KU Leuven
Biomedical Sciences Group
Doctoral School of Biomedical Sciences
Faculty of Medical Science
Department of Neurosciences

University of Lille
Doctoral School MADIS
UMR CRIStAL

DOCTORAL SCHOOL
BIOMEDICAL SCIENCES

A visual Brain-Computer Interface for gaze-free communication

ALGORITHMS & APPLICATIONS

Arne Van Den Kerchove

Jury:
Co-supervisor KU Leuven: *prof. Marc Van Hulle*
Co-supervisor Univ. Lille: *prof. François Cabestaing*
Chair examining committee: *prof. John Creemers*
Chair public defense: *prof. Koen Poesen*
Prof. Andrea Kübler
Prof. Fabien Lotte
Prof. Reinhold Scherer
Prof. Maarten De Vos
Prof. Adalberto Simeone

Dissertation presented in partial fulfillment of the requirements for the double degree of *Doctor of Biomedical Sciences* (KU Leuven) and *Doctor of Control Science and Signal Processing* (University of Lille)

Defended December 16, 2024

Université de Lille
École Doctorale MADIS
UMR CRIStAL

KU Leuven
Groupe des Sciences Biomédicales
École Doctorale des Sciences Biomédicales
Faculté des Sciences Médicales
Département des Neurosciences

Interface Cerveau-Ordinateur visuelle pour la communication sans le regard

ALGORITHMES & APPLICATIONS

Arne Van Den Kerchove

Thèse soutenue le 16 décembre 2024 en satisfaction partielle des exigences de double diplôme de Docteur en Traitement du Signal et des Images (Université de Lille) et Docteur en Sciences Biomédicales (KU Leuven) devant le jury composé de:

Président: *Prof. Koen Poesen, KU Leuven, Département des Neurosciences*
Rapporteuse: *Prof. Andrea Kübler, Univ. Würzburg, Institut de Psychologie*
Rapporteur: *Prof. Fabien Lotte, Inria Bordeaux, LaBRI*
Rapporteur: *Prof. Reinhold Scherer, Univ. Essex, École d'informatique et de Génie Électronique*
Examinateur: *Prof. Maarten De Vos, KU Leuven, Département de Génie Électrique*
Examinateur: *Prof. Adalberto Simeone, KU Leuven, Département d'Informatique*
Co-directeur de thèse (invité): *Prof. François Cabestaing, Univ. Lille, UMR CRIStAL*
Co-directeur de thèse (invité): *Prof. Marc Van Hulle, KU Leuven, Département des Neurosciences*

Abstract

Individuals with severe motor impairment, such as those with Locked-in Syndrome, face substantial challenges when using traditional, gaze-dependent brain-computer interfaces (BCIs). These systems require directing the gaze toward specific targets, which becomes unfeasible for individuals with limited or no eye control. This work addresses this limitation by developing gaze-independent BCI methods, focusing on improving covert visuospatial attention (VSA), where users can direct their attention to a target without corresponding eye movements. A key contribution is the compensation for event-related potential (ERP) latency jitter, a variability that negatively impacts decoding performance in covert VSA. By handling this jitter, the proposed method enhances communication accuracy without requiring gaze control, making BCIs more usable for individuals with motor impairments.

Beyond gaze independence, this work also advances general ERP decoding techniques by refining the structure of linear and multilinear estimators. These methods improve accuracy across a range of BCI settings, particularly when training data are limited. The introduction of structured regularization in both linear and multilinear models enhances interpretability and reduces training time and computational complexity. This allows for more efficient training and contributes to generalizability, allowing for more reliable systems that are adaptable to various contexts and user needs.

The proposed methods were validated in experiments involving both healthy individuals and seven individuals with severe physical impairment and impaired eye-motor control. These experiments demonstrated the robustness of the novel decoding methods, showing that the system could effectively decode covert attention even when direct gaze was impossible.

Keywords: brain-computer interface, electroencephalography, event-related potentials, visuospatial attention, (multi)linear decoding, physical impairment

Samenvatting

Personen met ernstige motorische beperkingen, bijvoorbeeld het insluitingssyndroom, ervaren aanzienlijke uitdagingen bij het gebruik van traditionele visuele brein-computer interfaces (BCI's). Deze systemen vereisen gerichte oogbewegingen wat onpraktisch is voor personen met beperkte of zonder oogbewegingscontrole. Dit werk ontwikkelt BCI-technieken onafhankelijk van de blik, met nadruk op bedekte visuospatiële aandacht (VSA) – aandacht zonder overeenkomstige oogbewegingen. Een belangrijke bijdrage is het compenseren van variabiliteit in de timing van event-gerelateerde potentialen (ERP's), een factor met een negatieve impact op de decodering van bedekte VSA. Door dit aan te pakken, verbetert de nauwkeurigheid van communicatie die onafhankelijk is van oogcontrole. Dit maakt BCI's bruikbaar voor personen met oog-motorische beperkingen.

