Localization of Brain Activity in Electroencephalography Data during Brain-Computer Interface Operation

Author:
Oskar Hjärtquist

Supervisors:
Dr. Fredrik Sebelius
Dr. Zoran Nenadic

M. Sc. Thesis

Department of Measurement Technology and Industrial Electrical Engineering
Division of Electrical Measurements
Lund University
Department of Biomedical Engineering
University of California, Irvine

LUNDS UNIVERSITET
Lunds Tekniska Högskola

UCIRVINE

December 2011

Abstract

In this Master's thesis I present a means of finding active sources of cortical electrical activity from electroencephalogram (EEG) data acquired during operation of a brain-computer interface (BCI). A novel subspace-based technique was used to suppress spatially correlated EEG interference sources, followed by a technique that estimates the source parameters with a near maximum likelihood performance. These sources are found to correlate with event-related potentials (ERPs) and are thus hypothesized to be responsible for the N200 and P300 ERPs. The source localization technique was tested on EEG data of 6 able-bodied subjects, and my analysis underlines consistencies and variation of brain activity locations both within and across subjects. Results are compared to literature and results using other techniques and the new methods show promise in localizing brain activity when dual-condition datasets are available.

Keywords: Localization, Dipole source, Null Space Projection (NP), Noise Subspace Fitting (NSF), Event-related potential (ERP, N200, P300), Electroencephalography (EEG), Brain-Computer Interface (BCI) Speller

Sammanfattning

I det här examensarbetet presenterar jag en metod för att hitta källor till elektrisk aktivitet i hjärnbarken i data som uppmätts vid användning av ett brain-computer interface (BCI) med Elektroencefalografi (EEG). En ny delrumsgrundad metod användes för att dämpa störkällor, följt av en metod som uppskattar källornas parametrar med nära maximum likelihood-precision. De funna källorna korrelerar med event-related potentials (ERPs) och antags ligga bakom N200 och P300 (ERPs). Källokaliseringstekniken testades på EEG data från 6 ej funktionshindrade individer, och min analys understryker överensstämmelser och variation, både inom och mellan dessa individer. Resultaten jämförs med litteratur och resultat med andra tekniker, och de nya metoderna verkar lovande i att lokalisera hjärnaktivitet när aktivitets- samt kontroll-data finns tillgängliga.

Contents

List of Figures

List of Tables

Introduction

Brain-computer interfaces (BCIs) is the term for a group of applications that allow an individual to obtain control over a device without any required motor function, i.e. by using signals from his or her brain. BCIs may be used to restore communication or a motor function that has been lost due to paralysis and thereby vastly improve the quality of life for humans, in particular for patients with some type of damage to the central or peripheral nervous system. Perhaps the most successful example of a BCI system is the so called P300 speller, which allowed a patient with locked-in syndrome (LIS, patients who are awake and conscious, but essentially unable to move anything but their eyes [26]) to communicate with the rest of the world [1]. Other BCI systems include control of wheelchairs [8, 14] and robotic orthoses [13, 19]. The ultimate goal of BCI technology is to integrate with functional electrical stimulation (FES) devices, thus enabling the brain and muscles to be reconnected. Ideally, these neurorehabilitive devices should let the user control a motor function in an intuitive manner, i.e. the BCI and/or user's brain will (co-)adapt so that when the user thinks: "move right", the device (FES controlled limb, prosthesis or wheelchair), will move right.

In addition to restoring communication or movements, BCIs represent a valuable scientific tool, as they allow brain activity to be studied in the context of feedback. By further understanding what happens in the brain when we are exposed to certain stimuli, or attempt a certain action, we can improve applications like the ones mentioned above by fine-tuning their movement, improving their rehabilitative abilities, or quickening the learning curve for using the device.

One important part of getting a deeper understanding of brain information processing is to be able to accurately and easily localize what part of the brain activity originates from. In this thesis, I have investigated methods that use data from electroencephalogram (EEG) BCI experiments to find sources of activity in the human brain after a sensory (in my case visual) stimulus. The questions I try to answer are:

1. Whether it is possible to localize sources of activity, and how to do this in a principled manner by using only 15-channel EEG data.

2. How the sources of activity differ between users with different levels of experience in using the BCI at hand (P300 speller).

3. If, and to what extent the naive user's brain or their approach seem to adapt as they become more familiar with the system.

Having the tools to localize brain signals accurately in three dimensions using a relatively simple method such as EEG, can be very useful in evaluating other BCIs, and may even become useful in clinical applications and diagnosis. Mapping the sources of activity to areas of the brain in this experiment allows us to better understand early processing of sensory stimuli and decision making. On the other hand, having more detailed insight into how the user actually operates a certain BCI system, and how the brain seems to adapt as it learns how to use that BCI, may be very helpful in developing new applications.

Before going into the details (methods, results, discussion) of my project, I will give a short introduction of the principles and methods that are considered routine in the field. Towards the end, I will talk in particular about the P300, and localization of this "event-related potential". For interested readers a very thorough review paper on these topics can be found in Polich [20].

1.1 Electroencephalography (EEG)

The most common and least invasive method for recording brain activity to be used in BCIs is electroencephalography (EEG). This method

is relatively simple, yet it offers a relatively high temporal resolution, making it "ideal" for real-time control of a computer or prosthesis. An electroencephalogram measures electrical potentials on the scalp, and the activity seen in EEG is believed to originate mainly from synchronized post-synaptic potentials in pyramidal cells in the cerebral cortex. Clinical uses of EEG include: diagnosing epilepsy, (epileptic patients will often have abnormal electrical activity), finding tumors and diagnosing brain death. In healthy well-rested subjects certain invariant oscillations (brain-waves) can be seen, named in order of their frequency from low to high: α, β, γ. These and other EEG waves (δ, θ, μ) are associated with physiological states such as sleep, relaxation and intense thinking.

The major drawback of EEG is the very poor spatial resolution, as electrical signals originating somewhere in the brain are distorted by the cerebrospinal fluid, the meninges and the skull, and then recorded on the outside of the scalp by electrodes often separated by at least a couple of centimeters. Additionally, the potentials seen in EEG reflect all kinds of different processes going on in the brain, and the signal-to-noise ratio is consequently quite low.

1.2 Event-Related Potentials (ERPs)

Event-related potentials (ERPs) are electrophysiological responses, often time-locked to some stimulus. By measuring scalp potentials with EEG, and averaging over many trials, one can see a pattern of positive and negative "humps". These are ERPs and are traditionally given names depending on their polarity (N or P) and latency (in ms, e.g. 200, 300) with respect to the stimulus. Some often prominent peaks are the N200 (sometimes referred to as N2) and P300 (or P3).

ERPs are typically recorded with the above mentioned EEG method, but can also be recorded with magnetoencephalography (MEG), which uses the magnetic fields from electrical currents in the brain, and the ERPs are hence referred to as event-related fields (ERFs) instead. Magnetic fields are much less distorted by the matter between the source and the measuring point, and MEG therefore has higher spatial resolution

than EEG. However, it is very expensive and requires a much bulkier apparatus, including a magnetically shielded room and active noise canceling magnetometers.

1.2.1 Oddball Paradigm and P300 Response

The discovery of the P300 event-related potential has been accredited to Sutton et al. (1965) [29] who were examining the response in humans after visual and auditory stimuli, mixed randomly and not mixed. They found that the more frequent stimulus gave a weaker ERP response, meaning that the unexpectedness of the stimulus is crucial. A more thorough investigation of the individual event-related potentials and the postulation of the "oddball" paradigm was then presented by Emanuel Donchin in 1978 [5]. The oddball paradigm is based on the presentation of an infrequent stimuli, and a subject's active response to that cue. They reported ERPs from both auditory and visual experiments. By averaging data from many trials, they were able to show later ERPs thought to be related to cognitive events, such as the N200 and P300, very clearly.

The oddball paradigm is in visual experiments often implemented by having the test subject focus on a screen and being asked to look for an infrequent cue (this is the "oddball"). The cue can be a certain shape [2], color [3] or simple recognition task (words vs. non-words or line drawings vs. scrambled drawings) [12]. The cue is then presented among other stimuli (non-oddballs or "evenballs") which the subject is asked to ignore. The surprise, or cognitive awareness that the user expresses when the oddball is presented, can be seen in the EEG as positive and/or negative ERPs. It is important that the oddball stimulus is infrequent and random, with an *oddball* : *evenball* ratio of typically lower than 1 : 7. If the oddball is more frequent, the response will become a lot weaker.

