to 33 ms prior to button push
Imagined L/R finger tap discrimination	0.77 ± 0.10	0.71 ± 0.08	9	90/90	64 EEG	400 ms before to 400 ms after synchronization
Response error/correct discrimination	0.79 ± 0.05	0.73 ± 0.05	7	40-80 /300	59 EEG	0 ms to 100 ms after response

**Table 1: Performance of the linear spatial integration method for the three
datasets. Mean and standard deviation (std.) are reported across N
subjects. N_e is the number of trails used to determine the best linear
classifier (#positive/#negative trials). The time window used for detection
is also specified.**

As seen in table 1, for all three datasets the number of trials for training is comparable

to the number coefficients to be trained. This typically leads to serious problems in

over-training. We mitigate these by including multiple training samples for each trial.[2]

We would expect that increasing the number of independent training samples (e.g.

trials) would similarly increase performance of the results presented below.

Figure 1 shows results for the dataset used to predict whether a subject will press a

button with their left or right hand by analyzing the MEG signals in a window prior to

the button push (left hand = 1, right hand = 0 in the logistic regression model). We use

an analysis window 100 ms wide centered at 83 ms prior to the button event, which at

300 Hz corresponds to T=30. Figure 1 shows the results for one subject (AT). In the

trial averaged results one can see significant separation of the means for left vs right

button push within the analysis window. Given that this separation is approximately

equal to one standard deviation, this suggests that single trial discrimination is possible.

Single-trial discrimination is shown in the ROC curve, which for this subject shows

good discriminability (A_z=0.93). Figure 1 also shows the sensor projection **a** and the

location of a dipole-fit for this projection. A single equivalent current dipole fits the data

with an accuracy of 64% using the least squares 'xfit' routine from Neuromag[3] and

assuming a spherical head model. This compares favorably with the 50% goodness of

fit which are typically obtained for somatosensory responses when using all 122 sensors

(Tang 2002). When considered with respect to the motor-sensory homunculus., these

results indicate that the discrimination component activity originates in the sensory-

---

[2] These samples are obviously not independent, however, they provide evidence for the natural variation of the data

and thus make the estimates much more robust. They were shown, through cross-validation, to improve estimated

generalization performance.

[3] See www.neuromag.com.

motor cortex corresponding to the left hand.

Figure 2 shows results for the second dataset, where the goal is to detect activity associated with a purely imagined motor response, a situation perhaps more relevant for a BCI system. Subjects are trained to imagine a tap with the left or right index finger synchronized to a brief, temporally predictable signal. Therefore there exists a known time window in which one can explicitly look for activity that discriminates between the left and right imagine conditions. We selected a 0.8 s time window around the time where the task is to be performed. 90 left and 90 right trials were available to train the coefficients of the 59 EEG sensors. The result for the best performing subject is $A_z = 0.90$, shown in Figure 2. The sensor projection of the 59 EEG sensors shows a clear left-right polarization over the motor area. In the context of BCI the metric of interest is the bit rate at which information can be transmitted with imagined motor activity. The information transmitted per trial is given by,

$$I = 1 + p \log_2(p) + (1-p)\log_2(1-p), \qquad (4)$$

where p is the fraction correct. For the subject shown in Figure 2 this corresponds to I =0.26 bits/trial. When averaged over the nine subjects the information transmitted is I=0.16 bit/trial. Note that with a repetition rate of 6s this experiment is not designed for an optimal transmission rate. Assuming equivalent performance with a repetition rate of 0.8s (corresponding to the time window used here for discrimination) we obtain an average bit rate of 12 bits/minute.

For comparison, we compare our approach with an alternative method, first described by Wolpaw et al. (1991), that is based on differences in the power spectrum in electrodes over the left and right motor cortex. Andersen et al. (1998) modify the approach by using six auto-regressive (AR) coefficients to model the power spectrum of each electrode within the analysis window and classify the imagined conditions using a linear discrimination on these AR coefficients. Following Penny et al. (2001), we used electrodes C3 and C4 (international 10/20 electrode placement system – see Towle et al. 1993) and obtain, $A_z = 0.65 \pm 0.09$, and fraction correct of $p = 0.62 \pm 0.07$, which corresponds to $I = 0.054$ bits/trial or a bit rate of 4 bit/minute. This is about a third of the bit rate produced by our method of linear spatial integration.