Naast onafhankelijkheid van oogbewegingen, verbetert dit werk ook algemene ERP decodering door de structuur van lineaire en multilineaire technieken te verfijnen. Deze methoden verhogen de nauwkeurigheid in verschillende BCI-toepassingen, vooral wanneer kalibratiegegevens beperkt zijn. Gestructureerde regularisatie van lineaire en multilineaire modellen verhoogt de interpretateerbaarheid en verlaagt de trainingstijd en computationele complexiteit. Dit draagt bij aan efficiëntere training en generalisatie, wat de ontwikkeling van betrouwbaardere systemen toelaat, aangepast aan verschillende contexten en gebruikersbehoeften.

De voorgestelde methoden zijn gevalideerd in experimenten met gezonde deelnemers en zeven individuen met een ernstige fysieke beperking en verminderde controle over oogbewegingen. De resultaten wijzen op de robuustheid van de nieuwe decodeermethoden, waarbij werd aangetoond dat het systeem bedekte VSA effectief kan decoderen, zelfs zonder directe blik.

Trefwoorden: brein-computer interface, electro-encefalografie, event-related potentials, visuospatiële aandacht, (multi)lineaire decodering, fysieke beperkingen

Résumé

Les individus ayant des handicaps moteurs sévères, tels que ceux atteints de syndrome d'enfermement, rencontrent des défis considérables lors de l'utilisation d'interfaces cerveau-ordinateur (BCI) traditionnelles. Ces systèmes nécessitent de orienter le regard, une tâche impraticable pour les personnes ayant un contrôle oculaire limité ou inexistant. Ce travail vise à pallier cette limitation en améliorant des méthodes de BCI indépendantes du regard et en optimisant la décodage de l'attention visuo-spatiale (VSA) cachée, où l' utilisateurs oriente son attention sans mouvements oculaires correspondants. Une contribution clé de ce travail est la compensation de la variabilité de latence des potentiels évoqués (ERP), laquelle impacte la performance de décodage de VSA cachée. En gérant cette variabilité, la méthode proposée améliore la précision de la communication sans nécessiter de contrôle du regard, rendant les BCI plus utilisables pour les personnes ayant des handicaps moteurs.

Au-delà de l'indépendance du regard, le travail fait également progresser les techniques générales de décodage ERP en affinant la structure des estimateurs linéaires et multilinéaires. Ces méthodes améliorent la précision de classification dans une gamme de conditions BCI, en particulier lorsque les données d'entraînement sont limitées. L'introduction de régularisation structurée dans les modèles linéaires et multilinéaires améliore l'interprétabilité des classificateurs et réduit le temps d'entraînement ainsi que la complexité computationnelle. Cela permet un entraînement plus efficace et contribue à la généralisation, ce qui permet le dévelopment des systèmes plus fiables qui s'adaptent à divers contextes et besoins des utilisateurs.

Les méthodes proposées ont été validées lors d'expériences impliquant à la fois des individus en bonne santé et sept individus avec des handicaps physiques sévères et un contrôle oculomoteur altéré. Ces expériences ont démontré la robustesse des nouvelles méthodes de décodage, montrant que le système pouvait efficacement décoder l'attention cachée même lorsque l'observation directe de la cible était impossible.

Mots-clés : interfaces neuronales directes, électroencéphalographie, potentiels évoqués, attention visuo-spatiale, decodage (multi-)linéaire, handicapés moteurs