1.2.2 Details on the Visual P300

The P300 arises whenever a subject has to discriminate between sensory stimuli. It is arousal dependent and results from very basic cognitive events. Johnson et al. [9, 10] predicted that the P300 arises from multiple

different neural generators by analyzing P300 and forming a "triarchic" model. Each of these generators would be responsible for their own neural processes. They said (in 1993) that our understanding for how to ellicit the P300 is far superior to what this ERP really means, and that the functions of the P300 may be much more complex than previously thought.

The idea that the P300 can be split up into subcomponents has since been generally accepted. The two most important, or only, of these subcomponents are P3a and P3b. Previously, theories of "no-go" P300 and "novelty" P300 that would be elicited by not only having a target stimulus, but also a distractor stimulus were popular, but these have later been shown to most likely be variations of the same neural mechanism, namely P3a [4, 22, 25, 27]. Note that the P3a may not be prominent in my experiments as it is typically seen in the above named experiments with distractor stimuli [2], although it is normally apparent in about 10 % of individuals [20].

Based on neurophysiological results, experimental findings and personality ERP variation, Polich hypothesized [20] that the P300 and the underlying subprocesses functional mechanism may involve inhibition of on-going processes where the first action is reallocation of attentional resources (P3a, frontally located) which may cause subsequent promotion of temporal-parietal located memory storage functions (P3b). The exact type of stimulus will determine what the resulting P300 ERP will look like, as the two subcomponents are very much over-lapping each other in time.

P300 latency is dependent on complexity of the task, and can vary from 300 ms up to as much as 1 second. [10].

P300, or the ERP response in general does not look the same between individuals. It has been shown however, that monozygotic twins have very similar P300s, and genetically close individuals in general have similar ERPs [17, 21]. Still, even within the same individual, ERPs will vary based on such seemingly small changes as ultradian rythms (natural processes in the brain that recur in cycles of around 90 minutes) [23]. Ravden and Polich found visual P300 habituation when measuring con-

sequent 10-minute trial blocks: P300 amplitude decreased over time at electrodes F_Z and C_Z (parietal lobe, see figure 2.2), possibly reflecting a biological variance in arousal state over time.

1.2.3 Brain-Computer Interface (BCI) Speller

The oddball paradigm can be utilized in P300 Spellers systems, which was first implemented by Farwell and Donchin in 1988 [6]. In this application, the user is paying attention to a screen showing a virtual keyboard, and the stimulus is the letters flashing. The P300 Speller utilizes the relatively strong response seen as primarily N200 and P300 ERPs. Farwell and Donchin's P300 speller reached a typing speed of 2.3 characters/min by collecting data from a flashing 6×6 grid and subsequently let a computer analyze it to figure out what letter the user tried to spell.

Since then, we have come a relatively long way. Faster computers allow for simultaneous feature extraction, and today's state-of-the-art P300 spellers can reach *online* information transfer rates (ITR) of over 3 bits/s, allowing a user to spell 12.75 characters/minute [31]. The effective rate of spelling sentences could be even further improved by adding auto-completion functions or word suggestions to the screen.

1.3 Localizing Brain Activity

An important step in getting a deeper understanding of the brain and BCI operation is to be able to accurately and easily localize what part of the brain electrical activity originates from. Today, this is typically done with functional magnetic resonance imaging (fMRI) or invasive methods such as Electrocorticography (ECoG), which both have significant drawbacks. ECoG is a very precise method, but it is highly invasive and is therefore used only in clinical applications to localize epileptogenic zones during presurgical planning for epilepsy. fMRI's most significant drawback relative to EEG, is that it is more expensive and complicated, and that it doesn't measure the electrical activity of the brain directly. fMRI measures the hemodynamic response of the brain, i.e. increase

of oxygenated hemoglobin, which presumably is correlated with higher metabolic requirements in regions that are electrically active. Since the time constant of the hemodynamic response is several hundreds of milliseconds, the results are also drastically smeared in time.

1.3.1 The EEG Inverse Problem

As previously mentioned, EEG signals are assumed to arise from post-synaptic potentials in pyramidal cells in the cerebral cortex. These potentials will then propagate through the brain, the meninges and the skull, distorting the signal widely. It can also be shown that reconstructing the sources does not have a unique solution. The reason for this is that there is always a possibility that some potentials cancel each other out, and this is known as the EEG inverse problem.

Instead, a forward model is used to solve the problem of localizing sources. The head is described using either a sphere or a standardized head model from magnetic resonance imaging (MRI), and the meninges and skull are given electrical and geometric parameters. Given this head model and the locations of the electrodes in the experiment, lead field vectors (LFVs) which will represent the response at the scalp from a unit amplitude dipole at position $\mathbf{r}$ and with dipole orientation Φ can be created. If the number of sources of interest is N_S then, mathematically, what is measured at the scalp can be modeled as:

$$x(t) = As(t) + n(t) \tag{1.1}$$

where $x(t)$ is an $m \times 1$ vector representing the measured potentials at the m EEG electrodes, at time t. $s(t)$ is the $N_S \times 1$ signal vector consisting of dipole moments of the N_S unit amplitude dipoles. A is an $m \times N_S$ lead field matrix. $n(t) \in \mathbb{R}^{m \times 1}$ is a term representing the noise, which includes not only sensor noise etc., but also actual EEG sources in the brain that are not of interest.

The parameters (location, dipole orientation and moment magnitude) are then estimated using some *localization* technique, such as multiple signal classification (MUSIC) [15, 24] which finds the $A(\mathbf{r})$ that is most

orthogonal to the noise, or a novel technique called noise subspace fitting (NSF) [32] which will be presented in the methods section.

1.3.2 Previous Work

The P300 cluster and its origin has been a research interest for a long time. However, success in localizing the actual sources for this cognitive function has been very limited. P300 has been defined as a positive ERP component with a maximum at C_Z or P_Z electrode (central part of parietal lobe, see figure 2.2) between 300 ms and 1 s after stimulus [10]. This quite wide definition may be one reason to the fact that many studies show quite different results for where activity actually originates from. Bledowski et al. [2] also found that many people have tried to localize P300 generating nuclei, with varying success. They found that there are significant discrepancies between results and explained it with the fact that there is no unique solution to the inverse problem.

One way to improve EEG localization is to do concurrent fMRI and EEG [16], to get a type of "seeding" for the EEG tomographic analysis, and to get a more accurate model for the head and brain. For example one could benefit from knowing the actual thickness of the cranium, instead of just assuming a uniform thickness which you have to do if you use a spherical head model. Simultaneous fMRI and EEG is hard to do practically, because the EEG system needs to be compatible with the strong magnetic field in the MRI machinery. To work around this, sometimes the subject first performs the experiment in the MRI, and then afterwards does the same routine with an EEG cap on. This, of course, imposes its own complications. (Difference in subject performance, previously named ultradian rhythms, etc.). The possibly most successful of these studies [2], found that the P3b was mainly produced in parietal and inferior temporal areas, and that P3a came from the frontal lobe and the insula.

Perhaps the most accepted tool for localization of brain activity is a free software called "low resolution brain electromagnetic tomography" (LORETA, described in methods section). LORETA, however, only eval-

uates a type of "heat map" of what regions of the brain are active, rather than pinpointing an actual dipole location.

To conclude this chapter: There are multiple separate components to the P300, and they are most likely located mainly in the parietal lobe, but maybe also in the frontal regions of the cortex. The P300 is related to simple cognitive processes, such as decision making and working memory. In the coming sections, I will present methods – and my results and experience in working with them – with the aim of finding sources of activity in brains that are operating the previously described brain-computer interface, a high-speed (superior information transfer rate) P300 speller.

Methods

In this chapter, I describe the methods: From experimental setup to data processing to how I get the results. Figure 2.1 summarizes all the steps of finding sources of activity in my project, as well as the P300 speller application itself (within dashed lines). A subject looks at the screen and in response to visual stimuli his/her EEG is recorded (1). The data is saved on a computer hard drive for subsequent offline analysis. The first stage of localization is to evaluate the most likely location for any given time point (2) and the last step is to figure out which of these sources actually contribute to shaping the ERPs (3). The details of the plots in this figure will be discussed in the Results chapter.

2.1 P300 Speller Experiment

The experimental setup to acquire data for the localization described in this thesis is in essence identical to that of a recently developed P300 speller system developed by Nenadic et al. [31]. The only difference is that I use 15 electrode EEG, instead of 8. This is visualized in figure 2.2. Electrode A2 (right ear tip) was used as reference point.

