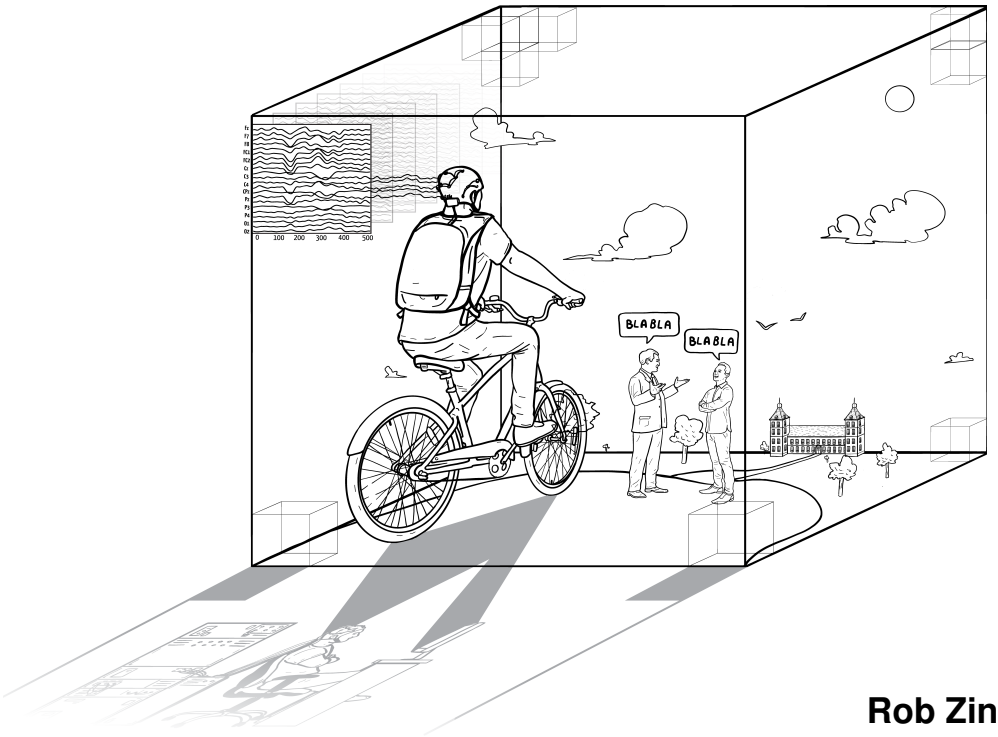
KU LEUVEN

ARENBERG DOCTORAL SCHOOL
Faculty of Engineering Science

Mobile EEG and tensor approaches for auditory attention analysis in real-life



Rob Zink

Supervisor:
Prof. dr. ir. S. Van Huffel
Prof. dr. ir. M. De Vos, co-supervisor
(Oxford University)

Dissertation presented in partial fulfillment of the requirements for the degree of Doctor of Engineering Science (PhD): Electrical Engineering

December 2017

Mobile EEG and tensor approaches for auditory attention analysis in real-life

Rob ZINK

Examination committee:
Prof. dr. ir. J. Berlamont, chair
Prof. dr. ir. S. Van Huffel, supervisor
Prof. dr. ir. M. De Vos, co-supervisor
(Oxford University)
Prof. dr. ir. L. De Lathauwer
Prof. dr. ir. M. Van Hulle
Prof. dr. ir. A. Bertrand
Dr. F. Smulders
(Maastricht University)
Prof. dr. P. Stiers
(Maastricht University)

Dissertation presented in partial fulfillment of the requirements for the degree of Doctor of Engineering Science (PhD): Electrical Engineering

December 2017

© 2017 KU Leuven – Faculty of Engineering Science
Uitgegeven in eigen beheer, Rob Zink, Kasteelpark Arenberg 10, B-3001 Leuven (Belgium)

Alle rechten voorbehouden. Niets uit deze uitgave mag worden vermenigvuldigd en/of openbaar gemaakt worden door middel van druk, fotokopie, microfilm, elektronisch of op welke andere wijze ook zonder voorafgaande schriftelijke toestemming van de uitgever.

All rights reserved. No part of the publication may be reproduced in any form by print, photoprint, microfilm, electronic or any other means without written permission from the publisher.

"If the human brain were so simple that we could understand it, we would be so simple that we couldn't"

Emerson M. Pugh, 1938

Abstract

In recent years there has been considerable interest in recording neurophysiological information from humans in natural environments. With the emergence of high quality mobile EEG equipment, new EEG applications may be within reach. However, to date, the number of studies using true mobile EEG recordings in natural scenarios is surprisingly limited, which questions the feasibility of recording reliable EEG in out-of-the-lab scenarios. Moreover, the cognitive functioning of humans in real-life scenarios is likely to deviate from artificially created lab environments. With the advent of real-life mobile EEG applications and real-time signal processing, current methods need to be re-evaluated, and new aspects of the EEG acquisition should be addressed. The effects of distractions, changes in cognitive load, physical engagement and subject behavioral variability in real-life scenarios are hypothesized to influence neurophysiological brain responses as described in traditional confined EEG experiments.

This thesis seeks to address the feasibility of applying mobile EEG for research grade auditory attention experiments in real-life scenarios. Auditory attention is widely recognized as a very important concept that plays a vital role in the way humans process auditory information. It is inherently related to the user's current environment, making it a very relevant subject of study with mobile EEG outside a lab environment. We evaluated several aspects of EEG recording, analysis and interpretation that are of major importance for the application of mobile EEG. Specifically, we evaluate the response to acoustic stimuli in three-class auditory oddball and auditory attention detection (AAD) in natural speech paradigms. The former relies on event-related potentials (ERP) in the EEG in response to artificial stimuli, i.e. P300, which is one of the most studied potentials in EEG, predominantly for brain-computer-interfaces (BCI). In contrast, AAD is based on tracking cortical EEG responses, in relation to attended natural speech, which holds potential for application in assistive devices such as hearing aids. The usage of regular speech stimuli strengthens the natural character of our experimentation.

The first part of this thesis focuses on the signal analysis in three-class auditory oddball paradigms. We introduce the concepts of canonical polyadic decompositions (CPD), and decompositions in multi linear $(L_r,L_r,1)$ terms (LL1)

of higher-order EEG data. We demonstrate their effectiveness in decomposing EEG datasets in a data-driven way, to obtain relevant components related to the P300. Additionally, we show that it is possible to eliminate the explicit subject-dependent calibration phase with a tensor-based decomposition (CPD/LL1) augmented with non-subject-specific templates, without sacrificing classification accuracy. This allows for instantaneous classification results that, on average, are similar to those of the subject-specific trained models. These tensor approaches lend themselves for use as data-driven classification methods of EEG that could conceivably lead to faster usage of BCI systems and provide meaningful information of the subject's performance from the mobile EEG in a more natural way.

Besides classification, we gained considerable insight with regard to the factors in real-life recordings that influence the neurophysiological responses such as the P300. We evaluate the ERP and single-trial characteristics of a three-class auditory oddball paradigm recorded in outdoor scenarios while pedaling on a fixed bike or cycling around freely. In addition, we also carefully evaluate the trial-specific motion artifacts through independent gyroscope measurements and control for muscle artifacts. This work was the first to successfully examine such aspects simultaneously in one study. Our findings suggest that cognitive paradigms measured in natural real-life scenarios are influenced significantly by increased cognitive load due to being in an unconstrained environment. Furthermore, our study paved the way for other free cycling studies; very recently our results were replicated by others. All in all, these results have strengthened our conviction that the lack of subject response is often the bottleneck in active BCIs and the attentional efforts of the subject need to be carefully evaluated.

In the last part we address the conscious attentional efforts in more realistic scenarios. To this end we evaluate mobile EEG recordings at-home for learning in an auditory context. We describe a closed-loop online analysis of AAD applied to natural speech in a cocktail party scenario. In addition, the effects of personalized training via neurofeedback are investigated. We conducted two experiments that took place in an office and home environment. The results prove the feasibility of AAD outside the lab, which is promising for future applications such as in auditory assistive devices. Moreover, the high variability between subjects in physiological responses as recorded with the EEG, highlight the importance of considering EEG training to increase the efficiency of the AAD. Preliminary evidence regarding changes in AAD performance during training was obtained and future studies are needed to examine these effects in more detail. Finally, this work suggests that multiple modalities, e.g. behavioral, physical and neurophysiological, need to be considered when evaluating users' cognitive performance exhaustively in real-life situations.

To conclude, even though our investigations have only touched upon a limited section of the wide variety of neurophysiological processes, our results demonstrate the feasibility of truly mobile EEG applications. The prospect of being able to achieve (online) application of the auditory oddball and AAD in out-of-the-lab experiments, serves as a continuous incentive for future research. Furthermore, our results encourage future mobile EEG studies to consider a holistic approach in order to extend, in the best possible way, the current lab-based knowledge of cognitive brain monitoring to real-life scenarios.

Beknopte samenvatting

Het gebruik van mobiele EEG voor het bestuderen van de hersenen in natuurlijke omstandigheden heeft recentelijk veel aandacht gekregen. De hardware is beschikbaar om deze mobiele metingen uit te voeren. Echter is het aantal studies dat daadwerkelijk mobiel meet tot op heden erg beperkt. Dit zet vraagtekens bij de haalbaarheid van deze mobiele EEG metingen. Bovendien is het cognitief functioneren van mensen in natuurlijke omgevingen waarschijnlijk anders dan in een gecontroleerde laboratoriumomgeving. Als we mobiele EEG metingen in natuurlijke omgevingen en situaties willen uitvoeren zullen bestaande signaalanalysemethoden herzien moeten worden. Daarnaast dienen nieuwe elementen in de experimenten te worden geëvalueerd: de wijze van gebruik, het omgaan met verlaagde cognitieve respons, hogere ruisniveau's, hogere mentale werkdruk en oncontroleerbare afleidingen door de open omgeving.

In deze thesis onderzoeken we de mogelijkheid om signalen van hoge kwaliteit te verkrijgen bij het toepassen van mobiele EEG in natuurlijke omgevingen. Hiervoor concentreren we ons op drie belangrijke aspecten: de classificatie, experimentele opstelling en het trainen van mentale processen gerelateerd aan auditieve aandacht. Auditieve aandacht is een fenomeen dat bij de mens een belangrijke rol inneemt bij de verwerking van auditieve informatie. Het is onlosmakelijk verbonden met de situatie waar de persoon zich in bevindt. Dit maakt het zeer relevant om te bestuderen met mobiele EEG in natuurlijke omgevingen. We bestuderen de reacties in het EEG op akoestische stimuli in de auditieve 'oddball'-taak en auditieve aandachtsdetectie (AAD). De oddball-taak maakt gebruik van event-related-potentials, te weten de P300, in het EEG om auditieve aandacht te meten, en is een van de meest bestudeerde eigenschappen van het EEG, in het bijzonder voor brein-computer-interfaces (BCI). De AAD-taak is gebaseerd op het interpreteren van EEG-respons in relatie tot natuurlijke spraak. Dit heeft veel potentieel voor het verbeteren van de werking van gehoorapparaten. Het bestuderen van aandacht voor natuurlijke spraak versterkt het natuurlijke karakter van de experimenten.

Bestaande classificatietechnieken gebruiken vaak niet optimaal de structuur in het EEG. Door het ontbinden van de EEG-data met tensormethoden verkrijgen we nieuwe en effectieve eigenschappen en classifiers voor het EEG. We bekijken de natuurlijke structuren aanwezig in het EEG en maken gebruik van de a

priori-kennis van paradigma en verwachte resultaten om zo betere en snellere methoden te ontwikkelen die bijvoorbeeld bij BCIs kunnen worden toegepast. We introduceren het concept van de 'canonical polyadic decomposition' (CPD) en decompositie in multi-lineaire (L_r,L_r,1) termen (LL1)van hogere-orde EEG data. We tonen de effectiviteit van deze methoden om zonder supervisie EEG data te onbinden in taak relevante componenten. Daarnaast tonen we de mogelijkheid om directe classificatie van EEG data uit te voeren door CPD en LL1-decomposities zonder aparte trainingsperiode. De resultaten hiervan zijn gelijkaardig aan die van persoon-specifiek getrainde modellen en maken het mogelijk sneller door middel van een BCI te communiceren. Tot slot tonen we alternatieve manieren om hogere-orde EEG-data te structureren en analyseren. De tensor-gebaseerde methoden hebben het potentieel om als basis te dienen voor volledig data-gedreven classificatiemethoden en het verkrijgen van additionele informatie ten opzichte van bestaande gesuperviseerde classificatiemethoden.

Er is een groot verschil in prestatie wanneer subjecten een bepaalde taak in een 'echte wereld'-omgeving doen vergeleken met sessies in een afgeschermd laboratorium. Om meer inzicht te krijgen in het hoe en waarom hiervan doen we onderzoek naar de factoren die deze prestatie kunnen beïnvloeden. Dit kan zowel cognitief zijn alsmede signaal-gerelateerd, doordat er bijvoorbeeld extra ruis in het signaal optreed in EEG-metingen die buiten plaatsvinden. In deze thesis hebben we naast het classificeren van EEG ook inzicht verkregen in zulke factoren in realistische metingen die de neurofysiologische processen, zoals de P300, beïnvloeden. We bestuderen ERP-karakteristieken op single-trial-niveau bij een auditieve oddball-taak opgemeten terwijl de proefpersoon fietst op een stationaire fiets alsmede terwijl deze volledig vrij rondfietst. Daarnaast bekijken we de invloed van bewegingsverstoringen en spieractiviteit-verstoringen op het EEG-signaal. Ons experiment is het eerste dat deze methoden combineert. Onze bevindingen suggereren dat cognitieve taken in realistische omgevingen worden beïnvloed door een hogere mentale druk door alle impulsen die men in zulke omgevingen te verwerken krijgt. In een gecontroleerde laboratoriumomgeving zijn zulke impulsen minimaal en kan alle aandacht aan de taak worden besteed. Onze bevindingen zijn recentelijk door een gelijkaardige studie bevestigd. Al met al suggereren deze resultaten dat een verminderde neurofysiologische respons in actieve aandachts-BCI's vaak een beperking vormt in de accuraatheid, en dat het monitoren van de aandachtsprocessen van de proefpersoon daarom van grote waarde kan zijn.

Om meer inzicht te verkrijgen in de bewuste aandacht voor spraak in de AAD-taak hebben we een online analaysesysteem voor mobiele EEG geïmplementeerd. Vervolgens hebben we de mogelijkheid van gepersonaliseerde training door middel van neurofeedback onderzocht. Deze experimenten vonden plaats in een kantoor- en thuisomgeving. De resultaten tonen dat AAD mogelijk is op plekken

buiten laboratoria, wat belangrijk is voor toekomstige toepassingen in dagelijkse situaties. De grote verschillen in accuraatheid voor verschillende proefpersonen tonen het belang om te onderzoeken of auditieve aandacht getraind kan worden. Preliminaire resultaten suggereren dat verbeteringen in accuraatheid tussen sessies in de AAD mogelijk zijn door herhaaldelijk oefenen van de taak met neurofeedback. Verdere studies zijn noodzakelijk om meer details te verkrijgen over deze (mogelijke) effecten. Tot slot geven de bevindingen in AAD aan dat het meten van diverse modaliteiten naast EEG belangrijk is voor de interpretatie van AAD in natuurlijke omgevingen.

Deze doctoraatsthese bevat zowel aspecten van signaalverwerking alsook mobiele EEG-data acquisitie. Zodoende trachten we de juiste balans te vinden tussen het ontwikkelen en testen van nieuwe methoden voor mobiele EEG toepassingen enerzijds, en het meten van innovatieve nieuwe datasets anderzijds. We hebben vooral naar auditieve processen gekeken met mobiele EEG. Echter, de resultaten en conclusies zijn breed toepasbaar in het hele mobiele EEG-onderzoeksdomein. Onze bevindingen stimuleren toekomstig onderzoek in natuurlijke omgevingen, en adviseren een holistische benadering om de huidige kennis van de mentale cognitieve processen verkregen in laboratoria tevens in natuurlijke omgevingen te toetsen.

List of abbreviations

Metrics

Hz	Hertz
s	seconds
ms	millisecond
μ**V**	microvolt

Abbreviations and acronyms

A	**AAD**	Auditory Attention Detection
	ANOVA	ANalysis Of VAriance
B	**BCI**	Brain Computer Interface
	BSS	Blind Source Separation
	BT	Baseline Template
	BTD	Block Term Decomposition
C	**CA**	Correlation to Attended stimulus
	CCA	Canonical Correlation Analysis
	CPD	Canonical Polyadic Decomposition
	CS	Calibration Session
	CUA	Correlation to Unattended stimulus

E	**ECG**	ElectroCardioGraphy
	ECoG	ElectroCorticoGraphy
	EEG	ElectroEncephaloGram
	EMG	ElectroMyoGram
	EOG	ElectroOculoGram
	EPSP	Excitatory PostSynaptic Potential
	ERP	Event Related Potential
F	**fMRI**	functional Magnetic Resonance Imaging
	fNIRS	functional Near-InfraRed Spectroscopy
	FUS	Follow-Up Session
G	**GUI**	Graphical User Interface
I	**IC**	Independent Component
	ICA	Independent Component Analysis
	IPSP	Inhibitory PostSynaptic Potential
L	**LDA**	Linear Discrimimant Analysis
	LL1	decomposition in multilinear $(L_{\mathrm{r}},L_{\mathrm{r}},1)$ terms
	LSL	LabStreamingLayer interface
M	**MEG**	Magnetoencephalogram
	MMSE	Minimum Mean-Squared Error
	MI	Motor Imagery

N	**NFB**	Neurofeedback
	NLS	Nonlinear Least Squares
O	**ODC**	Online Decoder Creation
	ODU	Online Decoder Update
P	**PCA**	Principal Component Analysis
R	**RMS**	Root Mean Square
S	**SC**	Subject Control
	SD	Standard Deviation
	SDF	Structured Data Fusion
	SF	Subject Feedback
	SNR	Signal-to-Noise-Ratio
	SO	Stimulus Onset
T	**Terp**	Target ERP
	TS	Training Session
	TT	Target Template
V	**VR**	Virtual Reality

Contents

Abstract vii

List of abbreviations xv

Contents xix

List of Figures xxiii

List of Tables xxxi

1 Introduction 1

1.1 Problem statement 1

1.2 Chapter-by-chapter overview 4

2 Background: Recording neural activity of the human brain 7

2.1 The human brain 7

2.2 At microscopic scale 10

2.3 Monitoring brain activity in humans 12

2.3.1 Basics of EEG 13

2.3.2 Other neuroimaging techniques 18

2.4 Measuring auditory attention 21

2.4.1 Attention to artificial tones 22

2.4.2 Attention to natural speech 23

2.5 Conclusion 24

3 Application of mobile EEG in cognitive neuroscience 27

3.1 Auditory attention in BCI 27
3.1.1 Real-life environments 27
3.1.2 Decoding approaches 28
3.2 Auditory attention detection to natural speech 30
3.3 Experimental EEG Data 33
3.3.1 Datasets 33
3.3.2 Data acquisition 34
3.4 Conclusion 36

4 Data-driven P300 extraction using tensor decompositions 37
4.1 From matrix to tensor 37
4.1.1 Matrix based Blind Source Separation 37
4.1.2 Structure of EEG data 39
4.1.3 Decomposing multidimensional EEG data 39
4.1.4 Practical Considerations 43
4.2 CPD based classification of the P300 in mobile EEG. 45
4.2.1 Classification in BCI 45
4.2.2 Data and methods 46
4.2.3 Results 49
4.2.4 Discussion 52
4.3 Chapter Conclusion 53

5 Innovative usage of tensor decompositions for data-driven classification of auditory mobile BCI data 55
5.1 Tensor-based single-trial classification of auditory mobile BCI without subject-specific calibration phases 56
5.1.1 Methods 57
5.1.2 Results 65
5.1.3 Discussion 68
5.2 Decomposition of multiple datasets using coupled CPD 72

5.2.1 Data Generation And Preprocessing 72
5.2.2 Tensor Based Models 74
5.2.3 Results 79
5.2.4 Discussion 82
5.3 Chapter Conclusion 83

6 Disentangling attentional and physical contributions to auditory attention tasks in real-life 85

6.1 Introduction 85
6.2 Methods 87
6.2.1 Participants 87
6.2.2 Stimuli and procedure 87
6.2.3 Data acquisition 88
6.2.4 Data analysis 89
6.3 Results 91
6.3.1 ERP 91
6.3.2 Noise levels 92
6.3.3 Classification 93
6.3.4 Gyro signals 94
6.3.5 Muscle artifacts 95
6.3.6 User metrics 96
6.4 Discussion 97
6.5 Chapter Conclusion 100

7 Online detection of auditory attention with mobile EEG: closing the loop with (neuro)feedback 101

7.1 Literature review 101
7.2 Methods 103
7.2.1 Participants 103
7.2.2 Data acquisition 104

7.2.3 Stimuli . . . 104
7.2.4 Procedure . . . 105
7.2.5 Analysis . . . 107
7.2.6 Neurofeedback Cue . . . 108
7.3 Results . . . 110
7.3.1 Study I: Closed-loop online AAD . . . 110
7.3.2 Study II: Longitudinal recordings at home . . . 115
7.4 Discussion . . . 119
7.5 Chapter Conclusion . . . 123

8 Conclusions and future work 125
8.1 Conclusions of the thesis . . . 126
8.1.1 Data-driven classification . . . 126
8.1.2 Real-life EEG recordings . . . 130
8.1.3 AAD system . . . 132
8.2 Future research perspectives . . . 134
8.2.1 Methodology . . . 134
8.2.2 Mobile Recordings . . . 137
8.2.3 Applications . . . 139

Appendices 145

Bibliography 159

Curriculum 183

List of publications 185

List of Figures

1.1 Overview of the chapters in the thesis. The blue-green route indicates the main route to navigate through the book. The Brown chapter provides supportive knowledge but are not necessary for those with a background in (cognitive) neuroscience and biomedical engineering. The green chapters represent work that is published in international conference and journal articles and varies in the degree of methodology and application potential as indicated. 6

2.1 The model of the mammalian 'triune brain' as was described by neurologist Paul MacLean in 1952. 9

2.2 A. Sagittal view of the four lobes of the human cortex. B. Example layer of the cortex with pyramidal cells. C. Schematic of two neurons. D. Illustrative signal that may be captured at the electrode if enough EPSPs have been trigger simultaneously 10

2.3 Different frequency bands of a random five seconds segment of mobile EEG data (i.e. from Chapter 6). From bottom to top: examples of 1. raw EEG signal; 2. delta rhythm (0.5-4Hz); 3. theta rhythm (4-8Hz); 4. alpha rhythm (8-13Hz); 5. beta rhythm (13-30Hz); and finally the gamma rhytm (>30Hz) 15

2.4 Schematic overview of the channel layout, CMS and GRL denote the reference and ground electrode respectively. Note that the electrodes were placed according to the standard 10/20 system and the colors are for illustrative purposes only. 17

2.5 A sequence of tones is presented to the subject. At random moments the tone deviates in pitch (i.e. higher or lower), one of which the participant is asked to focus on. Averaging the EEG instances per tone reveals a P300 effect around 300-500ms after the attended deviant stimulus was presented. This P300 effect is not present in the standard tones that are played frequently in the sequence, nor in the deviant tone that is ignored. 24

2.6 Illustration of the cocktail party effect. Two speech streams are presented simultaneously and the subject, i.e. the person in the front, tries to follow the conversation of only one stream (speaker) and ignore the other. 25

4.1 Schematic of a decomposition with CPD into R terms. 41

4.2 Example CPD outcome of a dataset from noise-level 2, illustrating the decomposition with the corresponding spatial, temporal and trial modes for a model of $R = 3$. The components depict the N100, P300 and noise signature from left to right, respectively. The first half of the trials corresponds to the target trials, the latter to non-targets. Component 2 is able to separate the classes with 92% accuracy, based on the factor weights of the third mode. 42

4.3 A) A single alpha component of the CPD model of subject 1 of dataset D. From top to bottom, the spatial, temporal and spectral modes are depicted. B) Example of the single-trial EEG data from (a) and the reconstructed alpha patterns by back-reconstructing the CPD component and applying an inverse wavelet transform on the filtered time-frequency matrix provides the cleaned EEG trial. 43

4.4 Overview of the most important CPD analysis steps 48

4.5 Average difference of the Target and NonTarget ERPs per subject at electrode Pz 49

4.6 Single subject (S3, rank 1) example of a CPDtime and CPDfreq component on the large dataset. The modes correspond to the spatial, temporal and trial dimensions from top to bottom, respectively. The first half of the trials correspond to the target stimuli, the latter to nontargets. This is also reflected in the trial factor loadings by CPD (75% and 78% accuracy, respectively). 50

4.7 Grand-average accuracy for different ranks of the CPD models. The swLDA accuracy is indicated as reference 51

4.8 Average accuracy per subject with a rank-1 CPD model. 52

5.1 Grand-Average (n-1) heatmap of the spatio-temporal pattern for each of the stimulus types Baseline, non-Targets and Targets, from top to bottom respectively. 59

5.2 Overview of the tensor construction. Every target (attended tone) and non-target (unattended tone) trial are concatenated with average ERP templates from other subjects of the target and baseline stimuli. In the end per binary command we constructed a 12 channels x 15 time points x 4 trials/templates tensor. . . . 60

5.3 Schematic diagram of the LL1 decomposition. 61

5.4 Overview of the tensor construction. Every target (attended tone) and non-target (unattended tone) trial are concatenated with average ERP templates from other subjects of the target and baseline stimuli. In the end per binary command we constructed a 12 channels x 15 time points x 4 trials/templates tensor. 61

5.5 Examples of decomposing a single trial pair tensor with (A) CPD, Illustrating the factor loadings in the spatiotemporal matrix and trial/template modes from top to bottom respectively. (B) LL1 decomposition with L=3,the obtained LL1 spatiotemporal matrix is derived from multiplying the first two modes of the LL1 estimates (not shown). The LL1 model estimates 2 distinct sources whereas the CPD estimates an average target-non-target effect. The trial/template estimates in the last mode will lead to classifying trial2 as the target trial since its value is close to the target template (TT) and further away for the baseline template (BT) as opposed to trial1. Note, in this particular example the subject focused on trial2. The LL1 model separates the target and non-target signature more accurately which is evident from a larger difference between trial1 and trial2. The opposite factor loadings for target and baseline are coherent with the ERP patterns of these stimuli c.f. Figure 5.1. Note: the spatiotemporal pattern as shown in (B) differs reasonably between trial-pairs. 63

5.6 Top: grand-average accuracies. Bottom: Relation between LL1 and rLDA with Pearson's correlation coefficient. Respectively in the seated condition (left) and walking (right). Significant differences and correlations with $p<0.05$ are indicated by an asterisk. 66

5.7 Average classification accuracy across subjects as a function of the available trials for training. Each point represents the average classification accuracy of all following trial-pairs. rLDAs and swLDA estimates are compared with the CPD and LL1 classification accuracies. Top: Seated condition. Bottom: Walking condition. The shaded areas mark significant differences between LL1 and rLDA at the 0.05 significance level. 68

5.8 Grand average accuracy as function of the non-specific subject used for training (LDA) or estimating templates (CPD/LL1) for the seated and walking condition left and right respectively. The shaded areas mark significant differences between LL1 and rLDA at the 0.05 significance level. 69

5.9 Overview of the dipole spatial locations (left) and temporal characteristics of each dipole on the right. The blue lines depict the temporal progression in the Target trials and the orange for the Non-Target trials. Image obtained from the BESA simulator (available online at `www.besa.de`) and modified to highlight the P300 source. 73

5.10 Average ERPs at Channel Pz for the four increasing levels of noise from left to right respectively. 74

5.11 Illustration of a transfer of a single component's spatial and temporal factor (CPD1) onto one component of a novel dataset and CPD model (CDP2). 75

5.12 Illustration of a coupled CPD of two datasets in which the spatial and temporal modes are shared within each component. 76

5.13 Joint analysis of multiple EEG datasets in which the trial mode is shared, for example in a case where multiple subjects execute the same task. 77

5.14 Average clustering percentage of the CPD for each of the four noise levels, dependent on the number of components considered in the models 79

5.15 Accuracies of CPD with a fixed factor for the highest noise condition. Each line represents a different noise level from which the best spatial and temporal factors were identified to be fixated in the highest noise dataset decomposition. 81

5.16 The CPD clustering accuracy of the highest-noise-level dataset if it was coupled to one of the lowest four noise levels (indicated as level 0-3) in contrast to the number of components in the decomposition. 81

5.17 Clustering percentage in relation to the number of coupled datasets in the trial-mode for the highest-noise-level dataset (SNR = 0.60). Each line corresponds to a different Rank. . . . 82

6.1 Schematic of the Still, Pedal and Move conditions. Note that in the Pedal conditions subjects sit on the fixed bike like the Still condition and pedal like they were biking in the Move condition. All recordings took place outdoors at or next to roads on the KU Leuven campus 89

6.2 Grand average ERPs to the Target, nonTarget and Baseline stimuli at electrode Pz for the Still, Pedal and Move conditions from left to right respectively. At the peak N100 and P300 ERP a topoplots indicates the distribution of the peak along all electrodes for the Target stimuli. Note: for illustrational purposes artifactual single-trials were removed before plotting. Reported statistics in the text are based on the full dataset. . . 92

6.3 Median RMS values of the baseline interval [-200-0]ms of the Target trials at each electrode for the Still, Pedal and Move condition. 93

6.4 A. Averaged (over 3 axes) gyro RMS per trial for the Still, Pedal and Move condition from left to right respectively. B. The average ERPs at Pz to eight divisions of the Move condition data based on the gyro RMS in A(most right figure). The color indicates the degree of motion and all degrees have equal number of trials. C. displays the peak P300 from B in contrast to the eight motion degrees; the P300 peaks seem to be unrelated to the degree of motion. Finally, D. shows the average classification accuracy in relation to the motion levels. 95

6.5 Median RMS values of the baseline interval [-200-0]ms of the Target trials at each electrode for the Still, Pedal and Move condition after EMG noise removal. 96

7.1 Overview of the recording sessions/days for both studies. Study I and Study II had different audio stimuli and each circle in the figure indicates one block of 6 minutes. After listening to one full story (2 blocks of 6 minutes), subjects switched attention to the opposite side, i.e. left or right speaker, to avoid introducing a lateralization bias in the decoder. Blocks in the shaded area represent that half of the subjects were presented with visual feedback (different per study) and at these blocks a pre-trained decoder was used based on the data of previous sessions. Study I consisted of two sessions of two stories (four blocks) each. Study II consisted of 8 sessions: 2 calibration (CS), 4 with feedback training (TS) and 2 follow up sessions (FUS). 106

7.2 Illustration of the Auditory attention closed-loop analysis. Values depicted are for illustrative purposes only. 110

7.3 Average and single subject accuracies per trial-length as calculated post-hoc. The dotted line indicates chance level as corrected for multiple comparisons. 111

7.4 Grand average and subject specific decoding accuracies depending on the number of electrodes for the 10s window length. The channel with the lowest decoder weight (after correction for channel variance differences) is removed in each iteration. Topoplots represent the average distribution for the least discriminative channels on the left and most discriminative on the right. The shaded area indicates the chance level. 112

7.5 Subject specific accuracies for a personalized and grand-average decoder (obtained from other subjects) for a trial length of 10s. The significance indicated by an asterix represents a $p < 0.0001$. 113

7.6 Adjacency matrices for the Neurofeedback session. The entry color indicates the normalized frequencies of a specific color cue (in the rows) that is followed by any other cue (columns). Cues: DG = Dark Green, LG = Light Green, LR = Light Red and DR = Dark Red. 114

7.7 Scatterplots illustrating the correlation between the decoder grand-average accuracy and the number of correct responses after each story in A and the User indicated task difficulty in B. A regression line has been added for illustrational purposes. . . 115

7.8 Evolution of a) session performance for both subjects and b) relative session performance of SF to SC. Dash lines indicate the on- and off-set of NFB (shaded area) to SF (left and right, respectively). 116

7.9 Evolution of CA-CUA variance for both subjects. Dashed lines indicate the on- and off-set of NFB to SF (left and right, respectively). 117

7.10 Topographic maps of the spatial distribution of decoder weights throughout the experiment, for a) SC and b) SF. (Red: high weight. Blue: low weight). Dashed lines indicate the on- and offset of NFB to SF (left and right, respectively). Correlations between consecutive spatial distributions are indicated, for both subjects. * denotes $p < 0.05$ and ** $p < 0.001$ 118

7.11 Block performances for both subjects and correlation between both subjects ($r = 0.46$, $p = 0.009$). 119

8.1 Decompositions of fifth-order EEG tensor that represents the subject average data of the bike experiment from chapter 6 (Dataset D). The subject × channel × time × stimulus × conditions are the dimensions from top to bottom. The left component predominantly represents the N100 and the right is characteristic of a P300 response. The subject factor loading correlates to the subject-average classification accuracies, obtained post-hoc, with the regular rLDA method in all three conditions as indicated by the scatterplots. In the other modes, similar trends to the findings of the conventional analysis can be noted (e.g. P300 shape, lower P300 in the Move condition, posterior focus of the P300, and opposite sign for the target stimulus as opposed to the baseline and non-target response). . 129

8.2 Correlation between the session performance of AAD and the three-class auditory oddball paradigm of Dataset F. 134

8.3 Left: Example of a single-subject tensor structure in which each entry represents the ERP for a specific combination of factors as illustrated by the three axes labels. Entries with a *?* are unknown, due to a lack of EEG data. Right: Based on the intra-subject properties between the various factors, and the similarities between subjects (subject properties), SDF will estimate missing elements in the tensor. 137

A.1 Example of the compute module for third-order EEG dataset of size [22 × 500 × 200] representing the channel, time and trial mode respectively. `Browse` allows to select any third order dataset. One can also load an example EEG datafile. `Compute` calculates the CPD models for all given parameters, an additional progress bar is shown. Finally, the `Save` button allows to save the calculated models and start the visualization module 146

A.2 Illustration of the visualization module for CPD visualization of third-order EEG data. Note that this is an alpha version and work in progress. 148

A.3 Example of two highly similar components that are indicative of an incorrect rank of the CPD model 149

B.1 Median RMS values of the EEG. 152

B.2 Grand-average classification accuracies as calculated with the same procedure as described in chapter 6 for each condition. * Indicates $p < 0.01$ and ** $p < 0.001$ 154

B.3 Grand-average ERPs for the still, pedal and move condition from left to right perspectively. S indicating the stimulus onset and each color represents the stimulus type. The line style represents the various artifact correction techniques. 154

C.1 Example of the simple and complex city visualization. Note that the cars in the complex city are moving. 156

C.2 Grand average ERPs to the Target, nonTarget and Baseline stimuli at electrode Pz for the Simple and Complex conditions from left to right respectively. 157

List of Tables

2.1 Comparison between neuroimaging methods. The table is adapted from [115] and modified to fit the scope of the current thesis and recent developments on EEG hardware. 20

3.1 An overview of the datasets used throughout the manuscript and sorted in chronological order. One external dataset was obtained to start instantaneously with proper mobile EEG data recorded in an ambulatory setting. During the course of the PhD new datasets were recorded (and one simulated). Ch. = chapter and Dis. = discussion. Note that Dataset C and F are summarized in the Appendix only. 35

5.1 Average accuracies of the rLDA, CPD and LL1 method for discriminating between target and non-target trials for each of the five different levels of noise. Values with an asterix were significantly lower as compared to rLDA ($p < 0.001$). 79

6.1 Grand-Average Session Specific accuracies (±SD). 93

6.2 Grand-Average Cross-Session accuracies (±SD). 94

7.1 Online AAD accuracies on training blocks and training sessions: range and average (±SD) for both subjects. 115

Chapter 1

Introduction

1.1 Problem statement

Decades of electroencephalography (EEG) research have focused on unraveling the cognitive functioning of humans in research labs and hospitals. This vast amount of studies established without doubt extremely useful (fundamental) theories and insights about the human brain in both medical (e.g. epilepsy, artificial prosthesis control) and the cognitive field (e.g. action readiness, attention). Without denying the need for fundamental EEG research, it should however be noted that there can be concerns on the (limited) societal validity of restricted lab experiments. Humans are likely to behave very different in such scenarios than for instance if they are at home or outside. Moreover, lab experiments usually follow the rationale of keeping as many factors as possible constant and only varying the specific factor to be investigated. Increased distraction and higher cognitive load (e.g. due to executing multiple tasks simultaneously) are expected to affect the results when such recordings take place in a real-world situation. The clearest evidence of this is reflected in the process of human attention allocation. Limiting attentional efforts solely to the task at hand is expected to introduce a bias in the outcome; the sterile laboratory environment creates a very specific context that potentially over-emphasizes top-down reasoning. The setting does not involve the same actions and perceptions that would arise in the real-world where attention is drawn additionally by external (unrelated) bottom-up influences.

New opportunities for out-of-the-lab measurements have come about with the introduction of mobile EEG technology. Its mobility allows both scientist and subject to shed the confinements of the laboratory and instead collect data on EEG paradigms in the real world, observing cognitive processes in their natural context. Recent studies highlight the interest in mobile recordings for cognition across various fields, for example environmental [12], sport [167, 33], cognitive

load [46], spatial cognition [138] and neuroergonomics [141]. In parallel to the increased interest from the research community, more and more people are using portable and wearable tech for health and cognitive monitoring (e.g. smartwatches). Therefore it is of little surprise that EEG research (and likewise industry) is experiencing a revival since the introduction of small portable devices a few years ago that allow to record EEG wirelessly. However, no new technology comes without its challenges, and mobile EEG is no different, requiring researchers to innovate along the whole EEG signal pipeline to deal with "the real world". In real life there are many uncontrollable factors which most of the current analysis tools are not evaluated for. To date, most EEG methods have been honed on EEG data from laboratories. With the advent of real world mobile EEG applications and real-time signal processing, however, current methods need to be re-evaluated and new elements must be considered in such unconstrained experiments: different ways of usage (e.g. quick deployment), reduced cognitive response of interest or increased influence of artifacts (i.e. lower SNR), increased cognitive load, or uncontrollable distractions.

In the current thesis we focus on the key aspects of EEG experimentation to advance auditory attention detection in real-life scenarios. Auditory attention is a very powerful mechanism that acts as a selector of salient information. In part this can be controlled voluntarily and in part this happens subconsciously, which makes it very sensitive to conscious effort of the subject and changes in the environment. In addition, auditory stimuli can be presented to subjects in real-life with ease (e.g. through headphones). These make auditory attention a very suitable phenomenon to investigate with mobile EEG in real-life. Purposely, we evaluate the response to acoustic stimuli in auditory oddball and auditory attention detection (AAD) in speaker detection paradigms. The former relies on event-related potentials (ERP) in the EEG in response to artificial stimuli (i.e. P300). This is one of the most studied potentials in EEG, predominantly for Brain-Computer-Interfaces (BCI) on the application side (e.g. [169, 79, 56, 91]). In contrast, AAD is based on tracking cortical EEG responses in relation to attended natural speech (e.g. [143, 62, 96, 191]); a relatively new area of research. Tracking the subjects' attention holds great potential for application in assistive devices for hearing aid improvement ([151]). Furthermore, the usage of very common speech as stimuli strengthens the natural character of our experimentation.

The number of studies using true mobile EEG recordings in natural scenarios is very limited. In order to increase the usage and acceptance of mobile EEG technology in daily-life we focus on three important aspects of its practice and analysis:

Obtaining reliable data True real-life EEG recordings are sparse, especially with research grade equipment and proper experimental designs. In order

to investigate the real-world influences on the auditory attention we measured mobile EEG data in a wide variety of experimental conditions and out-of-the-lab circumstances. These experiments range from purely artifact based (i.e. head motion) to completely unconstrained recordings outside and in-home (Chapter 3 summarizes the experiments conducted). We evaluate the effects of artifacts, distractions and cognitive load due to being in a natural scenario on the physiological responses related to auditory attention. This data provides the necessary basis to determine the feasibility of performing auditory BCI and AAD experiments outside the traditional lab environment.

Decoding strategies The analysis of mobile EEG data has to be considered from a multifaceted perspective. Focusing solely on one aspect of the analysis (e.g. classification accuracy, supervised learning) limits the interpretation of the data. Approaches that take full advantage of the naturally present structure of the EEG data and experimental setup (e.g. conditions) are expected to lead to novel insights. To this end we evaluate the usage of tensor decompositions in mobile BCI data. These models have been gaining momentum in biomedical signal processing and may be a beneficial addition to existing (supervised) analysis techniques. Given their data-driven nature, the tensor models are expected to account for the unknown nuisance signals that originate from unconstrained recordings and to reduce calibration time in P300 BCI applications. We aim to evaluate the efficiency and practicality of such methods as replacement or addition to the current standard.

Novel cognitive design The effectiveness of EEG applications depends besides the signal analysis also greatly on the (cortical) responses of the subjects themselves (e.g. BCI illiteracy, see [234]). It is known for cognitive paradigms that subjects' physiological responses can differ substantially, even within subjects when moving from restricted to more real-life scenarios [49]. Therefore it is necessary to evaluate which factors, related to both subject and the real-life recording environment, affect the responses. Furthermore, we explore ways to enhance auditory attention detection by improving the physiological signals related to the AAD through the use of training with neurofeedback (e.g. training P300 [17] and neurofeedback [233]). Therefore we established a closed-loop AAD analysis that allows for online assessment of auditory attention.

Through these series of experiments and objectives we aim to increase both the understanding and application of EEG based auditory attention detection (i.e. auditory BCI and speaker detection) in real-life situations.

1.2 Chapter-by-chapter overview

The following describes the chapters in the PhD manuscript and is schematically depicted in Figure 1.1. The work in each chapter can be related to one or more of the core objectives in the EEG experiment pipeline described in the previous section.

Chapter 2 provides background information on the structure and functioning of the human brain in an easy-to-read format and the basics of EEG as a neuroimaging method are explained. In the second half auditory attention paradigms are introduced. This chapter is aimed at readers for which the following words are not common: neuron, cortex, EEG, P300 and AAD.

Chapter 3 will provide a perspective of the current playing field of mobile EEG, auditory oddball BCI and AAD in auditory cognitive neuroscience. It elaborates more on the challenges that are posed in the previous section and describes the basics of the auditory tasks used in the following chapters. Finally it provides a synopsis of the mobile EEG data (both adapted and newly recorded) that is referred to throughout the thesis.

Chapter 4 introduces the concepts of higher-order EEG structures and Canonical Polyadic Decompositions (CPD) of EEG data. We demonstrate its effectiveness for decomposing full dataset tensors without training phase, to obtain relevant components related to the P300. This is illustrated on real mobile EEG data of an adapted three-class auditory oddball dataset [49] in which the subjects were seated outside and walked outside.

Chapter 5 extends the approach of Chapter 4 to single-trial BCI as is achieved through the use of CPD and decompositions in multi linear $(L_r,L_r,1)$ terms (i.e. LL1) of the mobile EEG data (i.e. the dataset of [49]). Calibration time of the BCI has been reduced through usage of structured information of the auditory oddball paradigm in the form of templates. The presented approach presents a different view on the analysis of the oddball data without reducing the accuracy as compared to reference methods using supervised training (i.e. Linear Discriminant Analysis). The tensor approaches hold potential to be used parallel to existing supervised classification methods to improve the interpretation of the BCI data as both methods seem to rely in part on separate elements in the EEG data. Furthermore, we extend the CPD tensor decompositions of single datasets to allow multiple datasets to be linked. We simulated a large number of datasets containing target and non-target trials, i.e. to mimic an auditory oddball scenario with various noise levels. This allows

to investigate whether coupling of high- and low-noise datasets (in various ways) can enhance data-driven clustering of the P300.

Chapter 6 demonstrates the feasibility of mobile BCI on a free bike ride outside. Through a novel study the effects of increased cognitive load on the P300 oddball effect are demonstrated. In addition, the role of muscle and head motion artifacts are discussed. This work is the first to successfully examine such aspects simultaneously in one study.

Chapter 7 introduces a closed-loop online analysis of AAD applied to natural speech of speakers in a cocktail party scenario. In addition, the effects of personalized training via neurofeedback are discussed. These results are based on two experiments that took place in an office and home environment. This work illustrates the high variability between subjects in physiological responses as recorded with EEG and the potential of EEG training to increase the efficiency of the AAD. Moreover, the results show the robustness of the analysis and feasibility to record outside the lab, which is promising for future applications such as in auditory assistive devices.

Chapter 8 a summary of the main findings of each chapter is given with respect to the objectives stated in the previous section. Results are put in perspective of the current mobile EEG field. Furthermore, future lines are considered including preliminary work to illustrate possible extensions of the current work.

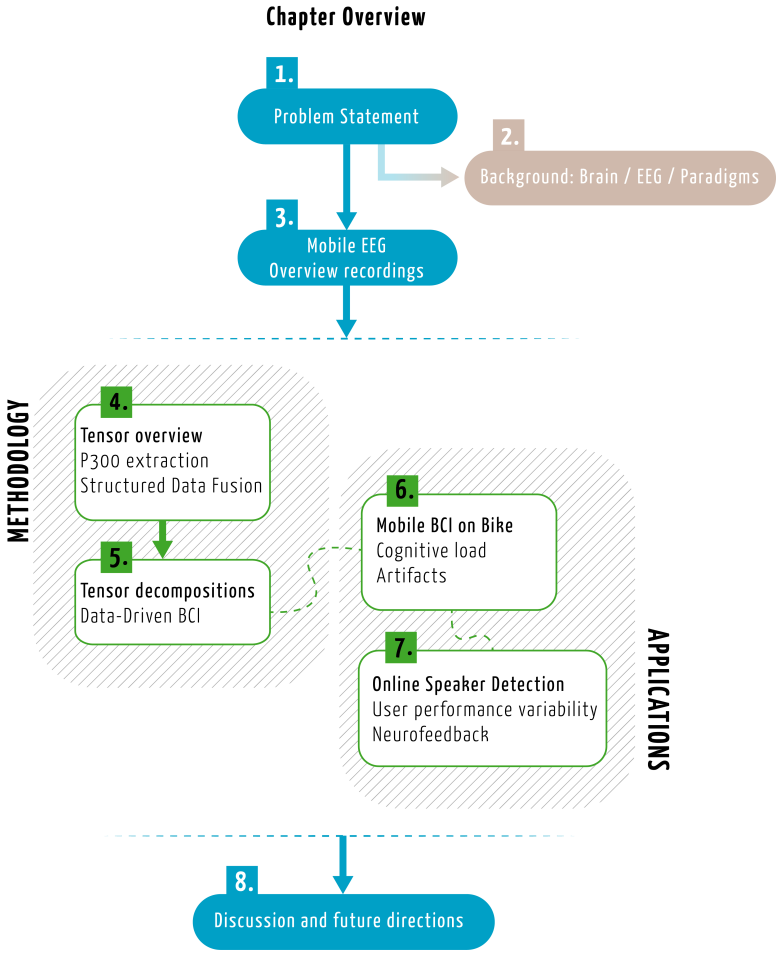


Figure 1.1: Overview of the chapters in the thesis. The blue-green route indicates the main route to navigate through the book. The Brown chapter provides supportive knowledge but are not necessary for those with a background in (cognitive) neuroscience and biomedical engineering. The green chapters represent work that is published in international conference and journal articles and varies in the degree of methodology and application potential as indicated.

Chapter 2

Background: Recording neural activity of the human brain

This chapter provides background information on the structure and functioning of the human brain on several scales. Consecutively the basics of EEG as a neuroimaging method are explained followed by a brief discussion of other non-invasive neuroimaging methods to place EEG in a broader context. The first part of the chapter is based to great extent on work from [160, 161, 229, 18] and aims to be of introductory level. In the last part we introduce the topic auditory attention and two important EEG paradigms that can measure such attentional processes.

2.1 The human brain

Depending on the criteria we use, the human brain can be divided into a different set of parts, for example, by anatomy or function. Throughout history many theories have been proposed to describe the structure of the brain and its functions. Over the course of time, most of these theories were criticized, discarded and subsequently replaced by a new view that was based on the latest experimental findings at the time. This trend of evolving theories and explanations give the impression that the human brain is either very complex to describe, or we (still) do not know the full workings of our brain or perhaps both. Explaining the fine-grained functioning and anatomy of the human brain in a four hundred paged book is already a challenge, let alone in the introductory paragraphs of this dissertation.

"If the human brain were so simple that we could understand it, we would be so simple that we couldn't" is a quote by Emerson M. Pugh in 1938 that nicely illustrates the invisible link between our brain and the person that we feel defines 'you' [173]. If you looked in the mirror this morning when you got out of bed, you might have had the feeling that what you saw (i.e. your face and body) was 'you'. However, one could argue that your body is just a biological machine controlled by a stiff jelly type of pudding in your head, the brain. To the human eye a brain looks rather funny and understanding its weirdness, you would not blame the ancient Egyptians, or Aristotle, or indeed many throughout history, for taking the view that the brain was basically irrelevant "cranial stuffing" (e.g. Aristotle held the view that the center of intelligence was the heart). Nevertheless, this strange-looking pudding is the universe's most complicated known object—1400 grams of one of the most structured, self-structuring and information-dense matter known to this day.

In order to provide some background knowledge on the human brain and pinpoint what parts of the brain are most researched in the current work we will start with an old and -mostly- deprecated theory, however one that still continues to attract the interest of the public due to its simplicity: the model of the mammalian 'triune brain'. This theory is based on the nature of evolution of different brain structures and differences between humans and other animals as was described by neurologist Paul MacLean in 1952. In his view, the brain can be separated into three parts forming a hierarchical chain of processing. Figure 2.1 illustrates the outline of the different parts as stated in the triune brain model:

The "reptilian brain" consists of the structures of the brain stem - medulla, pons, cerebellum, mesencephalon, the oldest basal nuclei - the globus pallidus and the olfactory bulbs. It is a reference to the structures of the brain connected with autonomic functions such as breathing and keeping balance and was referred to as "reptilian" behavior. These functions and brain parts were developed first in the evolution of the brain.

The "limbic brain" can be considered the part of the brain residing in the limbic system, which is concerned with emotions and instincts, feeding, fighting, fleeing, and sexual behavior. Maclean stated that these are the aspects and behaviors that developed during the mammalian age (i.e. the era that began about 65 million years ago and continues into the present).

The "neocortex" is a representation of a brain cluster structure that operates in the region of higher cognition, which includes modeling and planning. It relates to primate mammal brains (e.g. the macaque) and humans. The increased functions of cognition which distinguish animals from man reside in the cortex.

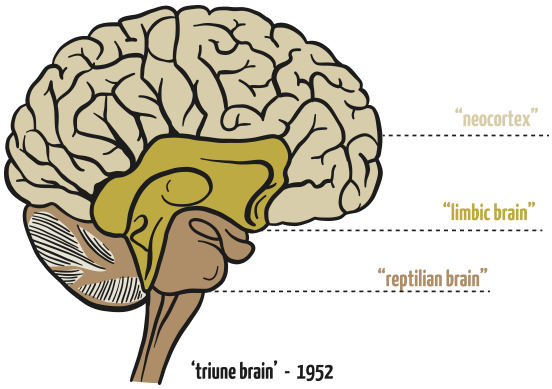


Figure 2.1: The model of the mammalian 'triune brain' as was described by neurologist Paul MacLean in 1952.

This coarse division of the human brain should be viewed as an oversimplified version of the scientific view of the brain to date (see [193] for a critical review of this model). The most often encountered criticism is a challenge to the assumption that the only elements of the human brain to experience changes throughout evolution were situated in the cortex. These issues arise from extensive research covering structural and functional connectivity which suggests strongly that the functionality a region of the brain serves is dynamic, and it can change over time which will be further elaborated in the next chapter on brain plasticity. Furthermore, others stated that the cortex is functionally connected to the limbic system and other 'lower' brain areas in such a way that we should evaluate human behavior as a combination and interaction of these brain parts, see [109]. As such, it may seem irrelevant to have mentioned this theory of the brain in the first place. However, it serves well to highlight the differences in structure and illustrate which parts of the brain this thesis will focus primarily on: the higher cognitive functions located in the cortex. From this point on the term brain and (neo)cortex will be used interchangeably to indicate the neocortex of the human brain. In addition, the cortex is generally divided into four distinct lobes, i.e. the frontal, parietal, occipital and temporal lobe, as illustrated in Figure 2.2A. The frontal lobe is important for cognitive functions and control of voluntary movement or activity. The parietal lobe processes information about temperature, taste, touch and movement. The occipital lobe is mostly concerned with the processing of visual stimuli whereas the temporal lobe is found to play an important role in the processing of audio. The next paragraph will zoom into the cells that the cortex is made of and their main working mechanism; the latter being the basis for what we call 'brain activity'.