The results, across the nine subjects, for predicting explicit finger taps from a window 300 ms to 100 ms prior to the taps is $A_z = 0.87 \pm 0.08$ with a fraction correct of 0.80 ± 0.08. As shown for subject PJM in Figure 2, sensor projections of the discrimination vector for explicit motor response are similar to the projections of the imagined motor response. This is consistent with previous finding in EEG and fMRI (Cunnington et al., 1996, Porro et al., 1996) and supports the approach of many current BCI systems— signals arising from the cortical areas that encode an explicit movement are also in some sense optimal for detecting the imagined movement.

Figure 3 shows the results for the target detection experiments where the goal is to detect the ERN on a single trial basis. The ERN has a medial-frontal distribution that is symmetric to the midline, suggesting a source in the anterior cingulate (Dehaene et al., 1994). It begins around the time of the perceived incorrect response and lasts roughly

100 ms thereafter. We use this time window for detection. 40 to 80 error trials and 300

correct trials were used for training and testing 64 coefficients. The sensor projection,

shown in Figure 3 for one subject, is representative of the results obtained for other

subjects and is consistent with the scalp topography and time course of the ERN. The

detection performance for this subject is $A_z = 0.84$ and is to be compared to $A_z = 0.63$

when detecting ERN from the front-center electrode where maximal activity is expected

(FCz in the 10/20 system).

Our results demonstrate the utility of linear analysis methods for discriminating between

different events in single-trial, stimulus driven experimental paradigms using EEG and

MEG. A particularly important aspect of our approach is that linearity enables the

computation of sensor projections for the optimally discriminating weighting. This

localization can be compared to the functional neuroanatomy, serving as a validation of

the data driven linear methods. In all three cases presented, we find that indeed the

activity distribution correlated with the component that optimizes single-trial

discrimination localizes to a region that is consistent with the functional neuranatomy.

This is important, for instance, to determine whether the discrimination model is

capturing information directly related to the underlying task-dependent cortical activity,

or is instead exploiting an indirect cortical response or other physiological signals

correlated with the task (e.g. correlations with the stimulus, eye movements, etc.).

Localization of the discriminating component activity, and its correlates, also enables

one to determine the neuroanatomical correlations between different discrimination

tasks, as was demonstrated for explicit and imagined motor responses in EEG.

In this work we have focused on spatial linear filtering of the magnitude of electrical or magnetic activity. This is consistent with the conventional concept of evoked responses, which capture magnitude differences under different experimental conditions. Note that in this approach the sign of the response is relevant for discrimination. Our approach does not consider temporal patterns of activity. An alternative and common approach is to consider signal powers after linear filtering in time (Wolpaw et al. 1991, Pfurtscheller et al., 2001, Anderson et al. 1998). When considering powers, (i.e. squared magnitudes) the sign of the activity is irrelevant and instead one typically captures oscillatory temporal patterns of activity. Both coherent evoked responses as well as oscillatory activity have been important analysis tools for encephalography, and work to combine optimal linear spatial and temporal filtering is beginning to emerge (Ramoser et al. 2001).

A potential disadvantage of the method we have described is that its utility is derived from integration across a large sensor array, which may be impractical for a realistic BCI system (e.g. may be too cumbersome and/or costly). In such cases, the method still offers utility as a means for identifying a reduced sensor configuration, since the computed sensor projection can be used to prune the array in a way that minimizes loss in discrimination performance.

We close by noting that the proposed method is applicable to other encephalographic modalities with linear superposition of activity, such as functional near infrared imaging (Boas et al., 2001).

References

Anderson, C., Stolz, E. & Shamsunder, S., Multivariate Autoregressive Models for Classification of Spontaneous Electroencephalogram During Mental Tasks. *IEEE Transactions on Biomedical Engineering*, **45**(3): 277-286, (1998).

Baillet, S., Mosher, J.C. & Leahy, R.M. Electromagnetic brain mapping. *IEEE Signal Processing Magazine*, **18**(6): 14-30, (2001).

Boas, D.A., et al. Imaging the body with diffuse optical tomography. *IEEE Signal Processing Magazine*, **18**(6): 57-75, (2001).

Coles, M. G. H. & Rugg, M.D. Event-related brain potentials: An introduction. In M. D. Rugg & M.G.H. Coles (Ed.), *Electrophysiology of Mind*. Oxford: Oxford University Press (1995).