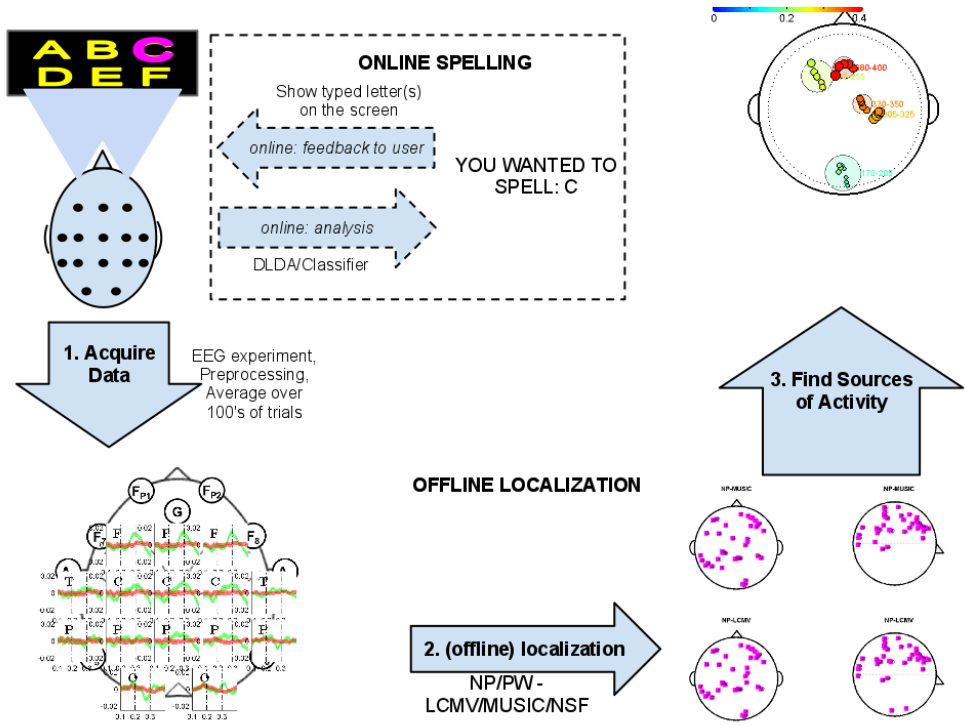

Figure 2.1: Process for localizing sources of activity in P300 speller EEG data, and schematic description of actual (online) application (within dashed rectangle).

2.1.1 Equipment

The subject is wearing an EEG cap (Compumedics USA, Charlotte NC), with 15 electrodes placed according to the 10-20 EEG standard. Signals are improved by lowering the contact impedance through a combination of conductive gel (Compumedics USA) application and a blunt needle skin abrasion. The cap is connected to an amplifier (Biopac, Goleta CA) which amplifies the signals and performs 16-bit A/D conversion. The sampling frequency is set to 200 Hz and an analog band-pass filter

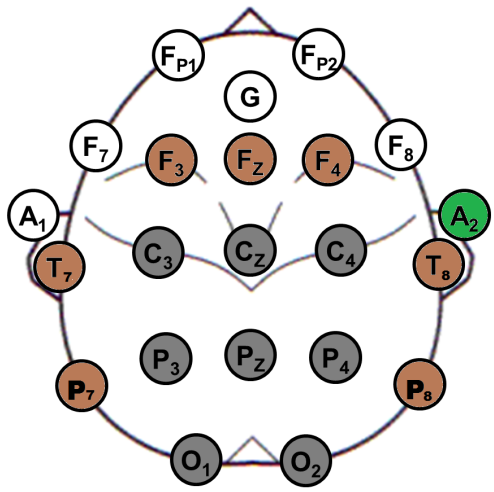

Figure 2.2: 15-channel EEG setup (brown and gray) and 8 channel setup used in [31] (gray only). In green is the reference electrode, on the right ear.

with frequency band $0.01-30\ Hz$ is used. A standalone ear clip Ag-AgCl (silver chloride) electrode on the right ear is used as a reference electrode.

2.1.2 Population Study

During an experiment a subject (typically) did 3 training sessions and 3 online sessions, with an alpha wave gauge and an impedance measurement between ground and each channel done at least once, before the first data acquisition. Impedances were kept at $< 5\ k\Omega$ at $30\ Hz$ and were recorded for later use. The three different sessions were done at three different interface speeds: low, medium and fast (referring to the *dwell time* between consecutive flashes: $400, 240\ and\ 170\ ms$ respectively). The significance of the dwell time parameter is discussed in the next section. In the online section, the subject was asked to spell a sentence to prove that he or she had purposeful control of the speller, and the offline dataset was discarded if the subject could not spell. An online

mode screen capture at fast interface speed can be found on the web [www.youtube.com/user/UCIBCI]. Below is a test subject in an experimental situation.

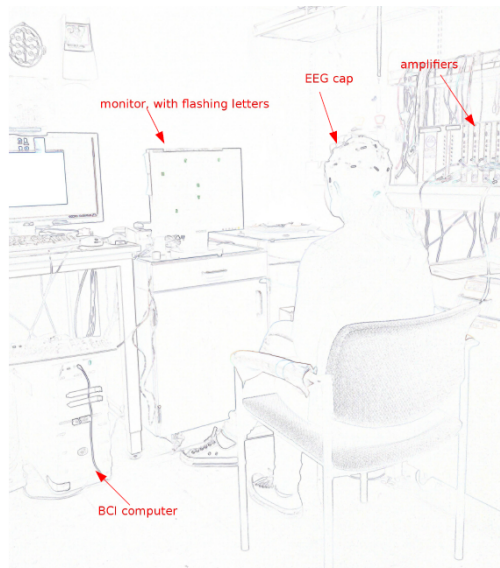

Figure 2.3: Sketch of expermimental situation. Seen is a subject, paying attention to flashing letters on a computer screen (center). Seen also is various lab equipment, such as BCI computer (bottom left) and amplifiers (right).

2.1.3 Offline "Training" Stage

The experiment starts with a training session, during which the subject is presented with a 7×6 character (A-Z, 0-9 and some miscellaneous characters) virtual keyboard on a computer monitor, and is asked to pay attention to one of these characters (called the "oddball", defined in subsection 1.2.1). 6 random letters at a time will then repeatedly flash, with an interval of 170, 240 or 400 ms (depending on dwell time

setting) for about 30 seconds. This process is repeated with 10 characters in varying locations on the screen, resulting in a dataset of 700 to 1200 trials (depending on speed setting) with an *oddball* : *evenball* ratio of 1 : 7. Data is recorded for 400 ms (referred to as *sample time*) after each presentation (oddball or non-oddball) and hence, for medium and fast speeds, since the sample time is longer than the dwell time, there is an overlap between trials. Data collected following an oddball is referred to as "activity" data and data following an evenball is then the "control" data.

The end result of data collection is essentially a: 15 channel x 80 time samples (400 $ms \cdot 200\ Hz$ (*sample time*$\cdot$*sample rate*)) x a few hundred to a thousand trials - dataset, of control and activity data lumped together. There is also a label corresponding to each trial, to keep track of the two types of data.

This data is used to train the BCI computer to recognize the pattern for the brain's reaction to an oddball stimulus, and a filter is created that extracts channels and time periods that are responsible for encoding the difference between "control state" (evenball) and "activity state" (oddball) data.

2.1.4 Online "Spelling" Stage

Once the computer has been trained a user proceeds with an online stage. Here, the user can determine what letter he or she wants to type. The letters will again randomly flash, with more common characters of the English alphabet, as well as expected letter combinations (based on built-in dictionary), flashing more prominently. Since the BCI computer knows the subject's pattern for an oddball stimulus, a classifier can quickly determine if the user was attending to any of the letters that just flashed and so figure out what character the user wanted to type.

During the online session, the subject was asked to spell the following sentence:

"The quick brown fox jumps over the lazy dog*",

a 44 character (counting spaces and *, which is the sign to exit the interface) English-language pangram.

In this study, the online spelling session was only used as a benchmark test of whether the subject had achieved purposeful control of the BCI or not. If a subject could not spell online, the corresponding offline data was determined to be of non acceptable quality.

2.2 Preprocessing and Denoising

The raw data acquired from each training session of the experiment can now be divided into control state data, X_c, and activity state data, X_a, with a number of trial ratio $X_c : X_a$ of about $7 : 1$. Due to the acquisition overlap for fast and medium speeds mentioned above, as well as the long response after an oddball (up to about 1 *second*) some control state data trials need to be removed, reducing the $X_c : X_a$ ratio to about $4 : 1$. Then, the rawdata needs to be adjusted for bias and impedance. It was also sent through one of two noise-suppressing algorithms (prewhitening or noise subspace fitting), all this is described in detail below.