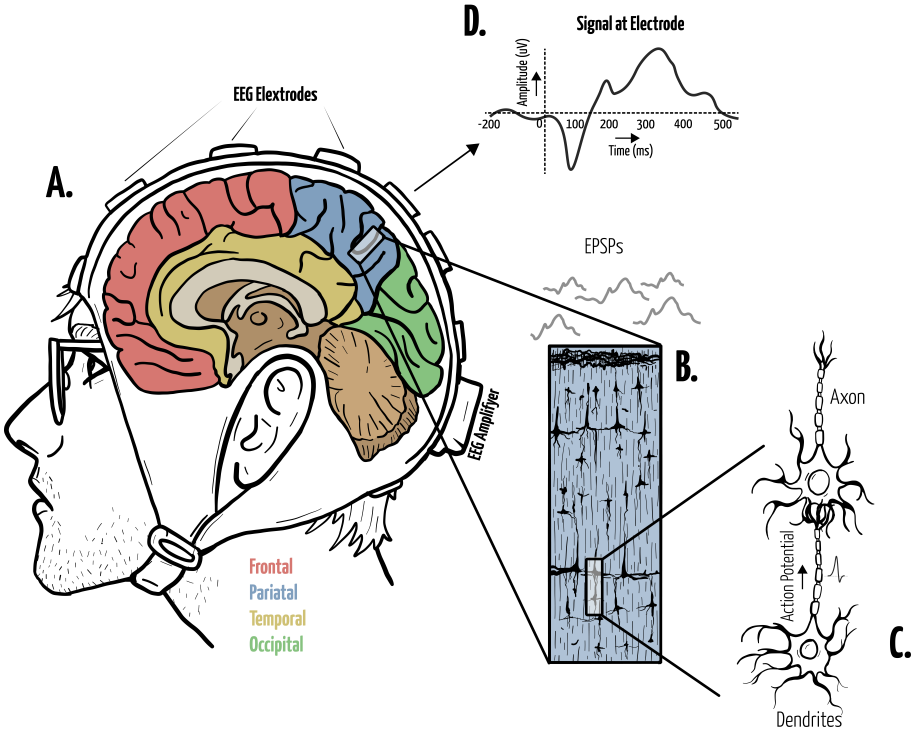


Figure 2.2: A. Sagittal view of the four lobes of the human cortex. B. Example layer of the cortex with pyramidal cells. C. Schematic of two neurons. D. Illustrative signal that may be captured at the electrode if enough EPSPs have been trigger simultaneously

2.2 At microscopic scale

Despite the fact that mankind discovered that the brain was the center of our intelligence centuries ago, it wasn't until the later stages of the 19th century that scientists realized precisely what the brain was made of. Camillo Golgi, an Italian physician, discovered an effective way to use staining techniques to reveal the structure of brain cells. Golgi discovered the neuron which deviated from other body cells in a remarkable way. Scientists realized that the neuron was the central unit in the wide communication network which constitutes the brain and the nervous system of nearly all animals. Nevertheless, scientists were not able to figure out how neurons interact with one another until the

1950s. An axon, the neuron's long strand that carries information, is usually microscopic in diameter. Too small for scientists in the early 20th century to examine, until J. Z. Young, a British Zoologist in the 1930s, accidentally discovered that the squid possesses an uncommonly large axon in their bodies on which experiments can be conducted more easily. Several decades later, making use of the squid's massive axon, scientists Andrew Huxley and Alan Hodgkin revealed how neurons transmit information: the action potential. A schematic illustration of a neurons axon, and action potential is shown in Figure 2.2C. Neurons share their functioning with computer transistors to some extent. They both use a binary language to transmit information in 1s (1 = action potential firing) and 0s (0 = no action potential firing). However, in contrast to computer transistors, the brain's neurons are continuously adapting, which is a crucial aspect of our brain's cells.

A neurons' capacity to modify itself in its structure, function, and chemical components, enables the neural network of our brain to adapt itself to the exterior world. This is known as neuroplasticity. Think for instance of the brains of babies: they are the most 'neuroplastic'. The moment a baby is given birth to, its brain does not exactly know if it should suit the existence of a musician who must develop very subtle muscle skills to play violin, or that of a cab driver who must have a tremendous spatial memory to memorize all streets of London, for example [135] or that of a (near) future human that needs to handle far reaching interactions with robots. However, the brain of the baby is so flexible it can develop over time a very wide range of expertises. As an example, if you desire to adjust your habits, you need a lot of determination to override the neural pathways of your brain. However, if you manage to apply yourself for long enough, it is likely your brain will eventually change those pathways and you would no longer have to exert determination for your new habit. Your brain will have physically developed the adjustments into a new pattern in the way the brain cells are structured and communicate.

Overall, the brain contains around 80-100 billion neurons that make up this extraordinarily extensive network—similar to the volume of stars in the sky and more than 10 times the number of human beings on the surface of the earth. About 15 to 20 billion of these neurons are located in the cortex, and the remainder are in the other areas of your brain (see previous section). The connections between these neurons are constantly strengthened or weakened to allow for very advanced information processing. The network structure even holds some resemblance to modern technological inventions such as the power grid in the United States, or the way social networks (e.g. Facebook) are structured.

A neuron consists of a soma—the body of the cell—one long axon, and short dendrites (See Figure 2.2C). The dendrites serve to increase the neuron's

receiving surface and build synapses with the axons of other neurons, which give them input through electrochemical impulses. A myelin sheath covers many of the axons which increases the impulse's propagation speed. Two visually identifiable areas within the nervous tissue are the white and grey matter. The grey matter consist mainly of the cell bodies, dendrites and axon terminals of neurons. The white matter is comprised of many (long) axons connecting different parts of grey matter to each other. The neurons communicate via electrochemical currents, with a -60 to -70 mV resting membrane potential in comparison to extracellular space, because of unequal cations and anions distribution. If a neuron fires an action potential it proceeds down the axon, releasing neurotransmitters near the axon terminal. These are picked up by the receptors (on the dendrites) of the next neuron in line, which causes a net in- or outflow of cations in the receptor. An excitatory postsynaptic potential (EPSP) is often described as: 'a net inflow of cations that occurs across the postsynaptic membrane, causing the depolarisation of the postsynaptic neuron'. By contrast, an IPSP (inhibitory postsynaptic potential) is generated and the postsynaptic neuron has 'a net cation outflow, resulting in the postsynaptic membrane hyperpolarizing'. If multiple action potentials proceed to the same synapse inside a short interval, there is a summing up of postsynaptic potentials. If the integrated EPSP and therefore the depolarization approaches a specific threshold, an action potential will be fired towards the postsynaptic recepter to influence the next neuron in line (c.q. Figure 2.2B&C, the action potential and EPSP). The EPSPs are likely the main generator of the signals we capture with EEG at the scalp which is further explained in the next section 2.3.

2.3 Monitoring brain activity in humans

Numerous techniques for human functional neuroimaging are available which can be divided in two main categories. Firstly, there are techniques for directly measuring electrical activity that is linked to neuronal firing (as described in the previous section), for example magnetoencephalography (MEG) and EEG. Second are neuronal activity indirect measuring techniques which adhere to the notion that neural activity has the support of metabolic activity and blood flow. These methods include functional magnetic resonance imaging (fMRI), and near-infrared spectroscopy (NIRS). Apart from these non-invasive methods several invasive methods exist that directly record electrical information from the human brain. For example, ECoG (electrocorticography), which is similar to EEG as it uses surface electrodes, albeit placed directly on the surface of the brain. This increases the signal quality but comes at the cost of a surgical operation to apply, aside from the ethical issues that one can easily think of. For conciseness, invasive recording methods are not further discussed in detail; we

kindly refer those interested to [70] for an overview of this area of neuroscience. In this paragraph we describe the basics of EEG recordings, followed by a short overview of the other most used non-invasive neuroimaging methods.

2.3.1 Basics of EEG

This section will provide a non-conclusive introduction to EEG, for a thorough description of the EEG principles and application fields we refer the reader to [161].

Brain activity

The brain maintains its electrical charge within its billions of neurons. As explained in the previous section neurons are "polarised" (electrically charged) by proteins that transport ions by pumping them across the membranes of the neurons. Neurons and the extracellular milieu, for example, constantly exchange ions, for the propagation of action potentials and also for the maintenance of resting potential. Similarly charged ions repel one another, and when multiple ions are ejected from multiple neurons simultaneously, they tend to move their neighbours, who in turn move their neighbors, and this continues as a wave, in a process called volume conduction. When the ion wave arrives at the scalp's electrodes, they will either pull or push electrons in the metal of the electrodes. As metal can easily control the pull and push of electrons, a voltmeter can be used to measure any differences in pull or push voltages. The EEG is produced by recording the voltages over a period of time. Each individual neuron's generated electric potential is too minute to be registered by EEG (as mentioned in the previous section), so EEG activity is mostly driven by EPSPs and the sum of the combined activity of millions (or thousands) of neurons which are spatially oriented in a similar way (see the illustrated example in Figure 2.2B and D.). Without a matching spatial orientation, the ions of the cells do not align to form detectable waves. The cortex's pyramidal neurons are believed to generate the greatest EEG signal since they usually fire together and are lined up well as illustrated in Figure 2.2B. Since the gradients of voltage fields reduce with the distance squared, the detection of deep source activity is harder than that near (or, 'nearer to') the skull.

EEG Activity

EEG activity on the scalp displays oscillations at multiple frequencies. Some of the oscillations have typical spatial distributions and frequency ranges, which

are linked to various brain function states (for instance, individual sleep stages). The following constitutes the major frequencies bands observable in the EEG waves of humans:

Delta band These waves lie within the range of 0.5-4 Hz, with variable amplitude. They are easily observable during deep sleep.

Theta band Theta waves are within the range of 4–8 Hz, containing an amplitude generally more than 20 uV. Theta band activity originates from stress and anxiety, particularly disappointment or frustration. Theta is also linked to accessibility to a meditative, drowsy, hypnotic or sleeping state.

Alpha band Mostly occurring between 8 and 13 Hz, with relative large amplitudes of 30-50 µV. Alpha waves are believed to signify relaxed awareness as well as inattention. Alpha band activity has been linked to various elements of attentional processing [16]. It is most prominent over the occipital cortex as well as over the frontal cortex. Alpha is the most dominant wave band in the brain activity as recorded through EEG. Alpha waves generally are observed when the person is relaxed and inattentive (such as in daydreaming). The spectra related to measuring mental workload through EEG are Alpha (8-13 Hz), and Theta (4-8 Hz), usually measured in the dorsolateral prefrontal cortex. Studies that use EEG to measure Theta and Alpha waves show an increase in Theta band power and a decrease in Alpha band power through memory tasks which are increasingly challenging, with a highly reliable measurement of mental workload [187].

Beta band Primarily between 13 and 30 Hz, with usual voltages of 5-30 uV. These brain waves are associated with active attention, active thinking, and concentration on the outer world or solving complex problems. Low beta components were found to affect attentional processing, as evident through biofeedback training [67]. Significant changes in alpha and beta activation were linked to increasing levels of engagement in virtual reality scenarios [223].

Gamma band Gamma waves are categorized as oscillations of 35Hz and above. It has been suggested that this band reflects mechanisms of consciousness.

Figure 2.3 displays examples of the different EEG rhythms belonging to each band. This was obtained on a segment of mobile EEG data recorded while the subject was seated without noticeable physical activity.

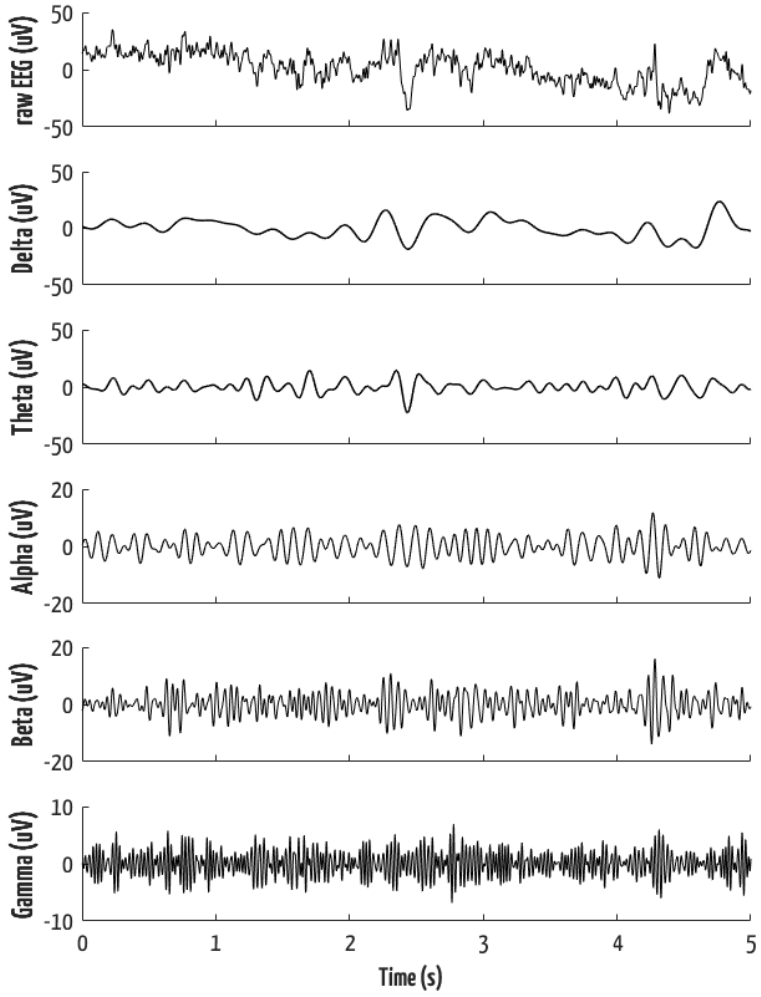


Figure 2.3: Different frequency bands of a random five seconds segment of mobile EEG data (i.e. from Chapter 6). From bottom to top: examples of 1. raw EEG signal; 2. delta rhythm (0.5-4Hz); 3. theta rhythm (4-8Hz); 4. alpha rhythm (8-13Hz); 5. beta rhythm (13-30Hz); and finally the gamma rhytm (>30Hz)

Electrode placement

The standard positioning of scalp electrodes (i.e. channels) for classical EEG recordings became common since the adoption of the 10/20 strategy. The basic principle of this technique is the distance in percentages of the 10/20 range between the Nasion-Inion and fixed points. The electrodes at these fixed points are labeled as Frontal (F), Central (C), Temporal (T), Parietal (P), and occipital (O) from the front to the back of the head respectively. The middle electrodes (i.e. along the direction of the nose) are labeled with a subscript z, this symbolizes zero. Other positions are indicated with a number suffix, Odd numbers are utilized as subscript for points on the left hemisphere, and even numbers for those on the right. For example, channel F8 indicates an electrode on the frontal area, right side, relative far from the midline (zero line). Figure 2.4 illustrate the channel positions that were recorded in the conducted experiments as presented in future chapters.

Since human hair or dead skin cells at the scalp can prevent proper functioning of the electrodes, each electrode is filled with a conductive gel that serves as an extension of the electrode. The gel maximizes skin contact and allows for a low-resistance recording through the skin. The impedance of the flow of the alternating current through the electrodes is used as measure for quality of the connection. The higher the impedance of the electrode, the smaller the amplitude of the EEG signal. Usually impedances lower than 10kOhm are preferred for EEG recordings.

Handling of Artifacts

The term artifact refers to all distortions in the EEG signal that do not originate from the human brain. Many different processes can cause artifacts in the EEG and can be categorized traditionally in two main classes:

- **Physiological**, that is, belonging to the subject (e.g. EMG, eye movement, ECG, head motion)
- **Extraphysiological**, coming from the body's exterior (e.g. cable movements, 50/60Hz artifacts, as well as electrode paste-connected)

However, the types artifacts that need to be taken into account in the setup and analysis vary with the application. During mobile out-of-the-lab experiments novel types of artifacts may arise that are not encountered in a confined lab setup. Especially, muscle artifacts and head motion are inherent part of the normal daily routine and need special attention in unconstrained EEG experiments.

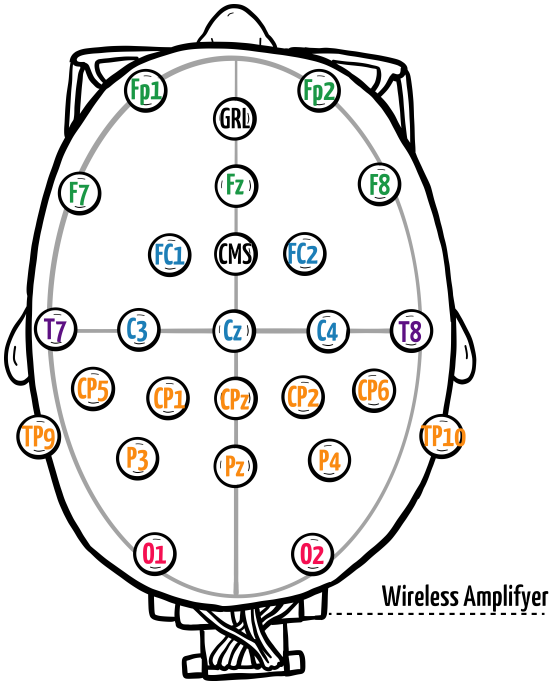


Figure 2.4: Schematic overview of the channel layout, CMS and GRL denote the reference and ground electrode respectively. Note that the electrodes were placed according to the standard 10/20 system and the colors are for illustrative purposes only.

Moreover, eye movement artifacts (e.g. blinking and horizontal movements) are expected to be more prominent in natural situations when the subject does not focus on a fixed point in space. Alpha power activity changes have been linked to changes in cognitive load and are sometimes seen as a special type of artifact. Even though they arise from the brain they might obscure the (tiny) cognitive effects of interest in the EEG.

Artifact reduction (i.e. removal) is the process of identifying artifact components in the EEG signal and separating them from the neuronal sources. We will utilize several techniques to correct for specific artifacts, especially those that increase during ambulatory recordings. A single method that performs best for all artifacts types does not exist to date. Appendix B provides exploratory data of head and body motion effects in the mobile EEG system that we used in the current thesis. Note that an extensive discussion and evaluation of artifact removal techniques is beyond the scope of the current thesis. For more detailed

analysis we refer the reader to recent publications, e.g. [149, 82, 112, 212]. Generally, it is of utmost importance to evaluate the workings of these methods on mobile EEG data as some of them were shown to remove, in addition to artifacts, also important brain related signals. Particularly in mobile EEG experiments where the characteristics of artifacts may be complex or even unknown, the effectiveness of these artifact handling techniques has not been validated.

2.3.2 Other neuroimaging techniques

Besides EEG several other neuroimaging techniques can be used, each with their own weaknesses and strengths. Since most of these methods may be used congruently with EEG (see [211] for an overview) we briefly introduce a couple of these methods. Note that this list does not include non-invasive imaging methods. Table 2.1 depicts the key aspects of each method as based on similar tables in [160] and [115].

fMRI and fNIRS

Functional Magnetic Resonance Imaging (fMRI) dates back to the 1970s and generates images of the brain using the changes in blood flow. The primary form of fMRI uses the blood-oxygen-level-dependent (BOLD) contrast. Areas of the brain that consume more energy (i.e. are more active) require more oxygen to be delivered, which causes an increased blood flow. The magnetic differences of oxygenated blood compared to de-oxygenated blood can be detected with fMRI. As a result of the increased strength of the magnetic fields used in the MRI hardware, the spatial resolution has decreased from 3mm3 to $<$ 1mm3 cubic volume pixels, so called voxels (see [129] for an overview). The brain has a volume of approximately 1.200.000mm3, so it follows that a high resolution fMRI generates an equal number of voxels in case of a resolution of one cubic millimeter. Nevertheless, it should be noted that although this sounds very small, the size of single neurons is much smaller, such that a single voxel still represents the average blood flow of roughly 40,000 neurons or more. fMRI has been very useful in a wide range of applications ranging from fundamental (e.g. [76, 182]) to applied research (e.g. [200]. Three major drawbacks are the high cost, lack of portability and low temporal resolution. Since fMRI has the inherent property of measuring blood flow rather than direct neural activity it is seen as an indirect measure of brain activity. Obtaining information on fast changing processes (e.g. in the sub-second range) is very challenging, although recent attempts have been made with success, the vast majority of fMRI studies focussed on slower processes in which timing is less important [68]. Finally, an

fMRI scanner is expensive to build and maintain as it is very sensitive to the configuration and imaging setting, which is often managed by a specialized crew. Subsequently it should be apparent that fMRI is not suitable for recording in mobile scenarios.

Functional Near-Infrared Spectroscopy (fNIRS) is an imaging method that is non-invasive, quantifying the resolution of chromophore concentration which comes from measuring (NIR) light phasic, temporal, or attenuation changes. The light spectrum of NIRS utilizes the optical window of 700-900 nm wherein bone, tissue, and skin do not absorb the light as good as deoxy-Hb (deoxygenated-hemoglobin) and Hb (hemoglobin). Variances in the absorption spectrums of oxy-Hb and deoxy-Hb facilitate the measurement of relative hemoglobin concentration by using multiple wavelength light attenuation. fNIRS is relatively cheap (as compared to the high costs of operating fMRI) as there is less expense associated with the systems themselves, and after purchase there are no extra running costs fMRI and fNIRS are often methods of comparison as they can sense similar physiologic variances. fNIRS offers a trade-off between the temporal resolution of EEG and fMRI's spatial resolution. There is a number of commercial fNIRS devices available on the market. For a review on fNIRS see [74].The main limitations of fNIRS are the low temporal resolution (compared to EEG) and low spatial resolution (compared to fMRI) which is estimated at 10–30 mm; only signals at the surface of the cortex can be retrieved. Besides mobile EEG it is considered the only serious option to record non-invasively cortical information outside hospitals or research centers [140], see [225] and references herein for an overview of recent developments on mobile fNIRS.

MEG

Magnetoencephalography (MEG) measures the magnetic field generated by neuronal electrical activity. For attenuation of the external noise generated by the system, the MEG is set within a room that has magnetic shielding. The sensors which record magnetic fields are known as gradiometers or magnetometers. MEG fields can pass through a human head with no distortion. This is a significant advantage of MEG over electroencephalography. MEG provides a high spatial and temporal resolution. EEG makes a record of electrical activity which orients perpendicular to the brain's surface, while MEG records activity parallel to the brain's surface. EEG is a measurement of pyramidal cell activity in cortical gyri and the sulci's depths, while MEG mostly captures pyramidal cell activity in the sulci's superficial areas [29]. While the signal-to-noise ratio (SNR) is often higher as compared to EEG, this technology is not portable which makes recording setups very restrictive.

Table 2.1: Comparison between neuroimaging methods. The table is adapted from [115] and modified to fit the scope of the current thesis and recent developments on EEG hardware.

		fNIRS	**fMRI**	**EEG**	**MEG**
Theoretical Background	Signal property	Scattered near-infrared light	Magnetic property	Collection of neural activity	Magnetic fields of brain's electrical activity
	Measurement area	Surface of the cortex	Whole brain region	Surface of the cortex	Surface of the cortex
	Time resolution (s)	1	2-3	0.01	0.01
	Spatial resolution (mm)	20	5	20	10
	Effect of extra-cortical tissue	Some	Little	Some	None
Recording setting	Invasiveness	No	No	No	No
	Head restraint	No	Yes	No	Yes
Instruments	Size	Small	Large, fix	Very Small	Large, fix
	Transportability	Yes	No	Yes	No
	Initial cost	Moderate	High	Low	High
	Measurement & maintenance cost	Moderate	High	Low	High

2.4 Measuring auditory attention

What we consider as sound is the pattern in which air molecules surrounding our head vibrate. When a sound is made by a guitar string, wind or a friend speaking, it is the result of the vibrations which push nearby air molecules into an identical vibration, that same pattern extends outwards in a spherical form, similar to the way a water surface extends outwards in a spherical ripple whenever an object comes in contact with it (e.g. throwing a rock in a still puddle). Your ear is an apparatus that transforms those air vibrations into electrical stimulations via the cochlear onto neurons in the auditory cortex. The specific manner in which the air is vibrating causes those nerves to fire a pattern of action potentials that send the code into your auditory cortex for processing. Your brain receives the information, and we call the experience of receiving that particular type of information "hearing." On the other hand, "listening" is based on "Attention" which is a reference to multiple cognitive and behavioral processes which affect sensory processing. Modulation of attention can be achieved through sensory-driven elements that are 'bottom-up', and through 'top-down' learned schemes, expectations and goals that are task-specific. Basically, attention is a process (or processes) of selection that focuses both cognitive and sensory resources on the sound-scape's most relevant events. The stimulus itself dictates relevance, (for example, an ambulance siren), or sometimes the current task dictates relevance (for example, listening to someone amidst a party) [104]. If such stimuli are used deliberately to induce specific neurophysiological response a brain-computer-interface (BCI) can be created. BCIs allow users to control a device by signals originating from the human brain alone, see [160] for a review. In its most general form BCIs provide a route for neural output that does not involve the neuromuscular system (Wolpaw et al., 2002; Lebedev, 2014). Although several modalities can constitute such an interface (e.g. fMRI, NIRS, MEG), EEG is the most popular one to date, mostly due to its relatively low cost, high temporal resolution (as mentioned in the previous paragraphs) and the availability of portable hardware. Active auditory BCIs would require the subject to focus his or her attention on a specific type of sound or audio stream that in turn generates a distinguishable pattern in the EEG.

Whilst eye-tracking information presents unbiased appraisal metrics for visual attention experiments, attention paradigms in audition have been plagued by an absence of distinct salience metrics. Many auditory attention studies utilize their very own verification information and metrics, e.g. human reactions or cautiously chosen sound activities and paradigms. Unfortunately, there is as of yet hardly a general agreement on the best method to investigate consequences of attention on auditory conception, whether it is task-directed or totally dependent on

salience. Nevertheless, if it were a prerequisite to have full understanding of the brain's functioning in order to solve problems or let humans interact with a BCI, we probably would never be able to do so. However, certain processes of auditory attention can be decoded from the brain without truly having to understand the dynamics of the brain's computation, some examples will be introduced in the next paragraphs. Technically, recording brain activity and processing the signals are engineering challenges. The understanding of the brains' development and functions in fine detail lies in the domain of neuroscience. In order to move forward these problems should not be approached separately but from an interdisciplinary point of view. Therefore, our work focuses on both aspects through distinguished auditory attention paradigms. The most simple form of measuring auditory attention is to investigate the neurophysiological responses as generated by simple artificial tones. One of the most reported artificial auditory tasks in the literature is the so-called auditory oddball paradigm. This is a task in which stimuli are presented in a continuous stream and participants must detect the presence of an oddball stimulus. The oddball is a stimulus that occurs infrequently (in time) relative to all other stimuli, and has distinct characteristics, e.g. a different pitch among auditory stimuli. The auditory oddball paradigm allows for systematic evaluation of the attentional effects. Contrastingly, a more natural and complex form of auditory stimuli is natural speech. Human speech is very important for the understanding of conversations and functioning in society. Being able to detect auditory attention to natural speakers can be promising for improving the speech intelligibility in complex situations such as multi-speaker scenarios, especially for users that have impaired hearing capacity.

We will examine both processes, i.e. the auditory oddball and auditory attention detection to natural speech, in this thesis. This way we are able to investigate auditory attention in both a systematic and natural way with mobile EEG. The next paragraphs explain the principles of these paradigms in more detail.

2.4.1 Attention to artificial tones

A three-class oddball auditory task as proposed by [91] consists of a standard tone (at 900Hz) and two deviant tones (at 600Hz, 1200Hz) that are presented binaurally to the subject. The standard tone plays repeatedly (with small random delays in between) and this pattern is occasionally interrupted by one of the two deviant tones. Subjects must detect the presence of one specific deviant tone (referred to as Target tone) and ignore the two others (i.e. the other deviant and standard tone). Target tones can be either the low (600Hz) tone or the high (1200Hz) tone and the subject can thus perform a binary action by focusing the attention to either of those tones in the oddball sequence. The attended

target tone constitutes an oddball-effect which results in a positive potential in the EEG around 300-500ms after stimulus presentation, referred to as the P300. This effect is the strongest detectable on the central posterior electrodes and attention allocation increases the P300 amplitude [86]. A suggested theoretical mechanism for P300 is neuro inhibition which is elicited when stimulus detection engages memory operations [169]. One complexity to this auditory oddball effect is the fact that it is generated by the subjects' conscious attention to the tone; therefore a target tone may not always result in a clear P300 (e.g. when the subject is distracted) during the task. There is high variability between P300s of different subjects and even within a subject across trials. Analysis of the P300 effect often averages the responses over a large number of trials, to obtain a robust event-related-potential (ERP). A schematic of the auditory oddball paradigm and corresponding P300 ERP is depicted in Figure 2.5). Note that another ERP component, the N100, is clearly visible for all stimuli types as a negative deflection around 100ms after stimulus presentation. This component is reflective of the physical aspects of stimulus processing that takes place automatic (i.e. unconsciously) and is therefore equally present in both types of tones. For an extended description of ERP types, analysis and interpretation we refer to [134].

2.4.2 Attention to natural speech

Understanding speech within a noisy or multiple-speaker environment is one of the auditory system's most difficult tasks. Nevertheless, humans have the astounding capacity to focus their attention to interesting sources of speech within complicated audio scenes. This allows listening to one speaker of interest while simultaneously ignoring other speakers and sound sources competing for attention. This problem was coined the 'cocktail party problem' by Colin Cherry in 1953 [34] and is schematically shown in Figure 2.6. The remarkable capability of the human brain to solve this problem seemingly easy motivated researchers to unravel the neural underpinnings to this phenomenon. Significant progress came from the finding that cortical oscillations phase lock to the envelope of the speech stream [3]. Consecutively, attended speech was found to be more prominently represented in electrophysiological signals as compared to unattended speech, especially at higher cognitive levels [143, 62, 96]. [191] illustrated that lower auditory areas of the brain encode both attended speech and background noise, whereas higher auditory areas have a preference for the attended speech stream only. Based on these findings it became possible to detect which speaker (i.e. in a two-speaker cocktail party scenario) was attended to from the recorded activity from EEG, MEG and ECoG [165, 4, 143]. This process will be referred to as auditory attention detection (AAD). These models followed mostly two approaches: define a regression model that predicts the neural responses based

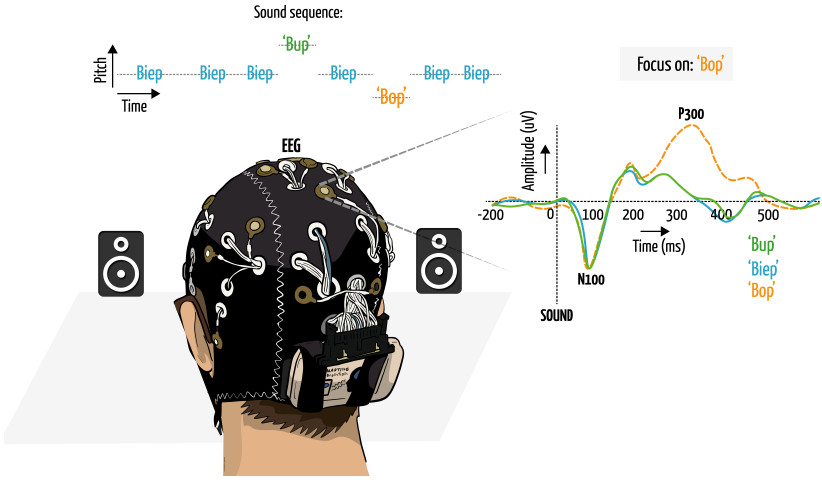


Figure 2.5: A sequence of tones is presented to the subject. At random moments the tone deviates in pitch (i.e. higher or lower), one of which the participant is asked to focus on. Averaging the EEG instances per tone reveals a P300 effect around 300-500ms after the attended deviant stimulus was presented. This P300 effect is not present in the standard tones that are played frequently in the sequence, nor in the deviant tone that is ignored.

on the sound, so-called forward modeling [6, 4, 5], or reconstruct the speech envelopes of the (un)attended speaker from the EEG or MEG neural data (backward modeling) [165, 42]. Others have experimented with alternative approaches, for example to extract features from the cross-correlation of neural response and speech envelope to train a linear classifier [90].

2.5 Conclusion

This chapter covered the basics of the human brain in term of structure and function. In addition, we highlighted the principles of EEG data, EEG recordings, and the relation of EEG to complementary measurement techniques in neuroscience. Finally we introduced two auditory attention tasks that will be employed in the current thesis: the auditory oddball and auditory attention detection in a cocktail party scenario.

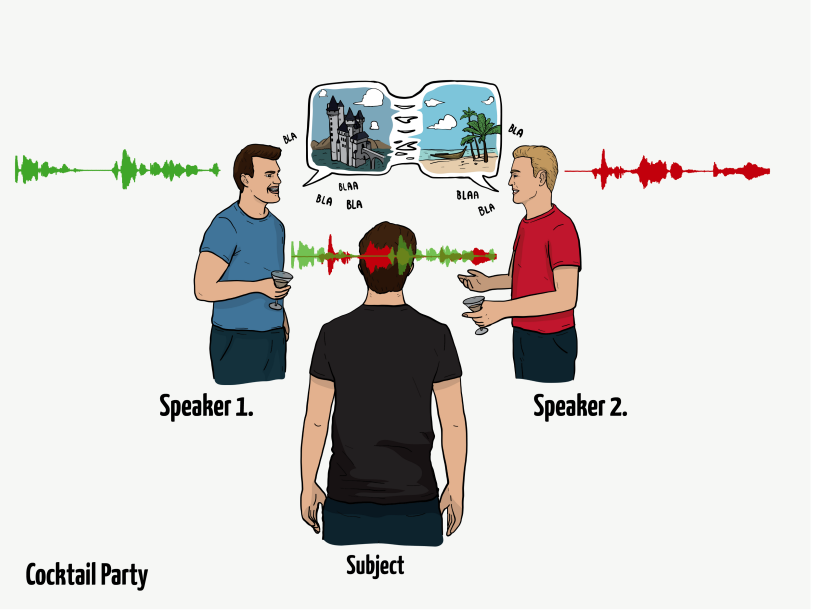


Figure 2.6: Illustration of the cocktail party effect. Two speech streams are presented simultaneously and the subject, i.e. the person in the front, tries to follow the conversation of only one stream (speaker) and ignore the other.

Chapter 3

Application of mobile EEG in cognitive neuroscience

In this chapter we describe the current state-of-the-art in the application of EEG for auditory attention analysis and highlight the current shortcomings in decoding strategies, experimental setups and choice of methods for out-of-the-lab applications. In the last part we provide an overview of all data used in this thesis and describe the hard- and software setup used in the experiments.

3.1 Auditory attention in BCI

3.1.1 Real-life environments

Research in the field of Brain-Computer-Interfaces has made significant progress in methodologies and applications. To date, the use of non-invasive EEG (electroencephalography) signals as primary input is most often used for such interfaces. The most studied feature in EEG-based BCI is the P300, generated in response to rare and task-relevant stimuli (e.g.[169, 79, 56, 91]). Originally, BCIs were proposed to provide accessibility to computers for locked-in patients [22]. More recently, they have also gained potential in applications for healthy users [232, 213]; especially when they are not bound to highly restricted traditional laboratory settings. In the past few years there has been a growing interest in studies that utilize mobile EEG in natural, real- life situations [57, 49, 95, 144, 12, 48, 243]. The success of these studies is closely related to the data quality observed with miniaturized portable EEG hardware such as Emotiv (`www.emotiv.com`) and SMARTING (`www.mBrainTrain.com`). For example the original Emotiv hardware was found to be inferior to a medical grade system [65, 66], however a modified setup allowed high quality recordings [57]. Moreover,

it was recently demonstrated that such mobile devices can deliver comparable waveforms as traditional EEG systems [50]. The presence of mobile EEG creates new possibilities for investigating cognitive mechanisms in natural environments in contrast to limiting and artificial laboratory setups [232, 213, 110]. Although real-life scenarios have been mimicked in labs [208, 30, 172, 114, 223], true real-life EEG recordings are sparse (e.g. [57, 49, 25, 55, 11]). These studies demonstrated the feasibility of investigating auditory attention while walking which mimics closely a practical application of auditory BCI in real-life [95, 144]. However, these studies face the challenge that the observed differences (e.g. decreased P300 while walking as compared to still conditions [57, 49]) can be attributed to either the higher cognitive demands or to the contributions of physical activity in such auditory attention tasks. In mobile real-life experiments the user's mental focus is less controllable compared to indoor recordings. Additionally, unconstrained setups may induce additional artifacts such as head and body motions or increased muscle activity [112, 77]. Walking outdoors while performing an auditory oddball task leads to significantly lower classification accuracies compared to sitting still indoors [57, 49]. De Vos et al. speculated that the subjects in the walking condition endured increased cognitive demands and distractions that caused the lower performance.

Taken together, in the past few years there has been a growing interest in studying brain functioning in natural, real-life situations. However, very little research exists that investigates the brain's functioning in real-life environments. In addition, mobile EEG allows to study the brain in real unconstrained environments but it faces the intrinsic challenge that it is impossible to disentangle observed changes in brain activity due to increase in cognitive demands by the complex natural environment or due to the physical involvement. In Chapter 6 we aim to disentangle the influence of cognitive demands and physical contributions that arise from such outdoor unconstrained recordings. We evaluate the ERP and single trial characteristics of a three-class auditory oddball paradigm recorded in outdoor scenario's while peddling on a fixed bike or biking freely around. In addition we also carefully evaluate the trial specific motion artifacts through independent gyro measurements and control for muscle artifacts.

3.1.2 Decoding approaches

Besides the limited validation in real-life scenarios, a second hurdle in real-life BCI is that most BCIs exploit supervised classification methods. This requires a separate training phase to calibrate the classifier function [132]. Essentially a substantial part of the data is discarded for model training rather than interacting with the BCI, which comes at the cost of consuming user's time

and effort. It is not uncommon that half of the experimental time is devoted to recording training data (e.g. [49]). The use of calibration-free classifiers could increase the application potential significantly by removing this specific training phase. Few studies have focused on developing such out-of-the box classifiers. Incipiently, an unsupervised approach for the P300 speller was presented offline in [108] relying on Bayesian statistics and exploiting constraints imposed by the BCI stimulation setup. The utilization of prior distributions of Target and Non-Target stimuli and repetitions was shown to lead to a fast adapting calibration free classifier. The same authors reported high performance in [107] extending the Bayesian approach with transfer learning techniques and re-learning of the classifiers by reason of language models. This approach was also successfully translated to an online scenario [106]. A novel classification approach for motor imagery BCIs that has gained momentum in the past two years exploits the spatial covariance matrices of EEG signals and rely on riemannian geometry to obtain accurate classification [13]. Successful, modifications of this approach were presented in [14, 40] to extend the framework for calibration free ERP (Event Related Potential) analysis. This allowed instantaneous P300 classification based on a special form of space-time covariance matrices and a riemannian distance-to-class-average as classifier function. See [41] for a recent review on the application of riemannian geometry for EEG-based brain-computer interfaces. Another pertinent approach aimed to incorporate information transfer between BCI sessions in the EEG by transferring session-specific changes between experimental conditions to new subjects [180]. The use of session specific changes in the EEG was shown to convey decisive information for classifier performance. Ultimately the aforementioned techniques allow to track classifier performance across subjects and sessions and grant instantaneous classification. All these approaches are of particular interest for bringing BCI applications into practical use.

In line with these views we provide new insights to utilizing the structure in the BCI through tensor decompositions of EEG data. In Chapter 4 and 5, we present a novel subject-specific calibration-free method based on tensor decompositions of the EEG applied to auditory P300 data (i.e. on data obtained from [49]). Since ERP data can be naturally represented as a channels $\times$ time $\times$ trials tensor, it can be advantageous to exploit such multidimensional structure in the analysis. Tensor-based methods such as the Canonical-Polyadic-Decomposition (CPD) have been used successfully in several biomedical applications for feature extraction or classification in a supervised approach (e.g.[1, 218, 36, 152, 53, 35]). CPD is a data-driven method that exploits multidimensional structure in data naturally present in single trial BCI datasets. It allows automatic derivation of multidimensional fingerprints, given that appropriate model parameters are set. Tensor based methods allow for intuitive interpretation of a structured dataset [117]. CPD has been applied in supervised ways to classify single trial ERP data

before [218]. We extend this work and propose an unsupervised CPD approach to automatically cluster auditory oddball mobile BCI data. Label information from prior trials is still utilized to evaluate the accuracy of the CPD models; however the actual CPD estimates are independent of any labeled data. We hypothesize that CPD is able to extract signal sources that allow discriminating target and distractor stimuli. Exploiting the spatial, temporal, and frequency information present in the decomposed EEG, discrimination between different stimulus types is obtained.

3.2 Auditory attention detection to natural speech

Online out-of-the-lab AAD

Natural speech is an important aspect in our way of functioning socially and it would be difficult to live without the ability to listen to other people. The process of listening is indissolubly linked to the focus of attention on the sound source of interest. Since attended speech was found to be more prominently represented in the human brain as compared to unattended speech (especially at higher cognitive level) detection through EEG of the attended audio source holds potential for use in assistive audiological devices, such as hearing aids [214, 151] (e.g. to steer a beamformer to the attended speaker). This is especially useful for humans who often encounter problems in complex acoustic scenes which in this thesis is reflected as auditory attention detection in a cocktail party scenario (referred to as AAD, see Chapter 2).

Merging the hearing aid technology with a mobile EEG device that could track the auditory attention of which speaker a user is listening to holds high potential to improve the speech intelligibility in real-life scenarios for the hearing-impaired. In the past few years, EEG has been employed for such AAD. A pre-trained decoder that makes a linear combination of the EEG and delayed versions of it allows to extract a signal with a significantly higher correlation to the attended speaker's envelope as compared to the unattended speaker's envelope. In dual-speaker scenarios, the subject's attention to a specific speaker could be reliably detected with high accuracy in research labs. However, these studies did not evaluate realistic recordings conditions, nor did they evaluate the differences between subjects more in depth. Finally, an online implementation of the EEG based AAD has not been evaluated, which is a key component for real-world application. Extending AAD to online decoding could enable the combination of hearing devices with cognitive state decoding, and this would allow the hearing aid device to adjust flexibly to the demands of the listener. In addition, the application of auditory attention detection may not be limited to hearing aids.

In a more industry-oriented approach, AAD could provide new information in the marketing and advertising industries. Furthermore, there are also other brain-computer interface (BCI) applications for which the auditory attention focus of a user may be used to control a device. As an example, one application could be for immersive virtual reality scenarios to integrate neural feedback to adapt the virtual world based on the focus of the subject.

Improving the neurophysiological response to natural speech

The effectivity of EEG-based AAD depends not only on the temporal resolution and the acoustic scenario, but also on the responses of the test subjects themselves (e.g. real and pseudo BCI illiteracy, see [234]). Large differences in accuracy between subjects were observed in the aforementioned studies. It is known for other cognitive paradigms (e.g. such as P300 oddball studies [242]) that subjects' physiological responses can differ substantially, even within subjects when moving from restricted to more real-life scenarios [242]. Providing feedback about the ongoing EEG signals was shown to be beneficial for other EEG paradigms to strengthen the brain responses (e.g. P300 ERP [17]) or induce specific neural patterns (e.g. motor imagery [233]). A similar reasoning could be applied for AAD, i.e. training users to elicit relevant brain responses with respect to the attended speech stream might increase the accuracy. To this end users might experience positive effects from such a feedback training to increase the performance of EEG-based AAD.

Changing the brain's neuronal activity in order to gain rewards is basically operant conditioning of neural activities. This is often referred to as neural biofeedback, or neurofeedback (NFB) whenever the brain activity of a human is the target. Neurofeedback training (e.g. through fMRI or EEG) has been defined as a method to self-regulate one's own brain activity [71, 192]. Adding a feedback in the typical BCI or AAD analysis closes the loop; measurements of neural activity are taken, and a real time representation of it is shown —whether tactile, auditory or visual— to the subject, thus facilitating neural activity self-regulation. Users can learn to control various aspects of their neural activity through a training process which displays online changes of relevant EEG features [157]. NFB training thereby engages several learning mechanisms in the brain of which operant conditioning is considered as the main principle [157, 71]. Changes in brain activity that reflect successful NFB training should be rewarded or positively reinforced. If the represented performance measure reaches a certain threshold, a rewarding stimulus should be presented. Typical visual feedback stimuli are a moving cursor, a bar or sphere varying in size or objects changing in color. Feedback is often presented in a game-like format, e.g. [157]. Such formats intend to increase and maintain the subject"s motivation and attention,

which helps facilitating self-regulation of neural activity and at the same time could cause increased behavioral task performance (i.e. create a positive state of mind). However, besides the behavioral control-input task, the subject thus has a second task to reach a specified goal regarding the NFB presentation. This results in a higher cognitive load and a division of cognitive resources between both tasks, which could lead to poorer BCI performance [101]. Therefore, the feedback signal should not distract the subject too much from the behavioral task to be performed [157]. For an in-depth review of the theories and models that have been proposed to explain the neural mechanisms of neurofeedback learning, we kindly refer the reader to a recent review by Sitaram et al. [192].

Training with neurofeedback has been utilized as a therapy tool for the normalization of brain activity (e.g. [174]), or as a tool for cognitive enhancement in the case of healthy subjects [71]. For example, NFB training for the improvement of attention ability [189], improvement of ADHD (attention deficit hyperactivity disorder) treatment [8], epilepsy [204] and post-stroke motor learning [233]-[122]. These studies assume a causal relationship involving specific motor function, behavior, cognition and brain oscillations-i.e., self regulating specific activities of the brain are meant to generate advantageous changes in motor, behavioral or cognitive function. The efficacy of NFB training can then be quantified through the use of two variables: (1) changes in EEG activity and (2) cognitive or behavioral changes of a targeted function [177]. Related studies have revealed that motor-imagery tasks can control EEG-based BCIs, in which users imagine specific body parts moving in order to control the BCI. After just a few training days, two brain states could be distinguished, (i.e. imaginary right- versus left-hand movements) with high accuracy [168]. In [233], participants were trained in the motor imagery (MI) task over four consecutive days during which the accuracy increased significantly and underlying brain patterns had evolved. In a related study [236] one subject received four weeks of motor imagery training on which the authors report of effectively altered EEG activity (reduction in contralateral event-related desynchronization of sensory motor rhythms) and increases in classification accuracy. They furnished a proof-of-concept for a mobile, low-density EEG system for effective NFB training at the home of the participant. Furthermore, recent case study reports by [235] strengthened the notion that MI training at home can lead to noticeable changes in brain patterns. Considering the AAD paradigm there is no reference study or standardized training protocol to date. Hence the method and extent that self-regulation of the neural responses can be applied to such an auditory attention task (i.e. speaker detection) remains unknown. Nevertheless, it should be noted that few reports exist of self-regulation of activity from human auditory areas facilitated by NFB training. For example the training of slow cortical potentials for suppression of tinnitus sensation [147]. In this longitudinal study one patient was provided with visual feedback on enhancing (and diminishing)

certain (low) frequency bands as recorded with EEG. This induced changes in resting state EEG patterns and lead to subjectively reported tinnitus suppression and reduction of tinnitus-related distress for the patient.

In Chapter 7 we implemented a fully automated closed-loop system that allows for convenient recording outside a lab environment and presents subjects with personalized feedback about their ongoing AAD performance. This way we investigate the interaction with an online AAD system and the potential for in-home neurofeedback training of the participants. To the best of our knowledge, neither of these have been reported on before.

3.3 Experimental EEG Data

3.3.1 Datasets

In order to investigate the signals and cognitive aspects of auditory attention as evident in mobile EEG in real life, it is necessary to measure EEG data in several conditions and (outdoor) circumstances. The current paragraph summarizes the experiments and datasets used in the following chapters. Table 3.1 provides an overview of the most relevant characteristics for each dataset. These datasets will be referred to with their corresponding label from this point forward.

At the start of the PhD, one of the only existing mobile EEG datasets [49] was obtained, i.e. made available by the authors, that possessed similar characteristics to the studies we planned to conduct. The authors evaluated the P300 and noise levels in a three-class auditory oddball task during walking outdoor and while seated outdoor. This way we were able to start with the exploration of our tensor based analysis. This dataset will be denoted with label A. In addition, we simulated three-class oddball data with a variety of noise levels and similar characteristics to systematically evaluate the benefit of exploiting structure in the paradigm and datasets through tensor decompositions (Label B). In parallel, we recorded novel EEG data to provide the necessary experimental data to deal with the cognitive and methodological issues we want to tackle. Psychological paradigms were welded into real-life scenarios (e.g. outdoor or in-home) with differentiations in the level of distraction and cognitive load. In addition, healthy volunteers were recruited over the course of the PhD to participate. This has led to three different types of data: Firstly, the mobile EEG system (SMARTING, mBrainTrain Serbia) signal quality was evaluated during head and body motion (label C). This allowed to observe the influence of specific head and body movement artifacts in unconstrained real-life recordings. Secondly, we aimed to analyze three-class auditory oddball

data in several outdoor conditions which allows for interpretation of the EEG related to the user's environment and physical engagement during biking (Label D). Thirdly, we aimed to create an online system that allowed for real-time monitoring of auditory attention to natural speech (i.e. speaker detection paradigm). In addition, this protocol was extended to provide personalized neurofeedback about the subject's performance (Labels E and G respectively). Finally, we recorded preliminary data from subjects that performed both the auditory oddball task as well as the AAD (Label F) to investigate the relation between these paradigms. Furthermore, this exploratory experiment was performed to examine visual distractions and physical engagement in more detail. Taken together, these experiments were conducted to provide a solid base for investigating the challenges described in the previous section. The next paragraph provides technical details that are (in part) shared between the experiments.

3.3.2 Data acquisition

All new data acquisitions were conducted with a SMARTING mobile EEG amplifier from mBrainTrain (Belgrade, Serbia, `www.mbraintrain.com`). This comprises a wireless EEG system running on a notebook (Dell Latitude E7440 until the end of 2015 and Dell XPS 9550 from 2016 on) using a small 24-channel amplifier with similar characteristics to a stationary laboratory amplifier ($<1\mu$V peak to peak noise; 500Hz sampling rate) and >4 hours runtime combined with a small lightweight design (<50 grams) which allows for capturing full brain dynamics with minimal restrictions to the user. The EEG was measured using 24 Ag/AgCl wet passive scalp electrodes (Easycap, www.easycap.com), placed according to the 10-20 standard system with positions: FP1, FP2, Fz, F7, F8, FC1, FC2, Cz, C3, C4, T7, T8, CPz, CP1, CP2, CP5, CP6, TP9, TP10, Pz, P3, P4, O1 and O2. These channel positions are shown in Chapter 2.3.1 Figure 2.4. Impedances were kept below 10 kOhm and an abrasive electrolyte-gel was applied in each electrode. In addition three-axis (x,y,z) gyroscope data was measured at similar sampling rate as the EEG. The gyroscope is included in the amplifier which is placed at the back of the head. Recent studies indicated that such a setup allows for comparable data quality to a state-of-the-art wired laboratory EEG system [50]. All experimental data acquisitions in this project was reviewed by the SMEC ethical commission of the Faculty of Psychology, Leuven.

The EEG data for experiments C, D, E, F and G were recorded through Openvibe and stored for offline analysis. Experiments E and G were analyzed online as well; the data were streamed consecutively from Openvibe to Matlab via the labstreaminglayer interface (LSL). Both online and offline analyses

Table 3.1: An overview of the datasets used throughout the manuscript and sorted in chronological order. One external dataset was obtained to start instantaneously with proper mobile EEG data recorded in an ambulatory setting. During the course of the PhD new datasets were recorded (and one simulated). Ch. = chapter and Dis. = discussion. Note that Dataset C and F are summarized in the Appendix only.