Cunnington R, Iansek R, Bradshaw JL, & Phillips JG. Movement-related potentials associated with movement preparation and motor imagery., *Exp. Brain Res.*, **111**(3):429-36, (1996)

Dehaene, S., Posner, M., & Tucker, D. Localization of a neural system for error detection and compensation. *Psychological Science*, **5**: 303-305, (1994).

Duda, R., Hart, P. & Stork, D. *Pattern Classification*. John Wiley & Sons, 2 edition, (2001).

Falkenstein, M., Hoorman, J., Christ, S. & Hohnsbein, J. ERP components on reaction errors and their functional significance: A tutorial. *Biological Psychology*, **51**:87-107,

(2000).

Gehring, W. J., Goss, B. **,** Coles, M. G. H. , Meyer, D. E. & Donchin, E. A neural system for error detection and compensation. *Psychological Science*, **4**(6):385-390, (1993).

Gehring, W.J. and Willoughby, A.R. The medical frontal cortex and the rapid processing of monetary gains and loss. *Science*, **295**: 2279-2282, (2002).

Haykin, S. *Adaptive Filter Theory*. Englewood Cliffs, NJ, Prentice-Hall, (1996).

Makeig, S. Bell, A. Jung, T. & Sejnowski T. Independent component analysis of electroencephalographic data. In *Advances in Neural Information Processing Systems*, **8**: 145-151. MIT Press. (1996).

Pfurtscheller, G. & Neuper, C. Motor imagery and direct brain-computer communication. *Proceedings of the IEEE,* **89**(7): 1123-1134, (2001).

Porro CA, Francescato MP, Cettolo V, Diamond ME, Baraldi P, Zuiani C, Bazzocchi M, di Prampero PE. Primary motor and sensory cortex activation during motor performance and motor imagery: a functional magnetic resonance imaging study. *J Neurosci.*, **16**(23): 7688-98 (1996).

Ramoser, H., Mueller-Gerking, J. & Pfurtscheller G. Optimal spatial filtering of single trial EEG during imagined hand movement. *IEEE Transaction on Rehabilitation Engeneering*, **8**(4): 441-446 (2000).

Schalk, G., Wolpaw, J., McFarland, D., & Pfurtscheller, G. EEG-based

18

communication: presence of an error potential. *Clinical Neurophysiology*, **111**: 2138-2144, (2000).

Swets JA. ROC analysis applied to the evaluation of medical imaging techniques. *Investigative Radiology* **14**:109-121, (1979).

Tang, A., Pearlmutter, B., Malaszenko, N., Phung, D. & Reeb, B. Localization of Independent Components of Magnetoencephalography in Cognitive Tasks. *Neural Computation*, in press (2002)

Towle V., Bolanos, J. et. al., The spatial location of EEG electrodes: locating the best-fitting sphere relative to cortical anatomy. *Electroenceph. Clin. Neurophysiol.,* **86**(1): 1-6, (1993)

Vigario, R., Sarela, J., Jousmaki, V., Hamalainen, M. & E. Oja. Independent component approach to the analysis of EEG and MEG recordings. *IEEE Transactions on Biomedical Engineering*, **47**(5):589-593 (2000).

Wolpaw, J.R., McFarland, D.J., Neat, D.J. and Fomeris, C.A., An EEG-based Brain-Computer Interface for Cursor Control. *Electroencephalography and Clinical Neurophysiology*, **78**: 252-259, (1991).

Acknowledgements

Appendix

Derivation of Sensor Projection $\mathbf{a}$

Suppose our observation vector is $\mathbf{x}$ and we have built a linear classifier, $y_1 = \mathbf{v}^T\mathbf{x}$, where y_1 is the binary number indicating some cognitive event that we are trying to detect. We assume that there are a number of such cognitive events occurring simultaneously. We represent these as a vector of binary indicators $\mathbf{y}$, with y_1 as its first element, and a matrix $\mathbf{A}$ that maps these to the observation vectors; i.e. $\mathbf{x} = \mathbf{A}\mathbf{y}$. We wish to identify this mapping, namely to find the first column of $\mathbf{A}$, which we call $\mathbf{a}$ and which is defined as the oberservation vector that we would obtain if only y_1 occurred. The least-squares estimate of $\mathbf{a}$ can be derived as follows. Let $\mathbf{X}$ be the observation matrix for many samples-i.e. the tth column is the observation for the tth sample. Let $\mathbf{y}_1^T$ be the corresponding binary column vectors across these samples given by $\mathbf{y}_1 = \mathbf{v}^T\mathbf{X}$. The analogous relationship with $\mathbf{a}$ is $\mathbf{X} = \mathbf{a}\mathbf{y}_1^T$. The least-square solution for $\mathbf{a}$ is given by $\mathbf{a} = \mathbf{X}\mathbf{y}_1(\mathbf{y}_1^T\mathbf{y}_1)^{-1}$.