2.2.1 Data Preprocessing

First thing to do is to divide all trials for each electrode by the impedance measurement from the experiment. In some cases, only one impedance measurement (from the beginning of the experiment) was available, and every dataset from that day was then divided by that measurement. In other cases each dataset had its own corresponding impedance measurement and was adjusted accordingly. Either way, I found that it did not much of a difference as impedances were quite stable throughout the experiment.

Next, some trials were removed. Since our data acquisition for control and activity state data is continuous, there will be some activity (ERPs) in what is captured and labeled as control data immediately preceding (fast interface speed) and following (slow, medium and fast interface

speeds) an oddball. In fact, the oddball response may be longer than the 400 ms that we use as sample time, and we are presented to a trade-off problem between removing trials following an oddball (and thereby minimizing activity contamination in control state data) and keeping control state data trials (and improving the statistical averaging).

Trials were then averaged, control data X_c and activity data X_a separately, resulting in two 15 (channels) by 80 (number of time samples) matrices.

Lastly, to adjust for sensor bias, for each sensor individually, the average value for all data of that sensor at $t = 0$ was subtracted from the whole 80 sample data such that each channel has a total average potential of 0 at time 0.

Once the data is preprocessed and ready to use, the last step is to find the "Number of Interfering Sources" (N_I) in the control state data, a parameter that is required for null space projection described below. This number was found as the number of singular values of X_c needed to satisfy a power ratio of .99, meaning 99% of the energy in the signal can be recovered by N_I sources.

2.2.2 Prewhitening (PW)

The whitening matrix is found by first estimating the covariance matrix of control state ("no activity") data, with the mean subtracted $\bar{X}_c(t)$ with n samples as:

$$\hat{R}_C = \frac{1}{n}\sum_{t=1}^{n} \bar{X}_c(t)\bar{X}_c^T(t) \tag{2.2}$$

where $\bar{X}_c(t) = X_c(t) - \frac{1}{n}\sum_{t=1}^{n} X_c(t)$. Whitening of the activity state data $X_a(t) = As(t) + n(t)$ (equation 1.1) is then achieved by multiplying with the whitening projection matrix, $P_w = \hat{R}_C^{-1/2}$:

$$X'_a(t) = P_w \cdot X_a(t) = \underbrace{P_w \cdot A \cdot s(t)}_{''unchanged''} + \underbrace{P_w \cdot n(t)}_{\text{white}} \tag{2.3}$$

The trick is that the noise vector is now (closer to being) white, i.e. any systematic information in the noise will be suppressed and thereby influence analysis of $X'_a(t)$ as little as possible.

2.2.3 Null Space Projection (NP)

A more sophisticated and newly developed method, null-space projection (NP) [32] relies on the presence of control state data, which is used to create a subspace that is orthogonal to the noise and that the activity state data then can be projected onto. This method requires knowledge of, or rather an estimation of, how many interfering sources, N_I are present. These interfering sources are not assumed to be static, but are assumed to be essentially the same in control and activity data.

$$Control : X_c = A_i S_{ic} + W_c \tag{2.4}$$

$$Activity : X_a = A_s S_s + A_i S_{ia} + W_a \tag{2.5}$$

where X_c and X_a now are the 15×80 control and activity data matrices described in the preprocessing section above. The noise is then broken up into sources of interference, S_{ic}, S_{ia} which are $N_I \times 1$ vectors, and unorganized noise terms, W_c, W_a. A is the m $(number\ of\ electrodes) \times N_S$ lead field matrix as described in equation 1.1.

The $m \times m-1$ nullspace H is then found by:

$$\hat{H} = \underset{H}{\operatorname{argmin}} ||H^T X_c||_F^2 \tag{2.6}$$

and with the decomposition of the control state measurements:

$$X_c = U\Sigma V^T = \begin{bmatrix} U_I U_N^C \end{bmatrix} \begin{bmatrix} \Sigma_I & 0 \\ 0 & \Sigma_N \end{bmatrix} V^T \tag{2.7}$$

the solution to equation 2.6 is $\hat{H} = U_N^C$ and the projection matrix can be calculated as $P_H = \hat{H}\hat{H}^T$. Finally, the new activity state data, X'_a is:

$$X'_a = P_H X_a = P_H A_s S_s + P_H W_a \tag{2.8}$$

which is "free" from interfering sources (S_{ia}), compare to X_a in equation 2.5.

2.3 Localization

After the data has been pre-processed, it will be sent to a localization algorithm. There is a multitude of options here, and I have used a couple of traditional methods: linearly constrained minimum variance (LCMV) beamforming [30] and multiple signal classification (MUSIC) [15, 24] as well as a newly developed method: noise subspace fitting (NSF) [32]. I also used a free software package called standardized low resolution brain electromagnetic tomography (sLORETA) which gives a distribution of origin of activity. MUSIC, LCMV and the fundamentals for NSF were all implemented and thoroughly described in [32] but are for quick reference briefly presented below.

2.3.1 MUltiple SIgnal Classification (MUSIC)

MUSIC [15, 24] finds the LFVs that are most orthogonal to some noise subspace. The location $\mathbf{r}$ for that LFV is then the most probable location of a source, and the corresponding Φ is the dipole orientation.

2.3.2 Linearly Constrained Minimum Variance (LCMV) Beamforming

LCMV beamforming [30] finds the sources by constructing beams by assuming unit strength from all possible source locations, and by then minimizing array output power from all other locations.

2.3.3 Noise Subspace Fitting (NSF)

Recently, a localization method originally developed for direction finding [28] has been adapted to find highly correlated sources in EEG [15, 24]. It is called noise subspace fitting, and is asymptotically equivalent

to maximum-likelihood approaches, but is much faster because dipole location and moment are decoupled.

Minimum Sample NSF

NSF was developed with the intention that a period of interest (in time) would already be available or loosely defined and localization would then be done with data from that time window. I found however, that using only one sample in time and localizing sources for consequent time samples, was quite efficient. It can also be shown that this does not violate the data model or calculations in the method. For N_S set to 1, as little as 1 time sample can be used, and for $N_S = 2$ at least 2 time samples must be used. Locating more than 2 sources was never relevant for me, not only because we expected 1 or maybe 2 sources to be present at any given time, but also because I had only 15 channels in the experiment, which would make locating 3 sources practically impossible.

Identifying Relevant Sources

Accompanied to the single sample NSF I developed an automatic cluster finder. We hypothesized that consecutive single sources located in close euclidean proximity to each other was a sign of a "true" source. (Remember that NSF always will give some location for the number of sources you are trying to find, but that it relies on the user to specify how many true sources there actually are, which in some cases can be 0.) The basic requirement is that a "source" must have a dipole moment of about 1/5 or more of that of the strongest dipole in the dataset. This to avoid very weak, probably randomly formed clusters. Then, the actual cluster finder relied on two simple euclidean distance requirements:

1. Two consequent sources must not be separated with more than a certain distance, and

2. At least 5 consecutive sources must satisfy (1).

2.3.4 Standardized Low Resolution Brain Electromagnetic Tomography (sLORETA)

sLORETA [7, 11, 18] is a tomographic tool for EEG data, available for free at [www.uzh.ch/keyinst/loreta]. It uses an averaged head model from real MRI scans to give a type of density function for what areas of the cortex are active. Loreta is quite accepted to be correct, and I used it as a type of validation for that the NSF method gave reasonable results.

Results

3.1 Raw Data

Each subject did 1-3 (depending on availability) sessions of the experiment described in the methods section 2.1.3. I had 6 subjects, a total of 10 experiments, and 24 datasets that passed the online spelling test, presented in the table below. Three of these subjects (A, E and F) were "experienced" users, meaning that they had spent at least a total of 20 hours using the P300 speller and other BCIs. The remaining subjects (B, C, D) had never used a BCI before this study, and are called "naive" users.

Table 3.1: Table over datasets. Datasets where the subject failed in the corresponding online session have been struck out. Note that only one subject completed the whole three sessions. *Subject A - session 2 was used for testing consistency. **Subject F was only available for completion of one single dataset.