External (Existing) Dataset

Label	Dataset	Paradigm	Main theme	Location	Setting	Posture	n =	Year	Ch.
A	M. De Vos et al.	3-class oddball	BCI outside	Oldenburg University (DE)	Outdoor	Sit-walk	20	2014	**4-5**

New Datasets

Label	Dataset	Paradigm	Main theme	Location	Setting	Posture	n =	Year	Ch.
B	Simulated	3-class oddball	Noise levels	n/a	n/a	n/a	15	n/a	**5**
C	Head and body motion artifacts	n/a	Artifacts	KU Leuven	Indoor/ Outdoor	Stand-walk	14	2014	**Dis.**
D	BCI during Biking	3-class oddball	Distractions	KU Leuven	Outdoor	(fixed) bike	15	2015	**6**
E	Online analysis of AAD	AAD	Online Analysis	Office setting KU Leuven	Indoor	sit	12	2016	**7**
F	Virtual city navigation with bike simulator	AAD & 3-class oddball	Distractions & subject variability	Lab setting KU Leuven	Indoor	Fixed bike	10	2017	**Dis.**
G	Long-term in home AAD training	AAD	Training	Brussels - Leuven	In-home	sit	2	2017	**7**

used custom-made Matlab scripts that were combined with functions from the EEGLAB13.4.3b [58], BCILAB1.1 [59] and TENSORLAB2.0-3.0 [196, 221] toolboxes that are publically available.

3.4 Conclusion

In this chapter we established the current state-of-the art in measuring auditory attention through mobile EEG and highlighted current shortcomings for which solutions are of high interest from both the research field and subject's point of view. Solutions for several of these challenges are proposed and serve as an introduction to the next chapters that will go more in depth into the methodology and obtained results. Additionally, we provided an overview of all the EEG data used in this thesis and presented in detail the characteristics of our data acquisition hardware and software interface.

Chapter 4

Data-driven P300 extraction using tensor decompositions

In Chapter 1 and 3 we identified long setup and calibration time as one of the key reasons that hinder adoption of mobile BCI. Moreover, in mobile scenarios the noise signatures in the data are often unknown. To this end, in the current chapter, we explore a novel approach for BCI data interpretation through data-driven tensor decompositions. First, we illustrate the CPD model and cover its basic concepts and parameters that are important in the analysis of higher-order EEG data. The simple and straightforward interpretation of the CPD is further outlined in the second part of this chapter in which CPD is applied to a mobile BCI dataset (Dataset A) to obtain data-driven separation of target and non-target trials in the auditory oddball paradigm. This work has been published in [240]. The next chapter builds on these findings and introduces more complex models for EEG BCI data.

4.1 From matrix to tensor

In this section we introduce decomposition techniques for higher-order datasets. This explanation is focussed on the application of these techniques for EEG data. For an extensive overview and description of higher-order tensor decompositions we refer to three comprehensive overview papers in this field, [117, 35, 190].

4.1.1 Matrix based Blind Source Separation

EEG data are often thought of as being in a two-way format (a matrix) representing the activation over time at several electrodes. The columns in this matrix would be referred to as the time mode and rows the spatial mode.

Matrix EEG data allows for convenient displaying and visual inspection of continuous EEG (e.g. artifacts) in which the horizontal axis usually displays the time dimension. Some of the most popular signal processing methods in EEG signal analysis are based on Blind Source Separation (BSS), i.e., of matrix EEG data. Generally, BSS describes the problem of processing multidimensional observations that are a linear mixture of distinct underlying sources with additive noise. In other words, let $x(t) = [x_1(t)...x_P(t)]^T \in \mathbb{R}^P$ be the observed signal at time instant t, and $s(t) = [s_1(t)...s_N(t)]^T \in \mathbb{R}^N$ the underlying sources. Then $x(t)$ can be written as

$$x(t) = As(t) \tag{4.1}$$

where A is an unknown mapping from $\mathbb{R}^N$ to $\mathbb{R}^P$. The goal of BSS is to find the sources $s(t)$ and the way they are mixed (i.e the mixing matrix A which provides information about the relative contribution of the sources) with respect to each observed signal in $x(t)$. For the matrix case, this problem is ill-posed and assumptions have to be made about the underlying sources in order to reach a unique solution [37]; uniqueness can be ensured by various constraints such as orthogonality or independence, among others. A frequently reported BSS method in EEG signal processing is Principal Component Analysis (PCA). It provides an estimate for $s(t)$ and A in equation (4.1) by the assumption that the sources $s(t)$ and mixing matrix A are orthogonal. Alternatively, PCA is often employed as dimensionality reduction technique; see [28] for an overview. Another common BSS method is Independent Component Analysis (ICA). ICA assumes that the EEG sources $s(t)$ are temporally independent; in ICA the EEG data is decomposed into so called independent components (IC); see [37] for an extensive overview. ICA has been proven a very effective and a popular signal processing tool in a large number of EEG applications. Some examples are, the process of artifact removal (e.g. of removal of Electrooculography (EOG), [103]) and feature extraction (e.g. N170 ERP detection [52]). Finally Canonical Correlation Analsysis (CCA) applied as BSS technique assumes sources $s(t)$ which are mutually uncorrelated and maximally autocorrelated in time. CCA was found to be useful in the extraction of electromyography (EMG) related sources (e.g. [43]).

The obtained solutions of these methods are still subject to two inherent ambiguities, namely the order and the scaling of the sources, as they can be compensated by the inverse scaling and permutation of the columns of the mixing matrix. We call a decomposition essentially unique if it is only subject to these indeterminacies. Finally, constraints like orthogonality and independence necessary in matrix BSS can be too strong assumptions or even lack physical meaning with respect to modeling the neural sources within the EEG [2, 153, 72].

4.1.2 Structure of EEG data

EEG data are often represented as higher-order data (i.e. >2). This may be due to specific choices in the analysis that extend the [channel × time data] for example to the frequency domain creating a three-dimensional tensor ([channel × time × frequency]) [155]. Additionally, EEG data are rarely analyzed as a continuous recording. Especially in cognitive paradigms data are epoched, usually before and after events, e.g. stimulus presentation or button press. Consequently, the data are structured per trial or event which results in a third-order tensor as well: [channel × time × trial]. Note that the time dimension could be equivalently replaced by frequency or by [time × frequency]. Furthermore, EEG data recorded from multiple subjects and in multiple conditions leads to tensors as well. One could stack all subjects and decompose the EEG patterns to find common structure between subjects' EEG signatures, for example, [subject × channel × time × frequency]. These are but a few of the natural ways that EEG data can be represented as a higher-order data tensor. In addition, new modes can be added artificially by means of tensorization techniques; for example, specific reshaping or duplication of EEG data to optimally exploit the structure present in the EEG (e.g. Hankel expansion for modelling ictal patterns in EEG [97]). See [54] for an extensive overview of tensorization techniques.

Analysis of multidimensional EEG data with matrix based methods requires a transformation of the EEG data into a matrix e.g. through tensor unfolding. However, such operations destroy the structural links between the modes of the EEG tensor. Alternatively, it may be advantageous to directly decompose the higher-order data to obtain information about the underlying sources without artificially unfolding the tensor to a matrix. These advantages will be highlighted in the next section.

4.1.3 Decomposing multidimensional EEG data

Whereas matrix decomposition techniques can be applied to perform BSS on matrix data, tensor BSS can be obtained through tensor decompositions. Decomposing higher-order data directly i.e. instead of unfolding to a matrix, allows for more relaxed constraints on the sources as opposed to those discussed in the matrix case. Moreover, as described in the previous section, EEG data can be often thought of as being naturally in a higher-order format. One of the simplest decomposition models in terms of computation and interpretation is the Canonical Polyadic Decomposition (CPD). It is widely used in many signal processing domains. Due to its straightforward interpretation CPD is a compelling example for introducing higher-order decompositions of EEG data.

More complex models can be used as well to achieve better performance in specific cases. In Chapter 5, this is demonstrated with a special case of Block Term Decomposition (BTD), i.e. a decomposition in multi linear (L_r,L_r,1) terms [45]. The latter will be referred to as LL1 decomposition from now on. In addition, we also consider coupled CPD models in the next chapter. This allows to decompose multiple EEG datasets at the same time which can be advantageous if there are many data available.

EEG data represented as a tensor can be decomposed as a sum of rank-1 terms [93]. A polyadic decomposition (PD) decomposes a data tensor into R components. If R is minimal the decomposition is canonical in which R is the rank and the decomposition is called CPD. As becomes evident in the next section it can be difficult to determine the rank in practice. Especially in the case of a low SNR we are satisfied with a low-rank approximation of the EEG as provided by the CPD. Nevertheless, this assumption will be considered valid in general for the CPD models as is often done in the literature. The extracted sources of a CPD should follow a rank-1 structure. However, this does not imply that the complete data tensor should consist of meaningful rank-1 terms; data patterns that do not follow structure in the tensor, e.g. specific types of artifacts which cannot be described well with a rank-1 (or low rank) component, are expected to be less likely captured by the CPD. Generally, the CPD aims to capture as much variance as possible in rank-1 terms and the residual is considered the error of the CPD model.

For the third-order case the CPD will decompose a tensor X as follows:

$$X = \sum_{r=1}^{R} a_r \circ b_r \circ c_r + \varepsilon \tag{4.2}$$

with R (i.e. the rank) representing the number of components, a_r, b_r, and c_r the signatures of every entry in each of the modes, and ε the model error. Each mode has a specific signature which characterizes the extracted component; in the third-order tensor representing a P300 BCI dataset as a channel $\times$ time $\times$ trial structure, the spatial distribution of the different entries would be contained in a_r, the time courses would be contained in b_r, and a strength of the space-time signature across trials would be given in c_r, this is schematically depicted in Figure 4.1. The CPD model is trilinear, i.e. each mode's vectors are proportional to each other within a rank-1 component. Hence, the CPD model assumes that the source maintains the same P300-waveform and topography within the observed trials.

An illustrative example of a CPD on EEG BCI data is depicted in Figure 4.2. A simulated EEG dataset (single example taken from Dataset B, as described

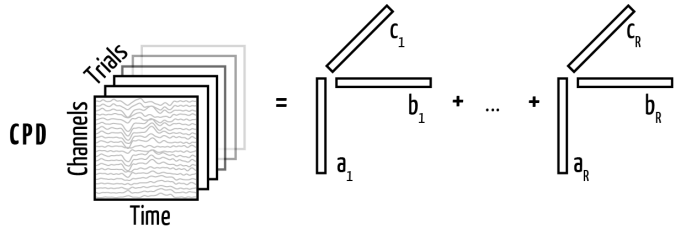


Figure 4.1: Schematic of a decomposition with CPD into R terms.

in section 5.2.1 in the next chapter) of a three-class auditory oddball paradigm is decomposed with CPD into $R = 3$ rank-1 terms. The artificially generated dataset has a relatively low noise level, i.e. high SNR of 1.75 (SNR represents the linear ratio of the maximum P300 peak and the pre-stimulus RMS of the EEG). The data tensor contains P300 ERPs in the first half of the trials (i.e. 1-100, target trials, in red in the figure) and not in the latter half (101-200 non-targets, blue color). The tensor dimensions are: [24 channels $\times$ 500 time points $\times$ 200 trials]. The first component extracts characteristics that relate to the N100 ERP with a frontal central focus, and the second to the P300 effect with a posterior topography. The last component reflects an alpha noise source. From the third mode, i.e. trial mode, it can be noted that the second component clearly differentiates between the two classes in the data. In this dataset the trials with P300 could be separated with 92% accuracy from the non-target trials by using a simple kmeans clustering on the trial mode vector (i.e. a threshold to obtain two classes).

Another example of CPD is illustrated in Figure 4.3 to extract specific time-frequency information of wavelet-transformed single-trial (mobile) EEG data (i.e. from Dataset D, published in [241]). Theta 4-8Hz, alpha 8-12Hz and Low beta 12-16Hz bands are considered in the analysis and the corresponding CWT scales are used for the decomposition resulting in a 22 channels $\times$ 40 time points $\times$ 17 frequency scales tensor per trial. R was set to 3, as this was shown to be adequate in previous work [217]. Frequency information was obtained through a wavelet transform of the EEG. Subsequently, the transformed EEG data is decomposed with CPD. This way we are able to derive estimates of the aforementioned oscillations (assuming they are present) in a data-driven way. Each component in the CPD can be characterized by an individual channel, time and frequency signature and is classified as theta, alpha and low beta band source if their spectral signature corresponds to the respective frequency bands. This way, relevant electrodes that convey similar spectral activity are clustered in a single component. In addition, given the nature of CPD, it is possible to remove selected sources. We reconstruct the frequency band-specific information

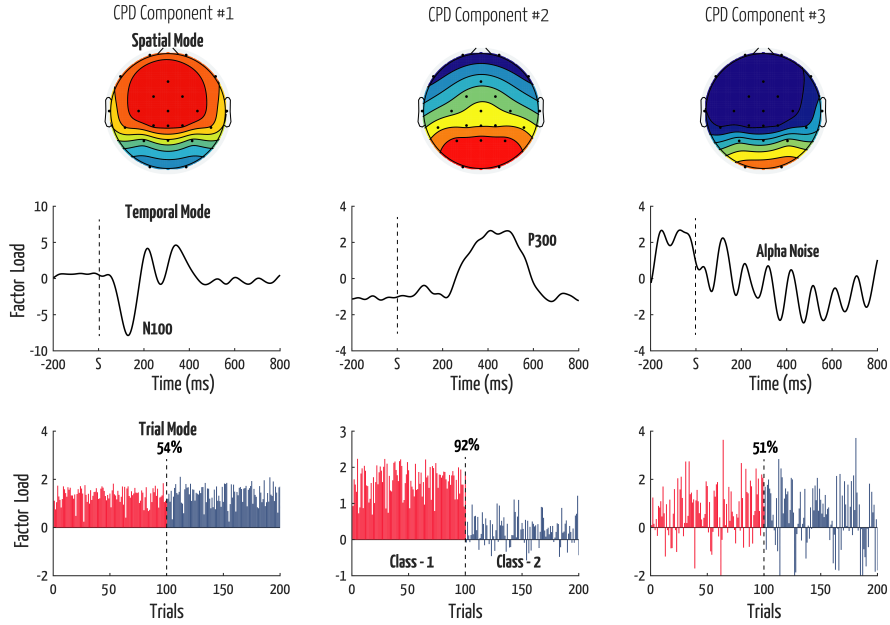


Figure 4.2: Example CPD outcome of a dataset from noise-level 2, illustrating the decomposition with the corresponding spatial, temporal and trial modes for a model of $R = 3$. The components depict the N100, P300 and noise signature from left to right, respectively. The first half of the trials corresponds to the target trials, the latter to non-targets. Component 2 is able to separate the classes with 92% accuracy, based on the factor weights of the third mode.

from the CPD component by multiplying the temporal and spectral mode and scaling it by the channel weight in the decomposition. This process generates per channel a matrix that estimates the source's contribution for a given trial. The result was subtracted from the original wavelet transformed data and an inverse wavelet transform was applied consecutively to obtain cleaned channel $\times$ time data in which the selected component is removed. An example of removing the alpha source from Figure 4.3A has been illustrated in Figure 4.3B.

The usage of tensor decompositions for biomedical data processing has gained interest recently in various domains. For example, supervised classification of EEG [163, 137], extraction of EEG frequency band information [15, 241, 217], ECG segmentation [84], analysis of white-matter structures in the brain using diffusion tensor imaging [203] and joint analysis of EEG-fMRI [98].

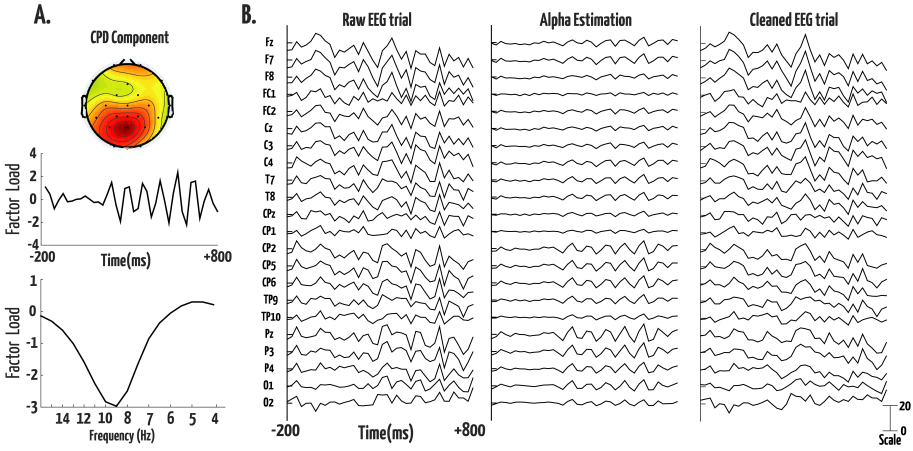


Figure 4.3: A) A single alpha component of the CPD model of subject 1 of dataset D. From top to bottom, the spatial, temporal and spectral modes are depicted. B) Example of the single-trial EEG data from (a) and the reconstructed alpha patterns by back-reconstructing the CPD component and applying an inverse wavelet transform on the filtered time-frequency matrix provides the cleaned EEG trial.

4.1.4 Practical Considerations

Obtaining an interpretable and useful decomposition depends on the handling of several aspects of CPD analysis. In this section we will discuss the initialization of the CPD, determination of the rank and visualization of the obtained sources.

Initialization

The solution of the CPD is computed through an optimization based algorithm. This requires an initial solution for each factor that is iteratively updated by the algorithm to obtain the appropriate guess. These initializations are usually generated randomly, although it is possible to specify specific patterns in one or more of the modes. It is not evident that CPD will return the same solution with different (random) initializations; the algorithm may converge to a local minimum and therefore not always give the optimal solution. Therefore it is good practice to evaluate the fluctuations in performance (e.g. classification accuracy) over a substantial number of random initializations.

Number of components

One of the most influential parameters that needs to be set is the rank (i.e. R, the number of components as mentioned in the previous section 4.1.3) of the model. For the matrix case BSS it is common practice to set the number of components (e.g. ICs in ICA) equal to the number of expected sources in the brain. Some have maintained an identical reasoning to estimate the rank of higher-order decompositions [38]. However, from experience, a lower number of components (i.e. as in the case of matrix based BSS) may be appropriate as long as the EEG patterns of interest contain structure as captured by the model. This differentiates our approach from that of choosing R as the number of sources in the brain. Since we aim to model the most structured high variance sources there are, to our experience, fewer of these. Artifacts often do not follow a low-rank pattern due to the fact they are not structured. That is the key of why this type of data driven approach can be useful and a low number of components can capture meaningful processes.

Automatic rank estimation methods (e.g. core consistency as applied in [97] or Rankest as part of Tensorlab [221]) were shown to be able to estimate the exact rank of a data tensor in specific cases. Unfortunately, they do not produce accurate results on data with a low SNR, such as single-trial EEG [39]. Subsequently for EEG tensor decompositions heuristics and visual exploration are key to obtain useful models. Similarly to ICA, plotting the obtained components factors (e.g. spatial, temporal, trials) is very informative for evaluation of the CPD (and other tensor decompositions). Even though this can be a tedious process, prior knowledge can help in such estimations. For example, let us consider the P300 classification problem discussed previously. It is expected to have a single component that captures the P300 spatio-temporal signature. If high correlations in multiple factors are observed in the outcome of CPD, one should check whether 2 or more terms nearly cancel each other. A phenomenon referred to as degeneracy in the decomposition and is indicative of a wrongly set rank.

Uniqueness

If R is low with respect to the size of the tensor the decomposition is usually unique up to permutation and scaling of the extracted components [117]. A more general framework for uniqueness has been recently presented in [63, 64]. This is one of the main differences between a matrix decomposition and CPD. For the latter there is no need to impose additional constraints on the factors such as for example independence in the case of ICA.

Visualization

In this thesis, the CPD is computed with the nonlinear least squares (NLS) algorithm in the publicly available Matlab toolbox Tensorlab [196, 221]. Tensorlab offers an extensive collection of algorithms and powerful tools for the analysis of higher-order datasets. In addition, the usage of tensor decompositions in the field of EEG signal processing would benefit from a graphical user interface (GUI) that allows to visualize the necessary steps for inspection of the model parameters and data exploration. To this end we created a proof-of-concept GUI that uses a third-order EEG tensor as input and lets the user visually interact with the different aspects discussed in this section (i.e among others: rank, initializations and degeneracy). We envision a future version of the GUI to be similarly intuitive as to how ICA solutions are plotted for removal of EOG activity in EEGLAB, which has been adopted by many researchers for manual inspection of the ICs [58]. In Appendix A the GUI layout is illustrated and a brief description of the GUI features is given.

4.2 CPD based classification of the P300 in mobile EEG.

In this section we will evaluate CPD models to classify real mobile EEG data in which the noise patterns and the variability of the P300 effect are unknown (i.e. as opposed to the simulated data example in Figure 4.2). We will focus on the use of CPD for decompositions of full datasets. This way we aim to provide a proof-of-concept for these types of methods to extract meaningful signatures with respect to the P300 and cluster the target from non-target trials in the auditory oddball data (Dataset A).

4.2.1 Classification in BCI

Currently most BCI systems involve 2 stages: training and testing. During the training stage, a supervised classifier is trained on the recorded data to discriminate between different stimulus classes. However, supervised classification methods require a reasonable amount of training data to optimize classifier performance [132, 131]. Acquiring separate train and test data increases the overall setup time. Moreover, nonstationarities due to varying signal quality and brain adaptations over time decrease the classifier's performance when train and test data are further apart in time. Especially in augmented reality scenarios in which only sparse interaction with a BCI is needed, supervised techniques still carry significant drawbacks. Unsupervised classification methods

can potentially address these issues and be advantageous for use in mobile BCI applications [106, 88]. We propose a new approach for unsupervised classification of P300 ERP data based on CPD. To illustrate our methodology, we first explore the relationship between the CPD parameters and performance on large datasets of 188 trials. An accurate CPD on smaller time-windows would allow for applications involving short-term interactions with a BCI (e.g. operating assistive technologies). To this end we present a method to classify single trials in a 10-trial window. The results are compared to those of stepwise LDA (swLDA), which is considered a prominent classifier for the P300 ERP. We also define future lines of research tackling the current limitations of CPD based classification for online unsupervised BCI.

4.2.2 Data and methods

Data acquisition and stimuli

Eight healthy subjects (mean age 24.6 years) performed a three-class oddball auditory task. These subjects are part of the data described in [49] and followed the previously described auditory oddball paradigm (See Chapter 2, Section 2.4.1); a standard tone (900 Hz) and two deviant tones (600, 1200 Hz) of 62-ms duration (10 ms rise/fall time) were presented in random order (ISI 1000 ms, Jitter 0-375 ms). The participants' task was to silently count the target tones (i.e. for 4 subjects the 600Hz tone and other 4 subjects the 1200Hz tone) and ignore the two other tones. 504 standards, 94 non-target deviants and 94 target deviants were presented in randomized order while the subject was sitting on a chair outside. OpenViBE software running on a notebook was used for stimulus delivery and experiment control. Data acquisition was performed with a modified Emotiv (www.emotiv.com) EEG system as described in [57, 50]. Sintered Ag/AgCl electrodes (wet) were placed at the 10–20 positions FPz, F3, Fz, F4, C3, Cz, C4, TP9, TP10, P3, Pz, P4, O1, and O2 (with common mode suppression ("online reference") at AFz and driven right leg ("ground") at FCz).

Pre-processing

The data were analyzed offline using EEGLAB [59] and MATLAB. Extended infomax ICA was used to semi-automatically attenuate contributions from eye blink artifacts [47]. EEG data were 20 Hz low-pass filtered, re-referenced to the average of Tp9 and Tp10, and epochs were extracted from -200 to 800 ms and baseline corrected (-200 to 0 ms). In order to speed up the CPD, the data were down-sampled from 128 Hz to 30 Hz. As input for the CPD we concurrently looked at the temporal signal as well as two derived signals: the normalized

time signal and the frequency power. For the former this was achieved by normalizing every epoch to z-scores and the latter by computing a fast-fourier transform of every epoch and computing the squared magnitude in a 2-14 Hz interval, as this is expected to capture most of the P300 waveforms [60]. CPD decompositions were computed on a post stimulus interval of 100-800ms.

CPD models

The CPD model as described in the previous Section 4.1.3 was used to decompose the datasets. The data is structured in three different tensors following the last three preprocessing steps. Firstly, the data is structured as [channels $\times$ time $\times$ trials], which is expected to capture the P300 waveform present in the data; this is referred to as CPDtime. Similarly we construct a tensor with the normalized time signals which is referred to as CPDtimeNorm. Normalizing the time course is expected to improve the CPD model's accuracy in detecting small amplitude changes on some channels in contrast to high amplitudes due to artifacts on others. For both described time models the fluctuation in magnitude of the P300 can be taken care of by CPD; it is reflected in the trial mode signature. However, shifts in the P300 latency between trials cannot be resolved thusly. In order to deal with these latency changes we apply a third CPD model on the frequency estimates from every trial in a [channels $\times$ frequency $\times$ trials] tensor, referred to as CPDfreq.

Model Parameters

The success of CPD depends on the appropriate setup of certain criteria in the analysis. For the full dataset analysis we discuss the most influential aspects: initialization and rank determination. Being a non-convex optimization method, CPD is not guaranteed to always find the optimal solution. Therefore we report our results based on 100 random initializations. This way we derive reliability estimates and present results based on the best initialization as evaluated by classification accuracy. The rank (R) is expected to differ with the complexity of the dataset. Without defining a specific rank beforehand, we evaluate the results for different ranks ranging from 1 to 12 (arbitrary choice).

Discriminating trials

In order to discriminate between target and distractor trials with CPD we focus on the factor estimates in the trial mode. If the rank equals 1, we obtain a single vector with trial weights. Two classes can be distinguished with opposite signs, allowing for a simple, yet effective separation technique, e.g. Figure

4.6. These opposing patterns are also recognizable between the average ERP patterns of the non-target and target stimuli (e.g. Figure 5.1). In the case of a higher rank we consider a k-means cluster algorithm (k-means++, [10]) to find 2 distinct clusters. Every trial mode of the R-components is normalized (z-score) and together given as input for k-means. Being dependent on semi-random initializations, the k-means procedure is repeated 999 times and the most frequent prediction vector is used as final prediction.

First we explore the relationship between the CPD parameters and performance on large datasets of 188 trials. Secondly, we investigate its performance on 10-trial subsets (with equal class distribution) to mimic an online scenario. Consequently, the 10-trial overall performance is based on 178 models per subject of 10 trials which are time-shifted by 1 trial for the consecutive model (i.e. sliding window of 1 trial). Single trials are analyzed together with the previous 9 trials in time, this way the added trial can be classified directly online. The number of correctly classified trials (max = 178, this deviates from the total of 188 trials because the first 10 trials are used as initial model) of the 10-trial models is reported as performance measure for each subject. Figure 4.4 summarizes the most important steps in the CPD analysis. Both the sign and kmeans clustering only allow for separation of trials into two classes; it does not identify which class corresponds to the target or distractor stimuli. Label information from prior trials is utilized to evaluate the accuracy of the derived CPD estimates in the 10-trial decompositions. This was achieved by comparing the factor weight in the trial mode of the newly added trial to that of the other trials with known labels in that window. This way we can identify which sign is associated with the target class. Future extensions for an adaptive way of determining the identity of the trials are presented in the discussion section.

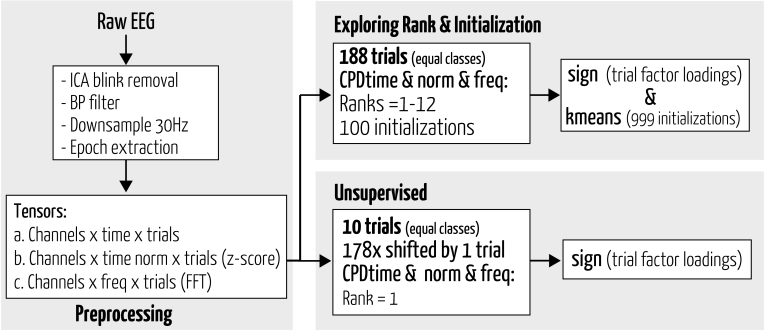


Figure 4.4: Overview of the most important CPD analysis steps

Evaluation of the CPD based accuracies is achieved by a comparison to those of swLDA, which is one of the widely used P300 classification algorithms. The

original feature set consisted of seventeen 47 ms data bins between 0 and 800 ms on all 12 electrodes. The swLDA method adds relevant features sequentially. A new feature is added to the final feature set if it statistically improves class discrimination (p_{in} <0.1). After adding a new feature, reanalysis of the current features could lead to removal of a redundant feature (p_{out} >0.15). Although the stepwise feature selection reduces the number of features used, shrinkage regularization as implemented in BCILAB [59] was applied to further reduce the risk of overfitting. Half of the trials are used for training and the other half for testing and vice-versa in the full dataset. Classification results for the small 10-trial subsets are obtained with a 5 fold cross-validation procedure.

4.2.3 Results

Electrophysiological results

The subject-average subtraction of target-distractor ERPs are presented in Figure 4.5. These averages illustrate different waveforms of the P300 and peak amplitude and latency between the eight subjects. For all subjects, the P300 of the target stimuli was significantly (at a 5% chance level) larger compared to the distractor stimuli at the P300 peak latency on the Pz electrode. Trial type classification for target and distractor resulted on average in 74.7% accuracy (Standard Deviation (SD) ±6.7) for swLDA across subjects

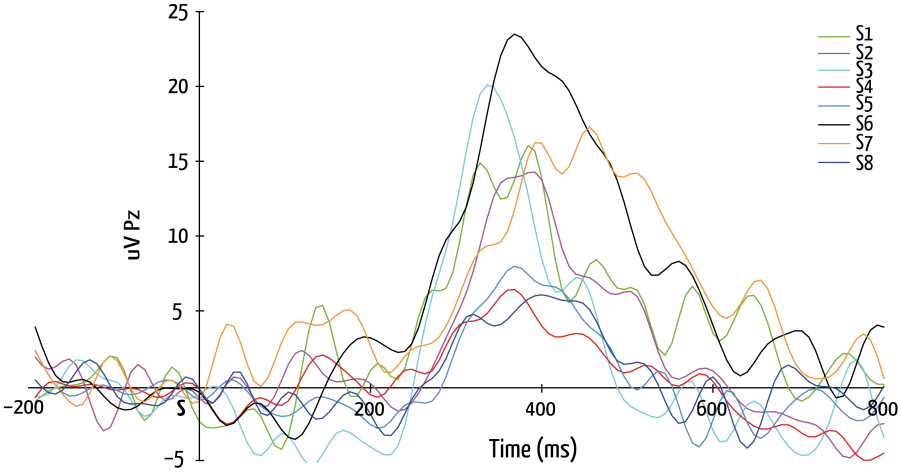


Figure 4.5: Average difference of the Target and NonTarget ERPs per subject at electrode Pz

CPD parameters on all trials

In order to illustrate the underlying single subject decomposition process, components from both a CPDtime and CPDfreq for a representative subject are shown in Figure 4.6. The modes correspond to the spatial, temporal and trial dimensions from top to bottom, respectively. The CPDtime component follows the average target ERP (red dots). Similarly, the CPDfreq component is congruent with the target stimulus' frequency spectrum. This leads to similar trial distributions in the third mode. Based on these trial factor loadings, 74.7 and 78.2% correct identification of the trials was obtained for the time and frequency model, respectively. These CPD components are uniquely reconstructed up to arbitrary scaling and permutation of the modes which explains the opposite signs for class-1 between the two CPD models. Finally, similar spatial maps can be observed with a focus on electrodes Cz and Pz.

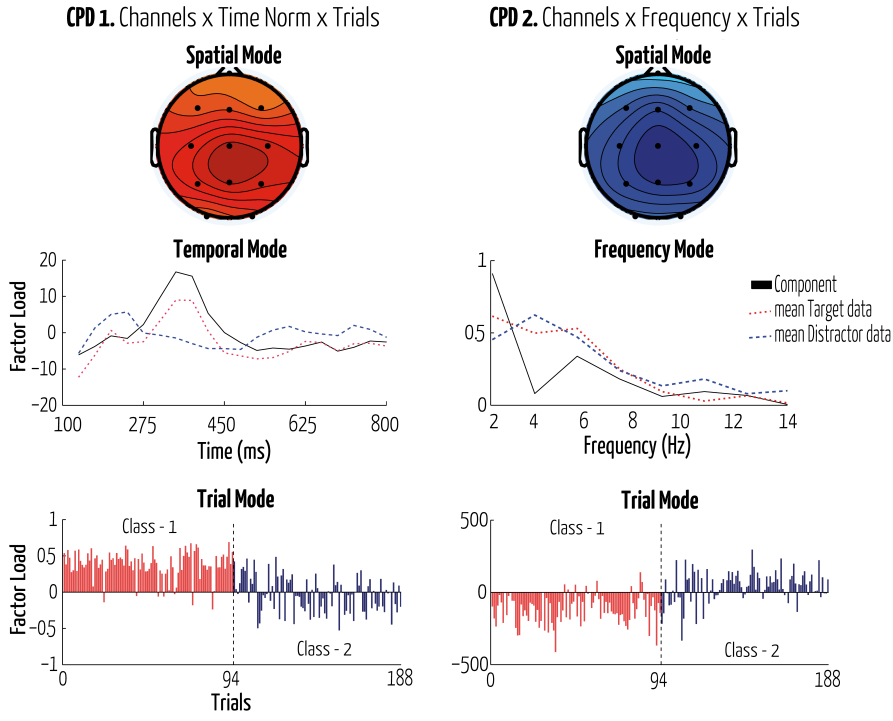


Figure 4.6: Single subject (S3, rank 1) example of a CPDtime and CPDfreq component on the large dataset. The modes correspond to the spatial, temporal and trial dimensions from top to bottom, respectively. The first half of the trials correspond to the target stimuli, the latter to nontargets. This is also reflected in the trial factor loadings by CPD (75% and 78% accuracy, respectively).

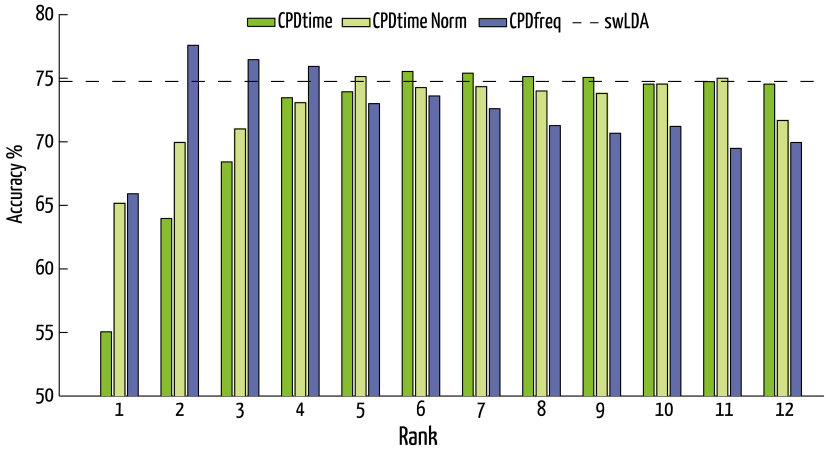


Figure 4.7: Grand-average accuracy for different ranks of the CPD models. The swLDA accuracy is indicated as reference

Figure 4.7 illustrates the average discrimination across subjects for rank 1 to 12 based on optimal initializations for CPD. On average across subjects, for all three CPD models (i.e. time, time-norm and frequency) we obtained up to 75.5% (±6.8), 75.0% (±5.7) and 77.5% (±5.4) discrimination respectively. No significant differences were observed between swLDA and neither the CPDtime for ranks > 3 nor the CPDfreq for ranks > 1. The CPDtime-norm and CPDfreq models outperform the time based models for lower ranks on this large dataset. However, the fluctuations between models with different initializations differ substantially. The CPDfreq models are considerably influenced by the initialization compared to the time models at lower ranks (<6). Further analysis showed that for low rank it often gets stuck at a local minimum which is not discriminative.

10 trial CPD models

Classification accuracies (±SD) on the 10-trial subsets for the swLDA are on average 62.1% (±4.7) across subjects. For the CPD models based on the time, normalized time and frequency tensor we obtained 56.4 (±2.1), 63.0 (±5.8) and 62.9% (±4.7) accuracy respectively. The normalized time and frequency CPD results are comparable to those of swLDA. Figure 4.8 illustrates the average accuracies for each subject. Further analysis showed that the CPDtime outcomes were only mildly influenced by the changes between random initializations. The frequency CPD models varied substantially, similar to

the full dataset results. Finally, we observed that components that correlated positively with the target ERP as opposed to the distractor ERP (or vice-versa) scored higher in discriminating power.

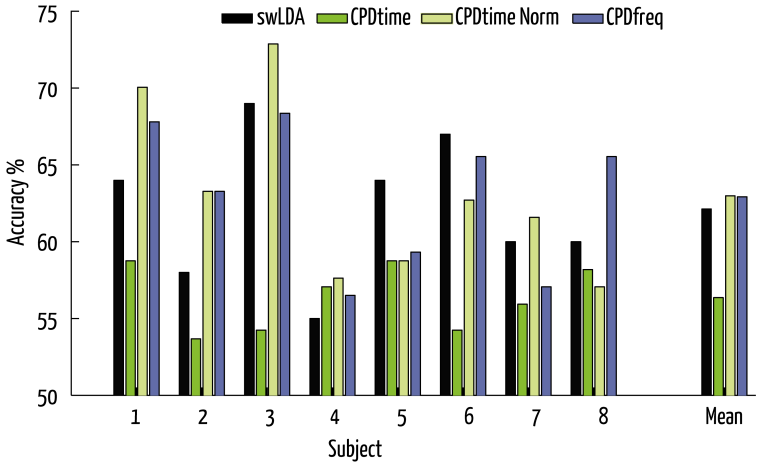


Figure 4.8: Average accuracy per subject with a rank-1 CPD model.

4.2.4 Discussion

The CPD as presented here illustrates a novel way to analyze auditory BCI data based on CPD. The benefits of CPD based discrimination of single trial ERPs in a 3-class auditory oddball task are shown. Even in small datasets, CPD is able to produce meaningful estimates. Applying CPD in an adaptive way, e.g. [216], so that it could update only the necessary structures is expected to further increase the results. A model initialized on only a few trials that can be updated over time so that it allows exclusion of old data when non-stationarity is occurring has potential for a successful zero-training algorithm. The principal drawback of the presented approach is the inability to identify extracted clusters without prior label information of some trials. Our CPD approach finds distinct clusters that could rely on either non-target or target responses. Future work should focus on the automatic detection of which cluster represents the target stimuli, e.g. by incorporating space-time heuristics about the P300. Addition of the known baseline trials in the CPD might solve this issue as the non-target response should be closely related to the baseline trials. Another solution might arise from the analysis of the 10-trial models which showed that the most discriminative models (80-100%) are strongly correlated to either the target or

distractor stimulus and oppositely to the other. Incorporating the results of the second mode in the CPD decomposition (i.e. time or frequency estimates) for identifying the class-stimuli relation might lead to correct labels.

Our CPDfreq results are noticeably influenced by the initialization of the CPD models. CPDnorm was found to be less prone to these instabilities and therefore more promising for future applications. Finding an objective measure to handle this instability would increase the application potential in noisy datasets. We considered using the residual error in the CPDs as a measure for model fitness to improve the stability. However, this measure proved unsuccessful in identifying well-classifying models.

Improvement of the classification results by combining the trial dimensions of the CPDtime and CPDfreq models in a joint classifier did not lead to an overall improved classification. This suggests that both time and frequency decompositions provide signal components that represent approximately similar underlying source activation. For the 10-trial analysis, the normalized tensor seems to be well decomposed by CPD. The normalization step decreases the amplitude differences across trials, making the optimization algorithm less prone to outliers due to artifacts. Apart from the CPD, our results can also be influenced by the clustering algorithm. Given the nature of k-means it is not always optimal in finding discriminative clusters. Classifying only on CPD components that represent the signal of interest can increase the results for ranks >1. However, this would require an additional (automatic) component selection step.

4.3 Chapter Conclusion

In this chapter we showed that EEG data can be naturally represented as a multidimensional tensor which grants possibilities for exploiting structure between the dimensions with tensor decompositions. We introduced the CPD as a data-driven method to represent the most prominent structure source. Data tensors can convey information about the conditions, subjects and trials in various forms (e.g. time or frequency). The most important parameters in the analysis were listed and ways to use the CPD in practice were discussed.

In the second part we focussed on CPD applied to real mobile BCI datasets. CPD seems to be able to, with a fairly simple model, separate signal (i.e. P300) and noise in single trial ERP data. Even in small near real-time datasets meaningful estimates are derived without the need for specific training data. Future work is needed to construct the datatensors and CPD models such that the identity of the derived clusters can facilitate online BCI applications.

Chapter 5

Innovative usage of tensor decompositions for data-driven classification of auditory mobile BCI data

The results of the previous chapter on data-driven clustering of the P300, with CPDs of full datasets, motivates to incorporate additional structure as present in the experimental design. The work in the current chapter extends the CPD approach to allow for classification at a single-trial level. In Section 5.1 we adapt the CPD approach to allow for simple extraction of the differences in P300 ERP between target and non-target trials. This is achieved by incorporating additional structured elements of the paradigm; the BCI tensors are constructed by concatenating ERP templates from other subjects to a target and non-target trial. This approach has been described in [239]. Besides CPD we introduce another tensor decomposition, the LL1 which decomposes data in multi linear (L_r,L_r,1) terms that allows additional flexibility in modeling the data. We applied these methods to the auditory oddball data of Dataset A and the results were published in [243]. The presented approach leads to a fast and intuitive classification with accuracies competitive with a supervised and cross-validated LDA approach. In the second part of this chapter (Section 5.2) we extend the CPD tensor decompositions of single datasets (i.e. as described in the previous chapter) to allow multiple datasets to be linked. We simulated a large number of datasets containing target and non-target trials (mimicking an auditory oddball scenario) of various noise levels. This allows to investigate whether coupling of high- and low-noise datasets (in various ways) can enhance data-driven clustering of the P300. This work was published in [238].

5.1 Tensor-based single-trial classification of auditory mobile BCI without subject-specific calibration phases

The majority of studies on BCI rely on supervised machine-learning techniques to distinguish task-relevant stimuli [132]. Essentially, a substantial part of the data is discarded for model training, rather than interacting with the BCI, which comes at the cost of consuming users' time and effort. The use of calibration-free classifiers could increase the application potential significantly by removing this specific training phase.

As became evident in the previous chapter, the data-driven nature of tensor decompositions lends these methods to be applied for unsupervised classification [240]. One of the drawbacks of these findings was the lack of class identification in the trial mode. Furthermore, the P300 effect we wish to model may be not strictly rank 1, but slightly higher. LL1 decomposition was recently demonstrated to be superior over CPD for certain EEG applications as it allows to model more variability [97]. Therefore in the current section we will explore both CPD and LL1 decomposition for unsupervised BCI classification at a single-trial level. By constructing a tensor from the trials we want to classify, and relying on structural properties in the data, we demonstrate that CPD and LL1 decomposition are able to provide accurate labels. The method elegantly exploits the presence of a particular spatiotemporal pattern underlying the target trials which is absent in the non-targets. We add non-specific ERP templates to the two deviant stimuli epochs to obtain a tensor and a subsequent decomposition identifies discriminative features between target and non-target trials. The decompositions create a uniquely adapted (i.e. from the templates) spatiotemporal pattern to separate target and non-target trials that can vary over time and sessions. In this way we achieve classification results without the need for a subject-specific calibration phase.

In order to demonstrate the potential of the new method, we compare the results to regularized LDA (rLDA), as this has been shown to be amongst the most effective algorithms for P300 classification [73]. As a cross-subject trained classifier would also be able to derive instantaneous classification of a new subject (in contrast to the subject-specific LDA-trained classifiers), we also compare our results to a cross-subject trained rLDA classifier [40]. These results are based on the data from Dataset A.

5.1.1 Methods

Participants, stimuli and data acquisition

The most important facts of the auditory attention dataset are summarized here, for more detail we refer to [49]. The used dataset consists of 20 healthy subjects (10 females) with a mean age of 24.6 years. The paradigm was a three-class auditory oddball task, where the participants were requested to count the target tones while ignoring the other two tones. The standard tone was 900Hz while the two deviant tones were 600Hz and 1200Hz respectively. Ten subjects were instructed to pay attention to the 1200 Hz tone and the other ten to the 600 Hz tone. The stimuli were presented randomly and binaurally through headphones. The mean inter stimulus interval was 1000ms and incorporated an evenly distributed jitter in the range of 0ms to 375ms. Per session the subjects were exposed to 94 target deviants, 94 non-target deviants and 504 standard tones that were randomly presented. Each subject had to undergo tests in two different outdoor settings. One test was conducted while walking along a planned route at the Oldenburg campus. The other test was conducted in a silent corner of the university in which the volunteer and examiner were seated on two chairs. It is important to note that a significant amount of unwarranted distraction could have occurred during each pair of recordings due to regular activities taking place both in and off the campus. The stimulus delivery and experimental control was carried out using OpenViBE software running on a laptop [176].

The data acquisition was conducted using an original Emotiv amplifier (emotiv.com) connected to a modern infracerebral electrodecap comprising of 14 channels (easycap.de). Further details regarding the modified Emotiv EEG system are described in [57]. The signals of the amplifier were bandpass-filtered between 0.16Hz and 45Hz. and were sampled at 128Hz. Fourteen Ag/AgCl sintered electrodes were fixed at 10-20 positions: F3, Fz, C3, FPz, Tp10, Cz, O1, O2, F4, C4, TP9, Pz, P4 and P3. The common mode suppression for the electrodes (online reference) was at AFz and they were grounded at FCz.

Preprocessing

The data were preprocessed offline using EEGLAB [58] and MATLAB (Mathworks Inc., Natick, MA). Eye-blink artifacts were semi-automatically attenuated by means of extended infomax ICA [47, 222]. EEG data were 20Hz low-pass filtered, and epochs were extracted from -200ms to 800ms with respect to stimulus onset and baseline-corrected (-200ms to 0ms) after re-referencing to the mean of TP9 and TP10. In order to reduce the complexity of the data, we

down-sampled the data from 128Hz to 30Hz for this approach. This decreases computation time of the tensor models and is yet still expected to capture most of the P300 waveforms [240, 60]. Prior to the downsampling a 15Hz low-pass filter was applied to avoid alliasing artefacts.

Binary classification

The aim of the classification process is to classify single trial pairs consisting of one 600Hz trial and one 1200Hz trial. It is to be expected that when subjects attend to one tone, these attended tones elicit a P300 oddball ERP pattern which is absent for the unattended and baseline tones. The subject's binary choice of focusing his/her attention on one of the two stimuli makes it logical to approach the analysis as classifying these trial pairs rather than individual stimuli. A BCI example could be to pre-code (i.e connect) the stimuli types to different hearing aid processing modes, allowing the user to switch between background processing or foreground processing by attending to the proper tone. The attended and unattended tones are referred to as "target" and "non-target" trials respectively. The paradigm's stimulation codes can be used to know when these sounds are presented and the selective attention of the subject causes one tone to elicit a P300 response. Therefore we approach the problem as distinguishing target and non-target trials in a set of 2 (i.e. trial-pair) rather than making a decision for individual trials as is traditionally done. This available paradigm information was not incorporated in [49].

A) Construction of the data tensor
The interpretability of the components extracted from a tensor decomposition depend on the construction of the data tensor prior to the decomposition. A window of 167ms − 633ms after stimulus onset (SO) is determined as the most discriminative window for the P300 signal. The windowed data is normalized by converting the time course to z-scores for every channel.

Tensor decomposition techniques are able to identify the most structured signature in a multidimensional data tensor. In order to enhance the likelihood of extracting a task-related signature, an average target and baseline ERP are added to the trial-pair data (i.e. one regular "target" and "non-target" trial). These ERP templates are the average ERP of all other subjects for the baseline and target stimuli. An example of these templates are depicted in Figure 5.1 for the baseline, non-target and target templates from top to bottom. Moreover, this inclusion of templates simplifies the identification of the class of each trial from the decomposed factors; which would be otherwise a non-trivial task [240]. Since some subjects might sometimes incorrectly focus to the non-targets, the baseline stimuli provide a cleaner non-target average template as compared to averaging all non-target trials. Taken together, for each trial pair we obtain a

$12 \times 15 \times 4$ data tensor in which the last mode consists of two trials (Target, non-Target of which we do not know the identity) and two templates (target, baseline). All analysis steps regarding the construction of the data tensor are conceptually visualized in Figure 5.2.

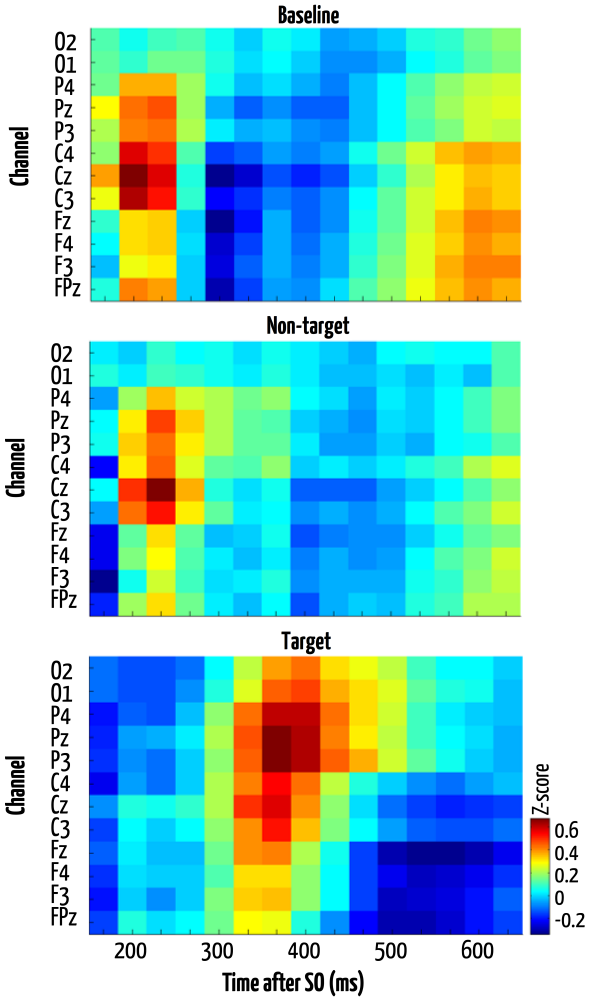


Figure 5.1: Grand-Average (n-1) heatmap of the spatio-temporal pattern for each of the stimulus types Baseline, non-Targets and Targets, from top to bottom respectively.

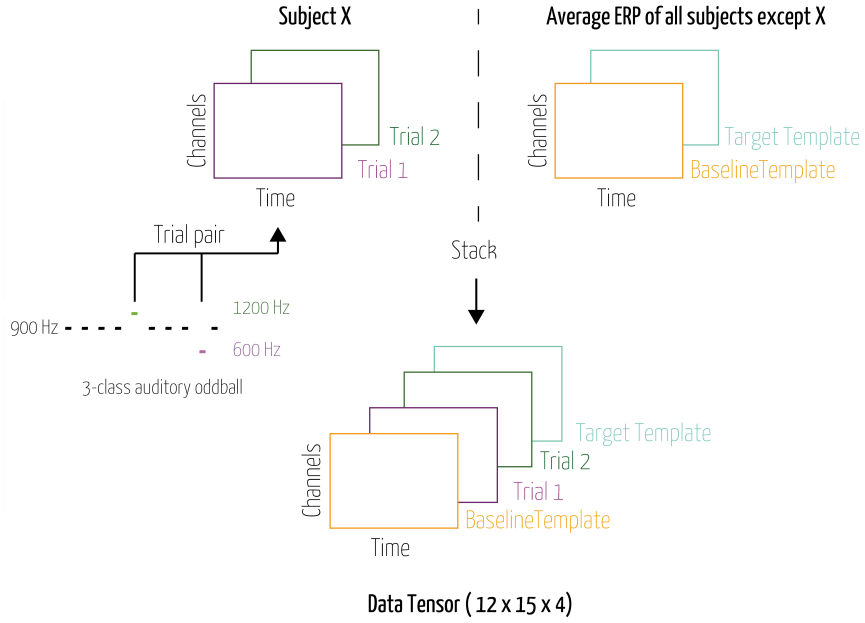


Figure 5.2: Overview of the tensor construction. Every target (attended tone) and non-target (unattended tone) trial are concatenated with average ERP templates from other subjects of the target and baseline stimuli. In the end per binary command we constructed a 12 channels x 15 time points x 4 trials/templates tensor.

CPD The higher-order data tensors are decomposed using CPD which has been introduced in the previous chapter (section 4.1.3).