We would like to determine the conditions under which the least-squares estimate of

$\mathbf{a}$ is actually proportional to the first column of $\mathbf{A}$. Let the matrix $\mathbf{Y}$ be the binary matrix of the simultaneous cognitive events across trials—i.e. the tth column is the cognitive events vector $\mathbf{y}$ for the tth trial. Since $\mathbf{X} = \mathbf{AY}$ we find that $\mathbf{a} = \mathbf{AY}\mathbf{y}_1 (\mathbf{y}_1^T \mathbf{y}_1)^{-1}$. Note that $\mathbf{Y}$ has dimensions of number of cognitive events (N) by number of samples (T), and that the quantity $\mathbf{YY}_1$ is the column vector of unnormalized correlations between the event indicators y_1 and the set of all cognitive events. If this is proporational to the Kronecker delta, $\delta_{i,1}$ (i.e. y_1 is uncorrelated with the indicators of the other events) then $a_i \propto \sum_j A_{i,j} \delta_{j,1} = A_{i,1}$ and therefore $\mathbf{a}$ is proportional to the first column of $\mathbf{A}$.
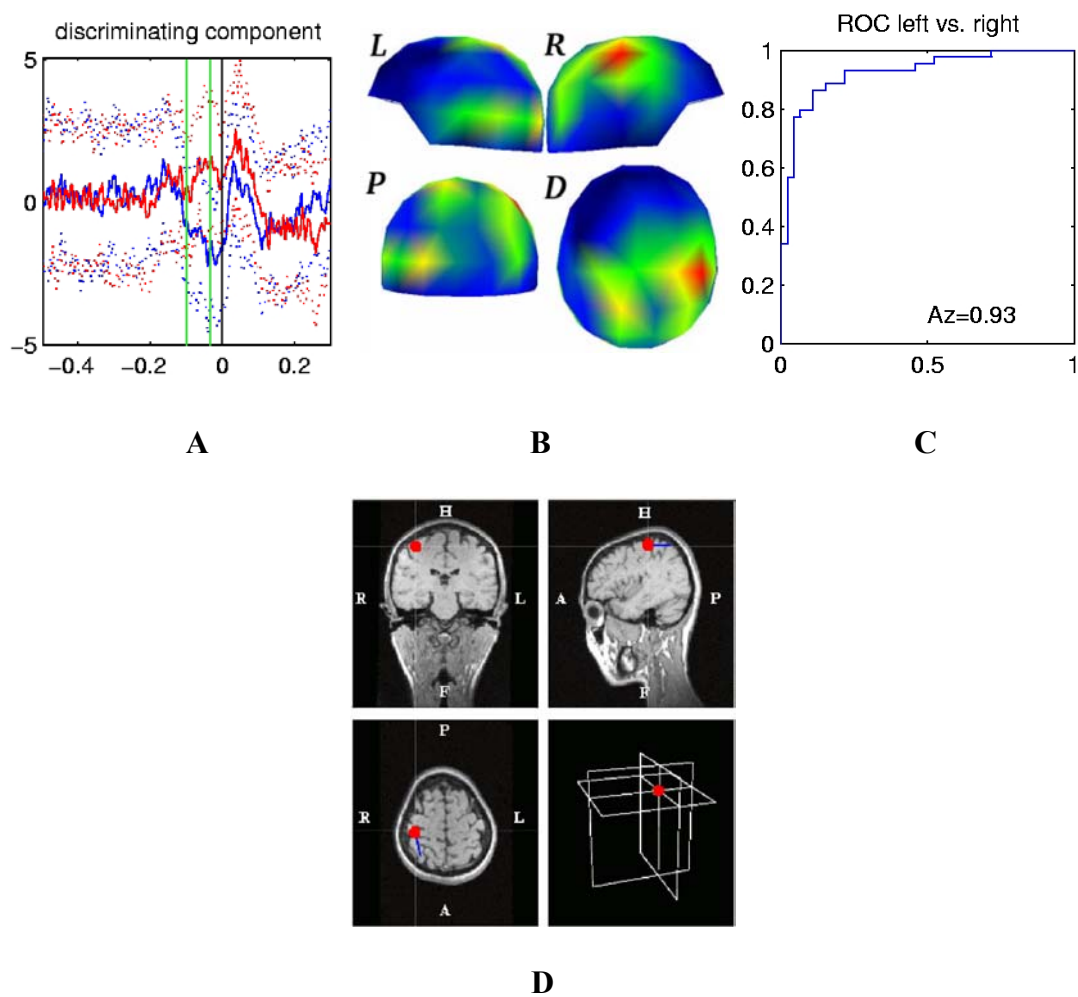
Figure 1: MEG left/right button push prediction. (A) Trial averages $\bar{y}_e(t)$ (solid curves) and standard deviations (dotted curves) of discriminating component for left (red) and right (blue) button pushes. Time is given in seconds. Black vertical line at t = 0 s indicates timing of button push. Green vertical lines indicate the discrimination window. (B) Sensor projections for discrimination vector. (C) ROC curve for left vs. right discrimination. Area under the curve (A_z) = 0.93. (D) Dipole-fit of a overlaid on MRI image.
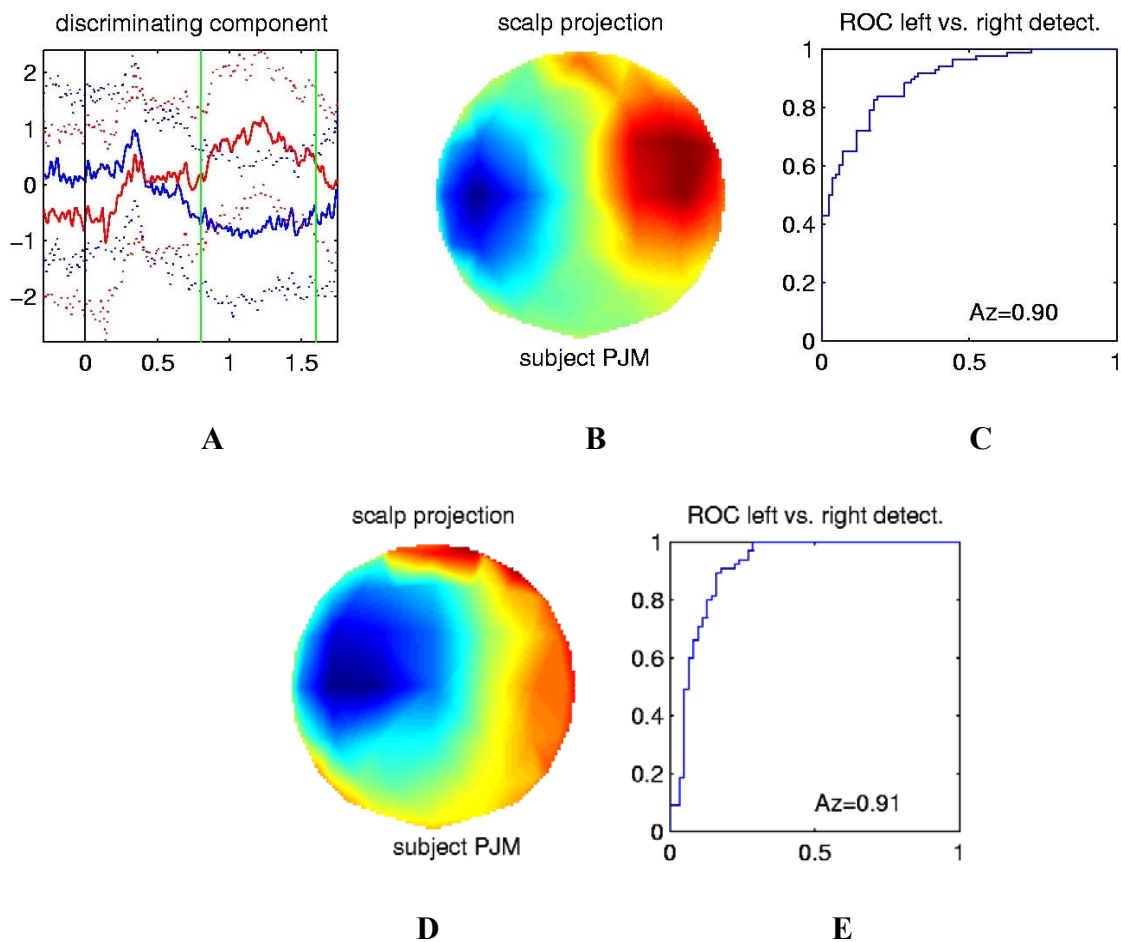
22