Subject	session 1	session 2	session 3
A	slow, med, fast	med, med, fast, fast*	
B	slow, med, fast	slow, med, fast	med, fast, slow
C	~~slow~~, ~~med~~, fast	med, ~~fast~~, slow	
D	~~slow~~, ~~med~~, ~~fast~~	~~med~~, fast, ~~slow~~	
E	slow, med, fast		
F	fast**		

After preprocessing as described in methods 2.2.1, the data for each individual channel can be plotted and the ERPs are clearly visible. To show difference between individuals, below are results from two experienced subjects (A and E), and the most successful of the naive users (B), all at medium interface speeds (figure 3.4). Note, in particular, the strong N200 (occipital channels) in both the experienced subjects, and the strong P300 (F_Z, C_Z) in the naive subject.

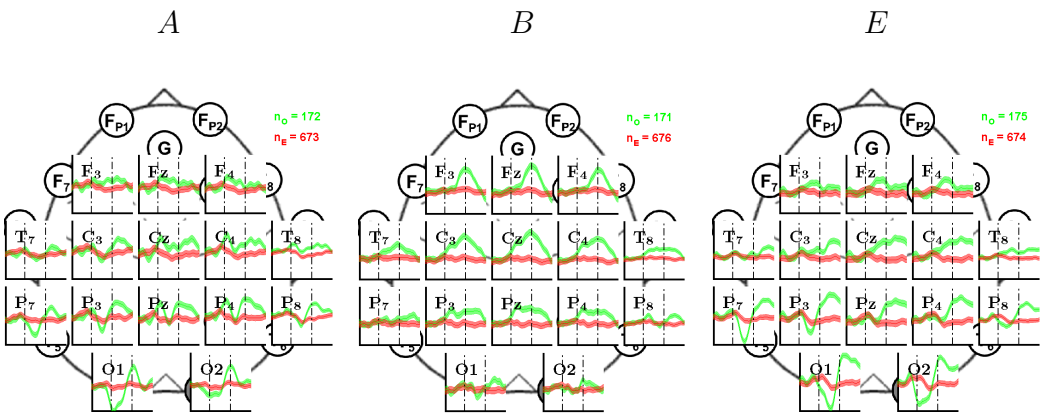

Figure 3.4: 15-channel ERP plots for three subjects (left: subject A, center: B, right: E). Each of the 15 subplots shows data for that particular channel from 100 to 400 ms after stimulus and spans 14 μV from top to bottom. Control state data is colored red, and activity data is green. Data seen is average of hundreds of trials with standard error added. (See n_O: number of oddball trials, and n_E: number of evenball trials.)

3.2 Case Study For Localization

Here, I present results for the different localization methods I used. Subject B, session 3, medium speed is studied. Please refer to the above ERP plot (3.4 center) to see all the data from this dataset. All results are

from the same time windows, namely 200-245 ms post-stimulus (N200) and 255-325 ms (P300), which were the N200 and P300 time windows automatically found by the source finder described in 2.3.3. (The results for this are presented further down 3.2.)

Scalp Topography

First out is a scalp topography, which was simply produced by taking the average values for each sensor over each time window, and interpolating values between the sensor positions. These values are then plotted as hot or cold depending on their deviation from 0. This is maybe simplest and most primitive way to localize activity and can be compared to just studying the ERP plot and manually estimate a source location. Note how there seems to be no activity in the occipital lobe even for N200.

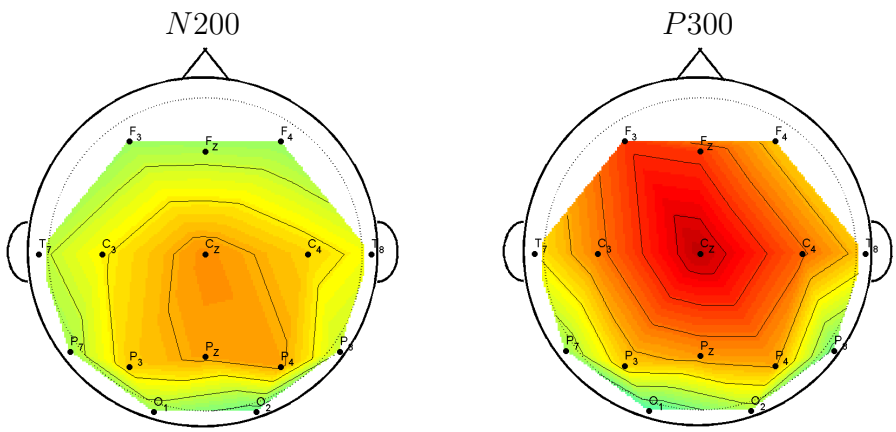

Figure 3.5: Scalp topographies for subject B: N200 (left) and P300 (right). Scale is the same in both plots.

NP-MUSIC

The below results are created with null space projection (NP, 2.2.3) and multiple signal classification (MUSIC, 2.3.1). The number of interfering

sources (NI) was set to 5, which corresponds to a ratio of 99 %, automatically set as described in methods (2.2.1). Results for LCMV are not shown as they yielded virtually the same results as MUSIC for $N_S = 1$, and because I chose to chiefly work with MUSIC. I also did localization for N_S set to 2, but it seemed to almost never be significant, i.e. there was always only one dominating source. Prewhitening results were in general not as good as NP and are not shown here.

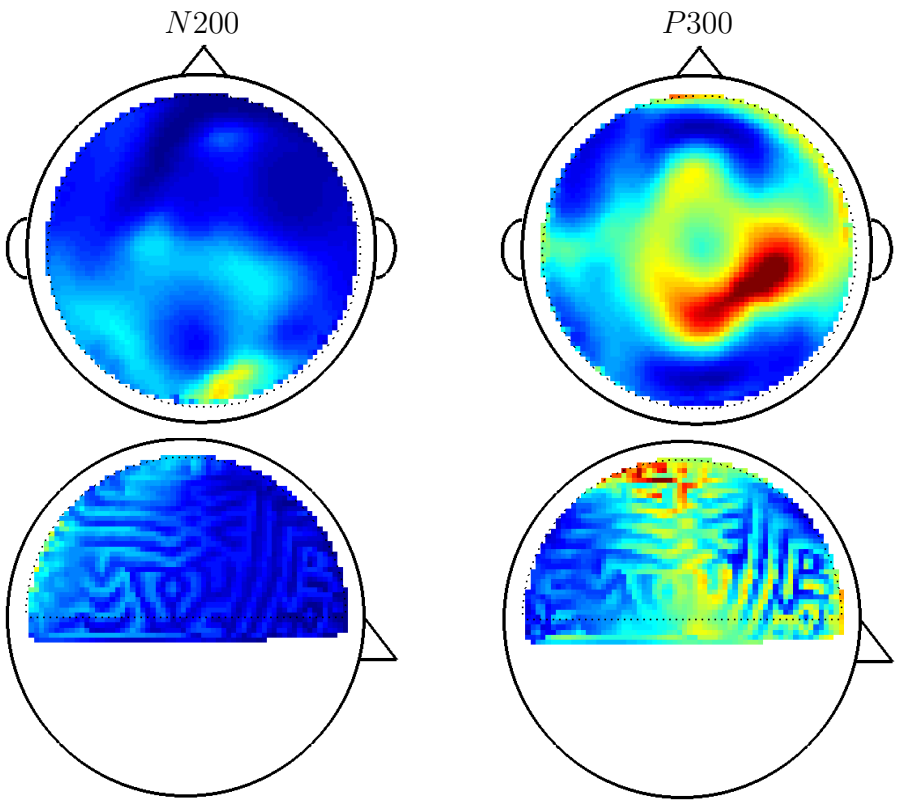

Figure 3.6: NP-MUSIC: N200 (left), P300 (right). Hot regions represent locations that are orthogonal to the noise subspace, which means they are likely to host a source.

Minimal sample NP-NSF

Finally, using NP, single time sample null subspace fitting (NSF, 2.3.3), and automatic clustering 2.3.3 I found three clusters (figure 3.7), whereof two correspond to N200 and P300. The time windows used in previous analysis (scalp topography, NP-MUSIC) were found using this method.