LL1 Although CPD provides interpretable components, the model can be too restrictive for some applications as it does not model all variability in the data [97]. A LL1 decomposition allows to model more variation in two modes. The LL1 decomposition approximates a third-order tensor by a sum of R terms, each of which is an outer product of a rank-L_r matrix and a nonzero vector [195, 45]. A third-order data tensor X can be decomposed by a LL1 as:

$$X = \sum_{r=1}^{R} (A_r \cdot B_r^T) \circ c_r + \varepsilon \tag{5.1}$$

The tensor is the sum of the outer products of a rank L_r matrix (the product of matrices A_r and B_r-transposed) and the component vector c_r, with R

representing the number of components and ε the model error. Similarly to the rank, L_r should be set a priori. We aim to model additional variance with the LL1 decomposition as opposed to the CPD. However, the exact mixture of the data tensor is unknown. The choice of L_r is therefore rather conceptual, depending on a priori knowledge about the task and preprocessing.

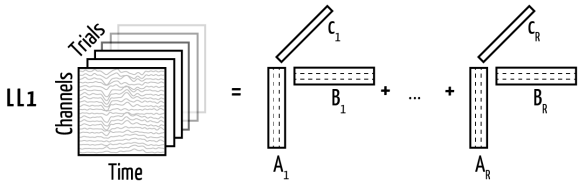


Figure 5.3: Schematic diagram of the LL1 decomposition.

Since our data-tensors are rather small (i.e. $12 \times 15 \times 4$), values of 2 to 5 for L_r were explored. This is expected to capture time and waveform variability more accurately between the target and non-target effects as constituted in the trials and templates. Similar to the CPD we utilize the NLS algorithm within the Tensorlab for the LL1 decomposition.

Both the CPD and LL1 models are extracted to retrieve a single component ($R = 1$) and are initialized randomly. Figure 5.4 illustrates the CPD and LL1 models in a single data tensor. In the case of a CPD decomposition it can be seen that the same waveform is linearly scaled over the channels (i.e rank-1), while in case of a LL1 solution the L_r 2 (or higher) spatiotemporal pattern allows more variation on the waveform on the different channels.

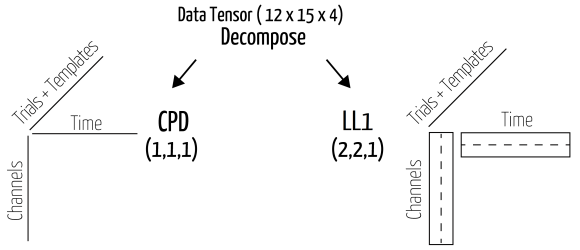


Figure 5.4: Overview of the tensor construction. Every target (attended tone) and non-target (unattended tone) trial are concatenated with average ERP templates from other subjects of the target and baseline stimuli. In the end per binary command we constructed a 12 channels x 15 time points x 4 trials/templates tensor.

B) Interpretation of the decomposition
Figure 5.5 illustrates a CPD and LL1 decomposition for a single trial pair constructed from real data. The modes depict the extracted spatial and temporal waveform together with the extracted weights in the third dimension. To classify a single trial pair we can utilize these weights as follows. A target trial is expected to differ most from the baseline template and differ the least from the target template; vice versa for a non-target trial. For every single-trial-pair augmented with the two templates a CPD and LL1 model is obtained. The third factor represents the trial weights accredited to the templates and the 2 unknown trials. In order to classify which trial is the target and non-target trial we calculate per trial in the third factor the absolute difference between the trial weight and the baseline template weight and from this we subtract the absolute difference of the trial weight and the target template weight. The trial with the largest obtained value is considered the target trial, the smallest the non-target.

Figure 5.5A illustrates the spatiotemporal pattern and trial mode of an extracted CPD component. In the trial mode the distance of the trials to the baseline and target template indicate that trial 2 is considered the target trial and 1 the non-target; the latter is closer to the baseline template (BT) weight and the former to the target template (TT) weight. However, the difference between trial 2 and trial 1 is marginal. Similarly we can decompose the data-tensor with LL1, for example a (2, 2, 1)-LL1 decomposition to derive similar estimates and derive class labels from the third mode estimates. It should be noted that the individual spatial and temporal signatures in the LL1 model are not unique in the extracted form [195, 45]; linear combinations of these solutions are equally plausible outcomes. In order to obtain the true unique spatiotemporal signature of the component we need to construct the combined spatiotemporal subspace (i.e. multiplying A with B^T in equation (5.1)). The result is visualized in Figure 5.5B. In this matter the LL1 models extract a more fine-tuned signature, namely a Parietal – Occipital shift of activity. This results in a similar correct classification of the trial pair (Figure 5.5B), however the better model also enlarges the distinction between trial 1 and trial 2 (c.q. Figure 5.5A, B the third mode).

Evaluating the CPD/LL1

Classification accuracy is an important factor to evaluate the CPD/LL1 classifiers. However, another important aspect is to understand subject-wise differences with the newly proposed method. Since the CPD/LL1 classification process exploits the morphologic structure guided by the ERP templates and the P300 is known to vary in both amplitude and latency between subjects [7], we hypothesize that the results might depend on the subjects' P300 morphology.

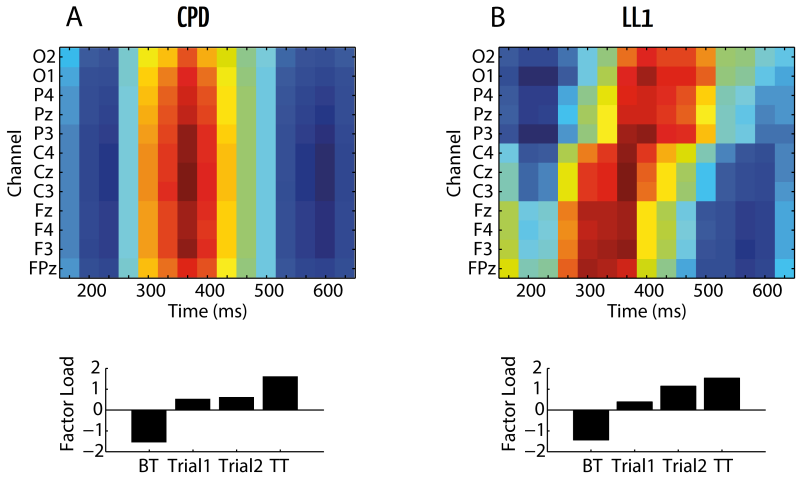


Figure 5.5: Examples of decomposing a single trial pair tensor with (A) CPD, Illustrating the factor loadings in the spatiotemporal matrix and trial/template modes from top to bottom respectively. (B) LL1 decomposition with L=3,the obtained LL1 spatiotemporal matrix is derived from multiplying the first two modes of the LL1 estimates (not shown). The LL1 model estimates 2 distinct sources whereas the CPD estimates an average target-non-target effect. The trial/template estimates in the last mode will lead to classifying trial2 as the target trial since its value is close to the target template (TT) and further away for the baseline template (BT) as opposed to trial1. Note, in this particular example the subject focused on trial2. The LL1 model separates the target and non-target signature more accurately which is evident from a larger difference between trial1 and trial2. The opposite factor loadings for target and baseline are coherent with the ERP patterns of these stimuli c.f. Figure 5.1. Note: the spatiotemporal pattern as shown in (B) differs reasonably between trial-pairs.

In order to investigate the influence of the added template ERPs, we define features that we correlate with the retrieved classification accuracies.

- Terp-Btemp = difference between: targetERP and the Baseline template (i.e. BT).
- Terp-Ttemp = difference between: targetERP and the Target template (i.e. TT).

The targetERP is defined as a matrix with the averages across all target trials per channel. The difference is calculated as the summed absolute difference

between the targetERP and template matrix. It should be noted that these ERP features are derived offline after the classification process for evaluation purposes only.

Linear Discriminant Analysis

In order to compare the results from the newly proposed method, we also compute supervised classification accuracies with different LDA variants as described here. Similarly to the CPD/LL1 method we classify single-trial-pairs allowing the classifier to discriminate between the target and non-target, rather than classify the individual responses. Traditionally this paradigm information is not included in the classifier function, therefore the LDA was adapted accordingly to allow for fair comparison between CPD/LL1 and the LDA. The included 'pair' information does not incorporate knowledge about which class (Target-NonTarget) the trial belongs to and is therefore a valuable addition to the classification of three-class oddball data.

Offline Subject Dependent
The basic LDA feature set comprised seventeen 47ms data bins on all twelve electrodes between 0−800ms. Shrinkage regularization as implemented in BCILAB [59, 23] is used for rLDA classification. Per subject the classifiers are trained based on five-fold cross-validation procedure.

Pseudo Online Subject Dependent
One way to evaluate the performance of rLDA without long calibration is to (re)train the classifiers for every single trial pair, in which all past trials of that subject serve as training input for evaluating the next trial under investigation. The accuracies reported at every trial pair are calculated as the number of correctly classified trials from that moment in time, until the end (i.e. of the session) with a model that is trained on the history of trials up to that point. It is expected that the performance will increase in time as the number of training data increases. In contrast, the CPD and LL1 decomposition are not expected to show such an increase as the method does not require an explicit subject-specific calibration phase.

LDA-based Cross-Subject Classification
As it is also possible to construct an LDA classifier on data from other subjects, we compare the performance of the new method to rLDA trained on cross-subject data. The feature formation and selection steps are identical to the steps described in the previous paragraph, with the exception that now, the model is trained on data from other non-test subjects. We compute accuracies as a function of the potential number of training subjects available, from 1-19. For all divisions of these 19 different scenarios the LDA's were trained and

tested on the remaining independent subject(s). In the case of training 1-18 subjects, the index of subjects chosen for training was selected at random from the 19 non-test subjects. For example, if we train on five subjects, many sets of five can be drawn out of the total 19 subjects that can be used for training. Therefore this process was repeated an arbitrarily chosen 25 times and the average accuracies were reported. As the CPD and LL1 performance is also influenced by the morphology of the templates added to the model, these used template averages are in turn also varied from 1-19 subjects in exactly the same way as for the LDA cross-subject results.

5.1.2 Results

The presented LL1 decomposition results are obtained with $L = 2$. However, for all considered values of L the LL1 decomposition results did not differ significantly in any condition.

Offline Comparison

Average classification accuracies per subject are presented in Figure 5.6. The grand average accuracies (and SD) are 74.4 (9.7), 71.7 (8.6) and 74.0% (8.5) in the seated condition for rLDA, CPD and LL1 respectively. The LL1 results were significantly higher compared to the CPD ($t_{19} = 3.34$, $p < 0.01$) and the rLDA difference to CPD was not significant ($t_{19} = 1.44$, $p = 0.17$). In the walking condition, grand average accuracies (and SD) were 66.4 (8.7) 64.8 (10.3) and 65.9% (9.9) for rLDA, CPD, LL1 respectively. It should be noted that although the means are very similar across methods, differences in LL1 and rLDA performance between subjects can be substantial (e.g. the subject indicated with a circle in Figure 5.6). When only including the Baseline Template, LL1 accuracy dropped 1.1 and 1.6 percentage points for the seated and walking condition, respectively.

In both conditions the CPD and LL1 decomposition estimates displayed a strong correlation in the seated (Pearson's $r = 0.93$, $p < 0.001$) and walking condition ($r = 0.97$, $p < 0.001$). The correlation between LL1 and rLDA is moderate for the seated condition ($r = 0.58$, $p < 0.05$) and strong in the walking condition ($r = 0.81$, $p < 0.001$). In order to understand how basic ERP features could predict classification performance, we computed the correlation between some Target P300 ERP features (described in section 2.3.5) and LL1 classification accuracies. The Terp-Btemp and Terp-Ttemp correlations with the LL1 accuracies were (Pearson's r=0.45, p=0.09) and ($r = -0.51$, $p < 0.05$) in the seated condition and ($r = 0.56$, $p < 0.05$) and (Pearson's $r = 0. - 67$, $p < 0.01$) for the walking condition. Because of this dependence on morphologic

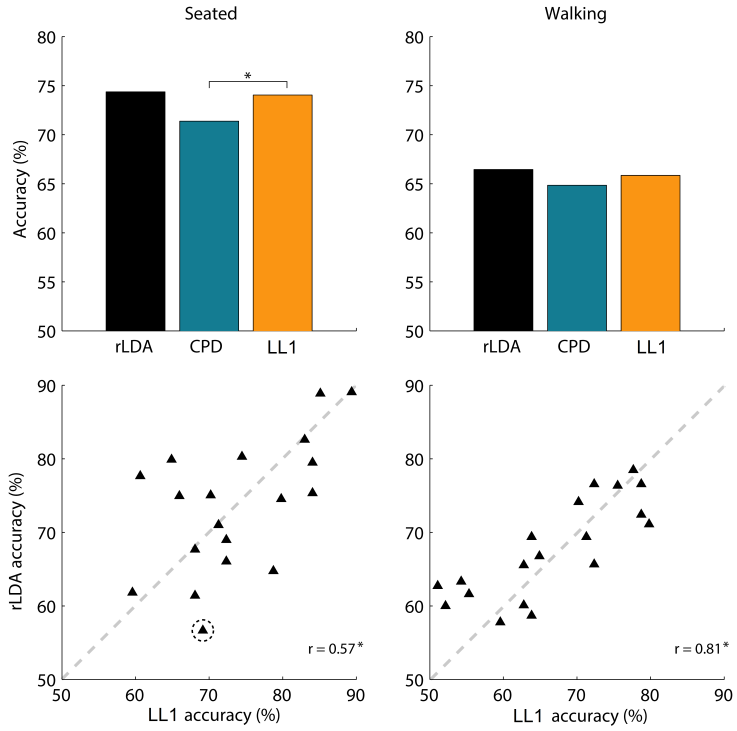


Figure 5.6: Top: grand-average accuracies. Bottom: Relation between LL1 and rLDA with Pearson's correlation coefficient. Respectively in the seated condition (left) and walking (right). Significant differences and correlations with p<0.05 are indicated by an asterisk.

features, we wondered if we could use these features to predict which of the two best performing methods, rLDA or LL1, would provide the best classification. The Terp-Ttemp feature is indeed mildly correlated with the difference of rLDA-LL1 accuracy in both the seated and walking conditions (Pearson's $r = 0.18$ and 0.53, respectively, $p = 0.43$ and $p < 0.05$). This implies that if the subject's average target ERPs are similar to the Target template, LL1 would perform better than rLDA. Post-hoc analysis showed that an arbitrary threshold can be defined on the Terp-Ttemp in order to indicate which method (rLDA or LL1) would perform best per subject. For the seated condition we could define two subjects and in the walking condition nine subjects for which the Target ERP differed substantially from the Target Template resulting in a higher rLDA accuracy as compared to LL1. Substituting the LL1 results of those subjects by their rLDA counterparts would lead to a mildly improved grand average

accuracy of +1.2 and +1.7% for the seated and walking conditions respectively. These results support the notion that morphology is an important factor in the classification procedure.

Pseudo Online Scenario

The previously mentioned rLDA results are based on training the classifiers on 80% of the subject specific data. In order to use this method online, in the best case only the past trials (and label information) at moment T can be used for classifier training whereas information from other subjects can be included in the CPD/LL1 approach without subject-specific data. First we present the results of the rLDA, CPD and LL1 method in such an online scenario. Secondly we present the results of the cross-subject trained rLDA, meaning that the classifier is trained on data from non-test-subjects.

Subject Dependent rLDA vs CPD/LL1

Our CPD and LL1 allow the classification process to start instantaneously. Evidently the rLDA needs sufficient training data in order to derive reliable estimates. Figure 5.7 displays this relation between training data for the LDAs and corresponding CPD/LL1 accuracies at similar sections of the data. The LL1 estimates outperform the rLDA estimates for the first 29% (shaded area) of the trials (p-values$<$0.05). In the walking condition similar results can be observed albeit more spread over time. For the first quarter of the trials the LL1 accuracy is significantly ($p < 0.05$) higher or marginally ($p < 0.1$) higher compared to rLDA.

Cross-Subject LDAs vs CPD/LL1

Consecutively we compare the results of the CPD/LL1 to that of an rLDA classifier trained on data from other subjects. Figure 5.8 shows the grand average accuracies depending on the number of available subjects for model estimation. It can be seen that the CPD/LL1 methods reach their optimal performance when 3 or 4 subjects are available, whereas the LDA based classifiers need data from more subjects to achieve similar accuracy to CPD and requires even more data when compared to LL1. The shaded area shows the significantly ($p < 0.05$) higher LL1 accuracy as compared to the rLDA in both conditions.

Finally, it is interesting to note that a cross-subject LDA trained on data from 10 subjects provides competitive performance to that of subject-specific LDA (c.q. Figure 5.8 and Figure 5.6).

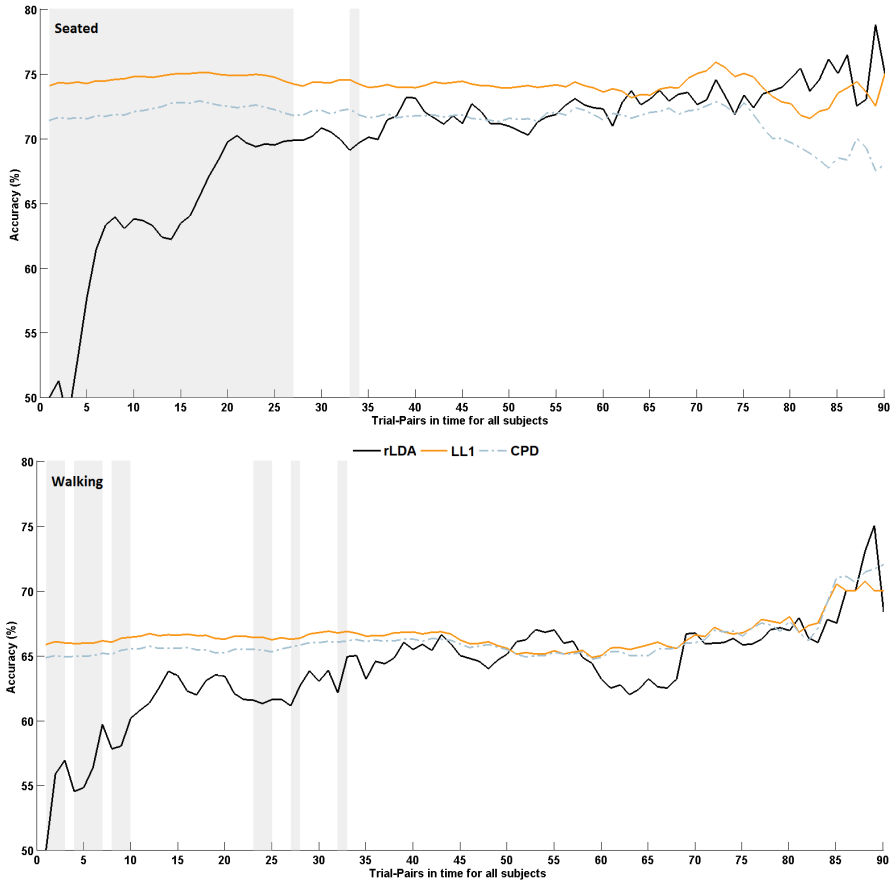


Figure 5.7: Average classification accuracy across subjects as a function of the available trials for training. Each point represents the average classification accuracy of all following trial-pairs. rLDAs and swLDA estimates are compared with the CPD and LL1 classification accuracies. Top: Seated condition. Bottom: Walking condition. The shaded areas mark significant differences between LL1 and rLDA at the 0.05 significance level.

5.1.3 Discussion

Recently it was shown that it is feasible to perform a binary auditory BCI classification in a real life environment [49]. However, the need for training supervised LDA classifiers on large parts of that data at the start of a session does not encourage applications outside experimental studies. In the current

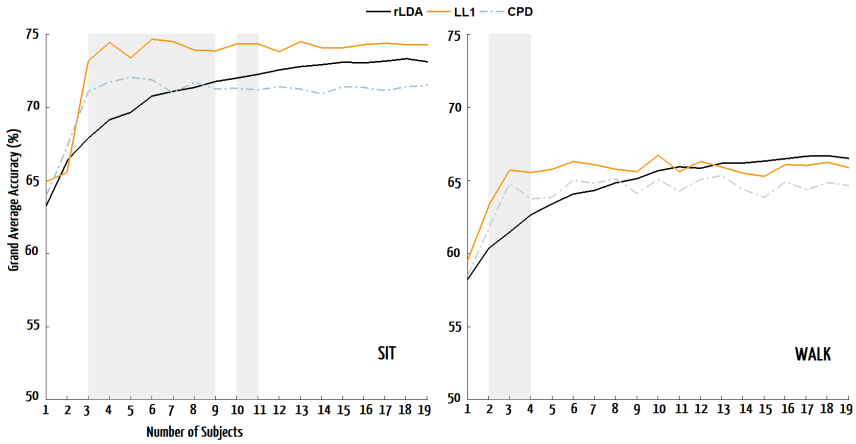


Figure 5.8: Grand average accuracy as function of the non-specific subject used for training (LDA) or estimating templates (CPD/LL1) for the seated and walking condition left and right respectively. The shaded areas mark significant differences between LL1 and rLDA at the 0.05 significance level.

section we showed that it is possible to remove that explicit subject-dependent calibration phase with a tensor-based decomposition (CPD/LL1) augmented with non-subject-specific templates without sacrifying classification accuracy as compared to the reference LDA approach. This allows for instantaneous classification results that on average are similar to those of the subject-specific trained models. This allows faster interaction with the BCI and is likely to increase the user interest and engagement. Interestingly, we also demonstrate that specific subject-related ERP features are predictive of the LL1 results. This direct relation between ERP characteristics and the accuracy of the LL1 models makes the classifier more easy to interpret, compared to LDA methods that do not select structured features and rely on additional post-processing steps for interpretation (e.g. [94]). The LL1 method presented here has a forthright link to the original P300 ERP signal.

We assumed that the differences between the target and non-target trials were fixed between subjects to some extent. Moreover, our method diminishes a certain amount of subject specific information by adding the average ERP templates to the trials prior to the low rank approximation by CPD/LL1. However, the obtained approximation is an efficient combination of the data and templates to classify the single trial pairs. Transferring additional information from subject to subject or session to session might increase the performance. For example, we noticed a small overall increase in accuracy in the walking condition

if the data tensors were constructed with templates from the seated condition. In addition the CPD/LL1 might improve with an ability to recognize uncommon ERP morphologies and subsequently update the templates. A similar approach in which the templates incorporate past subject specific data has been proposed in an Riemannian framework in [40]. Furthermore, a comparison to several prototypical responses instead of the 1-2 fixed templates as described currently might be a valuable line of future research.

We argue that the CPD/LL1 methods might not lead to the highest accuracy possible, but do lead to interesting insights regarding the EEG BCI signals at hand. For example in a feedback paradigm where the aim is to train a robust P300 waveform by the users, traditional LDA analysis might derive features that are not inherent to the P300 morphology as the proposed method.

The LDAs feature selection step selects very distributed features in time and space. Indeed as one expects the most chosen features are at central posterior electrodes in the 250-500ms range, however a substantial number of features is selected towards the extremes of the 0-800ms time window (not shown here). To investigate the effect of these features we limited the LDA feature selection to the same time-window as the CPD/LL1 and this did not lead to a significant difference in accuracies (results not shown here).

An important question is to what extent we can generalize our findings. First, the three class oddball paradigms allows for a clean non-target estimate based on the baseline stimuli; something that is not present in the other P300 paradigms (i.e. the frequently used P300-speller). Moreover, the results between the two conditions differ. The LL1 and rLDA differences are larger in the seated condition as compared to the walking condition. However, all methods performed significantly lower in the walking condition as compared to the seated one (Note: these significances are not indicated in Figure 5.6). Consecutively, the CPD/LL1 method distinctly needs a clear Target and non-Target template. Similar results with CPD were obtained with only the inclusion of a Baseline template in [239]. It is an open question if such reliable templates can be constructed for other -faster BCI experiments (e.g. speller) where the brain response to stimulation overlaps among stimulus presentations. Nevertheless, this is an important step towards constructing intuitive classification methods by exploiting data signatures from other subjects and structural paradigm information. Such approach has also proven useful and has been advocated by others [106, 14]. We encourage that BCI users really start using classifiers constructed on data from other subjects to maximize time where the BCI is actually used rather than spending significant amounts of time recording training data.

The CPD and LL1 performances are shown to be closely related.This was also evident from their high correlation described in the previous section. In

general the LL1 values outperform the CPD results in the seated condition but not in the walking condition. The CPD model approximates the data tensors as a rank-1 component which is shown to be too restrictive to achieve a discriminative target/non-target signature. The LL1 is better at modeling the related differences as is evident in the obtained accuracies and interpretation of the factors (c.q. Figure 5.5A, 5.5B). Similar results are obtained if L was increased to 5. This would suggest that the constructed tensors have prominent target and non-target differences that are already captured with a low rank model of L=2. It can be noted that the rLDA and LL1 values correlate strongly in the walking condition but only moderately in the seated. Conversely, the correlations between the ERP features and the rLDA-LL1 difference displayed the highest correlation (r>0.5) in the walking condition indicating that the LL1 method performs better than rLDA in case of highly prototypical ERPs.

Finally, besides the achievement of instantaneous classification, the cross-subject LDAs comparison to LL1 shows that fewer subjects are needed for LL1 in order to reach adequate classification. Even though substantial differences are observed between the two methods between subjects, only mild overall improvement can be achieved. It is striking that two completely distinctive methods are not able to improve the results excessively in either condition. This strongly suggests that the limiting factor in this case for the BCI is the lack of task related ERPs. Therefore, future work should focus also on the understanding of the fluctuations in brain responses on single trials and improving BCI paradigms to elicit strong responses rather than merely improving classifiers (e.g. [52, 156, 202]). This can lead to better results with an adaptive BCI approach, especially in real-life scenarios in which the distraction levels are high [123].

5.2 Decomposition of multiple datasets using coupled CPD

In the previous chapter we evaluated CPD models to derive meaningful P300 ERP-related components in a data-driven way from full datasets. In the current section we will explore the benefits of imposing a coupling between datasets as can be expressed by using one or more common factors in the different factorizations of the tensor models. This way it enables a priori knowledge of the data (e.g., noise levels, Expected ERP) and stimulus protocol to be incorporated in the analysis via constraints in the CPD model [221, 197]. The CPD models, coupled CPD models and the influence of noise on the derived clustering of the P300 are evaluated on simulated EEG data (i.e., 15 datasets of 200 trials) of four different noise levels (SNR-range: 3.73 – 0.60) which mimic realistic settings; for example, the auditory P300 in recent mobile EEG recordings [49]. The tensor decomposition results were compared to that of the widely utilized supervised rLDA classification (e.g. in [239, 49]). In the next section we explain the simulated data generation and characteristics of the signal and noise. Consecutively the preprocessing steps performed are elaborated, followed by a description of the (coupled) tensor models that are considered in the analysis. Finally, the results of the tensor models and reference method are presented and discussed.

5.2.1 Data Generation And Preprocessing

The simulated data were generated using the BESA simulator (freely available online: `www.besa.de/products/besa-simulator/`). This simulator utilizes a spherical four-shell head model to generate EEG data based on predefined dipoles [19]. The generators of the P300 event-related potential were obtained from the BESA software. Five dipoles were used for the generation of P300 ERPs as target trials, which are depicted in Figure 5.9. The two dipoles with a frontal focus correspond primarily to the N100 effect, a central dipole the P3a effects and finally two posterior dipoles that represent a P3b-like deflection around 450 ms after stimulus presentation. Consecutively, the two dipoles corresponding to the P300 (i.e., dipole 4 and 5 in Figure 5.9) were diminished in amplitude to obtain a second set of EEG data without explicit P300 effects, hereafter referred to as the non-target Trials. Differentiating between the target and non-target trials is the primary aim in BCI approaches, and the simulated data resembles a three-class auditory oddball paradigm that is used in other chapters of this thesis and congruent with other studies (e.g., [128, 49, 91]).

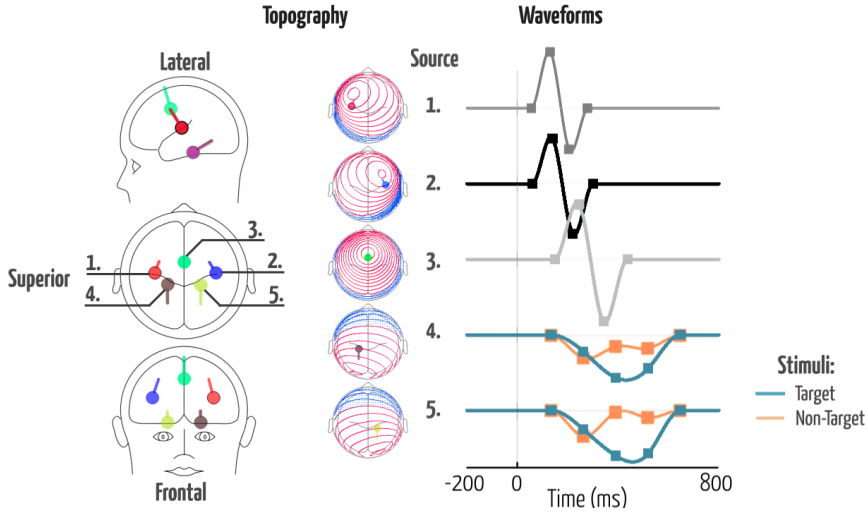


Figure 5.9: Overview of the dipole spatial locations (left) and temporal characteristics of each dipole on the right. The blue lines depict the temporal progression in the Target trials and the orange for the Non-Target trials. Image obtained from the BESA simulator (available online at `www.besa.de`) and modified to highlight the P300 source.

The EEG data were simulated at 24 channels corresponding to standard 10-20 locations: FP1, FP2, Fz, F7, F8, FC1, FC2, Cz, C3, C4, T7, T8, CPz, CP1, CP2, CP5, CP6, TP9, TP10, Pz, P3, P4, O1 and O2. The data were generated with a sampling rate of 500Hz. The topography and sampling rate are identical to that of recently published mobile EEG studies as presented in Chapter 6 and 7. To obtain realistic data, several types of noise were added to the model. First, the amplitude of the sources is modulated randomly to differ up to 50% in between trials. Second, EEG-like broadband noise (i.e., as implemented in the BESA simulator) was added to the dataset to simulate background activity unrelated to the simulated ERP events. The noise had a relative high correlation between signal amplitudes from electrodes that were close together. Noise-free data were generated for reference purposes, and in addition four different proportions of noise were added to the ERP effects. The noise amplitude is scaled by its root mean square (RMS) value to match noise values of 0.5, 1.0, 1.5 and 2.0 µV and alpha activity at 8-12Hz was generated and added to the signals with an amplitude that was 50% of the overall noise RMS value per trial. The four noise levels corresponded to SNR values of 3.73, 1.75, 0.97, 0.60 from high to low, respectively. The signal-to-noise ratio (SNR) was calculated following previously described procedures in the literature as the ratio of the maximum

P300 peak and the pre-stimulus RMS of the EEG. We simulated 15 datasets of 100 target and 100 non-target Trials for each of the four noise levels and the noise-free case. Differences between subjects was imposed by random variation in amplitude of the sources and random generation of noise and alpha-band activity. One trial lasted 1000ms in which the stimulus onset was simulated at 200ms. The data were preprocessed offline using EEGLAB and MATLAB (Mathworks Inc., Natick, MA). The EEG data were 0.5-20 Hz band-pass filtered and baseline-corrected (-200-0ms) after re-referencing to the mean of TP9 and TP10. To eliminate the effect of outliers in the data, we performed a z-score normalization per channel and trial. This way, the overall variance is more uniform before analyzing the data with the tensor models. The average ERPs at channel Pz for the target and non-target stimuli are presented in Figure 5.10 for each of the noise levels. The corresponding SNR values illustrate that the last two noise levels portray situations in which the noise is equal to or higher than the signal of interest (i.e., P300), mimicking realistic mobile EEG P300 scenarios [49].

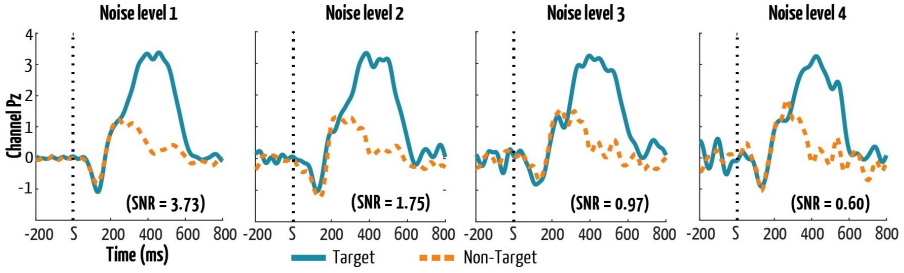


Figure 5.10: Average ERPs at Channel Pz for the four increasing levels of noise from left to right respectively.

5.2.2 Tensor Based Models

Single Dataset Decompositions

Multidimensional signals can be decomposed by the CPD as a sum of rank-1 terms [93] as was shown in the previous sections. The size of the data tensor per subject is $24 \times 500 \times 200$ for the channels, time and trial dimension, respectively. CPD is expected to derive meaningful signatures that separate the P300 signal from other (nuisance) sources. In the trial dimension we aim to separate the Target from NonTarget trials.

Fixed-Factor Decompositions

Tensor decompositions can be applied in a supervised way by transferring factors between models. For example, a spatiotemporal signature from one CPD model from a low noise dataset can be chosen and imposed on the decomposition of a high noise dataset (illustrated in Figure 5.11). This way, we aim to achieve a better clustering of the target trials in contrast to the non-targets in the high noise conditions. This so-called fixed-factor decomposition requires the datasets to be decomposed consecutively and relies on identifying the most useful component in the first decomposition that is to be transferred. This concept is illustrated in the schematic below. Note that the top decomposition is evaluated first and the best temporal-spatial pattern is transferred to the second model. The size of each data-tensor is $24 \times 500 \times 200$ (channel, time and trial), and each is decomposed independently of the other. Due to the transfer of the most discriminative component this decomposition is supervised and serves merely as an illustration of other (supervised) model types that can be explored for BCI.

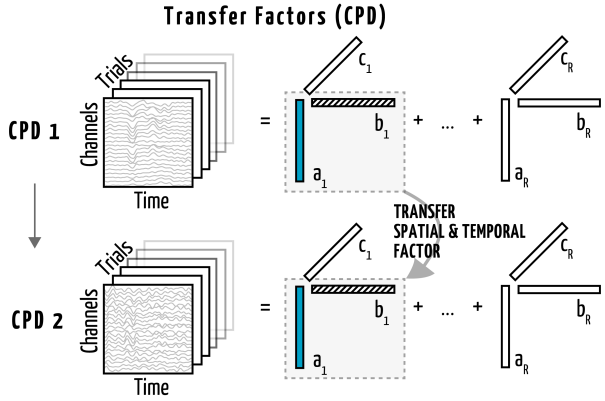


Figure 5.11: Illustration of a transfer of a single component's spatial and temporal factor (CPD1) onto one component of a novel dataset and CPD model (CDP2).

Coupled Decompositions

Instead of decomposing a single dataset, CPD can be used to express coupling between datasets as depicted in Figure 5.12. A so-called coupled decomposition of two (or more) datasets can be achieved in a framework known as Structured Data Fusion (SDF) which is a specific language to impose structure on factors in

the tensor decomposition [221, 197]; see [199, 198] for the theoretical background of coupled CPD. For example, we can impose equality between the spatial factors of two decomposed datasets of different SNR level. This is illustrated for the CPD case in the diagram below, indicating equality between both the spatial and temporal mode in the two datasets. In the current work, we studied the effectiveness on the P300 separation when the highest noise condition was coupled with a lower-noise-level dataset. This is based on the assumption that the underlying source model for the P300 in both datasets is (roughly) similar and only the variability over trials (over the course of the experiment) is different. For CPD, we evaluated the clustering on the lowest SNR dataset when jointly decomposed with higher SNR datasets. The decomposition of both datasets happens simultaneously and without any additional stimulus information. The size of the coupled model is comprised of 2 data tensors of dimensions [24 $\times$ 500 $\times$ 200] for the channel, time and trial dimension, respectively. Note that in this case all modes but one are shared and all components are shared; alternatively this can be achieved trough stacking the two tensors, i.e. concatenation along the trial mode and apply CPD on this tensor. The number of shared modes and components is a rather conceptual choice and depends on the expected effects in the data. If more than one mode is not shared or only a single (or a few, but not all) components are selected to be shared, stacking (i.e. concatenation into a single tensor) is not possible and coupled decompositions should be considered.

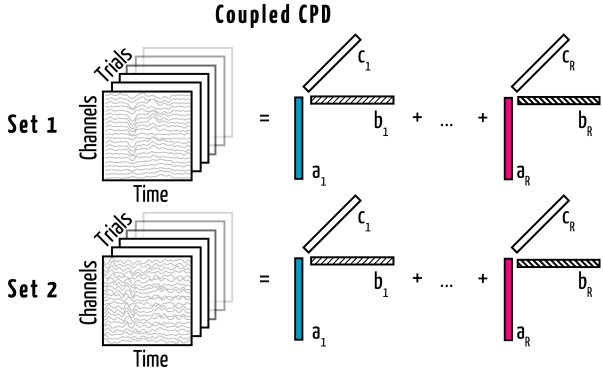


Figure 5.12: Illustration of a coupled CPD of two datasets in which the spatial and temporal modes are shared within each component.

Simultaneous Recordings (multi-user BCI)

Besides coupling or fixing factors of the spatial or temporal modes, one could opt for coupling the trial mode of two or more datasets as shown in Figure 5.13.

This could be useful for a specific type of paradigm that is gaining momentum in the past years in the field of EEG research and is comprised of simultaneous recordings of multiple subjects. One example of such paradigm is the control of a BCI Space Invaders game by two subjects [118] and another, the simultaneous analysis of EEG in response to video in a classroom environment [171]. Since these subjects may have different spatiotemporal patterns, the spatial and temporal modes are left unconstrained. However, the trial (i.e., time) dimension is assumed to be identical, as the subjects are doing the same task at the same time. The schematic below illustrates the constrained CPD model in which the third (trial) mode is shared between datasets. This way the CPD model derives the shared information between two or more of such datasets on the trial factor. The additional structure in the trial factor is employed to increase the overall SNR. The performance of this coupling is evaluated in the current study by combining up to five datasets of the highest noise scenario (i.e., SNR = 0.60) to increase the overall distinction between target and non-target trials. The size of data that is used in the decompositions is 1 to 5 times a single subject dataset of dimensions [24 *times* 500 *times* 200] that represent the channel, time and trial dimension, respectively.

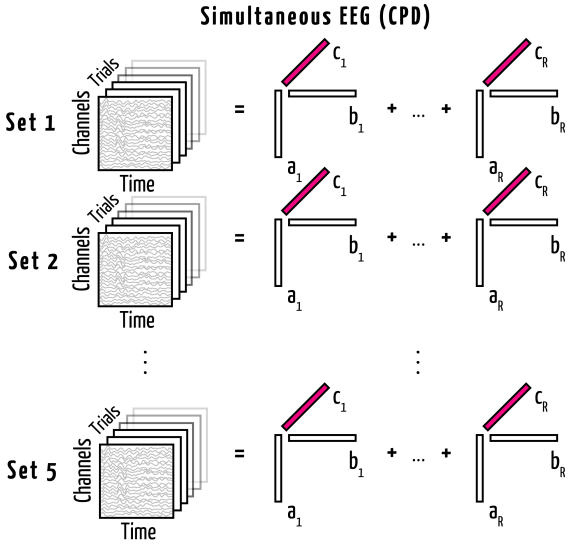


Figure 5.13: Joint analysis of multiple EEG datasets in which the trial mode is shared, for example in a case where multiple subjects execute the same task.

Decomposition Parameters

In this study, all CPD models were computed with the nonlinear least squares (NLS) algorithm in the publicly available Tensorlab 3.0 toolbox [221]. These methods are dependent on several parameters of which the following are evaluated: the initialization of the model, the number of iterations of the NLS algorithm, and the number of components (i.e., Rank R). Unless stated otherwise, all models were initialized randomly. By default, all computations were allowed up to 1000 iterations for the NLS. Post hoc evaluation of the number of iterations did not improve the results and these results were therefore not included here. The Rank of the models was evaluated explicitly for Ranks of 1 to 10 to explore the influence of R on the outcome of the models, without aiming specifically for the best approximation (i.e. optimal R).

Clustering and Classification

The factor loading on the trial dimension can be used to naturally obtain two clusters (e.g., Figure 4.2, Component 2). Accuracies were obtained by taking the median value of the trial factor as threshold for each component in the decomposition to obtain two clusters. The resulting classes are compared to the true labels, and this results in a clustering percentage for each component. This method only allows for separation of the trials into two classes; it does not identify which class corresponds to the target or distractor stimuli. Nevertheless, identifying the stimulus may be derived from the corresponding spatiotemporal signature in the other modes. Evaluation of the Tensor-based clustering is achieved by a comparison to those of rLDA, which is one of the widely used P300 classification algorithms (e.g., [49, 242]). This requires a separate training phase to calibrate the classifier function. Essentially a substantial part of the data is discarded for model training, rather than interacting with the BCI, which comes at the cost of consuming users' time and effort [132]. The basic LDA feature set comprised seventeen 47ms data bins on all 12 electrodes between 0–800 ms, which is reported repeatedly as an efficient feature set for classifying the P300 [49, 242]. Shrinkage regularization as implemented in BCILAB is used for rLDA classification. Per subject, the classifiers are trained based on five-fold cross-validation procedure.

Table 5.1: Average accuracies of the rLDA, CPD and LL1 method for discriminating between target and non-target trials for each of the five different levels of noise. Values with an asterix were significantly lower as compared to rLDA ($p < 0.001$).

	Average Clustering % (± SD)				
	Noise-Free	Noise 1 SNR=3.7	Noise 2 SNR=1.8	Noise 3 SNR=1.0	Noise 4 SNR=0.6
rLDA	100.0 (±0)	99.0 (±0.9)	91.8 (±1.3)	85.6 (±3.0)	78.9 (±2.9)
CPD	100.0 (±0)	99.6 (±0.5)	92.6 (±2.0)	85.0 (±3.0)	71.0 (±7.1)*

5.2.3 Results

Single Dataset Models

The grand average classification accuracies are summarized in Table 5.1. For the conditions 'Noise-Free' and 'Noise 1' (SNR = 3.73), all models performed near 100% for separating the target from non-target trials. In the case of an SNR of 1.8, the clustering was on average correct for 92% of the trials, for all methods. The CPD results were on par with the rLDA for the noise 3 case, even though they were lower in the highest noise data ($t14 = 4.53$, $p < 0.001$).

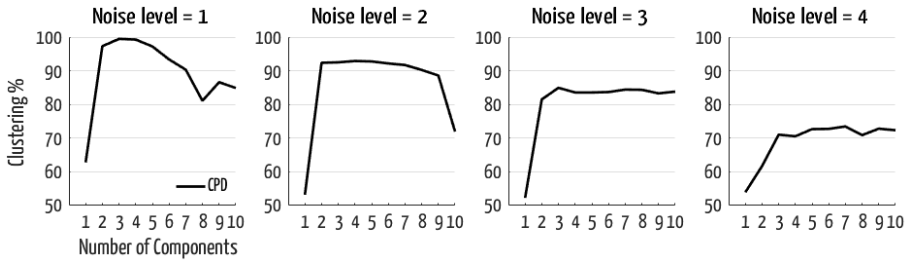


Figure 5.14: Average clustering percentage of the CPD for each of the four noise levels, dependent on the number of components considered in the models

The influence of the rank of the CPD models on the clustering percentage is shown in Figure 5.14. The number of components was found to be optimal

from 2 on for all noise levels except the noisiest datasets. The latter reached maximum discrimination between the stimuli from 3 components onwards.

Fixed-Factor Decompositions

Transferring the most discriminative spatial and temporal factor (based on classification accuracy of the trial mode) from a low-noise dataset to the decomposition of the highest-noise dataset improves the clustering for the CPD up to 83%. This procedure can be best compared to transferring knowledge of the same subject in between different conditions or recordings. Figure 5.15 illustrates the clustering results on the highest noise level and the origin of the fixed component: the noise-free case, or one of the other three noise levels. Surprisingly, the transferred spatial and temporal factor from the noise-free (i.e., noise level 0) component is not very efficient in the highest noise dataset, as is evident from the blue line in Figure 5.15. Note that these results are not dependent on the number of components. This is to be expected, as the first component is always fixed and already contains a good estimate of the P300 signal of interest, and therefore the other components are mostly filled by noise. For almost all presented fixed factor decompositions, the fixed component was the most discriminative. This removes the need for component selection but requires label information of the data tensor that the factors were taken from, rendering the method supervised.

Coupled models of low and high noise

With the coupled models, we aim to increase the clustering of the highest noise level, as here, the single-model results fall short of supervised alternatives such as the rLDA presented here. Clustering the highest noise condition, i.e. SNR = 0.60, through a structured coupled CPD with a dataset from one of the lower noise-level datasets resulted in a clustering percentage of 81.9% ($\pm$ 2.0), 80.3% ($\pm$ 2.7), 79.4% ($\pm$ 2.5) and 78.3% ($\pm$ 3.0) for the lowest to highest noise dataset. These values are significantly higher compared to the single CPD model (i.e., 71.0% ($\pm$ 7.1)). Figure 5.16 depicts the influence of the noise level of the dataset that was jointly decomposed and the dependency on the Rank of the models for CPD. Note that with the coupling a rank of 2 is sufficient to extract the most discriminative source (compared to 3 of the single model reference). In summary, these results are similar to the fixed-factor results presented in the previous paragraph without the need of specific label information.

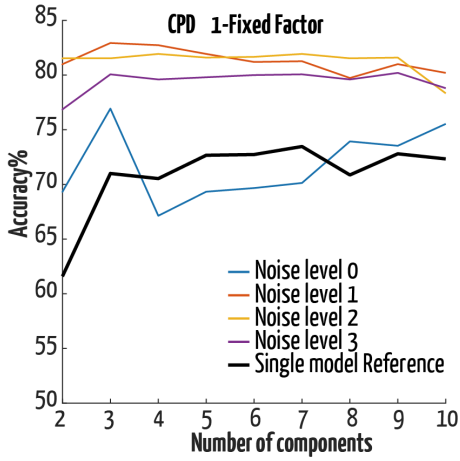


Figure 5.15: Accuracies of CPD with a fixed factor for the highest noise condition. Each line represents a different noise level from which the best spatial and temporal factors were identified to be fixated in the highest noise dataset decomposition.

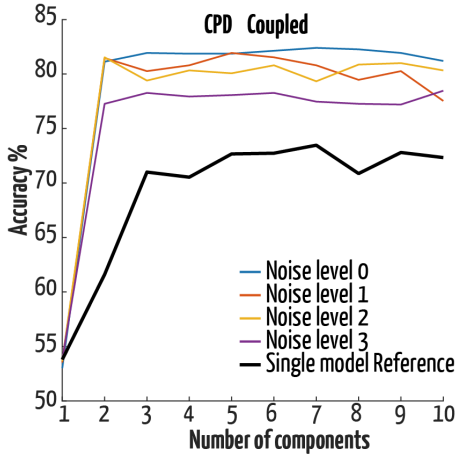


Figure 5.16: The CPD clustering accuracy of the highest-noise-level dataset if it was coupled to one of the lowest four noise levels (indicated as level 0-3) in contrast to the number of components in the decomposition.

Coupling in the Trial Dimension

The coupling of several datasets of the highest noise level with a shared trial mode results in improved clustering of the stimuli in the highest noise condition. Fig.

5.17 illustrates the clustering percentage in relation to the number of coupled datasets. Evidently there is a linear increase in the number of datasets that are combined and provide near 100% separation in the case of five simultaneously decomposed datasets. Moreover, these results are not dependent on the number of components as long as $R > 1$.

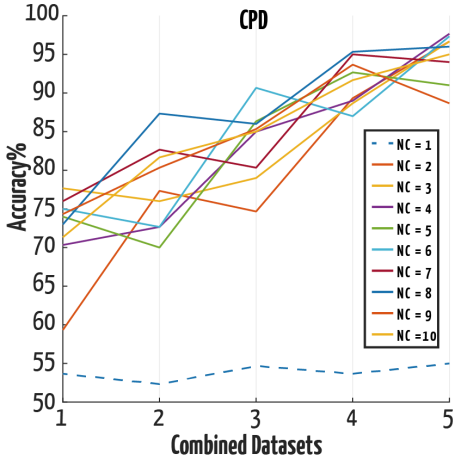


Figure 5.17: Clustering percentage in relation to the number of coupled datasets in the trial-mode for the highest-noise-level dataset (SNR = 0.60). Each line corresponds to a different Rank.

5.2.4 Discussion

In the current section we explored the incorporation of structural information into the analysis in several ways to increase data-driven clustering of target and non-target trials. The current study carefully simulated P300 ERP data from 15 subjects at 5 different noise levels, mimicking realistic settings. We showed that a simple single CPD is able, without supervision, to separate signal and noise in simulated single-trial P300 ERP data (e.g. the components that represent the N100, P300 and noise in Figure 4.2 in the previous chapter). At low noise levels, these models perform equally to that of supervised rLDA, which requires specific label and training phases (Table 5.1). In contrast, the CPD models lack the ability to achieve high performance for low SNR datasets. For the highest-noise dataset (SNR=0.60), we illustrated that imposing a coupling to datasets with a lower noise level can significantly improve the extracted clusters to classify target from non-target trials to achieve equal accuracy to

rLDA (Figure 5.16). These structured models function without a training phase or label information, although they require additional data. Transferring a derived spatiotemporal component from a low- to high-noise dataset resulted in a similar improvement of the clustering, albeit in a supervised way; the best component on the low noise dataset has to be determined beforehand. Finally, we illustrated the potential of coupled models for the analysis of simultaneous EEG recordings (e.g., multi-user recordings [118, 171]) in which the trial mode is shared between subjects. Without prior knowledge of the signal of interest, the tensor-based models successfully separated the two stimuli classes in the highest noise scenario up to 100% for coupling of five simulated subjects.

5.3 Chapter Conclusion

A method that removes the subject-specifc calibration phase for classification has been shown (Section 5.1) to have significant benefits over traditional supervised methods such as rLDA on a three-class auditory mobile BCI dataset. With structured CPD and LL1 decompositions of single trials and templates our estimates compared favourably to more complex model training in supervised ways. Future work should focus more on the understanding of the fluctuations in brain responses on single trials and incorporating structural information rather than merely improving classifier functioning.

In the second part of the chapter we hypothesized that imposing structure and linking the decompositions of higher dimensional data arrays called tensors could increase the classification accuracy. For the highest noise dataset (SNR=0.60), we demonstrated that imposing a coupling to datasets with a lower noise level can significantly improve the extracted clusters to classify target from non-target trials to achieve equal accuracy to the widely used rLDA. The coupled CPD models do not need a training phase or label information, although they require additional data(sets). Finally, we illustrated the potential of the tensor approach for the analysis of simultaneous EEG recordings in which the trial mode is shared between subjects. Without prior knowledge of the signal of interest, the tensor-based models successfully separated the two stimuli classes in the highest noise scenario up to 100% for coupling of five simulated subjects. These results highlight the benefits of exploiting prior knowledge and structure in the stimuli and experimental setup (e.g., conditions or subjects).

Chapter 6

Disentangling attentional and physical contributions to auditory attention tasks in real-life

In the previous chapter we introduced a novel approach to extract meaningful task related information, i.e. P300, from the EEG in a data-driven way. Generally, methods for auditory oddball BCI classification heavily rely on robust single-trial P300 ERPs. However, very little research has been done on the effects of being in a real-life environment onto the P300. Application of such BCIs in real-life scenarios is limited due to technological constraints and the large number of potential confounds. With the recent emergence of proper mobile EEG hardware we are able to conduct such experiments and examine the feasibility of real-life BCI. To this end we aimed to setup an experiment that pushes the boundaries of (mobile) EEG experimentation to allow unconstrained recordings outside while riding a bike, i.e. dataset D. Furthermore, we evaluated the influences of attention and physical contributions to the P300 as elicited in a three class auditory oddball task. Lastly, we draw conclusions on the real-life aspects and practicability of conducting mobile EEG experiments. This work has been published in [242].