Figure 2: Discrimination of imagined left/right finger taps. (A) trial averages $\bar{y}_e(t)$ **(solid curves) and standard deviations (dotted curves) of discriminating component for left (red) and right (blue) imagined. Time is given in seconds. Black vertical solid line at t = 0 s indicates timing of visual stimulus that defines the action to be performed (left or right imagine). The subjects are trained to execute the task at around t = 1:25 s. Green vertical lines indicate the discrimination window. (B) Dorsal view of sensor projections a. (C) ROC curve for left vs. right discrimination. For this subject the fraction of correct classification is** *p*=0.79 **which corresponds to an information transfer of 0.26 bits/trial. (D) Sensor projection for same subject for explicit finger tap. (E) ROC curve for same subject for explicit finger tap.**

discriminating component

scalp projection

subject 4_05

ROC correct vs. error detect.
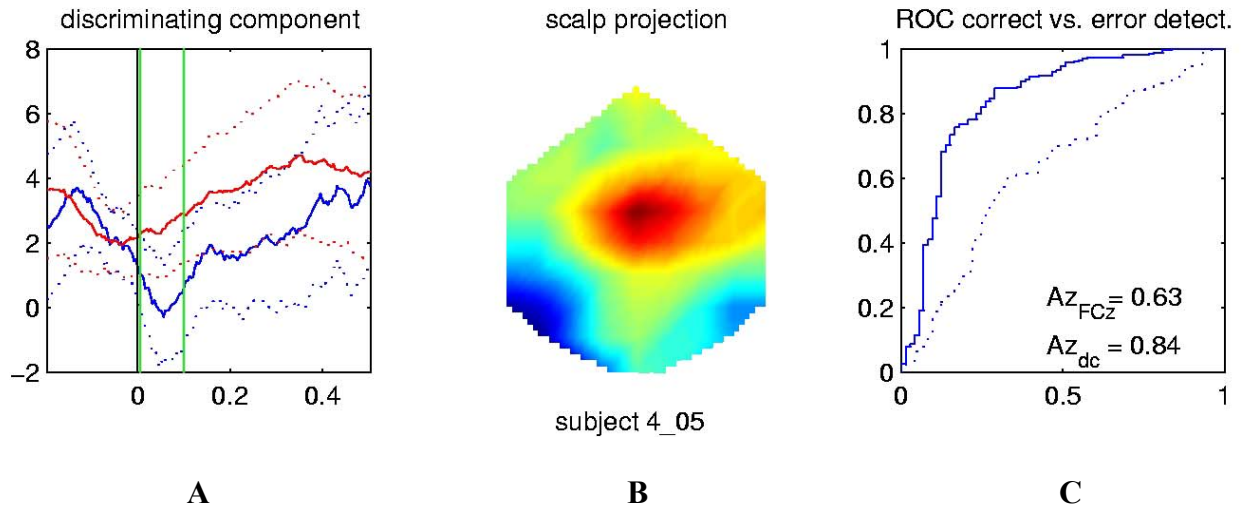
$Az_{FCz} = 0.63$

$Az_{dc} = 0.84$

A

B

C

Figure 3: Detection of decision errors with EEG. (A) trial averages $\bar{y}_e(t)$ (solid curves) and standard deviations (dotted curves) of discriminating component for correct and error trials. The negative deflection after a button push response at t = 0 s is the ERN. Green vertical lines indicate the discrimination window. (B) Dorsal view of sensor projections a. (C) ROC curve for error vs. correct trials. Solid curve corresponds to discrimination using Eq. (1) and dotted line to discrimination with center electrode (FCz).