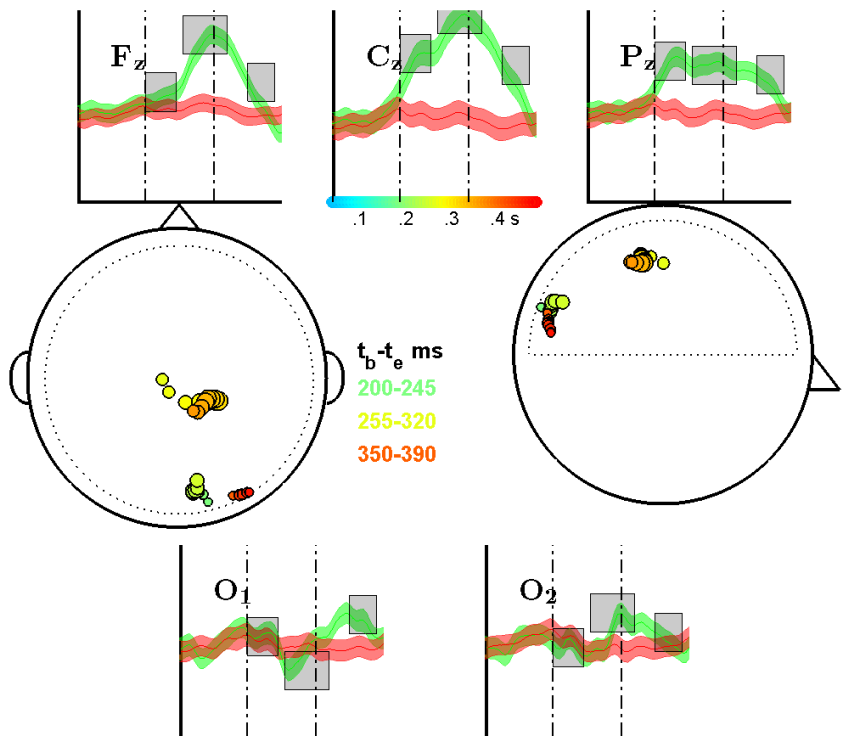

Figure 3.7: Minimal sample NP-NSF sources. Dots are color coded with respect to time and correspond to the location found by the method for that particular time sample. Dot sizes are related to the dipole moment magnitude. In the inserts that show the original data, the time windows are shaded to show what parts of the data the clusters correspond to.

sLORETA

For varification and comparison sLORETA is shown below. As sLORETA is evaluated for each time sample individually, I strictly used the sLORETA for the first time frame (200 ms (N200) and 255 ms post-stimulus for P300).:

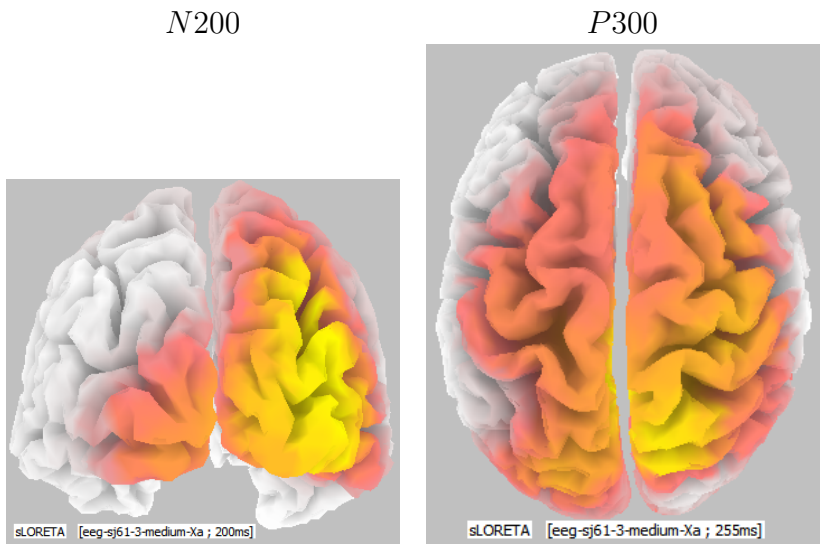

Figure 3.8: sLORETA: N200 (left, posterior view), P300 (right, top view). A standardized head model is colored with red and yellow for increasing neural activity.

3.3 Results For All Subjects

Using the methodology described above, with the difference that prewhitening (PW, 2.2.2) was used instead of NP, source clusters were evaluated for all 24 datasets. Sources were then picked out depending on time windows, matching N200 and P300. Not all datasets included potential N200 and P300 sources, and the dataset was in that case excluded. Here, as opposed to in the previous section, all the markers are the same size. This because dipole moments are not comparable between datasets. Instead, different markers are used for experiments performed in different sessions (different days). MUSIC heat maps were also evaluated for all datasets, but since multiple heat maps could not be combined in one plot, a representative one is shown. The situation is the same for sLORETA. Only one view is shown for MUSIC and sLORETA as the results were always confined to the cortex for both the methods, and hence the results can be visualized in two dimensions.

3.3.1 N200 source location

On the next page are all dataset with clusters starting around 200 ms post-stimulus (i.e. supposedly N200 source locations). Note the nice correlation between the three methods, especially for the experienced subjects, and subject B. Subject D is excluded here because no N200 source was found for the one dataset that passed the online test.

3.3.2 P300 source location

Doing the same for clusters that start around 300 ms (P300 locations) yielded the following plots for the 6 subjects. Results are not as unambiguous here, and this will be discussed in the last chapter.

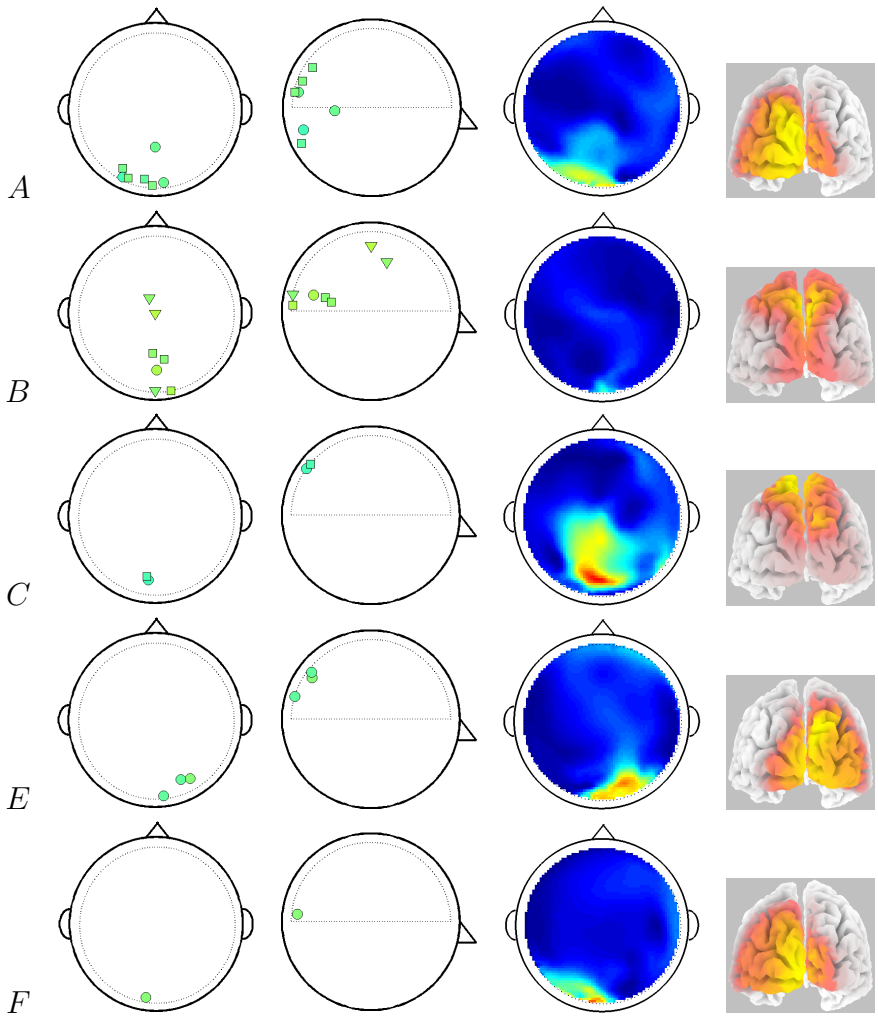

Figure 3.9: Combined results for N200 localization in subjects (left to right): A, B, C, E and F. From top to bottom: PW-NSF method top view, PW-NSF side view, PW-MUSIC typical result (single dataset), sLORETA typical result (single dataset).