6.1 Introduction

We created an outdoor experiment with the aim to disentangle the influence of cognitive demands and distractions that arise from such outdoor unconstrained

recordings and demonstrate the usability of a BCI while riding a bike. Subjects were presented with the same three-class auditory oddball paradigm in three different outdoor conditions to disentangle effects of physical activity from higher cognitive demand. In one condition subjects sat still on a regular bike fixed in a framework so that they did not move. Secondly subjects sat on the fixed bike and additionally pedaled at a user chosen comfortable pace. The third condition comprised of biking freely on a pre-explored course on the university campus. This was the most demanding condition as the subjects need to attend to the auditory task, execute motor patterns to pedal and deal with the complex environment as they move around the streets. The pedaling condition induces muscle and motor activity but probably no increased cognitive load as compared to actual biking around campus. Sitting completely still in a bike position can be seen as the most restricted condition (control) in which only environmental noise could distract the subjects. All conditions took place outside to ensure similar degrees of random environmental effects (i.e. weather conditions, cars passing by, people talking in the background).

Several EEG laboratory studies investigated specifically the effect of cycling exercise on the P300. Studies have reported an increase in P300 amplitude during or after physical pedaling on an exercise bike in response to visual attention tasks [32, 162, 170]. Vogt et al. contrastively reported no influence of exercise on the P300 in a virtual reality study [224]. Furthermore, Yagi et al. described decreased P300 amplitudes during exercise [231]. Potential reason for the decrease in P300 in these studies is reasoned to be the increase in cognitive load that diminishes the P300 due to lower attentional resources. Yet another study indicated increases in P300 only during 72 to 108 min of exercise [87]. These studies did not evaluate the fluctuations in P300 in a true real-life enviroment. Given that in our setup the cognitive demands will be higher as compared to the aforementioned lab-setups we would hypothesize for a decrease in P300 in the real biking conditions as compared to the fixed bike conditions. Additionally, very few mobile EEG studies explicitly controlled for the amount of movement while recording in real-life situations, although many systems allow to simultaneously record the level of movement. Lab studies usually try to minimize head and body motion. Furthermore, the presence of unwanted artifacts in the EEG are principally dealt with by trial rejection. In the current setup we have instantaneous correlates of head movement from a 3D gyro measurement (see Section 3.3.2) that allows to investigate these potential artifacts and derive measures to compare the motion to ERP and classification features. Being in less constrained scenarios, subjects will be more mobile, leading to notable EMG artifacts in the EEG (e.g. increased muscle activity to keep head up) [85, 127, 92, 175, 126]. For an overview on the methodological challenges of recording EEG during motion we refer the reader to [175]. Past experience has shown that CCA-based demixing of the data leads to a superior

removal of artifacts as compared to the wide-spread ICA-based demixing [43, 51]. We follow this approach to extract muscle related sources from the EEG of the bike study and quantify the influence of the EMG on our BCI.

To summarize, the main goal is to evaluate the influence of cognitive distraction and physical engagement on auditory attention correlates in unconstrained out-of-the-lab scenarios. We will investigate both ERP and single trial correlates recorded from subjects seated on a bike without pedaling, with pedaling and while biking in real life environments. Secondly, we will investigate the effect of different amounts of movement, as quantified with a 3D gyrometer signal on P300 features. Finally we compare the main P300 waveform results before and after correcting for EMG activity.

The remainder of this chapter is structured as follows: Firstly, we provide a brief overview of the stimuli, hardware, procedure and data acquisition with the mobile EEG. Secondly, we discuss the data analysis of the derived ERP features and classification methods used. Thirdly, the obtained results are presented in terms of ERP results, motion quantification and classifier output. Fourthly, the influence of EMG artifact removal on the aforementioned results are presented. Finally, we discuss the effect of real-life circumstances on auditory attention and the complementary information obtained from physical motion measures followed by future perspectives for the field of mobile EEG studies.

6.2 Methods

6.2.1 Participants

Fifteen subjects (mean age (±SD): 27.1 (±2.5), four women) participated in the current trial. Subjects reported normal hearing and no past or present neurological or psychiatric conditions. All participants signed informed consent prior to participation. The ethics committee of the KU Leuven approved the experimental setup. All subjects reported to be able to ride a bike confidently.

6.2.2 Stimuli and procedure

A three-class oddball auditory task was used as proposed by [91]. A standard tone (900Hz) and two deviant tones (600Hz, 1200Hz) of 62ms duration (10ms rise/fall time) were presented binaurally through low-cost consumer headphones (Sennheizer mx460) in random order. A mean inter-stimulus interval of 1000 ms (jitter [0-375] ms) was used. The participants' task was to silently count the target tones and ignore the two other tones. Eight subjects attended the lower

(600Hz) tone and the others attended the high (1200Hz) tone. The subjects' attention to a certain tone should elicit a distinguishable P300 (target) as compared to the non-attended tones (non-targets). 504 standards, 92 non-target deviants and target deviants were presented in the first condition. In the second recording a few more trials were presented in order to create a valid task silently counting targets (105 stimuli). In the third recording the stimulus numbers were equal to the first condition.

Due to technical issues, we recorded only 69 trials in 2 subjects. In all other subjects and conditions the trials were truncated, restricting each condition to 92 trials in order to ensure equal comparisons between conditions. The recordings took place outdoors on the campus of the KU Leuven. The subjects performed one recording while sitting still on a bike in a fixed standard facing nature and one recording while pedaling on the bike in the fixed standard. The third condition involved biking freely around on a 500m course on campus for the period of the experiment. An overview of the three conditions is presented in Figure 6.1. The biking route (move condition) was predefined and practiced before the start of the experiment. The order of the conditions was balanced between subjects. These three conditions will be referred to as Still, Pedal and Move from here on. During the experiment, participants carried an ultrabook laptop (weight <2 kg) in a backpack. Each condition lasted roughly 12 minutes. For the Pedal and Move conditions subjects were free to pedal at a comfortable pace. On average, subjects biked 12 km/h, as derived from the average time taken to bike once around the course. The course for biking in the Move condition involved regular streets with traffic, i.e., mostly pedestrians and cyclists, very few cars. Note that during all recordings, subjects were outdoors exposed to naturally occuring (sound) distraction (e.g. insects which seemed to be attracted to the gel, wind, car traffic noise and people passing by). Recordings took place during spring and summer of 2015; the temperature varied between 10 and 30 degrees Celsius. Recordings with strong wind >4Bft and precipitation were avoided. OpenViBE software [176] running on the ultrabook was used for stimulus delivery and experimental control. After the experiment, subjects were asked to rate (on a scale from 1 to 10) the three conditions according to the degree they enjoyed each of the recordings. This was included in order to obtain a coarse measure of motivation as this has been shown to influence the P300 waveform [17].

6.2.3 Data acquisition

The acquisition was conducted with a SMARTING mobile EEG amplifier from mBrainTrain (Belgrade, Serbia, www.mbraintrain.com) and 24 Ag/AgCl passive

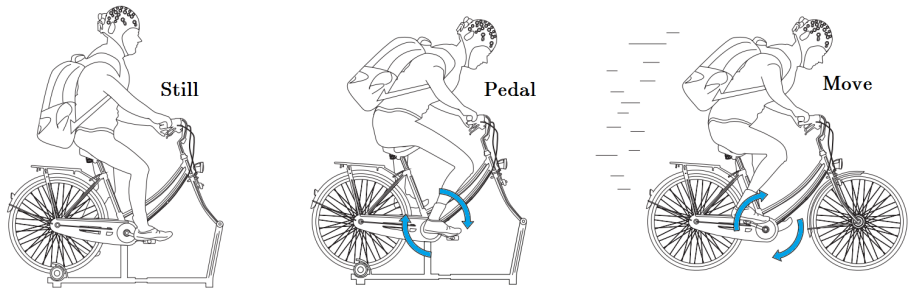


Figure 6.1: Schematic of the Still, Pedal and Move conditions. Note that in the Pedal conditions subjects sit on the fixed bike like the Still condition and pedal like they were biking in the Move condition. All recordings took place outdoors at or next to roads on the KU Leuven campus

scalp electrodes (Easycap, www.easycap.com). The hardware and software interface has been described in Section 3.3.2.

6.2.4 Data analysis

Preprocessing

Extended Infomax ICA was used to remove EOG activity by means of manual component selection that represented -blinks and lateral eye movements. As a pre-processing step for ICA all EEG data was high-pass filtered at 1Hz as this is shown to produce stable ICA results for artifact removal [52]. Consecutively the learned weights of the marked 'EOG' components were applied to a common band-pass filtered [0.5-20]Hz dataset. Signals were re-referenced offline to the mean of TP9 and TP10 and epochs were extracted 200ms before stimulus onset to 800ms after. The Tp9 and Tp10 channels were removed after re-referencing. A baseline subtraction was applied based on the [-200-0]ms interval. The gyro-signals were high-pass filtered at 0.5Hz as well. Average ERPs were calculated at the Pz electrode as this is the most prominent site for the P300 waveform. Visual representation of the ERPs was achieved after removing artifactual single-trials as implemented in EEGLAB [58]. Rejection criteria were based on abnormally distributed data and improbable data. The cutoff was defined as four standard deviations of the mean kurtosis value and probability distribution respectively. Note, all the analyses were based on the full dataset.

In addition we removed EMG activity by means of BSS-CCA. This assumes that EMG sources have significantly lower autocorrelation as compared to brain

sources [43]. Removal of the EMG was achieved through BSS-CCA on 10s non-overlapping windows. After extracting 24 components (per window), the power spectral density of each source was computed. Sources with a power ratio smaller than 2 between brain band (i.e. 3-15Hz) and muscle band (i.e. 25-50Hz) were removed. This approach is similar as presented in [43, 51].

Classification

For single-trial P300 classification we extracted one feature set which consisted of two different types of features: The first type comprised seventeen 47ms data bins on all 22 electrodes between [0-800]ms, which is similar to previously reported work [57, 49] resulting in 374 features. In addition we averaged all standard-stimulus trials of the non-test subjects to create an average template. This was shown to provide an accurate estimation of the non-target waveforms (see Figure 5.1 in chapter 5) and therefore the distance to this template can be used as discriminative feature to distinguish target from non-targets trials. We down-sampled the EEG data and template to 30Hz and subtracted per trial the EEG data from the template in eleven points in a 200-533ms window on all 22 electrodes, resulting in another 242 features. Both feature types were concatenated creating a total of 616 features per single-trial. We classify the EEG signals using a linear discriminant analysis as it is among the most used classifiers for the P300 [23, 136, 132, 131]. Shrinkage regularization as implemented in BCILAB [121] is used for the rLDA classification. Per subject and session the classifiers are trained and evaluated based on a five-fold cross-validation procedure. Finally, to evaluate differences between conditions further, we train per subject an rLDA classifier on data from one of the other two conditions of that same subject. For example train on the Still condition and Test on the Pedal condition (cross-session).

ERP features

In order to compare the ERP features we computed the maximum P300 at Pz for the Target and NonTarget (Distractor) tones for each subject in a [200-600]ms interval. Likewise we computed the minimum N100 at Pz for all tones in the [0-200]ms interval. This was achieved by averaging 22ms (i.e. 11 samples [-5 +5]) around the peak latency value at each electrode. The root mean square (RMS) in the [-200-0]ms baseline interval was used as noise measure for the EEG which is similar to [49]. Consecutively we took the average RMS over the three gyro axes in the [0-800]ms interval as measure of the amount of head motion per trial. Alpha band activity (8-12Hz) was calculated at electrode O2 as the average normalized power in this range for the Target trials.

Statistical analysis

One-way ANOVAs were performed to determine the significance of mean group differences between the three conditions. This was used for comparing EEG RMS, N100 amplitude, Gyro RMS and classification accuracies between conditions. For the P300 comparison a repeated measures analysis of variance ANOVA was conducted with the factors recording (Still, Pedal and Move) and tone (Target, NonTarget). This way we evaluate the elicted P300 differences between the target and distractor tone and explore impact of the recording condition on this effect. In all analyses, the significance level was set at 0.05. Pearson's correlation coefficient is used to identify linear relationships between our derived ERP and RMS features and classification results.

6.3 Results

The ERP, Noise level, Classification and Gyro results presented here are obtained on the data without explicit control for muscle artifacts. Subsequently, Section 6.3.5 displays the results after removal of muscle artifacts.

6.3.1 ERP

The average ERPs of the Baseline, Target and NonTarget tones are displayed in Figure 6.2 with topographies of the peak P300 and N100 for the Still, Pedal and Move condition. In all conditions a clear posterior focus of the P300 topography is visible whereas the N100 is more central. P300 amplitude in the Move condition was on average 31% and 26% lower as compared to the Still and pedal condition respectively. The repeated measures ANOVA recording (Still, Pedal, Move) $\times$ tone (Target, NonTarget) revealed a nearly significant main effect for recording, $F(2,14) = 2.89$, $p = 0.06$, reflecting larger amplitudes for the still (mean = 6.8 μV) and pedal (mean = 6.3 μV) compared to the Move condition (mean = 4.7 μV). In addition, a significant main effect for tone emerged, $F(2, 14) = 68.41$, $p < 0.001$, confirming larger P300 amplitudes for target tones (mean = 5.96 μV) compared to distractor tones (mean = 1.57 μV). The interaction recording $\times$ tone was far from reaching significance, $F(2, 14) = 0.58$, $p = 0.56$.

Contrastingly the peak N100 values did not differ significantly i.e. one way ANOVA ($F(2, 14) = 0.42$, $p = 0.66$), ($F(2, 14) = 0.21$, $p = 0.83$) and ($F(2, 14) = 1.20$, $p = 0.31$) across conditions for the target, nontargets and baseline stimuli respectively. A slightly higher P300 latency was observed in the Move condition; 445ms ($\pm$95) as compared to the Still 400ms ($\pm$56) and pedal 388ms ($\pm$89) latencies but this difference was not statistically significant ($F(2, 14) = 2.00$,

$p = 0.15$). Correlations between the subject average P300 across the still-pedal conditions was moderate ($r = 0.70$, $p < 0.001$), for the pedal-move ($r = 0.61$, $p < 0.001$) and still-move ($r = 0.55$, $p < 0.001$). Overall, subjects which had more consistent P300 waveform across conditions also had a higher P300 amplitude ($r = 0.74$, $p < 0.01$).

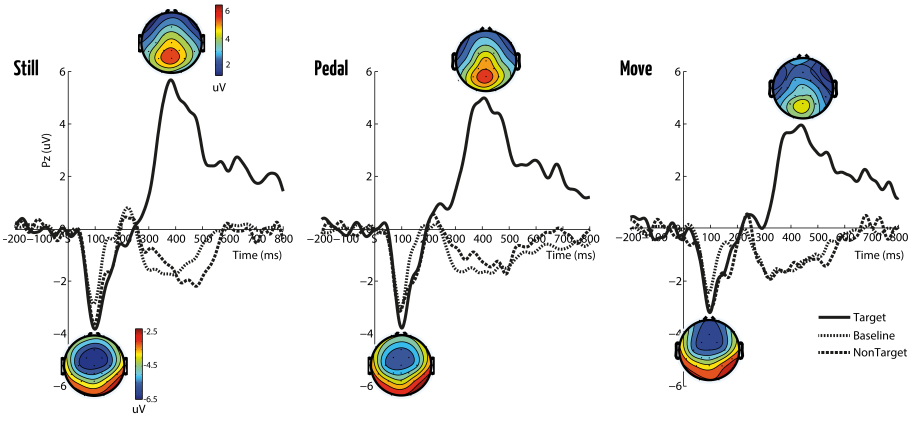


Figure 6.2: Grand average ERPs to the Target, nonTarget and Baseline stimuli at electrode Pz for the Still, Pedal and Move conditions from left to right respectively. At the peak N100 and P300 ERP a topoplots indicates the distribution of the peak along all electrodes for the Target stimuli. Note: for illustrational purposes artifactual single-trials were removed before plotting. Reported statistics in the text are based on the full dataset.

6.3.2 Noise levels

The mean RMS values of the target trials' baseline at Pz did not show significant differences overall between conditions ($F = 0.14$, $p = 0.87$). These RMS values at Pz are similar to the ones reported during Sitting and Walking in a previous study [49]. However, at different electrode sites differences between conditions can be observed. Figure 6.3 illustrates the median RMS of the pre-stimulus window for all channels in the Still, Pedal and Move condition. The Pedal condition RMS values were only slightly elevated as compared to the Still condition and the RMS topography does not change. In the Move conditions on the outer electrode sites clear differences can be observed as compared to the other two conditions. This was confirmed by (marginally) significant ANOVA results on channels: FP1 ($F = 3.09$, $p = 0.06$), FP2 ($F = 3.25$, $p < 0.05$), O1 ($F = 2.70, p = 0.08$) and O2 ($F = 4.50$, $p < 0.05$). Pairwise comparisons showed

a significant increase in RMS in the Move condition as compared to either the Still or Pedal condition. The artifactual pattern in the Move condition is similar to those presented by motion artifact studies that report increased artifacts especially at the outer electrode sites [77].

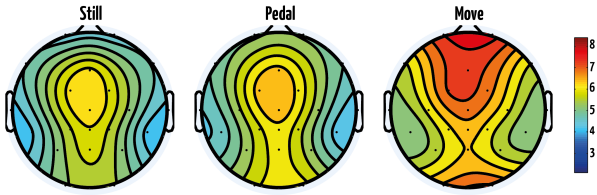


Figure 6.3: Median RMS values of the baseline interval [-200-0]ms of the Target trials at each electrode for the Still, Pedal and Move condition.

6.3.3 Classification

Grand average accuracies (±SD) on a single trial level are 76.6 (±7.0), 72.7 (±9.3) and 67.6% (±8.5) for the Still, Pedal and Move condition respectively which are presented in Table 6.1. These differences were deemed significant ($F = 4.45$, $p < 0.05$). The accuracies in the Move condition were significantly lower as compared to the Still ($t14 = 3.89, p < 0.01$) and Pedal ($t14 = 3.31$, $p < 0.01$) condition. The difference between Pedal and Still conditions was marginally significant ($t14 = 2.06$, $p = 0.06$). P300 ERP amplitudes were significantly correlated to the rLDA accuracy across subjects for all three conditions; still ($r = 0.71$, $p < 0.001$), pedal ($r = 0.80$, $p < 0.0001$) and move ($r = .64$, $p < 0.05$). This indicates that individuals with larger P300 ERPs had better classification accuracies. Similarly, correlations of the single-trial classification accuracies between the still-pedal ($r = 0.61$, $p < 0.05$) and pedal-move ($r = 0.78$, $p < 0.001$) were observed. This was not the case for the still-move comparison of accuracies (r = 0.33, p = 0.22).

Table 6.1: Grand-Average Session Specific accuracies (±SD).

	Still	Pedal	Move
Accuracy	76.6% (±7.0)	72.7% (±9.3)	67.6% (±8.5)

Classifying across sessions lead to accuracies of 71.4 (± 8.9) and 68.5 (± 7.6) for the Still condition when trained on the Pedal and Move data respectively. Accuracies in the Pedal condition were 71.3 (± 8.7) when trained on the Still recordings and 67.8% (± 9.4) for training on the Move condition. Finally the

Move condition accuracies were 66.4 (± 6.5) and 65.1 (± 9.3) when trained on the Still or Pedal condition respectively. The accuracies of all cross session comparisons are summarized in Table 6.2.

Table 6.2: Grand-Average Cross-Session accuracies (±SD).

Train On	Test Still	Test Pedal	Test Move
Still	n/a	71.3% (±8.7)	66.4% (±6.5)
Pedal	71.4% (±8.9)	n/a	65.1% (±9.3)
Move	68.5% (±7.6)	67.8% (±9.4)	n/a

6.3.4 Gyro signals

Average (±SD) RMS values of the 3-axis gyroscope signals are 2.4 (± 0.6) in the Still, 17.4 (± 9.7) in the Pedal and 27.6 (± 8.0) in the Move conditions. The single-trial motion distributions are displayed in Figure 6.4A. As expected the mean gyro motion RMS variables were higher in the Move condition as opposed to the Pedal and Still condition. Similarly the Pedal condition gyro RMS was higher compared to the Still condition. To examine the effect of the motion in the (real-life) Move condition as measured with the gyro to the ERPs, we split per subject the trials into eight bins of low to high motion of equal trial numbers. The gyro effect on the eight grand average ERPs is depicted in Figure 6.4B. The color coding indicates the eight different levels of motion. Finally in Figure 6.4C the peak P300 as seen in Figure 6.4B is plotted in order to reveal any effect of motion level on the P300 amplitude. No decrease in P300 amplitude can be observed and these results suggest that there is no negative impact of gyro motion on the ERP at Pz. Similar results were obtained by analysing the P300 waveforms on the gyro axes separately (not presented here). Similarly, we contrasted the grand-average accuracies (i.e. of all subjects combined) obtained with rLDA models per motion level. This did not lead to significant differences in the trials with the highest motion level, this is shown in Figure 6.4D as well (c.q. motion level 7 and 8). Finally, we investigated the relationship of the gyro motion on the accuracies as computed with the subject specific rLDA on subject level. The gyro motion level displays a moderate negative correlation to the rLDA accuracies for the Still ($r = -0.39$, $p = 0.15$), Pedal ($r = -0.51$, $p = 0.05$) and Move ($r = -0.33$, $p = 0.23$) conditions and lower correlations to the average P300 amplitude ($r = -0.36$, $p = 0.10$), ($r = -0.43$, $p = 0.11$) and ($r = -0.16$, $p = 0.58$) for the Still, Pedal and Move condition respectively.

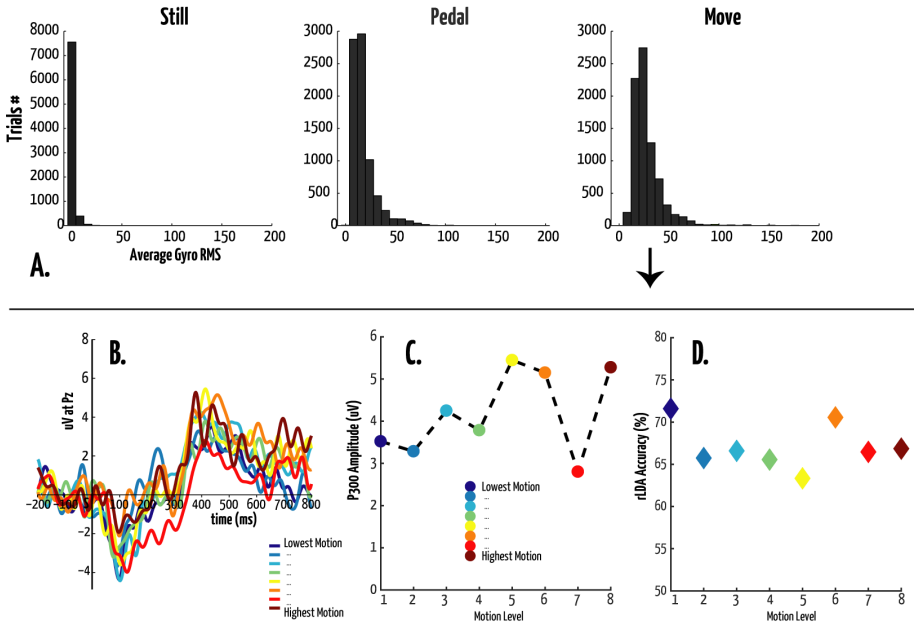


Figure 6.4: A. Averaged (over 3 axes) gyro RMS per trial for the Still, Pedal and Move condition from left to right respectively. B. The average ERPs at Pz to eight divisions of the Move condition data based on the gyro RMS in A(most right figure). The color indicates the degree of motion and all degrees have equal number of trials. C. displays the peak P300 from B in contrast to the eight motion degrees; the P300 peaks seem to be unrelated to the degree of motion. Finally, D. shows the average classification accuracy in relation to the motion levels.

6.3.5 Muscle artifacts

After removing muscle activity, the mean RMS values of the target trials' baseline at Pz showed no significant differences between conditions ($F = 0.59$, $p = 0.56$) which is similar as mentioned before in paragraph 3.2. Overall, the RMS topographies are more similar between conditions after EMG removal as the ones prior to removal of the EMG activity (cf. Figure 6.5 and Figure 6.3). The principal difference is evident as a strong decrease of the RMS values in the Move condition. In general the P300 amplitudes at Pz decreased slightly after removing the EMG activity. The repeated measures ANOVA recording (Still, Pedal, Move) $\times$ tone (Target, NonTarget) revealed a significant main effect for recording $F(2, 14) = 3.93$, $p =< 0.05$, reflecting significantly larger amplitudes for the Still (mean = 6.4 μV, $t14 = 4.08$, $p < 0.01$) and Pedal (mean = 5.9 μV,

($t14 = 2.84$, $p < 0.05$)) condition compared to the Move condition (mean = 4.1 μV). The difference between the Pedal and Still condition failed to reach significance ($t14 = 1.09$, $p = 0.29$). These differences in P300 are similar to the ones reported in section 6.3.1 without control for muscle artifacts.

The grand average accuracies (±SD) on single trial level after EMG correction are 75.5 (±8.2), 70.0 (±10.4) and 66.2% (±8.4) for the Still, Pedal and Move condition respectively. These differences were deemed significant ($F = 4.45$, $p < 0.05$). The difference between Pedal and Still conditions was deemed significant ($t14 = 2.47$, $p =< 0.05$) and the accuracies in the Move condition were significantly lower when compared to the Still ($t14 = 4.82$, $p < 0.001$) condition, but not in comparison to the Pedal condition ($t14 = 1.70$, $p = 0.11$).

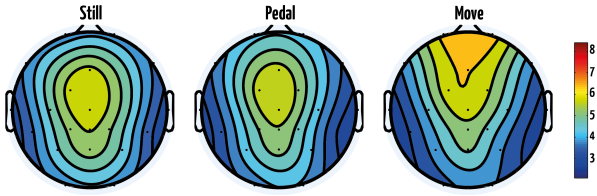


Figure 6.5: Median RMS values of the baseline interval [-200-0]ms of the Target trials at each electrode for the Still, Pedal and Move condition after EMG noise removal.

6.3.6 User metrics

We determined per subject the average normalized power in the alpha band (8-12Hz) for the three conditions prior to the muscle activity removal. Mean (±SD) normalized alpha power for the Still, Pedal and Move conditions are respectively 18.5 (± 4.2), 16.8 (± 3.6) and 14.7 (± 3.6). These values differ significantly ($F = 3.69$, $p < 0.05$) which was caused by a lower alpha power in the Move condition as compared to the Still (t14=2.51,p<0.05) and Pedal (t14=2.45,p<0.05) condition. The difference between the Still and Pedal conditions failed to reach significance (t14=1.37,p = 0.19).

Average motivational scores were highest in the Move condition with a mean 8.5 (SD ± 1.0) followed by the Pedal condition with 7.6 (± 1.5) and lastly the Still condition scored a 6.2 (±1.5). These were significantly different ($F = 11.28$, $p < 0.001$) and pairwise comparisons displayed a significantly higher motivation score for the Move condition compared to the other two. This indicates that the subjects enjoyed the Move condition the most. Finally, a nearly significant correlation ($r = 0.28$, $p = 0.07$) between the number of correctly counted tones and the single-trial classification accuracy suggested that those individuals

attending more to the task had better classification accuracy. Moreover, the percentage of correctly counted tones in the Still, Pedal and Move condition were very high, respectively 98.6 (± 1.2), 97.5 (± 3.3) and 97.7% (± 2.6).

6.4 Discussion

Mobile EEG allows to study the brain in real unconstrained environments. Previous (mobile) EEG studies did not control independently for the increase in cognitive demands by the complex natural environment or the increase in physical involvement. These effects have merely been studied in simulated environments (e.g. [32, 162, 170, 224, 231, 87, 85, 127, 89, 105]). We evaluated the ERP characteristics and classification performance of a three-class auditory oddball paradigm recorded in outdoor scenarios while biking on either a fixed bike or freely around. The Move condition was an unconstrained ride around the place, representing a fully realistic scenario. Above chance P300 single-trial classification in these highly dynamic real-life environments while biking outdoors was achieved. Certain artifact patterns were identified in the free biking condition, but we show that neither (See Figure 6.5) these nor the increase in movement (c.f. Figure 6.4) explain the reduction in classification accuracy and P300 waveform in the real-life biking condition as compared to the control conditions. Therefore this study provides strong evidence that the higher cognitive demands are the underlying cause of the reduced attentional brain responses to auditory stimuli. This reasoning corresponds to the intuitive explanation that biking around in a natural environment requires definitely more cognitive resources than the non-mobile conditions.

On the grand average a decreased P300 peak was observed in the Move condition as compared to the other (more restricted) conditions (Figure 6.2, P300 topo plots). Similar effects were notable after controlling for muscle artifacts. This decrease in P300 during outdoor moving conditions was similarly described by [57, 49], albeit without control for artifacts and movement. Interestingly, N100 peak values for all stimulus types did not differ significantly between our conditions. We reason that the P300 and N100 should be influenced equally by physical artifacts. Since predominantly the P300 is altered between conditions (as evident by the significances of the P300 and N100 amplitude ANOVAs) this suggests that the effect of the Move condition on the P300 is larger as compared to the N100.

Additionally a mildly increased latency of the P300 was found in the Move condition, although this was not deemed significant, it may be suggestive of increased cognitive load. Indoor exercise studies suggest an increase in P300 during or after physical pedaling on a bike [32, 162, 170]. In our ERP analysis

no clear distinction could be made between the Pedal and Still condition; similar P300 waveforms were observed which is similar to the results of [224]. These differences may originate from the paradigm that is used to generate a P300. The three-class auditory oddball task is dependent on the ability to attend to the acoustic stimuli and that may be more difficult with a dual task of listening and biking in a distractive environment. Another measure related to the level of cognitive functioning is the EEG power in the alpha band. A decrease of alpha power may be indicative of increases in task difficulty and mental workload [208, 194, 111]. The alpha power decrease in the move conditions suggests a lower focus on the auditory targets which can be related to increased difficulty due to the real-life circumstances as compared to the pedal and still conditions. An alternative explanation that we could not rule out is that the decrease in alpha power may be correlated with the extent of physical engagement.

The gyro motion quantification displayed a clear increase in head motion in the Pedal and Move conditions as compared to the Still condition (Figure 6.4A). However, similar ERPs were observed in the Still and Pedal condition illustrating that the wireless system used for recording and the unconstrained setup did not affect the EEG quality at channel Pz. In the real-life Move condition we split the trials to eight bins with increasing motion levels. This did not lead to a decreased P300 waveform or peak amplitude as was depicted in Figure 6.4B-C. Furthermore, the correlation between P300 amplitude and the head-motion was found to be minor, especially in the Move condition ($r = -0.16$). A smaller P300 while biking can be alternatively explained by that the brain is occupied with extra tasks over and above the experimental task, for example the evaluation of the dynamic acoustic and visual environment and ensuring that the bike stays on the road. Similar reasoning was put forward by [49].

The EEG RMS values between conditions were only significantly higher at the frontal and posterior electrodes (Figure 6.3) for the Move condition. Further analysis of these channels with higher RMS values in the Move conditions showed sporadic low frequency fluctuations (head motion) and short, higher frequency range artifacts that are indicative of muscle activity. This can be attributed to a mix of electrode movement and muscle tension which is more elevated while riding a bike as opposed to the fixed conditions. Similarly there was an increase in gyro motion for the Move condition (Figure 6.4A). The fact that central channels such as Cz and Pz remain almost unaffected indicated that the observed artifacts did not have a large impact on the ERP results. The increase in head motion in the Move condition was expected to increase the muscle acitivity as well. This was confirmed by the large decrease of baseline RMS in the Move condition after controlling for muscle artifacts. Generally, controlling for muscle artifacts did not lead to different insights. Finally, the RMS topographies and degree of motion artifacts varied across subjects which

is in line with recent views on motion artifacts and EEG [112].

Single trial characteristics have been assessed by computing rLDA classification results as is routinely done in BCI applications. For a practical BCI application, approaches that reduce the subject-specific calibration time have gained momentum for BCI processing [243, 14, 106]. Riemannian geometry (see [41] for a recent review) in particular has been shown to be very powerful for P300 classification and easy generalization over subjects [14]. Furthermore, the incorporation of additional structural properties present in the paradigm [243], or a reliance on stimulus presentation distributions through Bayesian statistics [106] could be beneficial. A valuable future line of research would be to see how these more profound methods would perform on real-life data such as the Move condition in the current study. A recent study on motor imagery suggested that methods which work well in a lab environment can perform poorly in realistic application scenarios [27]. The differences between conditions may be larger in real-life scenarios as compared to restricted lab setups. Differences between our conditions were also observed from the cross-session trained classifier models which performed consistently lower compared to session-specific classification (c.q. Table 6.1 and Table 6.2); when trained on the Move condition large differences in accuracy can be observed for the Still and Pedal condition. This is supported by the high average correlation ($r = 0.70$) between the subject P300 waveforms of the Still and Pedal condition, but lower for comparisons of these to the Move condition, ($r = 0.61$) and ($r = 0.54$) respectively. The observed differences between conditions are indicative of non-stationarities or unidentified artifactual influences in the signal and transferring these session-specific differences might improve classification [180]. Furthermore higher correlations were found between the gyro motion and the rLDA accuracy as between the gyro motion and the P300, suggesting that motion artifacts (at the outer channels) might have an influence of the rLDA results as compared to the P300 in our ERP analysis.

It is vital to understand that the precise extent by which participants could focus their attention on the target tone while blocking out the distractor stimulus is unknown. Future measurements of the environment by means of video and audio monitoring may provide better control of auditory distraction events. Future analysis may also look into different types of head motions in the gyro signals. Increased left-right motion patterns may be indicative of distracting events whereas nodding patterns occur more naturally during biking and walking. Several subjects that attended to the high pitched tones reported that they sometimes confused the baseline tones with the high tones. This may be controlled by measuring the hearing thresholds prior to the experiment. However, post-hoc analysis showed no effect of tone on the classification accuracy.

Indoor EEG experiments can be very long and monotonous for the subjects as external influences and distractions are minimized. High scores on the motivation question for the Move condition illustrated that subjects enjoyed the real-life condition more even though the distractions made it more difficult to focus on the auditory mental task. This is a promising finding as user experience is an important factor in BCI as well [231, 156, 17].

The current study presents several limitations to the analysis and setup. The self-reported motivation score is subject to bias and a standardized test such as used in [17] is likely to provide a more reliable measure. Subjects pedaled at a user-defined comfortable pace in the Pedal and Move conditions. There was no control of speed in our study, although in related treadmill walking studies the speed of walking was shown to be related to the number of artifacts in the EEG [30, 112]. However, forcing users to stay in a specific speed range might cause additional bias since subjects may be controlling (and checking) for speed constantly. Future work that combines heart-rate monitoring with the mobile EEG may provide complementary information regarding the degree of physical exercise [162]. Moreover, heart rate has been identified as an indicator for mental state switches that could have an influence on the EEG signals [223]. Finally users could have felt abashed as point of attention while biking outside which may have lowered the ability to focus on the auditory tones. Fortunately new hardware solutions are prominent to make the EEG equipment even less distractive and visible [130, 25, 55] which is a promising development.

6.5 Chapter Conclusion

In this chapter we evaluated the ERP and single-trial classification characteristics of a three-class auditory oddball paradigm recorded in outdoor scenario's while biking on either a fixed bike or freely around. We identified elevated noise levels in the natural biking condition affecting only the outer channels. The influence of the natural head motion on the ERP waveforms on Pz was minimal and controlling for muscle artifacts reduced the overall noise level but did not lead to different insights regarding the P300 differences between conditions. A reduction in P300 in the Move condition led to lower classification accuracies obtained with rLDA. The reduction in P300 in the Move condition can be attributed to increased cognitive load due to being in a real-life scenario. These findings suggest that cognitive paradigms measured in natural real-life scenarios, such as the auditory oddball results presented here, are influenced significantly by increased cognitive load due to being in an unconstrained environment.

Chapter 7

Online detection of auditory attention with mobile EEG: closing the loop with (neuro)feedback

The previous chapters focused on experiments with the P300 ERP in response to auditory oddball stimuli as a measure of auditory attention. These results led to important new insight regarding mobile EEG application for auditory attention analysis. In this chapter we consider a more natural auditory paradigm, i.e. AAD (see section 2.4), in which the subject's attention to natural speech is detected. Section 7.1 provides an overview of the current literature and highlights the key challenges that we will address here. Our work on AAD has been divided into two studies that were conducted successively; the first provides a closed-loop online analysis system for AAD (dataset E); the second study extends the usage of this system to illustrate the potential of longterm neurofeedback to improve user's AAD ability (dataset G). The results are presented for both studies together in section 7.3. These findings are discussed in section 7.4. This work is available online (technical report: [244]) and a conference precursor has been published in [237].

7.1 Literature review

As stated in chapters 2 and 3 auditory attention detection (AAD) is based on tracking cortical EEG responses in relation to attended natural speech (e.g. [143, 62, 96, 191]). Tracking the subjects' attention holds great potential for

application in assistive devices for hearing aid improvement ([151]). Furthermore, the usage of very common speech as stimuli enhances the natural character of our experimentation. In this chapter we focus on the online analysis of mobile EEG for AAD. More specifically on the backward modeling of AAD; a pretrained decoder that makes a linear combination of the EEG and delayed versions of it allows to extract a signal with a significantly higher correlation to the attended speaker's speech envelope as compared to the unattended speaker's envelope [165]. In such dual-speaker scenarios, the subject's attention to a specific speaker could be reliably detected with high accuracy in research labs (e.g. [165]). In the past years several aspects of the analysis pipeline have been optimized to enhance the estimation made by the decoder to increase detection accuracy. Different ways of decoder construction and envelope extraction methods have been evaluated [21, 69] whereas others focused on removing lateralization bias in the decoder and improving stimulus preprocessing to enhance the realism of the experiments [42]. Whereas most AAD studies assume availability of the clean speech envelopes, also more realistic scenarios have been studied where these per-speaker envelopes are not available. Successful extraction of per-speaker envelopes from noisy speech mixtures by means of blind source separation techniques has been demonstrated in [214], where AAD was used to steer an acoustic beamformer towards the attended speaker. Alternatively, deep neural networks have been employed to obtain a reliable AAD system based on mixed audio input [164]. Furthermore, the influence of noise on the audio waveforms used in the decoder was rendered insignificant to a certain extent [9]. Another study provided evidence that only a specific set of spatial regions is involved in the detection process and consequently the number of electrodes may be decreased drastically (e.g. from 96 down to 15) and still achieve robust classification accuracy, which would be especially meaningful for mobile experiments in which there are only limited number of electrodes available [151]. [78] showed that the focus of attention on one of two simultaneously presented speech streams was not influenced by including additional competing speakers from different spatial locations.

These studies on auditory attention detection were conducted in a lab setting, using clinical grade equipment. However, when considering the development of consumer-geared applications, the system needs to be robust to daily-life settings. A recent study addressed this issue by providing proof-of-concept AAD using EEG electrodes positioned around the ear (i.e. cEEGrids) that would allow for mobile AAD, although the classification accuracy was lower than to a traditional EEG cap, potentially due to suboptimal electrode placement [150]. Another recent study reported that the EEG response from a single in-ear electrode, hearing-aid-compatible configuration provides valuable information to identify a listener's focus of attention [75]. However, application of a true online EEG-based AAD system has not yet been evaluated to date.

The aim of the current work is twofold. Firstly we demonstrate the application of the AAD in a real-time setup. Twelve subjects were recorded in an office environment with mobile EEG hardware and stimuli were presented through consumer-grade headphones. In addition, we apply a neurofeedback scenario to provide a proof-of-concept for a fully automated closed loop AAD system. We implemented an online AAD system with a time resolution of 10 seconds with a backward decoding approach. Secondly, longitudinal data was collected from two subjects (of which one received the neurofeedback) to evaluate the feasibility of longterm neurofeedback training on AAD. The EEG was recorded at home in a 24 minute session for seven days (spread over four weeks). Note that although we report a case study a large number of sessions were recorded over time allowing to investigate longitudinal effects that have not been investigated in the literature to date. Together, these studies describe the feasibility of in home AAD recording and the potential of online AAD feedback for improving the AAD abilities of humans. In this setting we decoded the subjects' attention with high accuracy and report on the influential factors for training of the AAD protocol. In addition, we highlight aspects in the AAD experimentation that need deliberation in order to validate and compare AAD results between studies. These results pave the way to further investigate how subjects might be able to increase the AAD performance and thus improve the applicability of AAD in general.

The chapter describes both studies in parallel. Unless specified separately, both studies share the same specifications. The initial experiment that was used to setup the closed-loop online system will be denoted as 'Study I' and the in-home longitudinal case study as 'Study II'. In general the chapter is structured as follows: the most important experimental details and setup are described in the first part of the methods section, followed by a description of the online feedback procedure (which deviated slightly between the two studies). The results section firstly reports on the initial study (study I) with emphasis on the closed-loop system and neurofeedback presentation. The latter part of the results section highlights the findings related to longitudinal AAD training effects (over multiple days) and the role of the audio stimuli used in the current study (study II).

7.2 Methods

7.2.1 Participants

Twelve native Dutch-speaking subjects (mean age (±SD) 22.4 (±2.1) years, six women) participated in study I. In study II two native Dutch-speaking subjects

took part, aged 21 and 26, both female. All subjects reported normal hearing and no past or present neurological or psychiatric conditions. All participants signed informed consent forms prior to participation. The ethics committee of the KU Leuven approved the experimental setups.

7.2.2 Data acquisition

The acquisitions (study I & II) were conducted with a SMARTING mobile EEG amplifier and 24 Ag/AgCl passive scalp electrodes (Easycap), as described in in Section 3.3.2.

EEG data were recorded through Openvibe [176] (and stored for offline (reference) analysis) and streamed from Openvibe to Matlab via the labstreaminglayer interface (LSL) [120]. The audio stories were played via Openvibe and preloaded into Matlab for the online analysis part and synced via the Openvibe audio triggers. Every ten seconds, the data was retrieved from the LSL stream and analyzed in Matlab. Both online and offline analysis used custom-made Matlab scripts.

7.2.3 Stimuli

The audio stimuli of Study I consisted of four fairytale stories in Dutch (of approximately 12 minutes" length), narrated by four different male Dutch (Flemish) speakers, each story was split in two parts (blocks) of ±6 minutes (8 blocks in total). Study II consisted of 32 blocks (of similar ±6 minute length) of Dutch fairytale stories (not overlapping with Study I) narrated by different Dutch male speakers. A total of 48 min stimuli were prepared for Study I and 192 min for Study II.

Silences in the audio were truncated to 500ms and audio waveforms' RMS were normalized per story (i.e. 6 minutes) similarly as reported previously by [42]. Subjects listened to two stories presented simultaneously through low-cost consumer headphones (Sennheiser mx475). The two audio streams were filtered by head-related transfer functions leading to more realistic perception [42]. Subjects were asked to pay attention to only one story on the left or right side. After each block, multiple-choice questions were presented and subjects indicated the difficulty in listening (i.e., indicated on a ten-point scale) and likewise their interest in the current story. Post-hoc we evaluated potential relationships between these user metrics and the AAD performance.

7.2.4 Procedure

An overview of the procedures of both studies are schematically illustrated in Figure 7.1.

Study I

The experimental setup consisted of two sessions of two stories (four blocks) each. After listening to one full story (2 blocks of 6 minutes), subjects switched attention to the opposite side, i.e. left or right speaker, to avoid introducing a lateralization bias in the decoder [42]. The first session was used for calculating the decoder (calibration), the latter for evaluation (testing). Half of the subjects received visual feedback (feedback group) on the laptop screen for the second session of the experiment (as described in section 7.2.6), whereas the other subjects received no feedback (control group). Order effects were avoided by alternating the listening sides among subjects. The study took place in an office–like setting while seated on a chair at a desk.

Study II

The experiment consisted of 8 sessions: 2 calibration (CS), 4 training (TS) and 2 follow up sessions (FUS). Similarly to Study I each session consisted of four 6 min blocks and ear attendance was switched from right to left at mid-session. The calibration sessions were used to calculate the decoder that could be used for online detection during training session 1&2. A subject-specific decoder was constructed for both subjects (to allow for equal comparison of the subjects' results). This online decoder was updated one time, for both subjects, after training session 2, based on the subject-specific data collected during training session 1&2. During training phase, one participant, further referred to as SF (i.e. Subject Feedback), was presented with feedback about ongoing AAD performance (see section 7.2.6). The other participant, SC (i.e. Subject Control), did not receive any feedback throughout the entire study, but participated in as many sessions as SF. During follow up sessions, neither subject was presented with NFB and subjects' offline AAD performance was evaluated to assess for long-term training effects. Sessions took place at the subject's home, in a comfortable seated position in the living room.

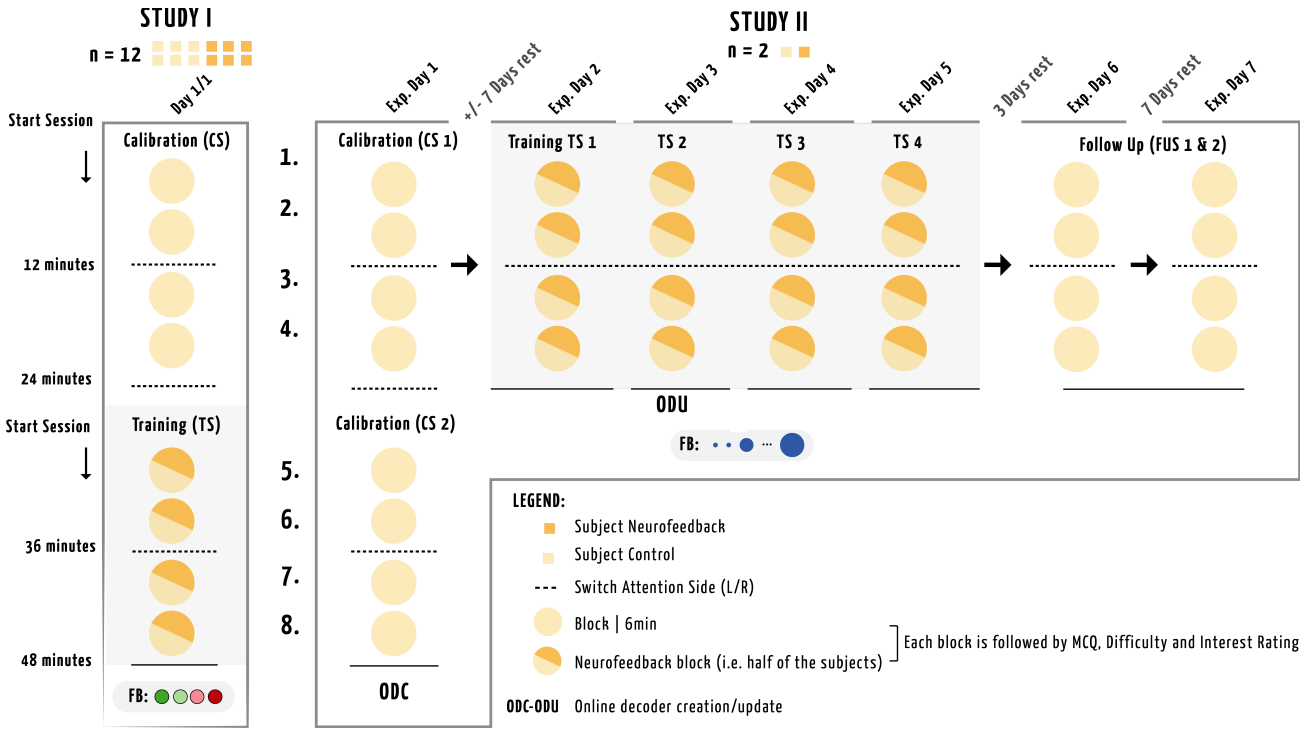


Figure 7.1: Overview of the recording sessions/days for both studies. Study I and Study II had different audio stimuli and each circle in the figure indicates one block of 6 minutes. After listening to one full story (2 blocks of 6 minutes), subjects switched attention to the opposite side, i.e. left or right speaker, to avoid introducing a lateralization bias in the decoder. Blocks in the shaded area represent that half of the subjects were presented with visual feedback (different per study) and at these blocks a pre-trained decoder was used based on the data of previous sessions. Study I consisted of two sessions of two stories (four blocks) each. Study II consisted of 8 sessions: 2 calibration (CS), 4 with feedback training (TS) and 2 follow up sessions (FUS).

7.2.5 Analysis

The following paragraphs provide further details on the closed-loop implementation of the AAD. It describes the preprocessing steps, decoder construction, detection (i.e classification) process and feedback procedure. These stages are schematically depicted in Figure 7.2.

Preprocessing

EEG data were bandpass-filtered at 1-8 Hz and consequently down-sampled to 20Hz as implemented in the EEGLAB toolbox [58] for each 10s segment of data. The absolute value of the audio waveforms with power-law compression with exponential 0.6 was taken to obtain the audio envelopes, and then an 8Hz low-pass filter was applied [21]. Envelopes were extracted from the clean audio signals for the separate speakers.

Decoder construction

Detection of auditory attention was based on the EEG-based method of stimulus reconstruction which was first presented by [165]. The goal is to reconstruct the envelope of the attended speech stream S based on neural recordings R via a linear spatiotemporal decoder D. In case of N electrodes, $R(t, n)$ represents the EEG signal recorded at electrode n at sample time t. A decoder $D(\tau, n)$ maps from $R(t, n)$ to the attended speech stream envelope $S(t)$ as follows:

$$\hat{S}(t) = \sum_{n} \sum_{\tau} D(\tau, n) R(t + \tau, n) \tag{7.1}$$

where $\hat{S}(t)$ denotes the estimated envelope of the attended speech stream and τ represents a time lag. $\hat{S}(t)$ is thus estimated from the EEG recordings at all N electrodes and time-lagged versions of those recordings. This way, information in the EEG at later time points is included in estimating the speech stream envelope at earlier time points. The decoder D is computed that minimizes the mean-squared error (MMSE) between the actual and reconstructed envelope. Analytically, D is found by:

$$D = C_{RR}^{-1} C_{RS} \tag{7.2}$$

where C_{RR} is the auto-correlation matrix of the EEG data (across all electrodes and time lags), and C_{RS} is the corresponding cross-correlation vector between the attended speech envelope and the EEG data (across all electrodes and time lags).

In both studies, six equidistant time-shifted versions of the EEG trial were obtained in a 50-300ms range after stimulus onset, i.e., τ ranges from 0 to 6 in equation 7.1. All EEG channels and their 6 delayed versions are then linearly combined using a pre-trained linear decoder following Equation 7.2; all the C_{RR}s and C_{RS}s computed over the training trials were averaged to create a single average covariance matrix and cross-correlation vector [21].

The decoders were trained per subject following an (offline) leave-one-trial-out structure. The decoders for the feedback parts of the experiment were computed solely on the calibration data of the previous session (Study I) or previous days (Study II, see 7.2.4) to allow for online classification. To make topographic plots of the decoder in Study II, single per-channel weights were obtained (for each session and for each subject separately) by averaging the decoder coefficients across all time lags which is similar to the procedure reported in [151]. The resulting topoplots are only used to compare decoders across different sessions or subjects, and should not be interpreted individually, as they do not necessarily highlight the activated brain regions, nor the importance of each channel in reconstructing the attended speech envelope. The latter would require an additional correction to take correlations across channels into account (see equations 25-26 in [20]), and the former requires a forward model instead of the backwards decoder used here.

Attention detection

For each test trial Pearson's correlation coefficient is employed to quantify the decoders' reconstructed envelope to the attended stimulus (CA) and unattended stimulus (CUA). The highest correlation value determines to which of the two speakers the subject was listening at the current trial. Offline analysis was done on trial lengths from 10s down to 1s. Real-time analysis was done solely on trial lengths of 10s. This trial length was based on results from a small pilot study (i.e. four subjects, not included in the current manuscript) conducted prior to Study I. The pilot study resulted in accuracies >80% for trial lengths of 10s which was deemed sufficient to provide accurate feedback to the subjects at an acceptable rate.

7.2.6 Neurofeedback Cue

The predictions of the AAD system were presented online to the subjects such that the feedback cue reflected the degree to which the neural signals of the current trial matched those in the decoder to detect the attended speaker. In our studies, we implemented two different versions of visual feedback on the center of the screen (to minimize eye movement related changes).

Study I

The six subjects who received feedback were presented with a colored circle in the center of the screen. After every test trial of ten seconds, the colored circle indicated the performance of the past ten seconds. Four different colors were used for performance indication: Dark red, light red, light green and dark green from worst to best respectively. Thresholds determining the colors were based on the training set (i.e. first session) in such a way that the CA and CUA difference would be equally divided for the correct and incorrect trials. This means that the median of all positive CA–CUA differences is used as threshold for dark green. Likewise, the median of all negative differences corresponds to the threshold below which the cue is dark red. Note that trials with a larger CA as compared to CUA (i.e., correct trials) are always green (light or dark) and the incorrect trials are always red (light or dark). Decoders for analyzing the neurofeedback sessions (i.e. second half of the data) were based on the individual calibration data as stated in the previous paragraph. Figure 7.2 illustrates the feedback cues for Study I, only one circle was shown for every trial in the center of the screen.

Study II

After evaluation of the feedback effects of study I, as presented in the Results section 7.3.1, the feedback appearance was modified to avoid potential effects of negative reinforcement penalty; after seeing a light red cue subjects more frequently performed incorrect in the subsequent trial. Therefore, incorrect trials (with CA< CUA) were not explicitly visualized in the feedback cue of Study II. Feedback was presented as a blue sphere positioned at the center of the computer screen. In case of above-threshold performance (CA-CUA>0), sphere radius grew proportionally to the performance measure (CA-CUA). In case of below-threshold performance (CA-CUA<0), the sphere did not grow nor shrink (with the subject not being explicitly informed what this meant). The experimental subject (SF) was instructed to maximize sphere size by focusing on the to-be-attended speech stream. Figure 7.2 depicts an example of the feedback cues used in this study.

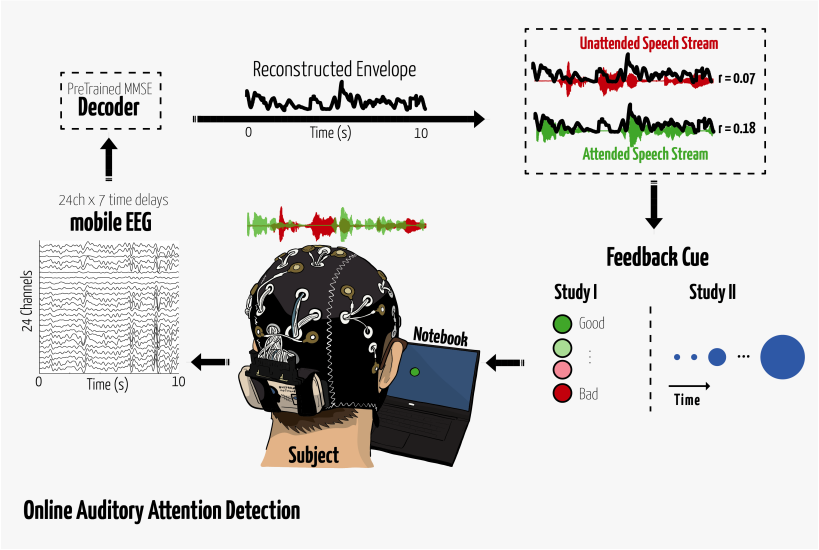


Figure 7.2: Illustration of the Auditory attention closed-loop analysis. Values depicted are for illustrative purposes only.

7.3 Results

7.3.1 Study I: Closed-loop online AAD

Classification performance

Grand average classification accuracy for the 10s trial length was 81.9% (SD =5.9%). Increasing the window length in the offline analyses to 60 seconds raised the accuracy up to 96.9% (SD=3.5%). All subjects scored above chance level (>57%) for the 10s window length.To investigate the potential for providing feedback at a higher temporal resolution we evaluated post-hoc the performance on trial lengths below 10 seconds. Figure 7.3 shows the classification accuracies for these trial lengths down to 1 second. It can be noted that there is a stronger decrease when moving from 3 to 2 and 1 second window length, although almost 70% of accuracy can be reached on average at a temporal resolution of 3s. For every trial length all subjects scored above chance level; chance levels were corrected for multiple comparisons (i.e. trial lengths) by a Bonferroni correction.

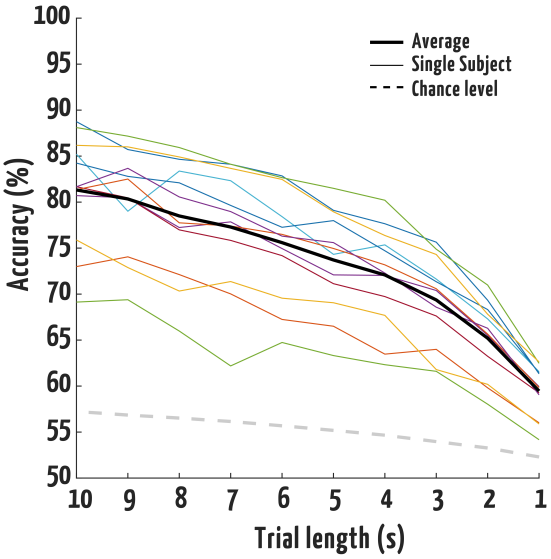


Figure 7.3: Average and single subject accuracies per trial-length as calculated post-hoc. The dotted line indicates chance level as corrected for multiple comparisons.

Approximation of spatiotemporal characteristics

Stable performance was achieved up to removal of 16 channels that contributed the least in the envelope-reconstruction model performance. With a reduced set of 8 channels, an accuracy of 78.8% (SD = 6.5%) was achieved. If we maintained only seven or fewer electrodes, the performance dropped significantly. This is depicted in Figure 7.4. To evaluate the most discriminative and most redundant channels, we plotted the average number of removal per channel, down to 8 channels. This is illustrated by the left topoplot in Figure 7.4. The red/yellow colors depict the channels that were removed without having a large influence on general accuracy. It can be noted that especially the frontal and posterior channels contributed the least to the decoder's performance. In contrast, the temporal electrodes (around the ear) are most important, as is depicted by the yellow colors in the right topoplot. The most discriminative single time-lag for the EEG was identified at 150ms (not depicted in the figure), resulting in an accuracy of 77.3% (SD = 8.2%).

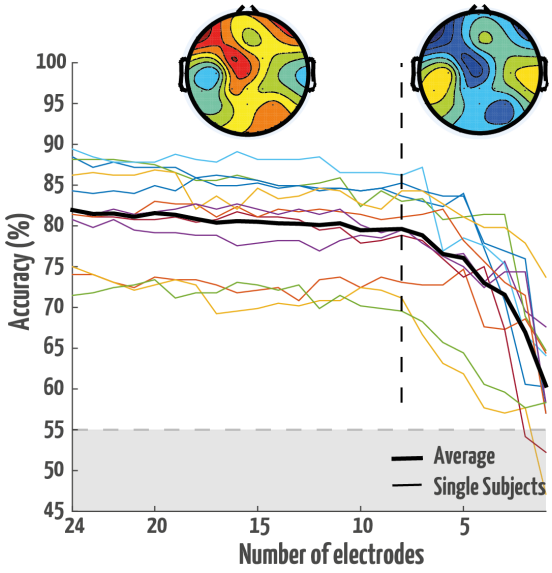


Figure 7.4: Grand average and subject specific decoding accuracies depending on the number of electrodes for the 10s window length. The channel with the lowest decoder weight (after correction for channel variance differences) is removed in each iteration. Topoplots represent the average distribution for the least discriminative channels on the left and most discriminative on the right. The shaded area indicates the chance level.

Grand average decoder

Individual differences in classification accuracy at the 10s trials length are depicted in Figure 7.5. Transferring decoders from one person to another can reduce the training time. Here we evaluated such approach by calculating a decoder based on the data from all other subjects. The grand average decoder resulted in lower accuracies on average of 71.8% (SD = 6.2%) as opposed to the subject-specific decoder 81.9% (SD =5.9%) which was deemed significant ($t_{11} = 11.49$, $p < 0.0001$). A personalized decoder resulted in higher accuracies for all subjects, although the effect of training time is not depicted.

Neurofeedback sessions

Average classification accuracy in the online decoding for the second half of the data, (session 2, see Figure 7.1) was 79.7% (SD = 7.0%) using the pre-trained decoder of session 1. When evaluating the offline leave-one-out decoder

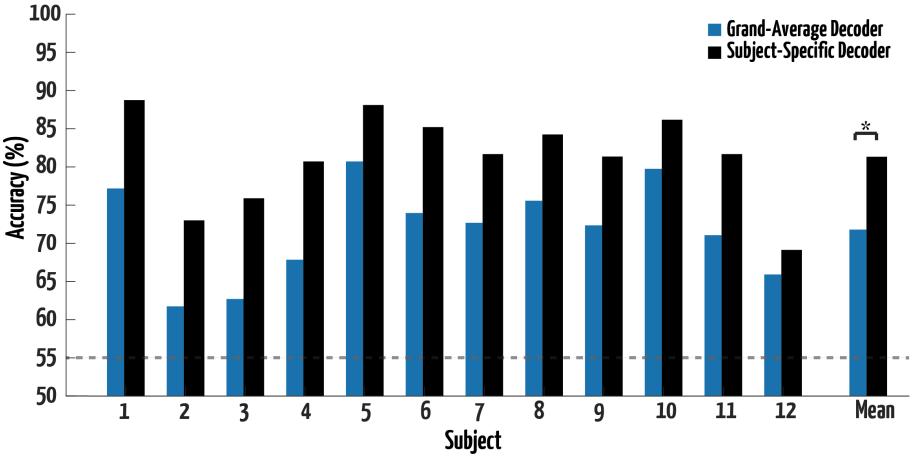


Figure 7.5: Subject specific accuracies for a personalized and grand-average decoder (obtained from other subjects) for a trial length of 10s. The significance indicated by an asterix represents a $p < 0.0001$.

for this part of the data, average accuracy was 83.0% (SD = 7.5%). This increase in accuracy was significant (two-sided paired sampled t-test, $t_{11} = 4.1$ $p < 0.01$), indicating that training data of the same session resulted in a small classification increase at the 10s windows. Differences in accuracy between the control and feedback group were not evaluated due to the high variability in AAD performance between the subjects (see Figures 7.3 and 7.5) and small sample size.

Comparing the accuracy within the feedback group between the feedback session and the calibration session revealed a slightly higher accuracy in the feedback session, +4,1 percentage points as calculated with the leave-one-out decoder when averaged over all subjects in the feedback group. Five out of six subjects scored higher in the feedback session, compared to the training session. For the group not receiving feedback, the difference between the second session and the training session was -0.2 percentage points. Surprisingly, no effect was found for increased or decreased CA or CUA changes in the feedback group with respect to the training session.

Relative occurrence of the four feedback cues was on average 41.8% dark green, 40.7% light green, 10.1% light red and 8.7% dark red. We evaluated the average (over subjects) temporal patterns at which the cues were evident in the feedback session. Entries in Figure 7.6 indicate, by the colors the relative frequency at which a colorcue in the rows was followed by the colorcues in the columns. It can be noted that, in the feedback session, participants shifted more frequently

to dark red after seeing light red. Contrastingly, after seeing a dark red cue, the following trial was less likely to be light or dark red - this may indicate awareness of the (very) bad performance and increased effort to compensate, although this remains speculation only. Note that the absolute CA-CUA difference for the light colors is smaller, compared to the dark colors. For example, light green denotes a correct decoding, but with less confidence than in the case of dark green (as described in section 7.2.6).

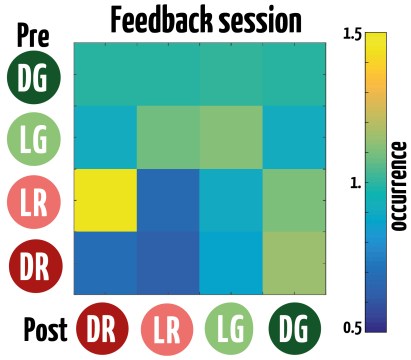


Figure 7.6: Adjacency matrices for the Neurofeedback session. The entry color indicates the normalized frequencies of a specific color cue (in the rows) that is followed by any other cue (columns). Cues: DG = Dark Green, LG = Light Green, LR = Light Red and DR = Dark Red.

User metrics

On average, subjects answered 84.2% (SD = 7.9%) of the questions correctly. We contrasted the number of correct responses of each subject to the general accuracy at the 10s window analysis. This revealed a strong positive correlation ($r = 0.72$, $p < 0.05$). One subject was removed, as its number of correct responses differed more than 2 standard deviations from the mean. Figure 7.7A displays the individual subjects' accuracy and number of correct responses. A regression line has been added for illustrational purposes. A moderate negative correlation ($r = -0.56$, $p = 0.059$) was found between the average accuracy and the subjects' reported difficulty in answering the questions and overall listening. This correlation is depicted in Figure 7.7B.

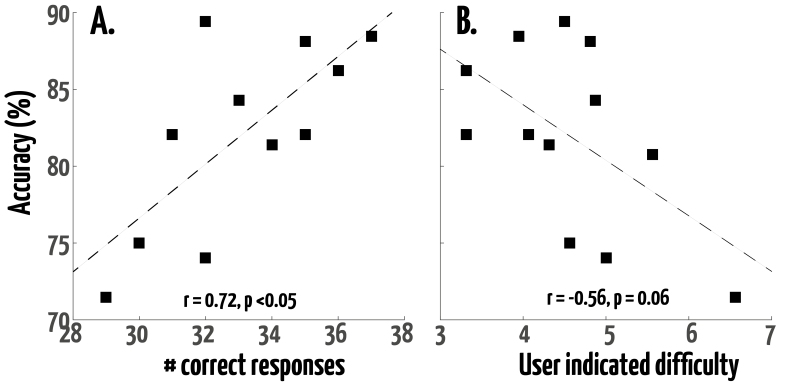


Figure 7.7: Scatterplots illustrating the correlation between the decoder grand-average accuracy and the number of correct responses after each story in A and the User indicated task difficulty in B. A regression line has been added for illustrational purposes.

7.3.2 Study II: Longitudinal recordings at home

Classification performance

Average offline AAD accuracy (and SD) across all days and sessions was 79.7 (±3.9%) for SC and 79.0 (±5.9%) for SF. Average online AAD accuracy across training sessions was 78.3 (±2.6%) for SC and 79.2 (±2.2%) for SF. Table 7.1 shows average and range of online AAD accuracies during the training at a block and session level, for both subjects.

Table 7.1: Online AAD accuracies on training blocks and training sessions: range and average (±SD) for both subjects.

Online		**Min**	**Max**	**average±SD**
Block	SC	72.2	88.9	78.3±5.7
	SF	66.7	88.9	79.2±7.0
Session	SC	75.7	80.6	78.3±2.6
	SF	77.1	82.0	79.2±2.2

Figure 7.8A shows online accuracy during training sessions (TS) and offline accuracy during calibration (CS) and follow up (FUS) sessions, for both subjects. Dashed lines indicate the on- and offset of the presentation of NFB to SF (left and right, respectively). No consistent performance increase can be noted

for either subject during the training days. These results would indicate the absence of a training effect (regardless of the neurofeedback). However, a strong correlation was found between the session performances of both subjects ($r = 0.81$, $p < 0.05$). This illustrates that a potential training effect might be masked by differences in the audio stimuli across sessions. Paragraph 7.3.2 highlights these results more in detail.

Comparing the relative performance of SF to SC removes the potential influence of the audio stimulus, and revealed a more specific effect of the feedback (Fig. 7.8B). While SF performed almost 6% worse than SC initially, SF became increasingly better than SC throughout training sessions (except for TS3). Furthermore, relative performance remained increased after training (at FUS1), but diminished over time.

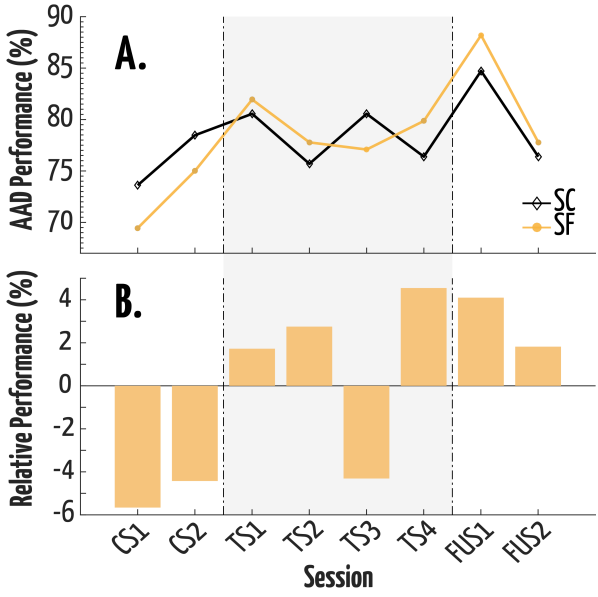


Figure 7.8: Evolution of a) session performance for both subjects and b) relative session performance of SF to SC. Dash lines indicate the on- and off-set of NFB (shaded area) to SF (left and right, respectively).

Relatively to SC, SF did not show a consistent increase in mean CA-CUA throughout training sessions. Figure 7.9 shows CA-CUA variance per session, for both subjects. Both subjects started with similar level of CA-CUA variance. While variance of SC was fluctuating around the initial variance, the variance of SF was consistently lower from the training sessions onwards. Relative CA-CUA

variance between the two subjects showed the exact inverse pattern of relative performance: consistently decreasing throughout training sessions (from -38% TS1 to -50% at TS4, with inconsistency at TS3), with a partly remained effect at FUS1 (-40%) that diminished towards FUS2 (-35%) (c.q. figure 7.9 and 7.8B).

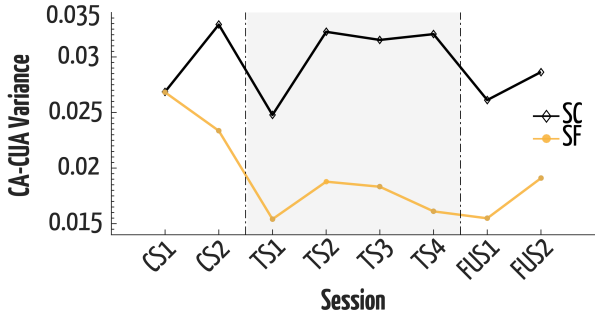


Figure 7.9: Evolution of CA-CUA variance for both subjects. Dashed lines indicate the on- and off-set of NFB to SF (left and right, respectively).

Decoder characteristics

Figure 7.10 depicts the spatial distribution of decoder weights (averaged across time lags, see 7.2.5) during the experiment, for both subjects separately. These topoplots show the decoder pattern as calculated on the indicated sessions. For visual comparison throughout the study, we averaged topographies per 2 sessions, this way the effect of the online decoders can be visualized. Note that during TS1&2 feedback was shown with a decoder based on CS1&2 and likewise TS3&4 with a decoder based on TS1&2 (See paragraph 7.2.4). Blue and red color indicate relatively low and high electrode weight, respectively. Note that colors are only relative within a topographic map, and dashed lines indicate the on- and offset of the presentation of NFB (shaded area) to SF (left and right, respectively). The spatial distribution of SF is more consistent as compared to SC throughout the experiment as is evident through higher and more stable correlations. For SF correlations between CS1&2-TS1&2 ($r = 0.66$, $P < 0.001$), TS1&2-TS3&4 ($r = 0.76$, $p < 0.001$) and TS1&TS2-FU1&FU2 ($r = 0.72$, $p < 0.001$) were deemed highly significant. These values indicate that during TS3&4 and to a lesser extent TS1&2 the subjects' spatial distribution resembled closely the decoder's pattern. During the follow-up sessions (FU1&2) patterns were still highly similar to the training days. These distributions for SC seem less consistent than that of SF based on lower correlations throughout the experiment resulting in lower correlations

which were to a lesser extent (marginally) significant, CS1&2-TS1&2 ($r = 0.30$, $p = 0.15$), TS1&2-TS3&4 ($r = 0.40$, $p = 0.51$) and TS1&TS2-FU1&FU2 ($r = 0.57$, $p < 0.05$)).

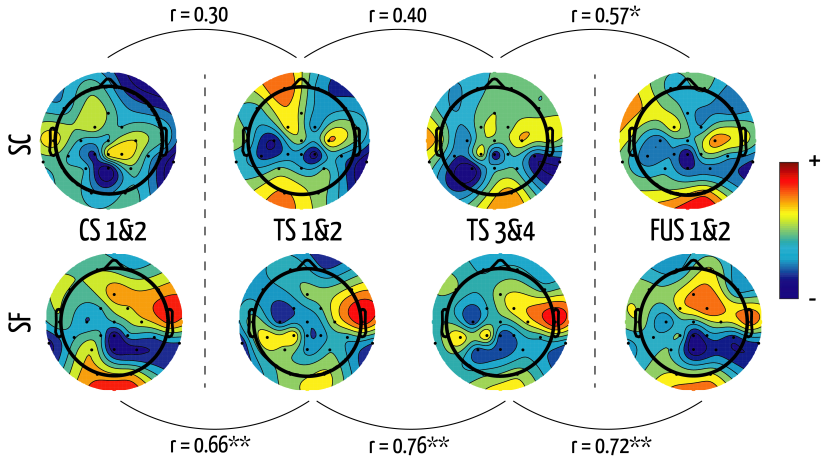


Figure 7.10: Topographic maps of the spatial distribution of decoder weights throughout the experiment, for a) SC and b) SF. (Red: high weight. Blue: low weight). Dashed lines indicate the on- and offset of NFB to SF (left and right, respectively). Correlations between consecutive spatial distributions are indicated, for both subjects. * denotes $p < 0.05$ and ** $p < 0.001$

User metrics

On average, SC correctly answered 86.7 (SD = ± 6.2%) of the multiple-choice questions that were asked during a session. For SF, this was 89.8 (± 5.7%). These values indicate that the subjects were successful in following the attended story. Average interest ratings per session were 5.8 (±0.8) for SC and 5.0 (±0.6) for SF. Average difficulty rating per session were 6.4 (± 0.7) and 4.7 (± 0.9) for SC and SF respectively. A negative correlation ($r = 0.46$, $p < 0.01$) was identified between the difficulty rating of the subjects and the block performance.

Audio stimuli

Finally we investigated the influence of the stimuli on the AAD performance measures in the current study. Due to the longitudinal setup of the study (i.e. >3 hours of unique audio stimuli, with 32 blocks per subject) it allows to look for effects that are related to the audio stories used. Both subjects

were presented with the same order of audio stimuli. Post-hoc we investigated the correlation between accuracy, CA-CUA mean and variance of the blocks, between both subjects. Figure 7.11 shows the block performances for both subjects throughout the experiment. A significant correlation was identified between the two subjects for block performances ($r = 0.46$, $p = 0.009$) and for mean CA-CUA ($r = 0.39$, $p = 0.029$, not shown in figure). In contrast, the CA-CUA variance did not correlate significantly between our subjects ($r = 0.24$,$p = 0.179$). Furthermore, we found a moderate correlation for mean CA ($r = 0.52$, $p = 0.002$), while mean CUA did not correlate significantly ($r = 0.25$, $p = 0.164$). These values indicate a dependency on the stimulus (i.e. story in each block) in which the AAD performance of both subjects were affected in similar ways.

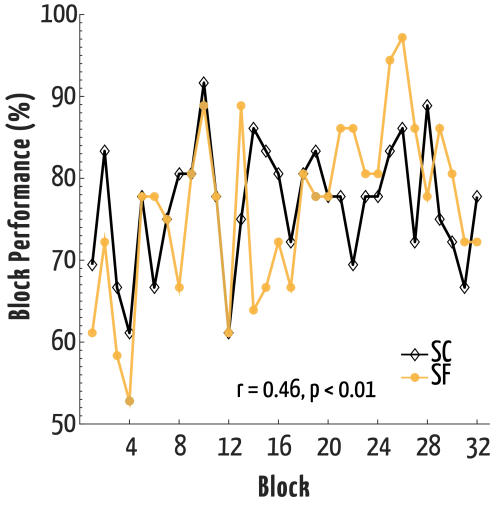


Figure 7.11: Block performances for both subjects and correlation between both subjects ($r = 0.46$, $p = 0.009$).

7.4 Discussion

In the current chapter, we evaluated the possibility of implementing an online closed-loop system for auditory attention detection and investigated the potential of providing feedback of the AAD performance to the subject. With window lengths of 10s, we obtained robust accuracies by detection auditory attention. We provided a fully working feedback system, and the online implementation did not significantly degrade the results. In the first study a slight improvement

in accuracy was observed when using neurofeedback; although the number of subjects was limited it is the first study reporting such effects. In the second study a more realistic scenario was established for long-term training; we evaluated the effects on various parameters of the AAD during training with feedback for several days. Changes were noted for the subject that received feedback training which had a positive effect on its AAD accuracy. The effect was mostly caused by a more stable neural response of the experimental subject to the auditory attention task (i.e. a more consistent CA-CUA). Due to the nature of our study, it cannot be determined with full clarity to what extent these effects were caused by true self-regulation of neural activity of the experimental subject; or whether these effects were caused by an increased motivation or effort of the experimental subject compared to the control subject. Overall, the two studies we presented achieved results that were similar and competitive to existing lab-studies, especially since we used mobile EEG equipment and natural recordings settings (i.e. office and home). This is particularly interesting for future application in other (long term) studies in real-life conditions. Moreover, the detected changes in AAD for the feedback subject during and after training (study II) suggest that changes in AAD may be achieved via training. Future studies should unravel the causes of these changes in more detail to pave the way for improved AAD. Finally, we revealed a significant relationship between the audio stimuli and the performance of the two subjects. Future studies should (re)examine and choose their audio stimuli with care to avoid the heterogeneity between speech stimuli observed in our Study II. A quantifiable approach to handle the stimuli or shared audio database could help in comparing between AAD studies.

The presentation of neurofeedback cues on screen showed overall, in both studies, no negative deflection in accuracy due to increased cognitive demands being in a dual-task scenario. For example, the visual cues could lead to increased distraction. Nevertheless, for the first study, subjects who saw light-red feedback were more prone to perform bad in the next trial as well. One explanation may be that when subjects became aware of an error (i.e. shift from green trial to light-red) this leads to a brief surprise effect which lowers the attentional response in the next 10 seconds. On the contrary, after seeing the worst option (dark red) subjects were less likely to perform bad on the next trial. This may have triggered an increased attentional effect or focus to perform better (since at the moment the subject cannot score worse). Faster feedback may be possible considering that at a trial length of 5 seconds above 70% accuracy (average of all subjects) can be obtained. This would mean that the subject sees twice as many feedback cues which may increase the potential of finding a true feedback training effect because the changes (and thus control of the training) are more instant. Alternatively, a sliding window approach (e.g. evaluate CA-CUA in the past 10s every x seconds) would allow more frequent updating of the feedback

cue. In addition, a paradigm in which the subjects switched more frequently the attentional side (i.e. left and right) mimics more closely a real-world application albeit more difficult for the user. This increased difficulty may increase the reliance on the neurofeedback. This factor is not well reflected in the current paradigm; subjects only switched attention after each story.

The true mechanism behind the changes of the AAD performance of the subjects that received online feedback remains up to speculation. Besides the aforementioned results and explanation, there is another area of change that may have occurred but remains hidden in the experiment; users may have adapted (un)consciously their listening behavior. In other words, different listening strategies, for example repeating words silently while listening, or focus particular on silences in between words. Experimental results of such changes are non existing to date and would be very interesting to investigate as such changes would be highly valuable for future users of auditory-assistive devices as these behavioral attention effects might boost the AAD analysis accuracy. Interestingly, the difficulty rating in both studies correlated to performance on subject (Study I) and block level (Study II). This might indicate that specific aspects of the stories lead to in- or decreased detection accuracy which would be an interesting line of future research.

The high accuracies in the present work were obtained with 24 electrodes and were found to be stable (not significantly different) up to removal of 16 channels in Study I (Figure 7.4). These results are in line with insights presented by [151] and are encouraging for future work in online processing and mobile EEG experiments. It would be particularly interesting to see an online implementation as presented here to be combined with the around- or in-the-ear electrodes as recently described in [150]. Moreover, meaningful accuracies were obtained even at window lengths as low as 3 seconds which opens up possibilities for fast feedback. To this end, it would also be interesting to see how these results relate to a state-space model, as the latter was shown to have a high temporal resolution in MEG recordings [4]. It is interesting that an increased AAD trend remains after training in Study II. Up to 10 days later, this trend could be noted in accuracy and variance of the signals (note that these effects seem to decrease from FU1 to FU2). The evaluation of neurofeedback effects lack follow-up sessions after a few weeks or even months. In real-life we would want long lasting (permanent) changes of the subject in relation to the AAD. In many neurofeedback studies the large amounts of recording sessions is seen as a hurdle. Nevertheless the option to use mobile EEG hardware facilitates more easily and convenient future recordings to overcome these limitations.

A general effect of performing repetitive AAD blocks, regardless of the presentation of feedback, could not be established due to the influence of the story stimuli. Being the first longitudinal AAD study many more stimuli

were used than previously reported in literature. This is likely to have caused increased variation in our stories that in turn influenced the performance. The significant correlation between the two subjects indicates these changes (i.e. in story content and/or acoustic features). They are not caused by a general increase or decrease (trend) as is evident from the highly scattered pattern in time in Figure 7.11A. Such variations in stimuli are not present in other auditory paradigms such as the P300 oddball (i.e. pure tones), or syllables without semantic meaning [80]. The latter holds potential to obtain complementary information on the segregating process of speech mixtures into distinct streams with respect to AAD studies that rely solely on natural speech stimuli. For AAD our results on the audio stimuli raise the question to what extent the stimuli in existing publications in the domain can be related to each other and maybe account for differences between studies (or within). Therefore we suggest to look into ways of quantifying more precisely the stimuli used for AAD. One influential factor may be the window length in RMS normalization. Shorter windows may lead to smaller differences between the two speech waveforms although it remains up to speculation to what these differences can be attributed; future work could look into the complexity and content of the stories, emotion levels, personal preference of the subject, among others. Another solution may be the sharing of stimuli, although most studies use stimuli in their subjects' native language, which makes this more difficult.

Another limitation in the current methodology is the lack of incorporating real-life audio signals, e.g. mixtures of speech with environmental noise. Recently there have been studies evaluating the effect of noisy reference signals and showing a negative impact on performance [214, 9, 164] Another aspect that remains unclear is whether the spatial location of the speaker will influence the classification accuracy and, to what extent a feedback procedure is reflected in the detection with multiple speakers. Finally, the nature of study II, being a case study with only two subjects, limits the generality of the results to certain extent. Nevertheless, promising results were obtained even in the limited setup paving the way for new experimental studies in the field of AAD.

All in all, we showed (to the best of our knowledge) the first closed-loop AAD system with a rather convenient recording setup and hardware. The described results are very promising to conduct future online experiments with (mobile) EEG hardware in order to obtain more robust conclusions about increasing AAD performance of the subjects through online feedback. Especially to investigate the effects of real-life distractions (e.g. cognitive load [242]) on the AAD in a cocktail party scenario would be very valuable as the brain signals in such scenarios may differ diminishing the application potential of existing AAD systems.

Acknowledgments

We would like to thank Annelies Baptist and Stijn Proesmans for their ongoing collaboration during the experiments described in the current chapter.

7.5 Chapter Conclusion

In the current chapter, we have implemented a fully automated closed-loop system that allows for convenient recording outside a lab environment to detect auditory attention in a dual speaker scenario. We achieved high auditory attention detection accuracies with a trial length of 10 seconds and provided subjects with online visual feedback on their ongoing performance. This exploratory study proves the feasibility of investigating the effect of neurofeedback in such a setting, paving the way for future studies.

Observed changes in the AAD have been linked to the presentation of the feedback process. Moreover these effects were observed a couple of days later as well. The degree of which this is related to the neurofeedback, increased motivation or attentional effort remains to be unraveled. Nevertheless, these results suggest that the AAD in healthy subjects is versatile (to changes). Future work should therefore focus on addressing the cause of these changes in more detail and opt for universal audio (speech) stimuli. If these changes can be made, AAD training could be highly beneficiary for future users and increase the application potential for subjects that seem to be less able to use the AAD paradigm at the moment.

Chapter 8

Conclusions and future work

Studying brain functioning in natural, real-life situations creates new challenges in terms of data acquisition, signal analysis and evaluation of cognitive models, as opposed to constrained lab recordings. One clear example of this is that free movement in real-life environments highly contrasts with restricted lab recordings; in the latter, non-task related disturbances are drastically minimized. The recent introduction of mobile EEG hardware means that new opportunities for out-of-the-lab recordings have emerged and researchers have hypothesized about applications of mobile brain monitoring in various domains. However, to date, still very few EEG experiments have been conducted in true out-of-the-lab settings. The effects of distractions, changes in cognitive load, physical engagement and subject behavioral variability in real-life scenarios are hypothesized to influence neurophysiological brain responses as described in traditional confined EEG experiments. Since auditory attention is very sensitive to the conscious effort of the subject and changes in the environment it may be a very valuable concept to investigate in unconstrained scenarios. Therefore, in the current thesis we focused on investigating the dynamics of auditory attention with mobile EEG in daily-life situations. We evaluated several aspects of EEG recording, analysis and interpretation that are of major importance for the usage of mobile EEG. Auditory attention detection with EEG was evaluated in two forms: attention to artificial tones, and natural speech. We investigated the feasibility of recording auditory attention in outdoor unconstrained experiments, and of disentangling the effects of increased distractions and physical engagement. Furthermore, we investigated the use of novel data-driven methodology to obtain meaningful information of the subject's performance from the mobile EEG in a more natural way. Another aspect of mobile EEG recordings is that they may enable at-home learning in an auditory context. To this end, a practical and promising application of auditory attention detection (AAD) is evaluated that is comprised of a fully working online AAD system, to increase the performance of auditory assistive devices. By means of a neurofeedback implementation we investigated the potential to train users in improving their auditory attention abilities.

In the first part of this chapter we provide the most important findings, and highlight future research directions that will improve the current work. Subsequently we put these views into perspective with the current mobile brain monitoring field. Finally, a note on the usage of mobile EEG technology in the (distant) future is given, and its implications for society.

8.1 Conclusions of the thesis

In chapter 3 we described three aspects that have a direct influence on the successful application of EEG paradigms to examine auditory attention in out-of-the-lab environments: data-driven signal-analysis, evaluation of the unconstrained recording scenario, and online AAD application. Here we summarize the main findings related to these aspects.

8.1.1 Data-driven classification

Recently, it was shown that auditory brain-computer-interface (BCI) classification can be performed in real life environments. The most relevant feature in BCIs with EEG is the P300 event-related potential (ERP) which is a positive deflection around 300-500ms in the EEG, produced as a response to both task-relevant and rare stimuli [169]. BCIs hold promise for use in assistive technology, especially if they are not restricted by the bounds of conventional laboratory scenarios [232]. However, the need for initial training of existing supervised classifiers on large parts of the data discourages practical application. Reducing calibration time of BCI analyses would benefit short-term interactions. More specifically, by removing the need for a supervised calibration phase, new users could potentially explore a BCI faster. In chapters 4 and 5 we aimed to remove this subject-specific calibration phase using tensor decompositions, and allowing direct classification of the P300 in a three-class auditory oddball BCI paradigm. As it is possible to naturally represent ERP data as a third-order tensor (i.e., channels $\times$ time $\times$ trials), it may be beneficial to exploit this multidimensional structure in the analysis. The existence of a specific spatiotemporal pattern which underlies the target trials, and which is not present in the non-targets, was elegantly extracted by means of tensor decompositions [243]. Canonical polyadic decomposition (CPD), and decomposition in multilinear rank $(L_r, L_r, 1)$ terms (LL1), were shown to derive meaningful P300 estimates that allowed the classification of single-trial ERP data in ways that are completely data driven [239]. This was illustrated on both simulated and real mobile EEG data of a three-class auditory oddball paradigm. More specifically, three separate

approaches were considered in order to capture meaningful signatures that display the P300 effect:

Data-driven P300 extraction We highlighted the natural ways to represent (mobile) EEG data as a higher-order tensor in Chapter 4. On mobile EEG data we were able to decompose the datasets in an elegant way in order to cluster target and non-target trials in a data-driven process. In both the time and frequency domain these models provided discriminative low rank terms related to the P300 ERP without prior information in any form. This work shows the simplicity and interpretability of CPD as an EEG data exploration tool. The main challenges involve identifying the most relevant component(s) and determining the number of components automatically. The extracted signatures of the CPD models are based on the structure in the dataset. Assuming a low-rank structure in the EEG which, in practice, is not apparent (e.g. due to high noise levels) is a potential pitfall; prior knowledge of the effect of interest is important for successful usage of tensor decompositions on EEG data.

Coupled decompositions In chapter 5 we showed the potential of structured data fusion for EEG analysis by incorporating experimental information of the datasets into the CPDs, which led to increased performance on the low SNR datasets. This was evident through coupling simulated datasets of various noise levels via structured data fusion and CPD while maintaining interpretable models. This coupling does not require additional label information, rendering it an elegant way to improve the data-driven extraction of the P300.

Singe-trial BCI classification In order to achieve single-trial classification, we extended the approach described in chapter 4 to incorporate templates that ease class identification (i.e., target or non-target). With incorporation of the baseline stimuli, which are usually ignored in the three-class auditory oddball paradigm, robust P300 based classification of real mobile EEG data was achieved in chapter 5. Furthermore, we showed that it is possible to increase classification accuracy by using LL1 models over CPD that allow slightly more variance to be modeled in the spatial and temporal mode. The results obtained from the tensor based models, for several subjects diverge from the standard reference LDA approach. This indicates that CPD/LL1 decompositions can provide, in specific cases, complementary task-relevant information to the LDA approach.

The previously described LL1 and CPD based classification approaches were applied post-hoc to the data of the outdoor bike study, i.e. dataset C (chapter 6). The CPD and LL1 decompositions for single-trial classification of the P300

target trials performed well for some subjects, but for most, these models did not reach the same performance as the (supervised) reference approach. This highlights the high variability between mobile EEG datasets. Consequently, it may be necessary to define differences in response patterns (e.g., P300 shape, timing and attentional level) of the subject on single-trial level to reveal non-stationarities and to quantify the degree of variability between sessions and subjects, e.g. [142]. These varying results motivated us in chapter 7 to focus on the understanding of the fluctuations in brain responses on single trials and also to focus on improving BCI paradigms in order to elicit strong responses, rather than merely improving classifiers, which has also been advocated by others recently, e.g., [52, 156, 202]. Nevertheless, even though data-driven tensor decompositions may not always result in the best classification of processes (such as the P300), compared to supervised approaches, they can provide additional information about the structure (e.g., spatio-temporal pattern over trials) that may be overlooked in cases when the data is unfolded to a matrix and analyzed solely with matrix based techniques. Similar reports can be found in the literature, e.g., [2, 153, 72]. Moreover, allowing instantaneous classification is inherent to faster usage of a BCI which has also been recently proposed by others, for example [14, 106]. Our work highlighted the need to look at the BCI datasets in multiple ways. Incorporating prior knowledge about the experimental design and usage scenario into the analysis can be very beneficial; the use of structured elements in the data has been employed in supervised classification as well, for example [14, 181].

Higher-order decompositions are not limited to third-order within-subject models. As mentioned in chapter 4, cognitive EEG data can be represented as higher-order data in many ways. An example of a tensor decomposition at study-level can be seen in Figure 8.1. This represents the CPD of a fifth-order tensor of subject-average data from dataset D's fifteen subjects. The subject $\times$ channel $\times$ time $\times$ stimulus $\times$ conditions are the dimensions from top to bottom in the figure and are of size 15, 24, 30, 3 and 3 respectively. The left component predominantly represents the N100 and the right is characteristic of a P300 response. The subject factor loading correlates to the subject-average classification accuracies, as obtained post-hoc with the regular rLDA method in all three conditions, as indicated by the scatter plots. Furthermore, in the other modes similar correlations to the findings of the conventional analysis can be noted: similar P300 shape; a lower P300 in the Move condition; a posterior focus of the P300; and opposite sign for the target stimulus as opposed to the baseline and non-target response. The compactness and simplicity of the CPD makes it an interesting tool which can be used to explore EEG data in a data-driven way; this has also been suggested by others [38].

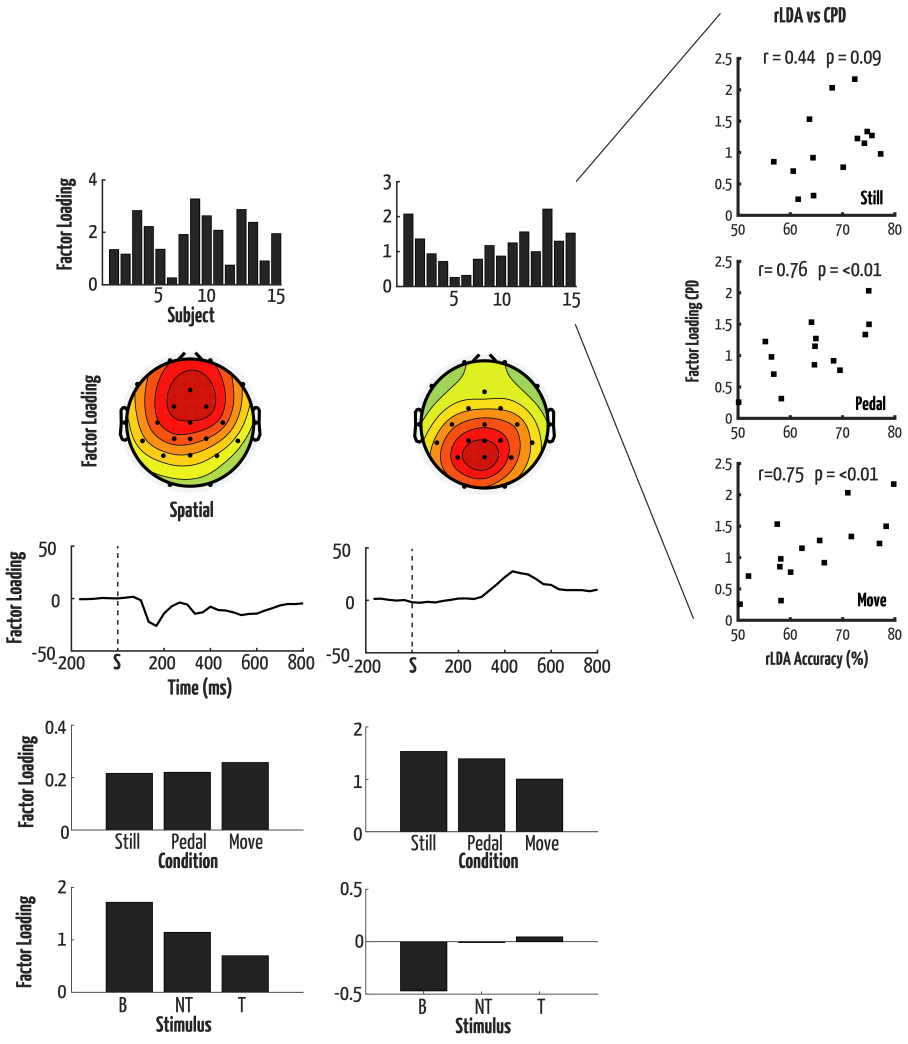


Figure 8.1: Decompositions of fifth-order EEG tensor that represents the subject average data of the bike experiment from chapter 6 (Dataset D). The subject × channel × time × stimulus × conditions are the dimensions from top to bottom. The left component predominantly represents the N100 and the right is characteristic of a P300 response. The subject factor loading correlates to the subject-average classification accuracies, obtained post-hoc, with the regular rLDA method in all three conditions as indicated by the scatterplots. In the other modes, similar trends to the findings of the conventional analysis can be noted (e.g. P300 shape, lower P300 in the Move condition, posterior focus of the P300, and opposite sign for the target stimulus as opposed to the baseline and non-target response).

To summarize, we conclude that flexible tensor decompositions allow for interpretable low rank approximations of the P300 effect in auditory mobile BCI data. The interpretability of the CPD/LL1 decompositions ensures that the models rely on neurophysiological features. The latter has not always been inherent in supervised classification approaches that strive to achieve maximum classification accuracy in BCI; it may be possible that obtaining the best classification accuracy in one session (or condition) might lead to less generalizability across sessions (i.e., overfitting), thus lowering the validity of the obtained results. Nevertheless, the importance of the classifier models' interpretability is gaining interest as is evident in recent developments of interpreting advanced classification models, for example deep neural networks [154]. Additionally, new evaluation metrics as an alternative to classification accuracy for BCIs have been proposed recently [133].

8.1.2 Real-life EEG recordings

Recently there has been a growing interest in studying brain function in natural, real-life situations. However, studying the human brain in real-life environments introduces several uncontrollable factors that could be different in a controlled lab scenario. Numerous distractions could be outlined that compete for attention with the task at hand when recording outside. Furthermore, in unconstrained experiments, subjects are free to adopt any preferred posture and move naturally, which requires additional mental effort that is usually avoided in lab-recordings. These processes are expected to reduce the performance of classification techniques described in chapters 4, 5 and others as described in the BCI literature. In chapter 6 we aimed to disentangle the influence of cognitive demands and distractions that arise from such outdoor unconstrained recordings. We evaluated the ERP and single trial characteristics of a three-class auditory oddball paradigm recorded in outdoor scenarios while pedaling on a fixed bike or cycling around freely. We also evaluated the trial-specific motion artifacts through independent gyro measurements and control for muscle artifacts. A decrease in P300 amplitude was observed in the free cycling condition, which had a direct influence on classification accuracies, compared to the fixed bike conditions. Nevertheless, above chance-level P300 single-trial classification in highly dynamic real life environments while cycling outdoors was achieved. Certain significant artifact patterns were identified in the free cycling condition, but neither these, nor the increase in movement (as derived from continuous gyrometer measurements), can explain the differences in classification accuracy and P300 waveform differences. In addition, a decrease of alpha power was identified in the free cycling which may be indicative of increases in task difficulty and mental workload. Our work has led us to conclude that the increased cognitive load in real-life scenarios plays a major role in the observed differences.

Our findings suggest that auditory oddball results measured in natural real-life scenarios are influenced mainly by an increased cognitive load due to being in an unconstrained environment. The head motion, as recorded through the gyroscope, did not influence the sensitivity of the analysis in detecting the P300 response. The ability to measure reliable ERPs in outside cycling conditions has been confirmed recently by others as well [184]. This may be a surprising finding as motion artifacts were seen as the biggest hurdle in general for mobile EEG experimentation [144]. However, the effect of recording in out-of-the-lab environments is highly variable depending on the setting, subject and electrode location [112]. This was confirmed in preliminary data on head and body motions that we recorded at the beginning of the PhD trajectory as presented in Appendix B. Fourteen subjects were asked to perform various motions with the mobile EEG system in place, as used in the recorded datasets for this thesis (the data in chapters 6 and 7. The central channels seemed less affected by motion artifacts as has been shown in the literature as well [112]. This suggests that motion artifacts are not always an insurmountable problem in mobile EEG recordings; such reasoning is also evident in recent literature [187]. For the bike data (dataset D), we applied CCA for the removal of EMG which lowered slightly, but consistently, the accuracies and P300 peak amplitude in all conditions. One could argue that the application of such automatic artifact removal techniques reduces the performance of the BCI. To this end we applied post-hoc several freely available (semi-)automatic artifact correction methods to examine the effect on the P300 amplitude and classification accuracy; this is presented in the second part of Appendix B. For all automated techniques, a decrease in either (or both) P300 peak and classification accuracy was observed. This suggests that traditional methods are not able to effectively remove only the noise, nor does a single solution exist for artifact handling in EEGs, which is in line with recent views in the literature [149, 112, 212].

Besides the influence of artifacts, performance of the BCI can be hindered by poor generation of task- related ERPs such as the P300 response. It is known that the P300 can vary between subjects and conditions [169]. The neurophysiological response varies with the focus of the user, especially in top-down attention based paradigms such as the three-class auditory oddball paradigm [231, 156, 17]. The results obtained in chapter 5 and 6 illustrate that it is very likely that an absence of P300 attention related response is the main cause of the upper bound of the classification accuracy. Diminished P300 responses during engagement in (mental) activities have been reported in the literature as well, for example during steering [185]. These findings highlight the need for further research into motivation, behavior and engagement of the subject, which are typically ignored in traditional EEG analysis. In addition it would be beneficial to investigate whether subjects can be trained to achieve more robust EEG patterns as shown, for example, in generating more robust

P300 responses after training [17].

To conclude, our fully unconstrained application of auditory BCI on a bike ride outside, paved the way for recent studies that evaluated ERPs in a similar way to our experiment, for example, investigating the effect of complex movements [183, 187] and outdoor studies such as [184]. The latter study achieved a similar BCI performance as our own work described in chapter 6 during outdoor cycling with a mobile EEG setup. Moreover, they confirmed our findings that the variations in P300 responses are not principally related to artifactual influences, and this was also hypothesized in earlier studies, e.g. [49]. These studies strengthen our conclusions that higher cognitive load and distractions due to being in a real-life environment influence the neurophysiological responses.

8.1.3 AAD system

The auditory oddball BCI measures auditory attention to artificial sinusoidal tones, which requires a relatively artificial mapping between stimulus and output, i.e., it does not incorporate semantics or language related aspects of the sound. A more realistic acoustic scenario is considered in the auditory attention detection (AAD) paradigm in chapter 7; attention to natural speech is detected in a cocktail party scenario. Such situations require stream segregation to disentangle the multi-speaker input into separate and comprehensive sound sources. AAD has several interesting applications, for example in auditory assistive devices (e.g. hearing aids) to detect the sound to which the user is attending. However, a large inter-subject variability shows that not everyone is able to achieve high performance (c.q. results of study I in chapter 7 or [151]). Being able to train subjects in achieving high AAD performance would greatly increase its application potential. In order to do so, an acceptable temporal resolution and online implementation are both essential prerequisites. Consequently, users of an online AAD can be presented with feedback about their performance. In chapter 7 we described two studies that demonstrate the feasibility of achieving online AAD, and we consecutively investigate the effects of AAD feedback. In the first study, we implemented a fully automated closed-loop system that allows for user-friendly recording environments. Subjects were presented online with visual feedback on their ongoing AAD performance. Following these results we implemented a longitudinal case study in which two subjects were presented with AAD sessions during four weeks. The results prove the feasibility of a fully working online (neuro)feedback system for AAD decoding. The detected changes in AAD for the feedback subject during and after training suggest that changes in AAD may be achieved via training. However, given the small sample size, caution must be applied. Nevertheless, this is early evidence of such training effects paving the way for future studies to

evaluate training of AAD in more detail. An interesting side-result emerged due to the need to use a huge number of sessions; a significant correlation between the stimuli (i.e., acoustic stories) and AAD performance was identified.

The online AAD setup was the first system to close the loop of AAD analysis. With the presentation of feedback, the subjects were given a sense of control. Furthermore, application in an online fashion allows for the use of AAD in a BCI format; focusing the attention on a specific speaker can be coupled to a command. In addition, the recording scenarios were out-of-the-lab as they took place in an office setting and at the subject's home, which is promising for practical usage.

Meaningful behavioral metrics were obtained in both AAD studies (see chapter 7). Significant correlations were found between the average accuracy, and the subjects' reported difficulty of listening and level of interest for the attended audio story. These measures gave rise to a preliminary hypothesis regarding the AAD experimentation: the lower performance during stories that were found to be less interesting suggests that the content of the stories influenced the AAD performance in which motivation or engagement acts as a confounding variable. In chapter 6 in the bike study the behavioral measures were shown to be correlated to P300 aspects, and we could rule out the possibility that a decreased P300 effect was caused by lower engagement, as this was contradicted with the behavioral results. The focus on various aspects of the subject (i.e., personality, cognitive abilities among others) to obtain complementary information about the AAD has gained interest in the BCI field as well, e.g., [102, 101]. The AAD paradigm and the P300, as captured in the auditory oddball paradigm, are both related to the subject's focus of attention. This raises the question of whether the performances of these paradigms are related. To this end we conducted an exploratory study. This work has not been described in the previous chapters but is listed as dataset F in Table 3.1 and outlined in Appendix C. This preliminary data provides, for the first time, evidence of a relation between the AAD paradigm and the three-class auditory oddball paradigm at subject level. Ten subjects were presented with 2 AAD sessions (similar to the ones described in study I of chapter 7) and 2 three-class auditory oddball sessions (identical to the setup used in chapter 6). The results show a significant linear correlation between the performance of the three-class auditory oddball task and the AAD performance ($r = 0.65$, $p < 0.05$), the scatterplot is shown in Figure 8.2. This suggests that both tasks, at least to certain extent, measure the same auditory attention abilities.