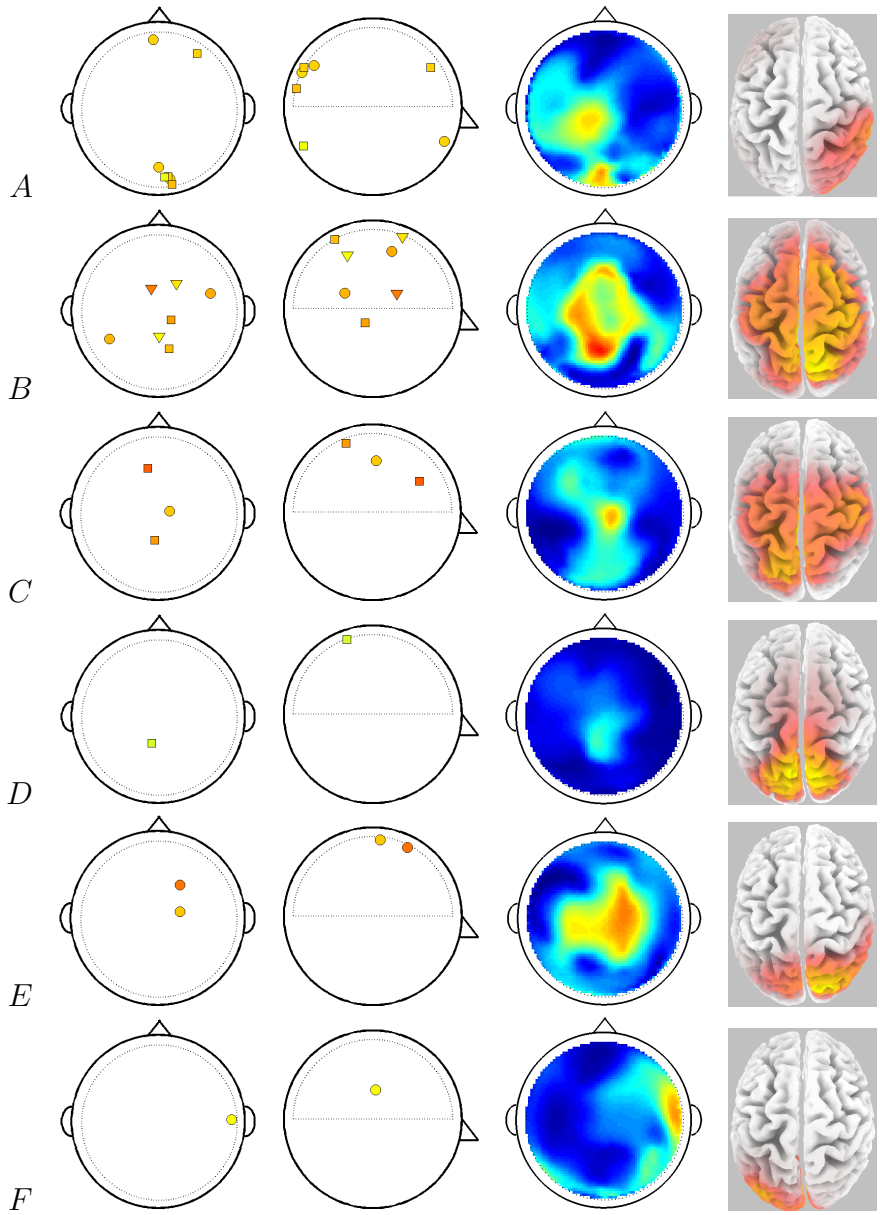

Figure 3.10: Combined results for P300 localization in all subjects (left to right): A, B, C, D, E and F. From top to bottom: PW-NSF method top view, PW-NSF side view, PW-MUSIC typical result (single dataset), sLORETA typical result (single dataset).

Discussion

4.1 Comments On Results

4.1.1 Localization Methods

I found, using previously described methods *null-projection*, *noise subspace fitting* with *euclidean time-dependent clustering* (2.3.3) that there seem to be clusters of activity that highly correlate with ERPs found in EEG, and additionally, they are located in reasonable areas of the brain, based on literature. More interestingly, the NP-NSF single sample-clustering method, found similar clusters in datasets that had no apparent ERPs (see figure 3.7, which has no visible N200), which speaks for the power of this method.

Shown in figure 3.7 is an example of locations that make sense. The last source of activity (red) might be something late that is going on, or just an artifact from data. The dots are smaller which means the evaluated dipole moment was smaller, which I have found to be typical for clusters that seem to have developed from non-interesting trends in data. In fact, as mentioned in the methods, I do discriminate for sources with too small dipole moments but this particular cluster passed the test, i.e. the 'dots' are "large enough".

Localizing the P300 can be hard when using NP, because the method is very effective in removing any activity that is shared between control and activity data. A few P300 peaks in control data, i.e. the subject falsely reacting to an evenball, may cause major problems for NP to locate P300, since the actual P300 then may be canceled out in the noise

subspace. This is one of the advantages prewhitening has against the more sophisticated null space projection. Another problem with NP was to estimate the NI parameter: The 99 % automatic method worked fairly well, but often results were better (more reasonable) if I tweaked the NI parameter + or - 1. An alternative would be to consider using PW instead, which does not require any input except for the control state data. Prewhitening seemed more "stable" with respect to small changes in data, but also far less powerful.

I also tried testing for consistency in localizing sources in same subject, same day (see * in table 3.1) and results were not shown because they were non-conclusive. Again, this may be because of the experiment. ERP data plots differed, maybe just enough to make locations deviate. Using only 15 channels may be too sensitive with respect to variation in measurements.

Lastly, it may be worth noting that by simply looking at the original ERPs, or the scalp topographies, one would hardly be able to figure out where the dipole sources that constructed the measurements were located. Proving the obvious need for more advanced localization.

4.1.2 ERP Locations

From figure 3.9 and 3.10 it is evident that N200 quite consistently was found in the occipital lobe and that no particular P300 location was found.

Maybe we just had a too complicated experiment, with too many parameters. We're doing this in an actual application environment, with more parameters than we can count. Just figuring out what was N200 or P300, among other processes that seem to be going on, was hard. By looking at the ERP plots in 3.4 this can be understood. These are some of the best datasets I had, and it is not necessarily clear were the different ERPs are. It is encouraging, that despite this, NP-NSF still managed to pick out a source around 200 ms, localized to the occipital lobe (figure 3.7).

As mentioned, it is quite clear that N200 is located in the occipital

lobe. As for P300, you need to try hard to see a pattern. I would explain this with an array of reason that are somewhat depending on each other: First of all, very few of the subjects or their datasets had clear P300s. In many of the ERP plots, it seems that there are multiple components and variations in the P300 topology, which would also be supported by the literature. As I wrote in the introduction, different neural nuclei are expected to "work together" to produce the P300, and that a source's individual strength is dependent on how the subject is responding to the oddball, which is not necessarily the same throughout our experiment. So, can the localization methods not deal with this? I think they could, there are a few problems however: To estimate more than one source with only 15 channel EEG can be very difficult and sensitive to interference or slight variances in data, for any reason. This, in a way, leads to the next problem: To use noise subspace projection (NP) which I found to be very effective in suppressing interfering sources, one has to estimate the N_I parameter (the number of interfering sources), which can be done automatically using the 99 % method described in 2.2.1. This method, however, is not perfect, and I found a few times that altering N_I by plus or minus 1, could improve (give more consistent) results. Lastly, it all comes down to the fact that we do not have access to the ground truth, i.e. we do not know when we have actually found good input parameters and when the localization works optimally. Additionally, same subject but different experiments will have an error from simple things such as the EEG cap not being in the exact same position between two experiments, and of course previously mentioned naturally occurring biological processes in the brain.

4.2 Conclusions

N200 was consequently located in the occipital lobe, which is expected and very satisfying, as that is where the visual cortex is located and the N200 results from physiological events that are in the very first few visual processing stages. I was not able to draw any conclusions for a definitive P300 location, but as said in the introduction, the P300 is expected to

have multiple neural origins. Taking into account the low number of electrodes (15) and the complexity of this experiment (not optimized for localization by any means, this is an actual BCI application) I would say that minimal sample NSF in conjunction with NP (with some tweaking) or PW, shows promise for localizing sources of activity in EEG (and MEG) data.

On the question of whether there is any difference between experienced and naive users, it is hard to say anything with such few (3) successful subjects. Given knowledge of earlier experiments [31], a hunch would be that experienced users have developed stronger early ERPs such as the N200, which my results would support (figure 3.4).

It is also hard to conclude anything about the adaption in naive users, partly because three sessions might be too few, and partly because I only had one naive subject that completed the three sessions. If anything, ERPs became stronger (higher amplitude) and sources became more clear as the subject used the interface more.

4.3 Future Work

To really test the ability of NSF, localization on data from a more controlled experiment, maybe even a saline model head with absolutely known source locations, number of interfering sources in the brain, etc, may be needed to learn more about how NSF works in EEG source localization, in practice. NSF was, when it was implemented, tested on digitally simulated EEG data, as well as real EEG (auditory) data[32], but this real-world application proved very difficult. The fact that I used only 15 electrodes was probably an issue as well, and using more electrodes should increase robustness with respect to data variance. A big problem in this study was probably that there were too many parameters in our experiment (stimulus, i.e. the letters, were spread out across the visual field, subject movements were not monitored, etc.), and that I did not know the ground truth.

Finding a better way to estimate NI, and maybe doing NP on subsets of the full control state data would probably be a good next step. I found

the NP method to be very powerful in suppressing interfering sources, but this also works against it, as in the case of the P300: Sources may be apparent, but weaker, in the control data, but still cancel out actual activity almost completely.

Acknowledgements

Thank you to Zoran Nenadic, Ph. D, for giving me the opportunity to do this project, and for teaching and mentoring me all the way through. Thank you to Shun Chi Wu for providing the Matlab code and helping me with further understanding of NSF. Thank you very much to the volunteers who did the experiments, and thank you Po T Wang, Christine E King and An Do, M. D, for helping me conducting the experiments, and for being great co-workers in general!

Thank you also to my roommates and teammates of the Billy B's Softball Club, for winning softball games and for driving me to work: Robbie Henetz, Steven King, Micaela "The Blind Pony" McGinnis, Mark Pleknic, Nikkie "Mansfield" Quinn, Mike "Andrew" Sevilla, and Ian Walsh. At last, a special thank you to Jun Wang, Dexter McCluster and Tim Tebow.

Bibliography

[1] N. Birbaumer, N. Ghanayim, T. Hinterberger, I. Iversen, B. Kotchoubey, A. Kubler, J. Perelmouter, E. Taub, and H. Flor. A spelling device for the paralysed. *Nature*, 398(6725):297–298, Mar. 1999.

[2] C. Bledowski, D. Prvulovic, K. Hoechstetter, M. Scherg, M. Wibral, R. Goebel, and D. E. J. Linden. Localizing p300 generators in visual target and distractor processing: a combined event-related potential and functional magnetic resonance imaging study. *J Neurosci*, 24(42):9353–9360, Oct 2004.

[3] M. E. Cano, Q. A. Class, and J. Polich. Affective valence, stimulus attributes, and p300: Color vs. black/white and normal vs. scrambled images. *International Journal of Psychophysiology*, 71(1):17 – 24, 2009.

[4] L. A. Combs and J. Polich. P3a from auditory white noise stimuli. *Clin Neurophysiol*, 117(5):1106–1112, May 2006.

[5] E. Donchin, W. Ritter, and C. McCallum. *Cognitive psychophysiology: the endogenous components of the ERP. In: Brain Event-Related Potentials in Man.* Academic Press, New York, 1978.

[6] L. Farwell and E. Donchin. Talking off the top of your head: toward a mental prosthesis utilizing event-related brain potentials. *Electroencephalography and Clinical Neurophysiology*, 70(6):510 – 523, 1988.

[7] M. Fuchs, J. Kastner, M. Wagner, S. Hawes, and J. S. Ebersole. A standardized boundary element method volume conductor model. *Clin Neurophysiol*, 113(5):702–712, May 2002.

[8] F. Galán, M. Nuttin, E. Lew, P. W. Ferrez, G. Vanacker, J. Philips, and J. D. R. Millán. A brain-actuated wheelchair: asynchronous and non-invasive brain-computer interfaces for continuous control of robots. *Clin Neurophysiol*, 119(9):2159–2169, Sep 2008.

[9] R. Johnson, Jr. A triarchic model of p300 amplitude. *Psychophysiology*, 23(4):367–384, Jul 1986.

[10] R. Johnson, Jr. On the neural generators of the p300 component of the event-related potential. *Psychophysiology*, 30(1):90–97, Jan 1993.

[11] V. Jurcak, D. Tsuzuki, and I. Dan. 10/20, 10/10, and 10/5 systems revisited: their validity as relative head-surface-based positioning systems. *Neuroimage*, 34(4):1600–1611, Feb 2007.

[12] A. Khateb, A. J. Pegna, C. M. Michel, T. Landis, and J.-M. Annoni. Dynamics of brain activation during an explicit word and image recognition task: an electrophysiological study. *Brain Topogr*, 14(3):197–213, 2002.

[13] C. E. King, P. T. Wang, M. Mizuta, D. J. Reinkensmeyer, A. H. Do, S. Moromugi, and Z. Nenadic. Noninvasive brain-computer interface driven hand orthosis. In *33rd Annual International Conference of the IEEE EMBS Boston, Massachusetts USA, August 30 - September 3, 2011*, 2011. http://cbmspc.eng.uci.edu/PUBLICATIONS/ceking:11.pdf.

[14] R. Leeb, D. Friedman, G. R. Müller-Putz, R. Scherer, M. Slater, and G. Pfurtscheller. Self-paced (asynchronous) bci control of a wheelchair in virtual environments: a case study with a tetraplegic. *Comput Intell Neurosci*, page 79642, 2007.

[15] J. C. Mosher, P. S. Lewis, and R. M. Leahy. Multiple dipole modeling and localization from spatio-temporal meg data. *IEEE Transactions on Biomedical Engineering*, 39(6):541–557, 1992.

[16] C. Mulert and L. Lemieux. *Chapter 5 What Can fMRI Add to the ERP Story? In: EEG-fMRI: Physiological Basis, Technique and Applications*. Springer, 2010.

[17] S. O'Connor, S. Morzorati, J. Christian, and T.-K. Li. Heritable features of the auditory oddball event-related potential: peaks, latencies, morphology and topography. *Electroencephalography and Clinical Neurophysiology/Evoked Potentials Section*, 92(2):115 – 125, 1994.

[18] R. D. Pascual-Marqui. Standardized low-resolution brain electromagnetic tomography (sloreta): technical details. *Methods Find Exp Clin Pharmacol*, 24 Suppl D:5–12, 2002.

[19] G. Pfurtscheller, C. Guger, G. Müller, G. Krausz, and C. Neuper. Brain oscillations control hand orthosis in a tetraplegic. *Neuroscience Letters*, 292(3):211 – 214, 2000.

[20] J. Polich. Updating p300: An integrative theory of p3a and p3b. *Clinical Neurophysiology*, 118(10):2128 – 2148, 2007.

[21] J. Polich and T. Burns. P300 from identical twins. *Neuropsychologia*, 25(1, Part 2):299 – 304, 1987.

[22] J. Polich and M. D. Comerchero. P3a from visual stimuli: typicality, task, and topography. *Brain Topogr*, 15(3):141–152, 2003.

[23] D. Ravden and J. Polich. On p300 measurement stability: habituation, intra-trial block variation, and ultradian rhythms. *Biological Psychology*, 51(1):59 – 76, 1999.

[24] R. Schmidt. Multiple emitter location and signal parameter estimation. *IEEE Transactions on Antennas and Propagation*, 34(3):276–280, 1986.

[25] R. F. Simons, F. K. Graham, M. A. Miles, and X. Chen. On the relationship of p3a and the novelty-p3. *Biol Psychol*, 56(3):207–218, Jun 2001.

[26] E. Smith and M. Delargy. Locked-in syndrome. *BMJ*, 330(7488):406–409, Feb 2005.

[27] K. M. Spencer, J. Dien, and E. Donchin. A componential analysis of the erp elicited by novel events using a dense electrode array. *Psychophysiology*, 36(3):409–414, May 1999.

[28] P. Stoica and K. C. Sharman. Maximum likelihood methods for direction-of-arrival estimation. *IEEE Transactions on Acoustics, Speech and Signal Processing*, 38(7):1132–1143, 1990.

[29] S. Sutton, M. Braren, J. Zubin, and E. R. John. Evoked-potential correlates of stimulus uncertainty. *Science*, 150(3700):1187–1188, 1965.

[30] B. D. Van Veen, W. Van Drongelen, M. Yuchtman, and A. Suzuki. Localization of brain electrical activity via linearly constrained minimum variance spatial filtering. *IEEE Transactions on Biomedical Engineering*, 44(9):867–880, 1997.

[31] P. Wang, C. King, A. Do, and Z. Nenadic. Pushing the communication speed limit of a noninvasive bci speller. (submitted), (submitted).

[32] S.-C. Wu, A. Swindlehurst, P. Wang, and Z. Nenadic. Mitigating interference and highly correlated sources in eeg localization. (in revision).