The work described in chapter 7 only investigated the focus of attention on two simultaneous speakers in a stationary setting. Therefore, the auditory attention dynamics in more complex situations, i.e., with multiple speakers or in busy environments, remains unknown. Moreover, outside recordings that allow

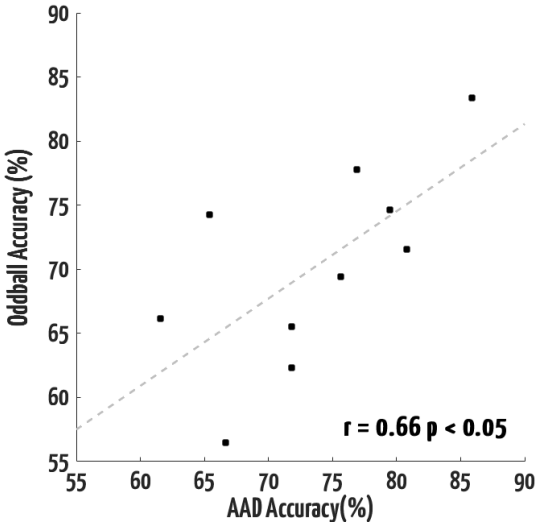


Figure 8.2: Correlation between the session performance of AAD and the three-class auditory oddball paradigm of Dataset F.

for free movement of the subject were not considered; the influence of motion artifacts on the AAD is not known. The picture of mobile AAD analysis is thus still incomplete, nevertheless, our work is the first step towards the application of online AAD and understanding the potential effects of feedback in improving the AAD.

8.2 Future research perspectives

8.2.1 Methodology

Extending the tensor analysis

In the previous section we highlighted the use of tensor decompositions for data driven classification. Future work can focus on expanding the models to incorporate additional structure, e.g., the simultaneously decomposed datasets in chapter 5. Recently, the acquisition of EEG data became accessible in a real-life scenarios, for example as presented in chapter 6. This enables new ways to apply EEG to larger numbers of people. In chapter 5 we illustrated the potential of coupled CPD models for the analysis of simultaneous EEG

recordings (e.g., multi-user recordings [118, 171]) in which the trial mode is shared between subjects.

Applying CPD or LL1 decomposition in an adaptive way so that it can update only the necessary structures, is expected to further improve the results. A model can be initialized prior to the session, or on only a few trials to start with, this allows exclusion of old data when non-stationarity is occurring. Such decompositions can be derived in a time- and memory efficient way by recent advances on adaptive tensor decomposition algorithms [216]. Handling of non-stationarities (e.g. due to fatigue or changes in cognitive load) could increase the application potential of such methods as EEG based BCIs are known to suffer from these changes in data recordings over time [188].

Besides a stand-alone application, the tensor decompositions can also be used as part of a supervised approach. It could be applied as a feature selection method, or artifact detection [72]. The latter removes non-relevant sources of variance before further in-depth feature extraction is performed. Another interesting approach would be to remove data samples that are likely to be contaminated by artifacts, in order to obtain a more reliable decomposition. Removal of specific channel or time-points from the EEG tensor still allows the derivation of meaningful components. The missing entries in the tensor are simply ignored in the model, putting more emphasis on the relevant (i.e., non-artifact related) activity. However, one needs to identify the noise patterns, which are often not known and very difficult to process. Nevertheless, if this approach turns out to be successful it would be an important finding, since single-trial EEG data are often contaminated with much noise.

General tensor framework

The ultimate goal of the current thesis is to reveal the dynamics of auditory attention in response to acoustic stimuli in real-life ambulatory circumstances. Drawing from the previous chapters, we have a proper data interpretation for the datasets at hand. However, for the interpretation of single subject performance we are restricted to the real-life influences (i.e., both environmental and cognitive) that took place during the actual recordings. As a result it is not possible to directly generalize our findings to other real-life scenarios for which no EEG data is available. A single subject data collection covering all influential factors would drastically increase the reliability of the analysis and interpretation. Since we can derive the complete personal effect of all influences we can increase the user performance of for example BCIs, or AAD assisted hearing aids. Unfortunately, for obvious reasons it is not possible to acquire data from 'all possible' real-life scenarios. In fact, experimental data is usually limited to a small subset of all variables that could influence user performance.

More data will improve the single subject information about auditory attention so that (our) methods can be better adapted to the extensive cognitive and environmental changes present in real life.

Since we cannot collect data in all conceivable situations, future work should aim to increase the interpretation potential of the recorded experimental data in a novel, methodological way. Estimates of how a subject would perform in an unexplored environmental setting would greatly improve the extrapolation possibilities and impact of the acquired results from the previous chapters. In order to achieve this, we propose a novel way of structuring the P300 waveforms (or likewise AAD related effects) in a tensor according to the level of many factors such as user performance, distractions, head motion, among others. A schematic drawing of this tensor is depicted in Figure 8.3. Each entry represents the average P300, although similar tensors can be constructed for other characteristics as well (e.g., classification accuracy). This way it may be possible to obtain a better understanding of how each factor, namely environmental audio level, motion level, and cognitive distraction, influence the waveform of the ERPs. Structuring the data this way has two advantages above interpreting classification accuracies alone. Firstly, we can apply a tensor decomposition on this higher-order array, which will allow us to investigate the relations between the factors simultaneously. Disentanglement of the factors allows for visualization of the underlying relation. Understanding the way these factors interact provides future studies with the ability to focus on these most influential aspects. Secondly, SDF can be applied in order to estimate missing data in our waveform-tensor. Since these models are able to handle tensors with only a few known elements [220], the line of research may be a valuable one. The structure can be extended to include results from all subjects as well. Estimation of the entries would be based on intra- (e.g., performance on related factor levels) and inter-subject (e.g., gender, age) properties. Assuming similar properties across subjects and factors it is possible to obtain ERP waveform estimates for a combination of experimental manipulations which were not recorded with the EEG (i.e., empty tensor entries). The possibility of estimating these missing entries can provide meaningful information about subjects and experimental factors without the need for expensive and time consuming data acquisition. This way we obtain general signatures that would relate to the factors used in constructing the tensor. It allows for faster and improved training of signal analysis models, as we have more (artificially generated) subject-specific information available. Note that the type of data described by the framework is not limited to the P300 feature; other paradigms and/or ERP components can be unraveled as well.

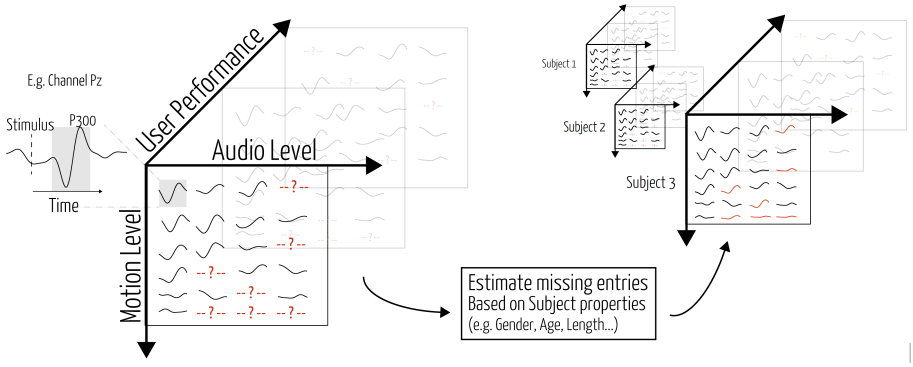


Figure 8.3: Left: Example of a single-subject tensor structure in which each entry represents the ERP for a specific combination of factors as illustrated by the three axes labels. Entries with a *?* are unknown, due to a lack of EEG data. Right: Based on the intra-subject properties between the various factors, and the similarities between subjects (subject properties), SDF will estimate missing elements in the tensor.

8.2.2 Mobile Recordings

Multimodal recordings

The value of mobile EEG as a method to obtain information about auditory attention has been outlined in chapters 6 and 7. The combination of EEG with other measures allows for improved interpretation of the neurophysiological processes of interest. For example, in chapter 6 we showed that the inclusion of gyroscope data allows us to investigate the relation of head motion to the P300. Similarly, in chapter 7 behavioral questionnaires revealed interesting links between user reported interest and difficulty to perform the task in response to the auditory attention performance. Future experiments are very likely to benefit from such data-fusion of various modalities to obtain additional insights in the mobile EEG data.

For outdoor studies, the incorporation of GPS data would provide information about the environment and, indirectly, about speed and steering. This is one of the limitations in the bike study described in chapter 6. Others have used GPS information to obtain information that allows investigation of the mental state in relation to the environment in urban and nature areas [12]. Likewise, another useful addition to outdoor studies would be to investigate physical engagement besides gyro and accelerometer data, with the incorporation of ECG recordings,

as this was shown to be related to EEG fluctuations as well, e.g. [162].

Perhaps the most informative addition to the aforementioned modalities could be achieved by coupling information of the visual domain to the work presented in the current thesis. Eye blinks [227, 228], eye movements and pupillometry [31] can be very informative of cognitive processes such as cognitive load and distractions, this has been proposed by others [124]; especially when recording in real-life environments EOG activity should not be seen merely as artifacts. We noticed an increased amount of eye blinking in the real cycling scenario compared to the other two scenarios (chapter 6). Interestingly, from our preliminary study, i.e., dataset F, it was found that during the auditory oddball, blinking was significantly less frequent as in the AAD sessions. More complex eye movements, e.g., gaze information, could not be obtained from our EEG recordings. Therefore we highlight the importance of merging eye tracker information with mobile EEG in future work to allow more detailed analysis and interpretation of the subject's cognitive state and intentions. In a similar fashion, the combination of mobile EEG with wearable fNIRS (as briefly explained in chapter 2) holds the potential to become a novel neuroimaging tool of cognition in natural environments. Recent reports describe novel hardware to enable wearable hybrid EEG/fNIRS [226]. The complementary character of fNIRS to EEG has been shown in several lab experiments, for example in improving the detection of driver drowsiness [159].

Hardware developments

Mobile EEG recording through a cap that covers the complete scalp is not very fashionable, and this should not be overlooked in the adoption of mobile brain monitoring. The aesthetics of the system are important for daily users, although in clinical usage these constraints are less important. Currently there are several types of mobile EEG hardware available that can be used for research and application, for example the commercially available Emotiv was shown to be useful for BCI [215], but also to be inferior to a medical grade system [65, 66]. Soon after, other solutions demonstrated that mobile EEG devices can deliver waveforms comparable to traditional EEG systems [50].

The next step in EEG hardware comes with the usage of miniaturized and customizable electrode layouts. More and more research is conducted in the field of unobtrusive EEG which was coined 'transparent EEG' recently [24, 25, 55]. Flexible electrodes behind the ear, so-called cEEGRIDs, have been shown to be very promising for auditory monitoring in AAD [150]. The authors performed an AAD task as described in chapter 7 albeit offline. The results showed the feasibility of recording around the ear with minimal obstruction for the user. This system uses the same SMARTING hardware as used in this thesis. The

miniaturized electrodes behind the ear were shown to correlate highly to those of full scalp EEG [26]. However, another study observed a lower SNR in the cEEGrids for the detection of ERPs, especially compared to occipital and central cap-EEG positions [166]. In addition to around the ear, in-ear electrodes have also been employed. They were shown to provide valuable information which was used to identify a listener's focus of attention [75]. Other studies found a large similarity between in-ear and scalp EEG, especially between the channels on the temporal regions [146]. All in all, the work described in the current thesis can be evaluated with the next generation of mobile EEG hardware. Nevertheless, the effects of motion and real-life environments have not been evaluated to date. Furthermore, processes that cannot be clearly detected at temporal electrode positions are likely to show a significant drop in performance. Future work should therefore look for application-specific hardware solutions, unfortunately this also requires case-specific research and evaluation.

8.2.3 Applications

Most of our focus in the current thesis was on auditory attention, but the implications of mobile EEG extends to many studies and applications. Here we highlight three cases that directly profit the described work and findings.

Auditory assistive devices

Although conventional prosthetic hearing devices can affect spatial selectivity and noise reduction by utilizing microphones that have directional capabilities, they cannot help with the segregation of speech-based systems, especially in surroundings with reduced spatial cues (such as in many restaurants or bars). Even if prostheses could filter sound sources to an accurate level, (which is a difficult and engaging computational process), we still have the issue of not knowing which of the filtered sounds the listener wants to focus on, and thus have the prosthesis amplify. Additionally, there is a need to examine all aspects of cognition, since hearing aid users can endure suffering due to an excess of cognitive demand. Chapter 7 described a working AAD using mobile EEG. Online detection of the attended audio source holds potential for use in assistive audiological devices, such as hearing aids [214, 151] (e.g., to steer a beamformer to the attended speaker). This is especially useful for humans who often encounter problems in the aforementioned cocktail party scenario. Extending AAD to online decoding could enable the combination of hearing devices with cognitive state decoding, and this would allow the hearing aid device to adjust flexibly to the demands of the listener. In addition, the application of auditory attention detection may not be limited to hearing aids.

In a more industry-oriented approach, AAD could provide new information in the marketing and advertising industries, for example to detect the type of messages people are paying greater attention to. Furthermore, there are also other BCI applications for which the auditory attention focus of a user may be used to control a device. As an example, one application could be for immersive virtual reality scenarios, integrating neural feedback to adapt the virtual world based on the focus of the subject.

Future studies on the topic of AAD are required to evaluate the meaning of the acoustic stimuli with care, in order to control for the effect of interest in the content of the attended story. A very interesting line of research would be to opt for universal speech stimuli as suggested at the end of chapter 7. Additionally, investigating the similarities between natural speech stimuli and systematic stimuli can increase the understanding of attentional focus, for example to discriminate between sound segregation and foreground/background separation c.q. the results of chapter 7 and [80].

Furthermore, future work on AAD should investigate similar problems to those addressed in chapter 6, i.e., the increased number of distractions and higher cognitive load in real-life environments. Being in a naturalistic environment increases the cognitive load and therefore limits the resources available for the AAD. The application of AAD in hearing aids is indissolubly linked to usage in daily out-of-the-lab environments. Therefore it is of utmost importance that similar neurophysiological patterns are present in these scenarios—something that has been avoided in the literature to date. We made a first attempt to investigate the effect of visual distractions and physical engagement on the AAD. This preliminary data was recorded while the subject was pedaling on a bike simulator and was watching a real-time simulated environment that was controlled by the pedaling and steering of the apparatus. A short description of this experiment, i.e. Dataset F, is given in Appendix 8.2.3. In short, the AAD performance in the most complex and distractive environment was only marginally lower, compared to a simple environment. Moreover, navigating through a virtual city and pedaling on a fixed bike still allowed for adequate performance in AAD, an average of 75.4% accuracy across subjects. Future experiments should consider the use of virtual reality (VR) to create a high level of immersion in the digital world. This may be one solution to bridge the large gap between lab and real-life environments, e.g. [209], especially since the adoption and quality of VR has increased tremendously in recent years.

Neuroergonomics

The feasibility of recording outside the lab and measuring cognitive load grants new ways to use mobile EEG. Neuroergonomics is a novel field that investigates

human mental states and workloads in order to improve the reliability of human performance. Ultimately, neuroergonomics improves the mental well being of humans in various environments. A recent example of the application of mobile EEG to assess mental workload and improve the focus of workers in workplaces can be found in [145]. Similarly, mental fatigue was evaluated using mobile EEG without disturbing the natural work situation [228]. These are but a few examples of the rapidly expanding field of neuroergonomics,; for an overview of this field we refer to [141]. In our view, the monitoring of auditory attention with mobile EEG is a valuable addition to the current studies. On the one hand it can be seen as a measure of workload, as it is sensitive to distractions and to the activity of the subject (e.g., decreased P300 in highly engaging scenarios as found in chapter 6). On the other hand, the (online) AAD analysis can serve as a direct measure of a subject's ability to follow audio messages in multi-speaker scenarios. For example in work places that require high alertness to incoming sounds, e.g., air traffic controller, the AAD can be employed to detect diminished attention to the salient speech stream.

Consciousness detection outside the hospital

Due to the historically cumbersome and impractical procedure of EEG monitoring, EEG has never been included as a part of routine health checks, and has also never been included in systematic research on human health status [144]. Reliable mobile data acquisition and attention detection, as shown in the current thesis, allow the expansion of this application to the database of patients that are not able to visit hospitals frequently. The detection of consciousness through assessing the auditory attention of patients in a vegetative state is a clear example for which such mobile solutions are promising. It is very difficult to assess whether these patients are conscious [125]. To date this is only tested in a very few instances, i.e. when they visit the hospital. Allowing recurring recordings and assessments of the level of consciousness at the patient's home would greatly improve the reliability of the decision that a patient is conscious; family members sometimes have the feeling that the patient is conscious at specific moments but this cannot be proven due to the lack of neurophysiological tests at those moments (i.e., at home). The ease of application of mobile EEG could increase drastically the amount of data recorded from such patients. Moreover, the AAD paradigm is expected to be a robust one, due to the high classification accuracy in general and higher levels of engagement, compared to (less engaging) systematic audio stimuli such as sine tones, and is to be tested for attentional efforts of the patient in a single subject BCI scenario. Neural responses to narrative speech audition on group level have been proven useful for the differentiation of healthy controls from patients with disorders of consciousness very recently [100].

Emerging applications

In addition to the aforementioned audio-based applications there is high potential for the application of mobile EEG in other (clinical) settings. Recently, there has been an increased interest in monitoring children's EEG activity as an indication of proper development [44]. This was mainly motivated by the increased prevalence of conditions such as attention-deficit hyperactivity disorder (ADHD) or autism spectrum disorders (ASD) among children [178]. Recording EEG has also been shown effective in estimating the level of visual attention [113] and group engagement [61] during real-life classroom activities. Similarly, it can be used to monitor cognitive decline in the elderly [205]. There are indications that detecting EEG changes can help in extracting early markers of Alzheimer's disease [219]. Moreover, EEG is a well-established technique in epilepsy diagnosis and monitoring [210, 99]. Its effectiveness could be drastically increased if the measurements could take place in daily-life, and deliver the same data quality as in stationary hospital settings. Specific clinical applications (e.g., BCI for amyotrophic lateral sclerosis (ALS) patients) and more general applications (e.g., routine health checks) will benefit from the current project, as it will take the necessary steps towards the application of (repeated assessments by) mobile EEGs outside of the hospital or research facility.

Most in-hospital solutions for monitoring and managing health are, by design, too cumbersome or too complex to be used outside hospitals, and require expert assistance. The only solutions available on the market that have entered users' homes are weight scales, blood pressure monitors, and activity monitors (i.e., advanced accelerometers). The vast increase of mobile health monitoring devices (e.g., obtaining a heart rate from a smartphone camera, or fitness bracelets) illustrates the growing urge to obtain biomedical data from humans in a natural environment. A clear medical exemplary mobile EEG solution that is currently in development is the SeizeIt project; this project aims to detect certain types of seizures in epilepsy through a personalized miniature EEG system placed on the epileptic onset region of the scalp [186]. Moreover, the interest in mobile brain monitoring devices by the general population is high and in constant growth as well, as evidenced by the recent introduction of several commercial ready-to-use hardware EEG-based products aimed at the consumer market, e.g. Mindset [148], Kokoon [116] and Brawas [139]. These very recent projects aim to provide sleep and mental state information from few dry EEG electrodes seamlessly integrated in headphones. While similar stand-alone commercial EEG products exist to date, e.g. Emotiv, Neurosky, none of the latter provide unobtrusive and invisible EEG technology. Furthermore, the connection between mobile EEG and other technologies like smartphones [206] and augmented reality [179] has emerged recently. In addition, popular and very powerful technology companies such as Facebook, and famous visionaries

like Elon Musk included recent neurophysiological recordings in their portfolio for research and development of (distant) future applications [207, 158]. Such news should however, be taken with a grain of salt as the ambitions of these projects seem to be unrealistic from a research point of view.

All developments and applications mentioned in this section (and thesis) have emerged in the past few years only. This should be convincing evidence for the fact that society will be faced with mobile brain monitoring technologies in the (near or distant) future. Therefore, as a last field of study, we would like to emphasize that the ethical aspects of these developments should not be overlooked, and applications, especially those targeted at large groups of people, should be evaluated with care. Similar views are expressed by many researchers in this field, e.g., [81] and [201].

Appendix A

Graphical User Interface for exploring CPD analysis of EEG data

GUI elements

A matlab GUI has been developed to allow for convenient visual inspection of third-order dataset decompositions with CPD. Note that this is in early development stage and improvements and additions are likely to be added in the next months. The current GUI allows for intuitive interaction with several aspects of the CPD analysis of third-order EEG data. The GUI utilizes the tensor algorithms from Tensorlab 3.0 [221] available at `www.tensorlab.net`. This appendix shows how to use the GUI to explore different decomposition parameters and evaluate the CPD outcome, for example as was shown in Figure 4.2 in Chapter 4.

The GUI is divided into two parts: the compute module and the visualization module. The former is necessary to define the parameters one wants to investigate and compute the CPD models. These results can subsequently interactively visualized in the visualization module. We will describe the most important aspects of these GUIs in the next sections. Note that references to specific buttons in the GUI are denoted as follows: `Button`.

The compute module: Setup

The compute module calculates all models with both predefined and custom parameters. This is shown in Figure A.1. Data should be stored in a *.mat* file and contain a third-order data structure. The maximum rank of the

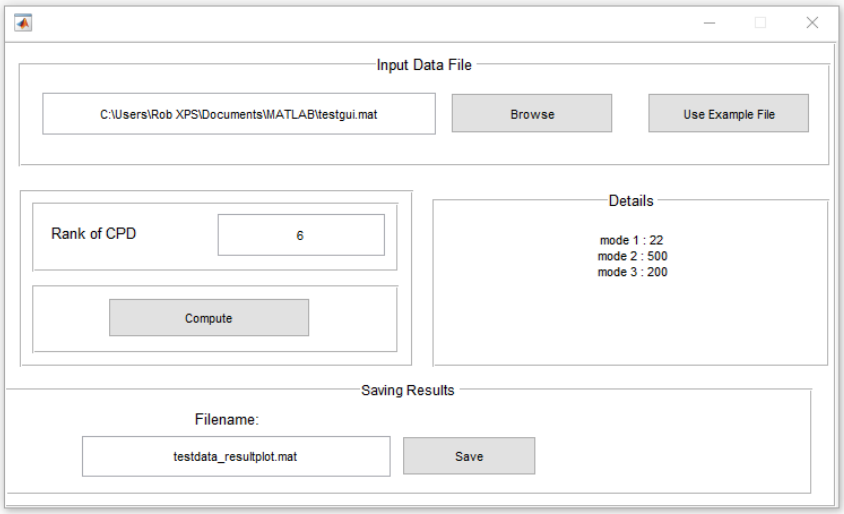


Figure A.1: Example of the compute module for third-order EEG dataset of size [22 × 500 × 200] representing the channel, time and trial mode respectively. `Browse` allows to select any third order dataset. One can also load an example EEG datafile. `Compute` calculates the CPD models for all given parameters, an additional progress bar is shown. Finally, the `Save` button allows to save the calculated models and start the visualization module

CPDs should be indicated. After pressing the `Compute` button all models are computed, after which they can be stored by pressing the `Save` button. This automatically stores the data for future reference and initializes the visualization module for inspection of the results. Note, at the moment all possible parameter combinations (e.g.different ranks and initializations) are calculated; large tensors can result in long computation times. In the current example we loaded data from 1 subject from Dataset B (as discussed in Chapter 5.2) which represents a single subject BCI dataset of 22 channels × 500 milliseconds × 200 trials. The first half of the trials represent target trials in a three-class auditory oddball paradigm, the latter half the non-target trials.

The visualization module: Interactive

Figure A.2 shows the layout of the visualization module. The panel on the left provides characteristics about the CPD error and multilinear singular values for inspection of MLSVD of the full dataset. The upper panel of the GUI allows

to choose the rank and iteration-number for plotting. The main panel plots the CPD decomposition of the set parameters. Different plot types (e.g. bar, topoplot, line) can be selected and specific segments can be highlighted with different colors. In the current example the target trials are indicated in orange and non-target in blue in the trial mode.

If two or more components in the model are highly similar a warning is shown above the components in question. This is designed to increase the awareness in case of degeneracy in the model. An example of two similar, or 'mirrored' components is shown in Figure A.3.

To summarize, the current early alpha version provides the following features:

1. Explore different initialization outcomes
2. See the influence of the number of iterations of the NLS (Note, this is heavily dependent on the initialization)
3. Vary the number of components.
4. Insights in the data through multilinear singular values.
5. Change plot style between topoplot, line and bar.
6. Highly similar components are detected (i.e. high correlation between the factors) and a warning is displayed for potential degeneracy.
7. Important parts in the signal/modes can be highlighted using colors.

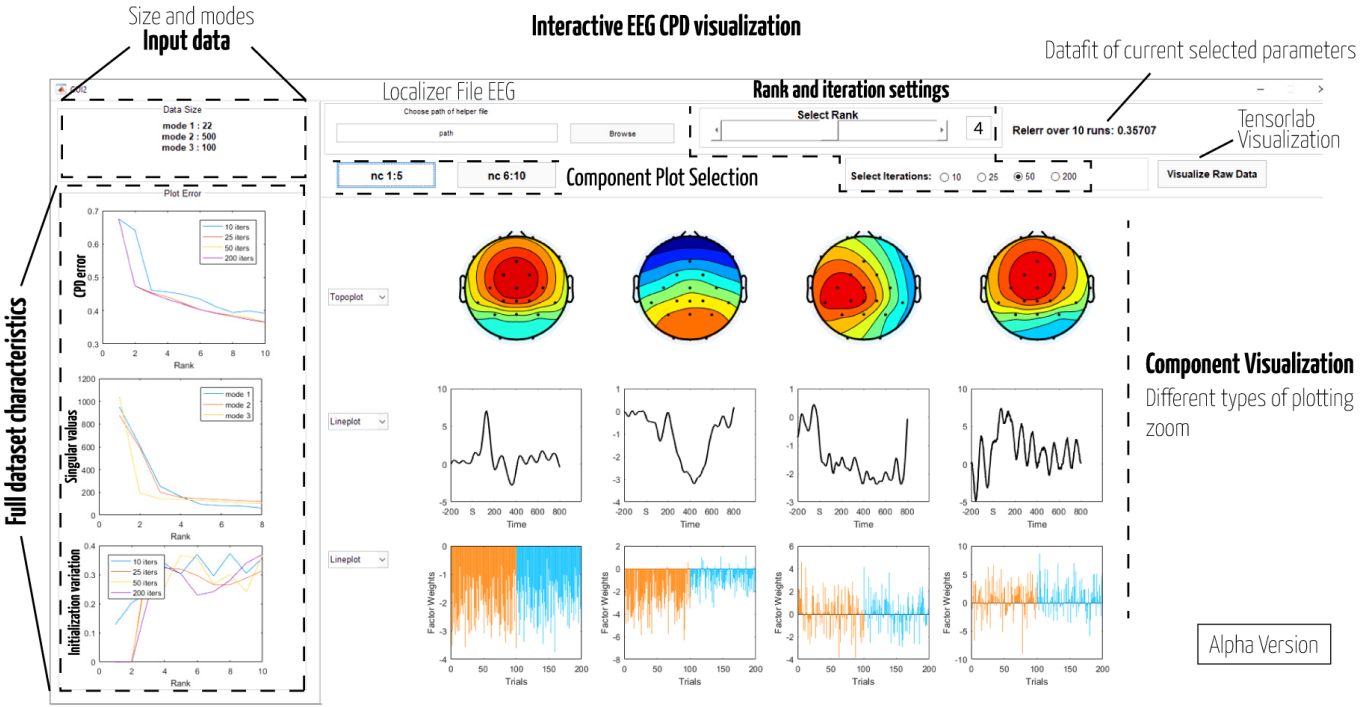


Figure A.2: Illustration of the visualization module for CPD visualization of third-order EEG data. Note that this is an alpha version and work in progress.

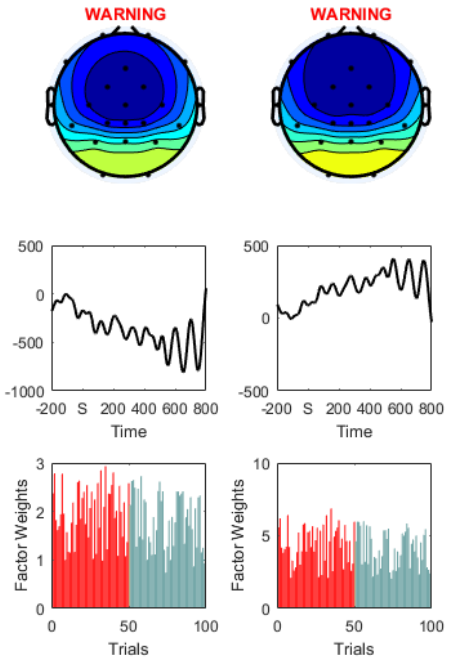


Figure A.3: Example of two highly similar components that are indicative of an incorrect rank of the CPD model

Future Versions

The GUI is still in the first stage of development and is not yet available publicly. Future versions are expected to extend the functionalities and flexibility of the GUI. In addition we aim to improve the user experience by providing step-by-step descriptions of the usage of the GUI and describe several examples. The source code will be made freely available online through the Biotensors project in the future.

Some suggestions for future improvements:

1. Allow wide choice of the parameters that can be calculated in the compute module (e.g. other type of decompositions such as LL1, constraints, weighted CPD).
2. Coupling when having multiple datasets.
3. Tensorization steps for matrix data (see [54] for an overview).

4. Interactive graphs that allows to click, for example to see a specific model that corresponds to a CPD error value in the general dataset characteristics.

5. Cluster the different CPD outcomes based on the initializations.

6. Incorporate custom code input to post-process the CPD outcome. For example to classify the trials in the trial mode with kmeans as was done in Chapter 4.

Appendix B

Head and body motion artifacts in mobile EEG, dataset C and D.

The influence of head and body motion on the mobile EEG.

At the beginning of the PhD trajectory we conducted a series of recordings to estimate the sensitivity to head and body motion artifacts of the EEG system (i.e. the SMARTING amplifier combined with the 24-channel wet electrode cap, as described in chapter 3.3.2). These can be considered as an exploratory study and does not provide conclusive evidence. However, it is interesting to compare these results to recent reports in the literature as the attention for motion artifacts in mobile EEG recordings is of high interest, e.g. [144]. The purpose of the current appendix is to provide additional experimental findings on these matters as derived from Dataset C. Furthermore, in the second part of the appendix we evaluate three out-of-the-box artifact correction techniques on the bike study, i.e. Dataset D as discussed in chapter 6. We refer the reader to [82, 112, 212] for a more in depth evaluation of (automatic) motion artifacts and handling techniques in (mobile) EEG. These results highlight the difficulty of defining a universal approach for handling these artifacts.

Noise levels in relation to motion

The term motion artifact refers to disturbances in the EEG, which are caused by human (head or body) motion. Fourteen subjects were recorded during

standing and performing simple movements with the head: nodding, tilting and turning. Subsequently the subjects were asked to move their body from front to back, left to right and up and down. Lastly the subjects were required to stand up and sit down (on a chair). Each motion type was executed for one minute and no external stimulation was applied. Data were high-pass filtered at 0.5Hz. This dataset is listed as Dataset C in chapter 3.3.1. The noise levels at each electrode was obtained by taking the RMS of section of data related to the motion type. The median noise values for each motion type are shown in Figure B.1 for a randomly chosen subject. Other subjects showed similar patterns.

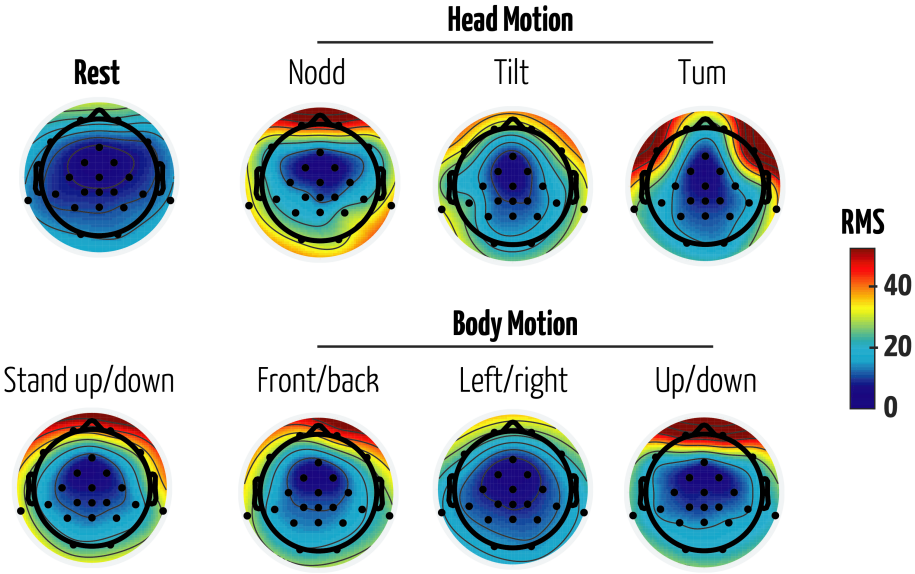


Figure B.1: Median RMS values of the EEG.

Mostly outer, and especially the front electrodes were affected by the motion of head and body. Note that the influence of eye blinks is naturally present in the Rest condition as well. These patterns are similar to the obtained topoplots in chapter 6 when biking outdoors. Furthermore, this lower susceptibility of the central electrodes to artifacts is in line with recent views on motion artifacts and EEG [112] and supports our conclusions drawn in chapter 6. Indirectly these findings provide evidence that EEG applications which use only centralized electrodes are likely to be successful in mobile scenarios.

Applying automated artifact removal techniques

Existing out-of-the-box artifact removal techniques are scarce. Nevertheless there are some listed in the EEGlab toolbox that may be applied. In chapter 6 manual ICA was applied for EOG activity removal and the influence of an automated CCA (from custom matlab scripts) procedure to remove EMG activity was evaluated. The automatic removal step lead to a small decrease in accuracy and P300 amplitude. These findings lead us to compare several off the shelf automatic artifact correction techniques on this data (Dataset D).

We considered the three most downloaded and freely available data processing extensions in the EEGlab toolbox:

1. MARA (v1.2): Multiple artifact rejection algorithm as described in [230].
2. AAR (v131130): ICA and CCA based automatic artifact removal [83].
3. ASR (v0.31): Artifact subspace reconstruction for artifact removal [119].

The performance of these methods is inherently related to the appropriate selection of parameters. Therefore we followed the recommended (i.e. mostly conservative) settings for the artifact removal and run the models with varying settings. Afterwards these data were analyzed in the same fashion as described in chapter 6. We compare the grand-average classification accuracies and inspect visually the grand-average ERPs to the three stimulus types in the three-class auditory oddball paradigm. Figure B.2 shows the average classification accuracy after applying automated techniques. It can be noted that the AAR and ASR methods significantly decrease the classification accuracies, i.e. $p < 0.01$ and $p < 0.001$. The MARA performed slightly lower in all conditions, nevertheless this difference was not deemed significant. To evaluate the influence on the P300 waveform Figure B.3 shows the grand-average ERP waveforms for the target, non-target and baseline trials for each of the methods. These results were only interpreted visually and similar effects were observed when plotting the median instead of the mean. It can be noted that similarly to the accuracy, the AAR and ASR seem too aggressive in the removal of P300 related information. The MARA approach seems more conservative and maintains most of the P300. Given that our findings are based on a limited amount of data and evaluation characteristics the results should consequently be treated with considerable caution. Nevertheless, other studies in this area support our belief that existing artifact correction methods cannot be universally applied and need proper evaluation per application [149, 112, 212]. In contrast to our exploratory and limited evaluation in this section, a recent study extensively evaluated several automated artifact correction techniques against removal of brain related signals [82].

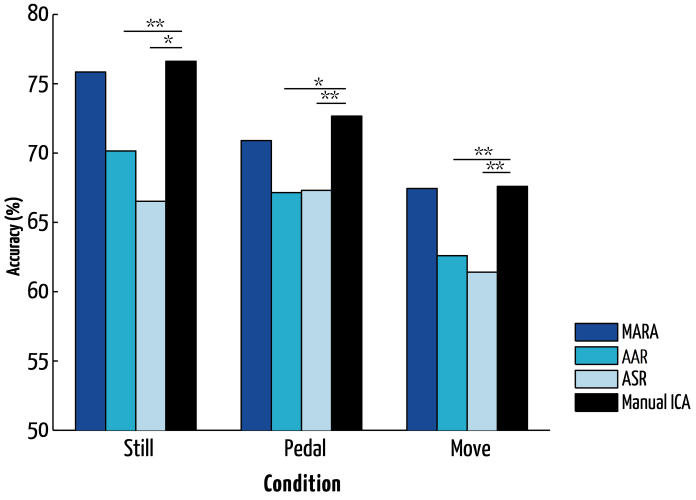


Figure B.2: Grand-average classification accuracies as calculated with the same procedure as described in chapter 6 for each condition. * Indicates $p < 0.01$ and ** $p < 0.001$

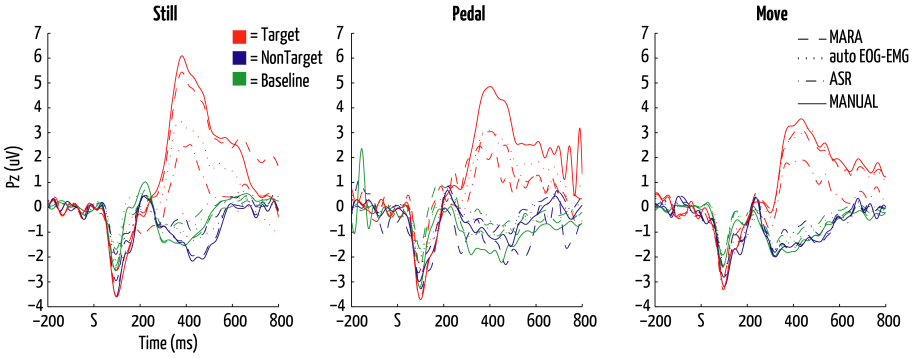


Figure B.3: Grand-average ERPs for the still, pedal and move condition from left to right perspectively. S indicating the stimulus onset and each color represents the stimulus type. The line style represents the various artifact correction techniques.

Appendix C

Virtual city navigation on a fixed bike, dataset F.

In this appendix we summarize the setup and most relevant findings of a preliminary study conducted to investigate the effects of visual distraction and physical engagement in AAD and auditory oddball tasks while navigating a virtual city on a fixed bike.

Methods

Subjects were required to pedal on a fixed bike and facing a projection screen in which a virtual city was presented. The virtual city was implemented in Blender and the pedaling and steering of the bike were recorded with LabVIEW. The bike controlled the movements in the simulated city. Ten native Dutch-speaking subjects (mean age (±SD) 24.9 (±2.2) years, two female) participated in the current study. The data acquisition hardware and protocol was similar as described in chapter 3. All subjects were presented with two AAD sessions (i.e. similar to the ones described in Study I, chapter 7) and two three-class auditory oddball sessions (i.e. similar as presented to the ones used in chapters 4, 5 and 6). The subjects navigated a path that was indicated on the roads in the city. Half of the sessions per paradigm took place with a simple city, meaning that textures were removed and only buildings are present. In the other half of the sessions a visually more complex city was shown on the screen. The complex city had detailed textures, stationary objects such as trees and moving cars. An example of the visuals in both conditions is shown in Figure C.1.

The EEG data were analyzed with the standard procedure as described in chapter 6 and 7 for the oddball and AAD respectively. In short, for the oddball data ICA was employed for the removal of EOG artifacts and BSS-CCA for the removal of muscle activity. Eye blink activity was captured by thresholding the

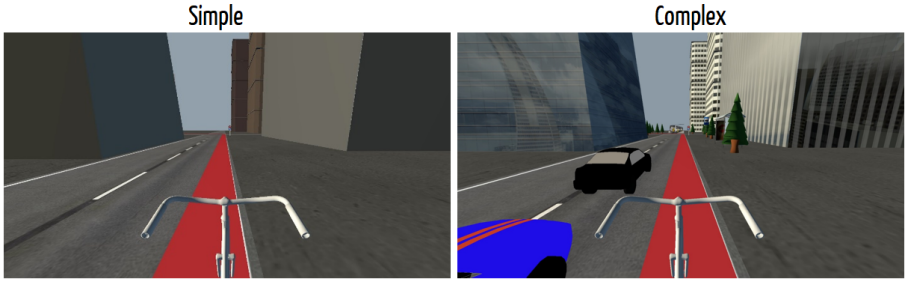


Figure C.1: Example of the simple and complex city visualization. Note that the cars in the complex city are moving.

IC representing eye blinks. The AAD analysis was not corrected for artifacts as is similar to the literature. Trial lengths of 10s were considered for the AAD.

Results and discussion

Oddball

The grand average accuracy (±SD) for the oddball paradigms were 70.9% (±8.2) and 69.4% (±9.1) for the simple and complex condition respectively. This differences was far from significant and there was a high correlation between the simple and complex session across subjects, ($r = 0.85$, $p < 0.05$). Six out of the ten subjects showed a slightly higher P300 amplitude, i.e. at Pz, in simple conditions as compared to the complex condition 9.5μV (±2.4) and 8.6μV (±2.4) respectively. However, this difference was not deemed significant ($p = 0.41$). The grand average ERPs for both conditions are shown in Figure C.2. Slightly higher latencies were observed in the complex condition as opposed to the simple condition, but also this difference was not insignificant ($p = 0.62$). The P300 amplitude correlated positively with the accuracy ($r = 0.57$, $p < 0.01$). The application of the BSS-CCA lowered slightly the accuracies as obtained with rLDA. This is similar as compared to the findings in chapter 6 and Appendix 8.2.3. Interestingly, a negative correlation was identified at subject level between the accuracy and average speed of biking ($r = -0.51$, $p < 0.05$). One possible explanation could be that faster movement leads to a faster changing visual input and requires quicker handling of the bike as compared to biking slowly. Such effects could not be investigated in chapter 6 for the lack of detailed GPS/speed measurements. In the latter study (Dataset D) most subjects biked

at similar speed as was approximated by timing the number of tours on the campus.

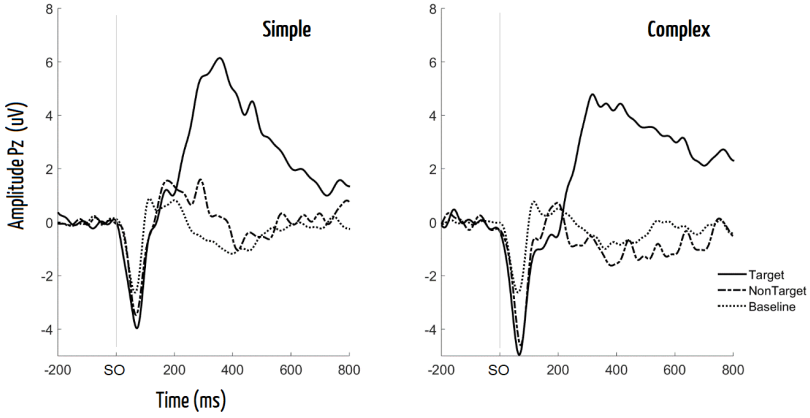


Figure C.2: Grand average ERPs to the Target, nonTarget and Baseline stimuli at electrode Pz for the Simple and Complex conditions from left to right respectively.

AAD

For the AAD the grand average accuracies in the simple and complex condition were 75.1% (±6.74) and 72.1% (±10.1) respectively. This difference was not deemed significant, ($p = 0.21$). The CA and CUA did not differ significantly between the simple and complex condition. No effects have been found of changes in speed, steering or head motion on the AAD. The accuracy is 7-10% lower as compared to that reported in Study I of chapter 7. Since these studies used the same auditory stimuli these results may provide preliminary evidence that being in a simulated real-life environment affects the AAD to great extent.

Other

Average eye blink rate for the AAD sessions (19.8/min±5.9) was significantly higher ($p < 0.01$) as compared to the oddball sessions (15.4/min±4.5). In addition, a negative correlation between oddball accuracies and blink rates ($r = -0.62$, $p < 0.01$) was found, but not between eye-blinks and AAD ($r = -0.19$, $p = 0.4$).

Finally, the results show a significant linear correlation between the performance of the three-class auditory oddball task and the AAD performance ($r = 0.65$, $p < 0.05$), the scatter plot was shown in Figure 8.2 in Chapter 8. This suggests that both tasks, at least to a certain extent, measure the same auditory attention abilities.

Acknowledgment

We would like to thank Markos Kapeliotis and Yi Wang for the technical support and ongoing collaboration during this experiment.

Bibliography

[1] Acar, E., Aykut-Bingol, C., Bingol, H., Bro, R., and Yener, B. Multiway analysis of epilepsy tensors. *Bioinformatics 23*, 13 (2007), i10–i18.

[2] Acar, E., and Yener, B. Unsupervised multiway data analysis: A literature survey. *IEEE transactions on knowledge and data engineering 21*, 1 (2009), 6–20.

[3] Aiken, S. J., and Picton, T. W. Human cortical responses to the speech envelope. *Ear and Hearing 29*, 2 (2008), 139–157.

[4] Akram, S., Presacco, A., Simon, J. Z., Shamma, S. A., and Babadi, B. Robust decoding of selective auditory attention from MEG in a competing-speaker environment via state-space modeling. *NeuroImage 124* (2016), 906–917.

[5] Akram, S., Simon, J. Z., and Babadi, B. Dynamic estimation of the auditory temporal response function from MEG in competing-speaker environments. *IEEE Transactions on Biomedical Engineering* (2016).

[6] Alickovic, E., Lunner, T., and Gustafsson, F. A system identification approach to determining listening attention from EEG signals. In *24th European Signal Processing Conference (EUSIPCO)* (NEW YORK, 2016), European Signal Processing Conference, Ieee, pp. 31–35.

[7] Anderer, P., Semlitsch, H., and Saletu, B. Multichannel auditory event-related brain potentials: effects of normal aging on the scalp distribution of N1, P2, N2 and P300 latencies and amplitudes. *Electroencephalography and Clinical Neurophysiology 99*, 5 (1996), 458–472.

[8] Arns, M., de Ridder, S., Strehl, U., Breteler, M., and Coenen, A. Efficacy of neurofeedback treatment in adhd: the effects on inattention, impulsivity and hyperactivity: a meta-analysis. *Clinical Eeg and Neuroscience 40*, 3 (2009), 180–189.

[9] Aroudi, A., Mirkovic, B., De Vos, M., and Doclo, S. Auditory attention decoding with EEG recordings using noisy acoustic reference signals. In *2016 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)* (2016), IEEE, pp. 694–698.

[10] Arthur, D., and Vassilvitskii, S. k-means++: The advantages of careful seeding. In *Proceedings of the eighteenth annual ACM-SIAM symposium on Discrete algorithms* (2007), Society for Industrial and Applied Mathematics, pp. 1027–1035.

[11] Askamp, J., and van Putten, M. J. Mobile EEG in epilepsy. *International journal of psychophysiology 91*, 1 (2014), 30–35.

[12] Aspinall, P., Mavros, P., Coyne, R., and Roe, J. The urban brain: analysing outdoor physical activity with mobile EEG. *British journal of sports medicine* (2013), bjsports–2012.

[13] Barachant, A., Bonnet, S., Congedo, M., and Jutten, C. Multiclass brain-computer interface classification by Riemannian geometry. *IEEE Transactions on Biomedical Engineering 59*, 4 (2012), 920–928.

[14] Barachant, A., and Congedo, M. A plug&play P300 BCI using information geometry. *arXiv preprint arXiv:1409.0107* (2014).

[15] Barzegaran, E., Vildavski, V. Y., and Knyazeva, M. G. Fine structure of posterior alpha rhythm in human EEG: Frequency components, their cortical sources, and temporal behavior. *Scientific Reports 7* (2017).

[16] Başar, E., et al. A short review of alpha activity in cognitive processes and in cognitive impairment. *International Journal of Psychophysiology 86*, 1 (2012), 25–38.

[17] Baykara, E., Ruf, C., Fioravanti, C., Käthner, I., Simon, N., Kleih, S., Kübler, A., and Halder, S. Effects of training and motivation on auditory P300 brain–computer interface performance. *Clinical Neurophysiology 127*, 1 (2016), 379–387.

[18] Bear, M. F., Connors, B. W., and Paradiso, M. A. *Neuroscience*, vol. 2. Lippincott Williams & Wilkins, 2007.

[19] Berg, P., and Scherg, M. A fast method for forward computation of multiple-shell spherical head models. *Electroencephalography and clinical neurophysiology 90*, 1 (1994), 58–64.

[20] Bertrand, A., and Moonen, M. Efficient sensor subset selection and link failure response for linear MMSE signal estimation in wireless sensor networks. In *Proc. European Signal Processing Conference (EUSIPCO)* (2010), pp. 1092–1096.

[21] Biesmans, W., Das, N., Francart, T., and Bertrand, A. Auditory-inspired speech envelope extraction methods for improved EEG-based auditory attention detection in a cocktail party scenario. *IEEE Transactions on Neural Systems and Rehabilitation Engineering 25*, 5 (2016), 402–412.

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

[23] Blankertz, B., Lemm, S., Treder, M., Haufe, S., and Müller, K.-R. Single-trial analysis and classification of ERP components—a tutorial. *NeuroImage 56*, 2 (2011), 814–825.

[24] Bleichner, M. G., and Debener, S. Concealed, unobtrusive ear-centered EEG acquisition: cEEGrids for transparent EEG. *Frontiers in Human Neuroscience 11* (2017).

[25] Bleichner, M. G., Lundbeck, M., Selisky, M., Minow, F., Jäger, M., Emkes, R., Debener, S., and De Vos, M. Exploring miniaturized EEG electrodes for brain-computer interfaces. an EEG you do not see? *Physiological reports 3*, 4 (2015), e12362.

[26] Bleichner, M. G., Mirkovic, B., and Debener, S. Identifying auditory attention with ear-EEG: cEEGrid versus high-density cap-EEG comparison. *Journal of Neural Engineering 13*, 6 (2016), 066004.

[27] Brandl, S., Hohne, J., Muller, K.-R., and Samek, W. Bringing BCI into everyday life: Motor imagery in a pseudo realistic environment. In *Neural Engineering (NER), 2015 7th International IEEE/EMBS Conference on* (2015), IEEE, pp. 224–227.

[28] Bro, R., and Smilde, A. K. Principal component analysis. *Analytical Methods 6*, 9 (2014), 2812–2831.

[29] Bunge, S. A., and Kahn, I. Cognition: An overview of neuroimaging techniques. *Encyclopedia of neuroscience 2* (2009), 1063–1067.

[30] Castermans, T., Duvinage, M., Cheron, G., and Dutoit, T. About the cortical origin of the low-delta and high-gamma rhythms

observed in EEG signals during treadmill walking. *Neuroscience letters 561* (2014), 166–170.

[31] Causse, M., Peysakhovich, V., and Fabre, E. F. High working memory load impairs language processing during a simulated piloting task: an erp and pupillometry study. *Frontiers in human neuroscience 10* (2016).

[32] Chang, Y.-K., Pesce, C., Chiang, Y.-T., Kuo, C.-Y., and Fong, D.-Y. Antecedent acute cycling exercise affects attention control: an ERP study using attention network test. *Frontiers in human neuroscience 9* (2015).

[33] Cheron, G., Petit, G., Cheron, J., Leroy, A., Cebolla, A., Cevallos, C., Petieau, M., Hoellinger, T., Zarka, D., Clarinval, A.-M., et al. Brain oscillations in sport: toward EEG biomarkers of performance. *Frontiers in psychology 7* (2016).

[34] Cherry, E. C. Some experiments on the recognition of speech, with one and with two ears. *The Journal of the acoustical society of America 25*, 5 (1953), 975–979.

[35] Cichocki, A., Mandic, D., De Lathauwer, L., Zhou, G., Zhao, Q., Caiafa, C., and Phan, H. Tensor decompositions for signal processing applications: From two-way to multiway component analysis. *Signal Processing Magazine, IEEE 32*, 2 (2015), 145–163.

[36] Cichocki, A., Washizawa, Y., Rutkowski, T., Bakardjian, H., Phan, A., Choi, S., Lee, H., Zhao, Q., Zhang, L., and Li, Y. Noninvasive BCIs: Multiway signal-processing array decompositions. *Computer 41*, 10 (2008), 34–42.

[37] Comon, P., and Jutten, C. *Handbook of Blind Source Separation: Independent component analysis and applications*. Academic press, 2010.

[38] Cong, F., Lin, Q.-H., Kuang, L.-D., Gong, X.-F., Astikainen, P., and Ristaniemi, T. Tensor decomposition of EEG signals: a brief review. *Journal of neuroscience methods 248* (2015), 59–69.

[39] Cong, F., Nandi, A. K., He, Z., Cichocki, A., and Ristaniemi, T. Fast and effective model order selection method to determine the number of sources in a linear transformation model. In *Signal Processing Conference (EUSIPCO), 2012 Proceedings of the 20th European* (2012), IEEE, pp. 1870–1874.

[40] Congedo, M., Barachant, A., and Andreev, A. A new generation of brain-computer interface based on riemannian geometry. *arXiv preprint arXiv:1310.8115* (2013).

[41] Congedo, M., Barachant, A., and Bhatia, R. Riemannian geometry for eeg-based brain-computer interfaces; a primer and a review. *Brain-Computer Interfaces* (2017), 1–20.

[42] Das, N., Biesmans, W., Bertrand, A., and Francart, T. The effect of head-related filtering and ear-specific decoding bias on auditory attention detection. *Journal of Neural Engineering 13*, 5 (2016), 056014.

[43] De Clercq, W., Vergult, A., Vanrumste, B., Van Paesschen, W., and Van Huffel, S. Canonical correlation analysis applied to remove muscle artifacts from the electroencephalogram. *Biomedical Engineering, IEEE Transactions on 53*, 12 (2006), 2583–2587.

[44] De Haan, M. *Infant EEG and event-related potentials*. Psychology Press, 2013.

[45] De Lathauwer, L. Decompositions of a higher-order tensor in block terms-part II: Definitions and uniqueness. *SIAM. J. Matrix Anal. & Appl. 30*, 3 (2008), 1033–1066.

[46] De Sanctis, P., Butler, J. S., Malcolm, B. R., and Foxe, J. J. Recalibration of inhibitory control systems during walking-related dual-task interference: a mobile brain-body imaging (mobi) study. *Neuroimage 94* (2014), 55–64.

[47] De Vos, M., De Lathauwer, L., and Van Huffel, S. Spatially constrained ICA algorithm with an application in EEG processing. *Signal Processing 91*, 8 (2011), 1963–1972.

[48] De Vos, M., and Debener, S. Mobile EEG: towards brain activity monitoring during natural action and cognition. *International journal of psychophysiology: official journal of the International Organization of Psychophysiology 91*, 1 (2014), 1–2.

[49] De Vos, M., Gandras, K., and Debener, S. Towards a truly mobile auditory brain-computer interface: Exploring the P300 to take away. *International Journal of Psychophysiology 91*, 1 (2014), 46–53.

[50] De Vos, M., Kroesen, M., Emkes, R., and Debener, S. P300 speller BCI with a mobile EEG system: comparison to a traditional amplifier. *Journal of neural engineering 11*, 3 (2014), 036008.

[51] De Vos, M., Riès, S., Vanderperren, K., Vanrumste, B., Alario, F.-X., Huffel, V. S., and Burle, B. Removal of muscle artifacts from EEG recordings of spoken language production. *Neuroinformatics 8*, 2 (2010), 135–150.

[52] De Vos, M., Thorne, J. D., Yovel, G., and Debener, S. Let's face it, from trial to trial: comparing procedures for N170 single-trial estimation. *Neuroimage 63*, 3 (2012), 1196–1202.

[53] De Vos, M., Vergult, A., De Lathauwer, L., De Clercq, W., Van Huffel, S., Dupont, P., Palmini, A., and Van Paesschen, W. Canonical decomposition of ictal scalp EEG reliably detects the seizure onset zone. *NeuroImage 37*, 3 (2007), 844–854.

[54] Debals, O., and De Lathauwer, L. The concept op tensorization. *Internal Report 17*, 99 (2017).

[55] Debener, S., Emkes, R., De Vos, M., and Bleichner, M. Unobtrusive ambulatory EEG using a smartphone and flexible printed electrodes around the ear. *Scientific reports 5* (2015).

[56] Debener, S., Kranczioch, C., Herrmann, C. S., and Engel, A. K. Auditory novelty oddball allows reliable distinction of top–down and bottom–up processes of attention. *International Journal of Psychophysiology 46*, 1 (2002), 77–84.

[57] Debener, S., Minow, F., Emkes, R., Gandras, K., and Vos, M. How about taking a low-cost, small, and wireless EEG for a walk? *Psychophysiology 49*, 11 (2012), 1617–1621.

[58] Delorme, A., and Makeig, S. EEGLAB: an open source toolbox for analysis of single-trial EEG dynamics including independent component analysis. *Journal of neuroscience methods 134*, 1 (2004), 9–21.

[59] Delorme, A., Mullen, T., Kothe, C., Akalin Acar, Z., Bigdely-Shamlo, N., Vankov, A., and Makeig, S. EEGLAB, SIFT, NFT, BCILAB, and ERICA: New Tools for Advanced EEG Processing. *Computational Intelligence and Neuroscience 2011* (2011), 1–12.

[60] Demiralp, T., Ademoglu, A., Istefanopulos, Y., Basar-Eroglu, C., and Basar, E. Wavelet analysis of oddball P300. *International Journal of Psychophysiology 39*, 2-3 (2001), 221–227.

[61] Dikker, S., Wan, L., Davidesco, I., Kaggen, L., Oostrik, M., McClintock, J., Rowland, J., Michalareas, G., Van Bavel, J. J., Ding, M., et al. Brain-to-brain synchrony tracks real-world

dynamic group interactions in the classroom. *Current Biology 27*, 9 (2017), 1375–1380.

[62] Ding, N., and Simon, J. Z. Emergence of neural encoding of auditory objects while listening to competing speakers. *Proceedings of the National Academy of Sciences of the United States of America 109*, 29 (2012), 11854–11859.

[63] Domanov, I., and De Lathauwer, L. On the uniqueness of the canonical polyadic decomposition of third-order tensors—part i: Basic results and uniqueness of one factor matrix. *SIAM Journal on Matrix Analysis and Applications 34*, 3 (2013), 855–875.

[64] Domanov, I., and De Lathauwer, L. On the uniqueness of the canonical polyadic decomposition of third-order tensors—part ii: Uniqueness of the overall decomposition. *SIAM Journal on Matrix Analysis and Applications 34*, 3 (2013), 876–903.

[65] Duvinage, M., Castermans, T., Dutoit, T., Petieau, M., Hoellinger, T., Saedeleer, C. D., Seetharaman, K., and Cheron, G. A P300-based quantitative comparison between the Emotiv Epoc headset and a medical EEG device. *Biomedical Engineering 765* (2012), 2012–764.

[66] Duvinage, M., Castermans, T., Petieau, M., Hoellinger, T., Cheron, G., and Dutoit, T. Performance of the emotiv epoc headset for P300-based applications. *Biomedical engineering online 12*, 1 (2013), 56.

[67] Egner, T., and Gruzelier, J. H. EEG biofeedback of low beta band components: frequency-specific effects on variables of attention and event-related brain potentials. *Clinical Neurophysiology 115*, 1 (2004), 131–139.

[68] Eichele, T., Specht, K., Moosmann, M., Jongsma, M. L., Quiroga, R. Q., Nordby, H., and Hugdahl, K. Assessing the spatiotemporal evolution of neuronal activation with single-trial event-related potentials and functional mri. *Proceedings of the National Academy of Sciences of the United States of America 102*, 49 (2005), 17798–17803.

[69] Ekin, B., Atlas, L., Mirbagheri, M., and Lee, A. K. An alternative approach for auditory attention tracking using single-trial EEG. In *Acoustics, Speech and Signal Processing (ICASSP), 2016 IEEE International Conference on* (2016), IEEE, pp. 729–733.

[70] Engel, A. K., Moll, C. K., Fried, I., and Ojemann, G. A. Invasive recordings from the human brain: clinical insights and beyond. *Nature reviews. Neuroscience 6*, 1 (2005), 35.

[71] Enriquez-Geppert, S., Huster, R. J., and Herrmann, C. S. Eeg-neurofeedback as a tool to modulate cognition and behavior: A review tutorial. *Frontiers in Human Neuroscience 11* (2017), 19.

[72] Estienne, F., Matthijs, N., Massart, D., Ricoux, P., and Leibovici, D. Multi-way modelling of high-dimensionality electroencephalographic data. *Chemometrics and Intelligent Laboratory Systems 58*, 1 (2001), 59–72.

[73] Farquhar, J., and Hill, N. Interactions between pre-processing and classification methods for event-related-potential classification. *Neuroinformatics 11*, 2 (2012), 175–192.

[74] Ferrari, M., and Quaresima, V. A brief review on the history of human functional near-infrared spectroscopy (fnirs) development and fields of application. *Neuroimage 63*, 2 (2012), 921–935.

[75] Fiedler, L., Woestmann, M., Graversen, C., Brandmeyer, A., Lunner, T., and Obleser, J. Single-channel in-ear-EEG detects the focus of auditory attention to concurrent tone streams and mixed speech. *Journal of Neural Engineering 14*, 3 (2017), 036020.

[76] Formisano, E., Kim, D.-S., Di Salle, F., van de Moortele, P.-F., Ugurbil, K., and Goebel, R. Mirror-symmetric tonotopic maps in human primary auditory cortex. *Neuron 40*, 4 (2003), 859–869.

[77] Frolich, L., Winkler, I., Muller, K.-R., and Samek, W. Investigating effects of different artefact types on motor imagery BCI. In *Engineering in Medicine and Biology Society (EMBC), 2015 37th Annual International Conference of the IEEE* (2015), IEEE, pp. 1942–1945.

[78] Fuglsang, S. A., Dau, T., and Hjortkjær, J. Noise-robust cortical tracking of attended speech in real-world acoustic scenes. *NeuroImage in press* (2017).

[79] Furdea, A., Halder, S., Krusienski, D., Bross, D., Nijboer, F., Birbaumer, N., and Kübler, A. An auditory oddball (P300) spelling system for brain-computer interfaces. *Psychophysiology 46*, 3 (2009), 617–625.

[80] Gandras, K., Grimm, S., and Bendixen, A. Electrophysiological correlates of speaker segregation and foreground-background selection in ambiguous listening situations. *Neuroscience* (2017).

[81] Goering, S., and Yuste, R. On the necessity of ethical guidelines for novel neurotechnologies. *Cell 167*, 4 (2016), 882–885.

[82] Goh, S. K., Abbass, H. A., Tan, K. C., Al-Mamun, A., Wang, C., and Guan, C. Automatic eeg artifact removal techniques by detecting influential independent components. *IEEE Transactions on Emerging Topics in Computational Intelligence 1*, 4 (2017), 270–279.

[83] Gómez-Herrero, G., De Clercq, W., Anwar, H., Kara, O., Egiazarian, K., Van Huffel, S., and Van Paesschen, W. Automatic removal of ocular artifacts in the eeg without an eog reference channel. In *Signal Processing Symposium, 2006. NORSIG 2006. Proceedings of the 7th Nordic* (2006), IEEE, pp. 130–133.

[84] Goovaerts, G., Vandenberk, B., Willems, R., and Van Huffel, S. Automatic detection of t wave alternans using tensor decompositions in multilead ecg signals. *Physiological measurement 38*, 8 (2017), 1513.

[85] Gramann, K., Gwin, J. T., Ferris, D. P., Oie, K., Jung, T.-P., Lin, C.-T., Liao, L.-D., and Makeig, S. Cognition in action: imaging brain/body dynamics in mobile humans. *Reviews in the Neurosciences 22*, 6 (2011), 593–608.

[86] Gray, H. M., Ambady, N., Lowenthal, W. T., and Deldin, P. P300 as an index of attention to self-relevant stimuli. *Journal of experimental social psychology 40*, 2 (2004), 216–224.

[87] Grego, F., Vallier, J.-M., Collardeau, M., Bermon, S., Ferrari, P., Candito, M., Bayer, P., Magnié, M.-N., and Brisswalter, J. Effects of long duration exercise on cognitive function, blood glucose, and counterregulatory hormones in male cyclists. *Neuroscience letters 364*, 2 (2004), 76–80.

[88] Grizou, J., Iturrate, I., Montesano, L., Oudeyer, P.-Y., and Lopes, M. Calibration-free BCI based control. In *AAAI* (2014), pp. 1213–1220.

[89] Gwin, J. T., Gramann, K., Makeig, S., and Ferris, D. P. Removal of movement artifact from high-density EEGs recorded during walking and running. *Journal of neurophysiology 103*, 6 (2010), 3526–3534.

[90] Haghighi, M., Moghadamfalahi, M., Akcakaya, M., and Erdogmus, D. EEG-assisted modulation of sound sources in the auditory scene. *arXiv preprint arXiv:1612.00703* (2016).

[91] Halder, S., Rea, M., Andreoni, R., Nijboer, F., Hammer, E., Kleih, S., Birbaumer, N., and Kübler, A. An auditory oddball brain–computer interface for binary choices. *Clinical Neurophysiology 121*, 4 (2010), 516–523.

[92] Hallez, H., De Vos, M., Vanrumste, B., Van Hese, P., Assecondi, S., Van Laere, K., Dupont, P., Van Paesschen, W., Van Huffel, S., and Lemahieu, I. Removing muscle and eye artifacts using blind source separation techniques in ictal EEG source imaging. *Clinical Neurophysiology 120*, 7 (2009), 1262–1272.

[93] Harshman, R. A. *Foundations of the parafac procedure: models and conditions for an "explanatory" multimodal factor analysis.* 1970.

[94] Haufe, S., Meinecke, F., Görgen, K., Dähne, S., Haynes, J.-D., Blankertz, B., and Biessmann, F. On the interpretation of weight vectors of linear models in multivariate neuroimaging. *Neuroimage 87* (2014), 96–110.

[95] He, B., Coleman, T., Genin, G. M., Glover, G., Hu, X., Johnson, N., Liu, T., Makeig, S., Sajda, P., and Ye, K. Grand challenges in mapping the human brain: NSF workshop report. *IEEE Trans. Biomed. Engineering 60*, 11 (2013), 2983–2992.

[96] Horton, C., Srinivasan, R., and D'Zmura, M. Envelope responses in single-trial EEG indicate attended speaker in a ''cocktail party". *Journal of neural engineering 11*, 4 (2014), 046015.

[97] Hunyadi, B., Camps, D., Sorber, L., Paesschen, W., De Vos, M., Huffel, S., and Lathauwer, L. Block term decomposition for modelling epileptic seizures. *EURASIP Journal on Advances in Signal Processing 2014*, 1 (2014), 139.

[98] Hunyadi, B., Dupont, P., Van Paesschen, W., and Van Huffel, S. Tensor decompositions and data fusion in epileptic electroencephalography and functional magnetic resonance imaging data. *Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery 7*, 1 (2017).

[99] Hunyadi, B., Signoretto, M., Van Paesschen, W., Suykens, J. A., Van Huffel, S., and De Vos, M. Incorporating structural information from the multichannel eeg improves patient-specific seizure detection. *Clinical Neurophysiology 123*, 12 (2012), 2352–2361.

[100] Iotzov, I., Fidali, B. C., Petroni, A., Conte, M. M., Schiff, N. D., and Parra, L. C. Common neural responses to narrative speech in disorders of consciousness. *bioRxiv* (2017).

[101] Jeunet, C., N'Kaoua, B., and Lotte, F. *Advances in user-training for mental-imagery-based BCI control: Psychological and cognitive factors and their neural correlates*, vol. 228 of *Progress in Brain Research*. Elsevier Science Bv, Amsterdam, 2016, pp. 3–35.

[102] Jeunet, C., N'Kaoua, B., Subramanian, S., Hachet, M., and Lotte, F. Predicting mental imagery-based BCI performance from personality, cognitive profile and neurophysiological patterns. *Plos One 10*, 12 (2015), 21.

[103] Jung, T.-P., Humphries, C., Lee, T.-W., Makeig, S., McKeown, M. J., Iragui, V., and Sejnowski, T. J. Extended ica removes artifacts from electroencephalographic recordings. In *Advances in neural information processing systems* (1998), pp. 894–900.

[104] Kaya, E. M., and Elhilali, M. Modelling auditory attention. *Philosophical Transactions of the Royal Society B-Biological Sciences 372*, 1714 (2017), 10.

[105] Kerick, S. E., Oie, K. S., and McDowell, K. Assessment of EEG signal quality in motion environments. Tech. rep., DTIC Document, 2009.

[106] Kindermans, P., Schreuder, M., Schrauwen, B., Müller, K., and Tangermann, M. True zero-training brain-computer interfacing - an online study. *PLoS ONE 9*, 7 (2014), e102504.

[107] Kindermans, P., Tangermann, M., Müller, K., and Schrauwen, B. Integrating dynamic stopping, transfer learning and language models in an adaptive zero-training ERP speller. *Journal of Neural Engineering 11*, 3 (2014), 035005.

[108] Kindermans, P., Verstraeten, D., and Schrauwen, B. A Bayesian model for exploiting application constraints to enable unsupervised training of a P300-based BCI. *PLoS ONE 7*, 4 (2012), e33758.

[109] Kiverstein, J., and Miller, M. The embodied brain: towards a radical embodied cognitive neuroscience. *Frontiers in human neuroscience 9* (2015), 237.

[110] Kleih, S., Kaufmann, T., Zickler, C., Halder, S., Leotta, F., Cincotti, F., Aloise, F., Riccio, A., Herbert, C., Mattia, D., et al. Out of the frying pan into the fire–the P300-based BCI faces real–world challenges. *Progress in brain research 194* (2011), 27–46.

[111] Klimesch, W. Eeg alpha and theta oscillations reflect cognitive and memory performance: a review and analysis. *Brain research reviews 29*, 2 (1999), 169–195.

[112] Kline, J. E., Huang, H. J., Snyder, K. L., and Ferris, D. P. Isolating gait-related movement artifacts in electroencephalography during human walking. *Journal of neural engineering 12*, 4 (2015), 046022.

[113] Ko, L.-W., Komarov, O., Hairston, W. D., Jung, T.-P., and Lin, C.-T. Sustained attention in real classroom settings: An eeg study. *Frontiers in Human Neuroscience 11* (2017).

[114] Kohli, S., and Casson, A. J. Towards out–of–the–lab EEG in uncontrolled environments: Feasibility study of dry EEG recordings during exercise bike riding. In *Engineering in Medicine and Biology Society (EMBC), 2015 37th Annual International Conference of the IEEE* (2015), IEEE, pp. 1025–1028.

[115] Koike, S., Nishimura, Y., Takizawa, R., Yahata, N., and Kasai, K. Near-infrared spectroscopy in schizophrenia: a possible biomarker for predicting clinical outcome and treatment.

[116] Kokoon. Kokoon: The ultimate sleep sanctuary - sleep sensing headphones. `https://kokoon.io/`.

[117] Kolda, T., and Bader, B. Tensor decompositions and applications. *SIAM Rev. 51*, 3 (2009), 455–500.

[118] Korczowski, L., Barachant, A., Andreev, A., Jutten, C., and Congedo, M. " brain invaders 2": an open source plug & play multi-user BCI videogame. In *6th International Brain-Computer Interface Meeting (BCI Meeting 2016)* (2016), pp. 10–3217.

[119] Kothe, C. Artifact subspace reconstruction EEGLAB plug-in. `http://sccn.ucsd.edu/eeglab/plugins/ASR.pdf`.

[120] Kothe, C. Lab streaming layer (lsl). *https://github. com/sccn/labstreaminglayer. Accessed on October 26* (2014), 2015.

[121] Kothe, C. A., and Makeig, S. BCILAB: a platform for brain–computer interface development. *Journal of neural engineering 10*, 5 (2013), 056014.

[122] Krucoff, M. O., Rahimpour, S., Slutzky, M. W., Edgerton, V. R., and Turner, D. A. Enhancing nervous system recovery through neurobiologics, neural interface training, and neurorehabilitation. *Frontiers in Neuroscience 10* (2016), 23.

[123] Kübler, A., Holz, E., Sellers, E., and Vaughan, T. Toward independent home use of brain-computer interfaces: a decision algorithm for selection of potential end-users. *Archives of physical medicine and rehabilitation 96*, 3 (2015), S27–S32.

[124] Ladouce, S., Donaldson, D. I., Dudchenko, P. A., and Ietswaart, M. Understanding minds in real-world environments: toward a mobile cognition approach. *Frontiers in Human Neuroscience 10* (2016).

[125] Laureys, S., Owen, A. M., and Schiff, N. D. Brain function in coma, vegetative state, and related disorders. *The Lancet Neurology 3*, 9 (2004), 537–546.

[126] Leutheuser, H., Gabsteiger, F., Hebenstreit, F., Reis, P., Lochmann, M., and Eskofier, B. Comparison of the AMICA and the InfoMax algorithm for the reduction of electromyogenic artifacts in EEG data. In *Engineering in Medicine and Biology Society (EMBC), 2013 35th Annual International Conference of the IEEE* (2013), IEEE, pp. 6804–6807.

[127] Lin, Y.-P., Wang, Y., and Jung, T.-P. Assessing the feasibility of online SSVEP decoding in human walking using a consumer EEG headset. *Journal of neuroengineering and rehabilitation 11*, 1 (2014), 1.

[128] Linden, D. E. The p300: where in the brain is it produced and what does it tell us? *The Neuroscientist 11*, 6 (2005), 563–576.

[129] Logothetis, N. K. What we can do and what we cannot do with fmri. *Nature 453*, 7197 (2008), 869.

[130] Looney, D., Kidmose, P., Park, C., Ungstrup, M., Rank, M. L., Rosenkranz, K., and Mandic, D. P. The in-the-ear recording concept: User-centered and wearable brain monitoring. *IEEE pulse 3*, 6 (2012), 32–42.

[131] Lotte, F., Bougrain, L., and Clerc, M. Electroencephalography (EEG)-based brain–computer interfaces. *Wiley Encyclopedia of Electrical and Electronics Engineering* (2015).

[132] Lotte, F., Congedo, M., Lécuyer, A., and Lamarche, F. A review of classification algorithms for EEG-based brain–computer interfaces. *Journal of neural engineering 4* (2007).

[133] Lotte, F., and Jeunet, C. Online classification accuracy is a poor metric to study mental imagery-based BCI user learning: an experimental demonstration and new metrics. In *7th International BCI Conference* (2017).

[134] Luck, S. J. *An introduction to the event-related potential technique*. 2014.

[135] Maguire, E. A., Gadian, D. G., Johnsrude, I. S., Good, C. D., Ashburner, J., Frackowiak, R. S., and Frith, C. D. Navigation-related structural change in the hippocampi of taxi drivers. *Proceedings of the National Academy of Sciences 97*, 8 (2000), 4398–4403.

[136] Manyakov, N. V., Chumerin, N., Combaz, A., and Van Hulle, M. M. Comparison of classification methods for p300 brain-computer interface on disabled subjects. *Computational intelligence and neuroscience 2011* (2011), 2.

[137] Matic, V., Cherian, P. J., Koolen, N., Naulaers, G., Swarte, R. M., Govaert, P., Van Huffel, S., and De Vos, M. Holistic approach for automated background EEG assessment in asphyxiated full-term infants. *Journal of neural engineering 11*, 6 (2014), 066007.

[138] Mavros, P., Austwick, M. Z., and Smith, A. H. Geo-EEG: towards the use of EEG in the study of urban behaviour. *Applied Spatial Analysis and Policy 9*, 2 (2016), 191–212.

[139] mBrainTrain. Brawas. `https://brawas.co/`.

[140] McKendrick, R., Parasuraman, R., and Ayaz, H. Wearable functional near infrared spectroscopy (fnirs) and transcranial direct current stimulation (tdcs): expanding vistas for neurocognitive augmentation. *Frontiers in systems neuroscience 9* (2015), 27.

[141] Mehta, R. K., and Parasuraman, R. Neuroergonomics: a review of applications to physical and cognitive work. *Frontiers in human neuroscience 7* (2013).

[142] Melnik, A., Legkov, P., Izdebski, K., Kärcher, S. M., Hairston, W. D., Ferris, D. P., and König, P. Systems, subjects, sessions: To what extent do these factors influence EEG data? *Frontiers in human neuroscience 11* (2017).

[143] Mesgarani, N., and Chang, E. F. Selective cortical representation of attended speaker in multi-talker speech perception. *Nature 485*, 7397 (2012), 233–U118.

[144] Mihajlovic, V., Grundlehner, B., Vullers, R., and Penders, J. Wearable, wireless EEG solutions in daily life applications: what are we missing? *Biomedical and Health Informatics, IEEE Journal of 19*, 1 (2015), 6–21.

[145] Mijović, P., Ković, V., De Vos, M., Mačužić, I., Jeremić, B., and Gligorijević, I. Benefits of instructed responding in manual assembly tasks: an erp approach. *Frontiers in human neuroscience 10* (2016).

[146] Mikkelsen, K. B., Kidmose, P., and Hansen, L. K. On the keyhole hypothesis: High mutual information between ear and scalp EEG. *Frontiers in Human Neuroscience 11* (2017).

[147] Milner, R., Lewandowska, M., Ganc, M., Ciesla, K., Niedzialek, I., and Skarzynski, H. Slow cortical potential neurofeedback in chronic tinnitus therapy: A case report. *Applied Psychophysiology and Biofeedback 41*, 2 (2016), 225–249.

[148] Mindset. Mindset: smart headphones that improve your concentration. `https://www.kickstarter.com/projects/mindset/headphones`.

[149] Minguillon, J., Lopez-Gordo, M. A., and Pelayo, F. Trends in EEG-BCI for daily-life: Requirements for artifact removal. *Biomedical Signal Processing and Control 31* (2017), 407–418.

[150] Mirkovic, B., Bleichner, M. G., De Vos, M., and Debener, S. Target speaker detection with concealed EEG around the ear. *Frontiers in Neuroscience 10* (2016).

[151] Mirkovic, B., Debener, S., Jaeger, M., and De Vos, M. Decoding the attended speech stream with multi–channel EEG: implications for online, daily-life applications. *Journal of neural engineering 12*, 4 (2015), 046007.

[152] Miwakeichi, F., Martinez-Montes, E., Valdés-Sosa, P., Nishiyama, N., Mizuhara, H., and Yamaguchi, Y. Decomposing EEG data into space-time-frequency components using parallel factor analysis. *NeuroImage 22*, 3 (2004), 1035–1045.

[153] Möcks, J. Decomposing event-related potentials: A new topographic components model. *Biological psychology 26*, 1 (1988), 199–215.

[154] Montavon, G., Samek, W., and Müller, K.-R. Methods for interpreting and understanding deep neural networks. *arXiv preprint arXiv:1706.07979* (2017).

[155] Mørup, M., Hansen, L. K., Herrmann, C. S., Parnas, J., and Arnfred, S. M. Parallel factor analysis as an exploratory tool for wavelet transformed event-related EEG. *NeuroImage 29*, 3 (2006), 938–947.

[156] Myrden, A., and Chau, T. Effects of user mental state on EEG-BCI performance. *Frontiers in Human Neuroscience 9* (2015), 308.

[157] Neuper, C., and Pfurtscheller, G. *Neurofeedback Training for BCI Control*. Frontiers Collection. Springer-Verlag Berlin, Berlin, 2010, pp. 65–78.

[158] Neuralink. Neuralink. https://www.neuralink.com/.

[159] Nguyen, T., Ahn, S., Jang, H., Jun, S. C., and Kim, J. G. Utilization of a combined EEG/NIRS system to predict driver drowsiness. *Scientific Reports 7* (2017).

[160] Nicolas-Alonso, L. F., and Gomez-Gil, J. Brain computer interfaces, a review. *Sensors 12*, 2 (2012), 1211–1279.

[161] Niedermeyer, E., and da Silva, F. L. *Electroencephalography: basic principles, clinical applications, and related fields*. Lippincott Williams & Wilkins, 2005.

[162] Olson, R. L., Chang, Y.-K., Brush, C. J., Kwok, A. N., Gordon, V. X., and Alderman, B. L. Neurophysiological and behavioral correlates of cognitive control during low and moderate intensity exercise. *NeuroImage* (2015).

[163] Onishi, A., Phan, A. H., Matsuoka, K., and Cichocki, A. Tensor classification for p300-based brain computer interface. In *Acoustics, Speech and Signal Processing (ICASSP), 2012 IEEE International Conference on* (2012), IEEE, pp. 581–584.

[164] O'Sullivan, J., Chen, Z., Herrero, J., McKhann, G. M., Sheth, S. A., Mehta, A. D., and Mesgarani, N. Neural decoding of attentional selection in multi-speaker environments without access to clean sources. *Journal of Neural Engineering 14*, 5 (2017), 056001.

[165] O'Sullivan, J. A., Power, A. J., Mesgarani, N., Rajaram, S., Foxe, J. J., Shinn-Cunningham, B. G., Slaney, M., Shamma, S. A., and Lalor, E. C. Attentional selection in a cocktail party environment can be decoded from single–trial EEG. *Cerebral Cortex* (2014), bht355.

[166] Pacharra, M., Debener, S., and Wascher, E. Concealed around-the-ear EEG captures cognitive processing in a visual simon task. *Frontiers in Human Neuroscience 11* (2017), 290.

[167] Park, J. L., Fairweather, M. M., and Donaldson, D. I. Making the case for mobile cognition: Eeg and sports performance. *Neuroscience & Biobehavioral Reviews 52* (2015), 117–130.

[168] Pfurtscheller, G., and Neuper, C. Motor imagery and direct brain-computer communication. *Proceedings of the Ieee 89*, 7 (2001), 1123–1134.

[169] Polich, J. Updating P300: an integrative theory of P3a and P3b. *Clinical neurophysiology 118*, 10 (2007), 2128–2148.

[170] Pontifex, M. B., and Hillman, C. H. Neuroelectric and behavioral indices of interference control during acute cycling. 570–580.

[171] Poulsen, A. T., Kamronn, S., Dmochowski, J., Parra, L. C., and Hansen, L. K. Eeg in the classroom: Synchronised neural recordings during video presentation. *Scientific Reports 7* (2017).

[172] Presacco, A., Goodman, R., Forrester, L., and Contreras-Vidal, J. L. Neural decoding of treadmill walking from noninvasive electroencephalographic signals. *Journal of neurophysiology 106*, 4 (2011), 1875–1887.

[173] Pugh, G. E. *The biological origin of human values*. Basic Books: New York, 1977.

[174] Quaedflieg, C. W., Smulders, F. T., Meyer, T., Peeters, F., Merckelbach, H., and Smeets, T. The validity of individual frontal alpha asymmetry EEG neurofeedback. *Social cognitive and affective neuroscience 11*, 1 (2015), 33–43.

[175] Reis, P. M., Hebenstreit, F., Gabsteiger, F., von Tscharner, V., and Lochmann, M. Methodological aspects of EEG and body dynamics measurements during motion. *Towards a New Cognitive Neuroscience: Modeling Natural Brain Dynamics* (2014), 9.

[176] Renard, Y., Lotte, F., Gibert, G., Congedo, M., Maby, E., Delannoy, V., Bertrand, O., and Lécuyer, A. OpenViBE: an open-source software platform to design, test, and use brain-computer interfaces in real and virtual environments. *Presence: teleoperators and virtual environments 19*, 1 (2010), 35–53.

[177] Rogala, J., Jurewicz, K., Paluch, K., Kublik, E., Cetnarski, R., and Wrobel, A. The do's and don'ts of neurofeedback training: A review of the controlled studies using healthy adults. *Frontiers in Human Neuroscience 10* (2016), 12.

[178] Russell, G., Rodgers, L. R., Ukoumunne, O. C., and Ford, T. Prevalence of parent-reported asd and adhd in the uk: findings from the millennium cohort study. *Journal of autism and developmental disorders 44*, 1 (2014), 31–40.

[179] Saboor, A., Rezeika, A., Stawicki, P., Gembler, F., Benda, M., Grunenberg, T., and Volosyak, I. Ssvep-based bci in a smart home scenario. In *International Work-Conference on Artificial Neural Networks* (2017), Springer, pp. 474–485.

[180] Samek, W., Meinecke, F., and Müller, K. Transferring subspaces between subjects in brain-computer interfacing. *IEEE Transactions on Biomedical Engineering 60*, 8 (2013), 2289–2298.

[181] Samek, W., Nakajima, S., Kawanabe, M., and Mueller, K.-R. On robust parameter estimation in brain-computer interfacing. *Journal of Neural Engineering* (2017).

[182] Santoro, R., Moerel, M., De Martino, F., Valente, G., Ugurbil, K., Yacoub, E., and Formisano, E. Reconstructing the spectrotemporal modulations of real-life sounds from fmri response patterns. *Proceedings of the National Academy of Sciences* (2017), 201617622.

[183] Scanlon, J. E., Sieben, A. J., Holyk, K. R., and Mathewson, K. E. Your brain on bikes: P3, MMN/N2b, and baseline noise while pedaling a stationary bike. *Psychophysiology 54*, 6 (2017), 927–937.

[184] Scanlon, J. E., Townsend, K. A., Cormier, D. L., Kuziek, J. W., and Mathewson, K. Taking off the training wheels: Measuring auditory P3 during outdoor cycling using an active wet EEG system. *bioRxiv* (2017), 157941.

[185] Scheer, M., Bülthoff, H. H., and Chuang, L. L. Steering demands diminish the early-p3, late-p3 and ron components of the event-related potential of task-irrelevant environmental sounds. *Frontiers in human neuroscience 10* (2016).

[186] SeizeIT. Discrete and personalized detection of epileptic seizures. `https://www.imec-int.com/nl/imec-icon/research-portfolio/seizeit`.

[187] Shea, K., and Tung, J. System identification of motion artifact: Noise in EEG headsets from locomotion. In *H-Workload 2017: The first international symposium on human mental workload* (2017), Dublin Institute of Technology.

[188] Shenoy, P., Krauledat, M., Blankertz, B., Rao, R. P., and Müller, K.-R. Towards adaptive classification for BCI. *Journal of neural engineering 3*, 1 (2006), R13.

[189] Shin, M.-S., Jeon, H., Kim, M., Hwang, T., Oh, S. J., Hwangbo, M., and Kim, K. J. Effects of smart-tablet-based neurofeedback training on cognitive function in children with attention problems. *Journal of child neurology 31*, 6 (2016), 750–760.

[190] Sidiropoulos, N. D., De Lathauwer, L., Fu, X., Huang, K., Papalexakis, E. E., and Faloutsos, C. Tensor decomposition for signal processing and machine learning. *IEEE Transactions on Signal Processing 65*, 13 (2017), 3551–3582.

[191] Simon, J. Z. The encoding of auditory objects in auditory cortex: Insights from magnetoencephalography. *International Journal of Psychophysiology 95*, 2 (2015), 184–190.

[192] Sitaram, R., Ros, T., Stoeckel, L., Haller, S., Scharnowski, F., Lewis-Peacock, J., Weiskopf, N., Blefari, M. L., Rana, M., Oblak, E., Birbaumer, N., and Sulzer, J. Closed-loop brain training: the science of neurofeedback. *Nature Reviews Neuroscience 18*, 2 (2017), 86–100.

[193] Smith, C. U. The triune brain in antiquity: Plato, aristotle, erasistratus. *Journal of the History of the Neurosciences 19*, 1 (2010), 1–14.

[194] Smith, M. E., Gevins, A., Brown, H., Karnik, A., and Du, R. Monitoring task loading with multivariate EEG measures during complex forms of human-computer interaction. *Human Factors: The Journal of the Human Factors and Ergonomics Society 43*, 3 (2001), 366–380.

[195] Sorber, L., Van Barel, M., and De Lathauwer, L. Optimization-based algorithms for tensor decompositions: Canonical polyadic decomposition, decomposition in rank-(l_r,l_r,1) terms, and a new generalization. *SIAM Journal on Optimization 23*, 2 (2013), 695–720.

[196] Sorber, L., Van Barel, M., and De Lathauwer, L. Tensorlab v2.0, 2014.

[197] Sorber, L., Van Barel, M., and De Lathauwer, L. Structured data fusion. *IEEE Journal of Selected Topics in Signal Processing 9*, 4 (2015), 586–600.

[198] Sorensen, M., and De Lathauwer, L. Coupled tensor decompositions for applications in array signal processing. In *Computational Advances in Multi-Sensor Adaptive Processing (CAMSAP), 2013 IEEE 5th International Workshop on* (2013), IEEE, pp. 228–231.

[199] Sørensen, M., and De Lathauwer, L. D. Coupled canonical polyadic decompositions and (coupled) decompositions in multilinear rank-(l_r,n,l_r,n,1) terms—part i: Uniqueness. *SIAM Journal on Matrix Analysis and Applications 36*, 2 (2015), 496–522.

[200] Sorger, B., Reithler, J., Dahmen, B., and Goebel, R. A real-time fmri-based spelling device immediately enabling robust motor-independent communication. *Current Biology 22*, 14 (2012), 1333–1338.

[201] Specker Sullivan, L., and Illes, J. Beyond 'communication and control': towards ethically complete rationales for brain-computer interface research. *Brain-Computer Interfaces 3*, 3 (2016), 156–163.

[202] Spinnato, J., Roubaud, M., Burle, B., and Torresani, B. Detecting single-trial EEG evoked potential using a wavelet domain linear mixed model: application to error potentials classification. *Journal of Neural Engineering 12*, 3 (2015), 036013.

[203] Stamile, C., Cotton, F., Maes, F., Sappey-Marinier, D., and Van Huffel, S. White matter fiber-bundle analysis using non-negative tensor factorization. In *International Conference Image Analysis and Recognition* (2016), Springer, pp. 650–657.

[204] Sterman, M. B., and Egner, T. Foundation and practice of neurofeedback for the treatment of epilepsy. *Applied Psychophysiology and Biofeedback 31*, 1 (2006), 21–35.

[205] Stomrud, E., Hansson, O., Minthon, L., Blennow, K., Rosén, I., and Londos, E. Slowing of eeg correlates with csf biomarkers and reduced cognitive speed in elderly with normal cognition over 4 years. *Neurobiology of aging 31*, 2 (2010), 215–223.

[206] Stopczynski, A., Stahlhut, C., Larsen, J. E., Petersen, M. K., and Hansen, L. K. The smartphone brain scanner: a portable real-time neuroimaging system. *PloS one 9*, 2 (2014), e86733.

[207] Strickland, E. Facebook announces "typing-by-brain" project. `http://spectrum.ieee.org/the-human-os/biomedical/bionics/facebook-announces-typing-by-brain-project`, Apr 2017.

[208] Thompson, T., Steffert, T., Ros, T., Leach, J., and Gruzelier, J. EEG applications for sport and performance. *Methods 45*, 4 (2008), 279–288.

[209] Tromp, J., Peeters, D., Meyer, A. S., and Hagoort, P. The combined use of virtual reality and EEG to study language processing in naturalistic environments. *Behavior Research Methods* (2017), 1–8.

[210] Tzallas, A. T., Tsipouras, M. G., and Fotiadis, D. I. Epileptic seizure detection in eegs using time–frequency analysis. *IEEE transactions on information technology in biomedicine 13*, 5 (2009), 703–710.

[211] Uludağ, K., and Roebroeck, A. General overview on the merits of multimodal neuroimaging data fusion. *Neuroimage 102* (2014), 3–10.

[212] Urigüen, J. A., and Garcia-Zapirain, B. EEG artifact removal—state-of-the-art and guidelines. *Journal of neural engineering 12*, 3 (2015), 031001.

[213] Van Erp, J. B., Lotte, F., and Tangermann, M. Brain–computer interfaces: beyond medical applications. *Computer*, 4 (2012), 26–34.

[214] Van Eyndhoven, S., Francart, T., and Bertrand, A. EEG-informed attended speaker extraction from recorded speech mixtures with application in neuro-steered hearing prostheses. *IEEE Transactions on Biomedical Engineering 64*, 5 (2017), 1045–1056.

[215] van Vliet, M., Robben, A., Chumerin, N., Manyakov, N. V., Combaz, A., and Van Hulle, M. M. Designing a brain-computer interface controlled video-game using consumer grade EEG hardware. In *Biosignals and Biorobotics Conference (BRC), 2012 ISSNIP* (2012), IEEE, pp. 1–6.

[216] Vandecappelle, M., Vervliet, N., and De Lathauwer, L. Nonlinear least squares updating of the canonical polyadic decomposition.

[217] Vanderperren, K., et al. Reduction of alpha distortion in event related potentials. In *4th European Conference of the International Federation for Medical and Biological Engineering* (2009), Springer, pp. 1298–1301.

[218] Vanderperren, K., Mijovic, B., Novitskiy, N., Vanrumste, B., Stiers, P., Van den Bergh, B., Lagae, L., Sunaert, S., Wagemans, J., and Van Huffel, S. e. a. Single trial ERP reading based on parallel factor analysis. *Psychophysiol 50*, 1 (2012), 97–110.

[219] Vecchio, F., Babiloni, C., Lizio, R., Fallani, F. V., Blinowska, K., Verrienti, G., Frisoni, G., and Rossini, P. M. Resting state cortical eeg rhythms in alzheimer's disease: toward eeg markers for clinical applications: a review. *Supplements to Clinical neurophysiology 62* (2013), 223–236.

[220] Vervliet, N., Debals, O., Sorber, L., and De Lathauwer, L. Breaking the curse of dimensionality using decompositions of incomplete tensors: Tensor-based scientific computing in big data analysis. *IEEE Signal Processing Magazine 31*, 5 (2014), 71–79.

[221] Vervliet, N., Debals, O., Sorber, L., Van Barel, M., and De Lathauwer, L. Tensorlab 3.0, 2016. *Available online, Mar* (2016).

[222] Viola, F. C., Thorne, J., Edmonds, B., Schneider, T., Eichele, T., and Debener, S. Semi-automatic identification of independent components representing EEG artifact. *Clinical Neurophysiology 120*, 5 (2009), 868–877.

[223] Vogt, T., Herpers, R., Askew, C. D., Scherfgen, D., Strüder, H. K., and Schneider, S. Effects of exercise in immersive virtual environments on cortical neural oscillations and mental state. *Neural plasticity 2015* (2015).

[224] Vogt, T., Herpers, R., Scherfgen, D., Strüder, H. K., and Schneider, S. Neuroelectric adaptations to cognitive processing in virtual environments: an exercise-related approach. *Experimental brain research 233*, 4 (2015), 1321–1329.

[225] von Lühmann, A., Herff, C., Heger, D., and Schultz, T. Toward a wireless open source instrument: functional near-infrared spectroscopy in mobile neuroergonomics and BCI applications. *Frontiers in human neuroscience 9* (2015).

[226] von Lühmann, A., Wabnitz, H., Sander, T., and Müller, K.-R. M3ba: A mobile, modular, multimodal biosignal acquisition architecture for miniaturized EEG-NIRSbased hybrid BCI and monitoring. *IEEE Transactions on Biomedical Engineering 64*, 6 (2017), 1199–1210.

[227] Wascher, E., Heppner, H., and Hoffmann, S. Towards the measurement of event-related EEG activity in real-life working environments. *International Journal of Psychophysiology 91*, 1 (2014), 3–9.

[228] Wascher, E., Heppner, H., Kobald, S. O., Arnau, S., Getzmann, S., and Möckel, T. Age-sensitive effects of enduring work with alternating cognitive and physical load. a study applying mobile EEG in a real life working scenario. *Frontiers in human neuroscience 9* (2015).

[229] Why, W. B. Neuralink and the brain's magical future. `https://waitbutwhy.com/2017/04/neuralink.html#`, Jul 2017.

[230] Winkler, I., Brandl, S., Horn, F., Waldburger, E., Allefeld, C., and Tangermann, M. Robust artifactual independent component classification for BCI practitioners. *Journal of neural engineering 11*, 3 (2014), 035013.

[231] Yagi, Y., Coburn, K. L., Estes, K. M., and Arruda, J. E. Effects of aerobic exercise and gender on visual and auditory P300, reaction time, and accuracy. *European journal of applied physiology and occupational physiology 80*, 5 (1999), 402–408.

[232] Zander, T. O., and Kothe, C. Towards passive brain–computer interfaces: applying brain–computer interface technology to human–machine systems in general. *Journal of neural engineering 8*, 2 (2011), 025005.

[233] Zich, C., De Vos, M., Kranczioch, C., and Debener, S. Wireless EEG with individualized channel layout enables efficient motor imagery training. *Clinical Neurophysiology 126*, 4 (2015), 698–710.

[234] Zich, C., Debener, S., Kranczioch, C., Bleichner, M. G., Gutberlet, I., and De Vos, M. Real-time EEG feedback during simultaneous EEG–fMRI identifies the cortical signature of motor imagery. *Neuroimage 114* (2015), 438–447.

[235] Zich, C., Debener, S., Schweinitz, C., Sterr, A., Meekes, J., and Kranczioch, C. High-intensity chronic stroke motor imagery neurofeedback training at home: Three case reports. *Clinical EEG and Neuroscience 0*, 0 (2017), 1550059–417717398. PMID: 28677413.

[236] Zich, C., Schweinitz, C., Debener, S., Kranczioch, C., and Ieee. Multimodal evaluation of motor imagery training supported by mobile eeg at home: A case report. In *IEEE International Conference on Systems, Man, and Cybernetics (SMC)* (LOS ALAMITOS, 2015), IEEE International Conference on Systems Man and Cybernetics Conference Proceedings, Ieee Computer Soc, pp. 3181–3186.

[237] Zink, R., Baptist, A., Bertrand, A., Van Huffel, S., and De Vos, M. Online detection of auditory attention in a neurofeedback application. In *Proc. 8th International Workshop on Biosignal Interpretation* (2016), vol. in press.

[238] Zink, R., Hunyadi, B., De Vos, M., and Van Huffel, S. Data-driven clustering of P300 EEG data using coupled tensor decompositions. In *International Scientific Conference on Information Technology and Data Related Research* (2017), vol. in press.

[239] Zink, R., Hunyadi, B., Van Huffel, S., and De Vos, M. Classifying the auditory P300 using mobile EEG recordings without calibration phase. In *Engineering in Medicine and Biology Society (EMBC), 2015 37th Annual International Conference of the IEEE* (Aug 2015), pp. 1777–1780.

[240] Zink, R., Hunyadi, B., Van Huffel, S., and De Vos, M. Exploring CPD based unsupervised classification for auditory BCI with mobile EEG. In *7th International IEEE/EMBS Conference on Neural Engineering (NER)* (April 2015), pp. 53–56.

[241] Zink, R., Hunyadi, B., Van Huffel, S., and De Vos, M. Alpha and low-beta oscillatory patterns extracted with canonical polyadic decomposition relate to lda classifier performance in real-life mobile EEG. In *Proc. 8th International Workshop on Biosignal Interpretation* (2016), vol. in press.

[242] Zink, R., Hunyadi, B., Van Huffel, S., and De Vos, M. Mobile EEG on the bike: disentangling attentional and physical contributions to auditory attention tasks. *Journal of Neural Engineering 13*, 4 (2016), 046017.

[243] Zink, R., Hunyadi, B., Van Huffel, S., and De Vos, M. Tensor-based classification of an auditory mobile BCI without a subject-specific calibration phase. *Journal of neural engineering 13*, 2 (2016), 026005.

[244] Zink, R., Proesmans, S., Bertrand, A., Van Huffel, S., and De Vos, M. Online detection of auditory attention with mobile EEG: closing the loop with neurofeedback. *bioRxiv 218727* (2017).

Curriculum

Rob Zink was born in Heerlen, The Netherlands on the 28th of July, 1989.

He obtained two M.Sc degrees in the field of Cognitive and Clinical Neuroscience in 2011 (cum laude) and 2013 at Maastricht University in Maastricht, the Netherlands. During the master thesis projects he worked on *'Accurate timing of LRP events using single-trial EEG'* and *'Single-trial analysis of Face and House processing in multimodal EEG-fMRI data'* under the supervision of Dr. Fren Smulders and Prof. dr. Peter Stiers respectively. The latter work was conducted at STADIUS Center for Dynamical Systems, Signal Processing and Data Analytics, Department of Electrical Engineering (ESAT), KU Leuven with additional supervision of Prof. dr. ir. Sabine Van Huffel and Prof. dr. ir. Maarten De Vos.

In September 2013 he joined STADIUS (ESAT, KU Leuven) in Belgium as a PhD student under the supervision of Prof. dr. ir. Sabine Van Huffel and Prof. dr. ir. Maarten De Vos. He worked on the transition from laboratory EEG recordings to real-life scenarios by focusing on both signal and cognitive analysis of mobile EEG data. His interdisciplinary work focused on novel EEG data acquisition, EEG tensor decompositions and neurofeedback to investigate auditory attention outside a lab environment. In 2017 his work was internationally recognized as he was awarded the *"Best paper award in the field of Mobile Brain/Body Imaging research between 2014 and early 2017"*.

Besides this scientific endeavor, he has large interests in digital art, video/VR and science communication; the latter aimed at explaining research attractively to non-experts (age range 7-77) with modern visualization technologies. He contributed to numerous (inter)national science outreach activities such as workshops, demonstrations and contests. Moreover, he is selflessly involved in science communication projects in Belgium.

www.robzink.be

List of publications

Articles in internationally reviewed journals

1 Zink, R., Hunyadi, B., Van Huffel, S., & De Vos, M. (**2016**). Mobile EEG on the bike: disentangling attentional and physical contributions to auditory attention tasks. *Journal of Neural Engineering*, 13(4), 046017.

2 Zink, R., Hunyadi, B., Van Huffel, S., & De Vos, M. (**2016**). Tensor-based classification of an auditory mobile BCI without a subject-specific calibration phase. *Journal of Neural Engineering*, 13(2), 026005.

Article in preparation

1 Zink, R., Proesmans. S., Van Huffel, S., & De Vos, M. (**2017**). Online detection of auditory attention with mobile EEG: closing the loop with neurofeedback. - *2017/bioRxiv: 218727*

Papers in proceedings of international conferences

1 Zink, R., Hunyadi, B., Van Huffel, S., & De Vos, M. (**2017**). Data-Driven Clustering of P300 EEG Data Using Coupled Tensor Decompositions. SINTEZA - *International Scientific Conference on Information Technology and Data Related Research.* Belgrade, Serbia.

2 Zink, R., Baptist, A., Bertrand, A., Van Huffel, S., & De Vos, M. (**2016**). Online detection of auditory attention in a neurofeedback application. In Proc. *8th International Workshop on Biosignal Interpretation* Osaka, Japan.

3 Zink, R., Hunyadi, B., Van Huffel, S., & De Vos, M. (**2016**). Alpha and low-beta oscillatory patterns extracted with Canonical PolyadicDecomposition relate to LDA classifier performance in real-life Mobile. In Proc. *8th International Workshop on Biosignal Interpretation* Osaka, Japan.

4 Zink, R., Hunyadi, B., Van Huffel, S., & De Vos, M. (**2015**). Exploring cpd based unsupervised classification for auditory BCI with mobile EEG. In Proc. *7th International IEEE/EMBS Conference on Neural Engineering (NER)* (pp. 53-56). IEEE.

5 Zink, R., Hunyadi, B., Van Huffel, S., & De Vos, M. (**2015**). Classifying the auditory P300 using mobile EEG recordings without calibration phase. In Proc. *37th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)* (pp. 1777-1780). IEEE.

6 De Vos, M., Zink, R., Hunyadi, B., Mijovic, B., Van Huffel, S., & Debener, S. (**2013**). The quest for single trial correlations in multimodal EEG-fMRI data. In Proc. *35th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)* (pp. 6027-6030). IEEE.

Abstracts in proceedings of (inter)national conferences

1 Detection of auditory attention near real-time with mobile EEG – R. Zink *7th Dutch Bio-Medical Engineering Conference*, 26-27 January **2017**, Egmond aan Zee, The Netherlands

2 Exploring auditory BCI in natural outdoor scenaros with mobile EEG – R. Zink *14th National Day on Biomedical Engineering*, 4 March **2016**, Brussels, Belgium

3 Unsupervised BCI classification – Proof of concept, R. Zink *5th Dutch Bio-Medical Engineering Conference*, 22-23 January **2015**, Egmond aan Zee, The Netherlands

4 CPD based feature extraction in mobile EEG data. R.Zink *13th National Day on Biomedical Engineering*, 28 November **2014**, Brussels, Belgium

5 Face/House detection in single trial concurrent fMRI/EEG. R. Zink *12th National Day on Biomedical Engineering and Annual Symposium of the IEEE EMBS Benelux Chapter*, 6-7 December **2013**, Brussels, Belgium



FACULTY OF ENGINEERING SCIENCE
DEPARTMENT OF ELECTRICAL ENGINEERING
STADIUS - CENTRE FOR DYNAMICAL SYSTEMS, SIGNAL PROCESSING AND DATA ANALYTICS
Kasteelpark Arenberg 10
B-3001 Leuven
rob.zink@esat.kuleuven.be
www.robzink.be