Contents

List of figures

List of tables

List of acronyms

ALS Amyotrophic Lateral Sclerosis

BCI brain-computer interface

BTTDA Block-Term Tensor Discriminant Analysis

CBLE Classifier-based Latency Estimation

CCA canonical correlation analysis

CP cerebral palsy

CSP common spatial patterns

cVEP code-modulated visually evoked potential

DMD Duchenne's Muscular Dystrophy

ECoG electrocorticography

EEG electroencephalography

EOG electrooculogram

ERP event-related potential

FRDA Friedreich's Ataxia

HODA Higher Order Discriminant Analysis

HOSVD Higher Order Singular Value Decomposition

ICA independent component analysis

ISI inter-stimulus interval

ITR information transfer rate

LCMV linearily constrained minimum-variance

LDA linear discriminant analysis

LiS Locked-in Syndrome

LOOCV leave-one-out cross-validation

MI motor imagery

MS Multiple Sclerosis

mVEP motion-onset visually evoked potential

PCA principal component analysis

ROC-AUC area under the receiver-operator characteristic curve

RSVP rapid serial visual presentation

SMA Spinal Muscular Atrophy

SNR signal-to-noise ratio

SSPGI severe speech, physical and gaze impairment

SSPI severe speech and physical impairment

SSVEP steady-state visually evoked potential

STBF spatiotemporal beamformer

STBF-emp STBF with empirical covariance estimation

STBF-shrunk STBF with shrunk covariance estimation

STBF-struct structured spatiotemporal beamformer

SVM support vector machine

TBI traumatic brain injury

tLDA block-Toeplitz linear discriminant analysis

UCD user-centered design

VSA visuospatial attention

WCBLE Classifier-based Latency Estimation with Woody iterations

Chapter 1

Brain-computer interfaces

People with severe speech and physical impairment (SSPI) due to neurodegenerative disease, stroke, traumatic brain or spine injury become entirely reliant on caregivers or family members in their day-to-day routine. Would it not be great if we could help them regain some quality of life by making them more independent?

The most basic ability any independent human being needs is interaction with their environment. In its core essence, this is the ability to communicate intentions, emotions, frustrations, or thoughts such that they might be acted on. For an able-bodied person, physical interaction happens through body movements, while communication is usually done through speech and body language. Both require sending signals through the body's nervous system to control muscles. Not everyone has this capability.

The key problem for severely paralyzed individuals is that the mind wants to go where the body cannot follow. Could we therefore not design a system that directly interacts with the mind? This system can then form an interface between the mind and any kind of technology, like a robot arm or speech synthesizer, interacting with the world on the user's behalf. Such a device is the brain-computer interface (BCI)[1], a system that directly couples actions in the 'real world' to the mental state of the user.

BCIs started as a speculative concept in the 1970's [277]. Now, they have gradually evolved from methods developed in the 1990's to establish basic communication from minimal brain activity (yes/no, left/right, ...) [299] into sophisticated interaction schemes that restore or enhance many capabilities by bypassing defective parts of the nervous or muscle system. By now, classic BCIs methods are starting to mature outside the lab setting and can be used as assistive communication technology by individuals with severe speech and physical impairment [298, 166]. Impressive cutting-edge experimental examples include a brain-spine interface allowing a paraplegic individual to walk again [154], a BCI

[1] Sometimes also termed brain-machine interface (BMI) when coupled to a physical actuator, like a robot arm or an exoskeleton. The term BCI is usually preferred for assistive technology and communication devices.

translating internal speech to a facial avatar speaking the imagined text [165], and fast communication through decoding imagined handwritten symbols [289]. Recently, the idea has also gained popular traction through Elon Musk's NeuraLink [182] initiative, which imagines a multipurpose BCI.

In general, devices processing direct inputs from the central nervous system are useful for rehabilitation and medical diagnosis or treatment. Additionally, they bring a new paradigm to human-computer interaction, especially when paired with virtual or augmented reality [171]. Ultimately, they have always been especially promising as assistive technology for people limited in their communication ability. It is this BCI-controlled communication where our focus lies, as it offers a transformative means for individuals with severe motor impairments to regain functional independence to write and speak through assistive technology. This allows them to talk to their loved ones and caregivers, exercise hobbies, and makes many aspects of daily life more accessible.

1.1 A direct interface to the brain

Simply put, a BCI records the user's brain activity, extracts some relevant output from this brain activity, and couples this output to a function of a device as illustrated in fig. 1.1. Optionally, the user can then observe the action of the device and adapt their behavior or brain activity accordingly, closing the loop[2].

Let's break this definition down and focus on its separate parts. First of all, we need to identify a signal that is a direct representation of what is going on in the brain. Multiple signals can be recorded from the brain, including blood flow and magnetic or electrical activity. For several reasons laid out in section 1.2, the electrical modality is especially interesting.

The recorded brain activity forms only one part of the interaction scheme presented in fig. 1.1. It would be inefficient to search all brain activity randomly for the desired output. Instead, knowing the specific activity to target would help significantly. A BCI operates under a specified *paradigm*, which defines how interaction is conducted. This typically means instructing the user to perform a task, like attending to a flickering stimulus. Background knowledge from neuroscience research allows us to couple specific brain activity with the information or actions conveyed via the BCI. The brain signal now *modulates* this information. In turn, detecting this signal can trigger an action, like typing the intended letter A or B in fig. 1.1. The BCI paradigm includes both the method of stimulation (if applicable) and the user's task, and is often closely tied to the feedback mechanism in a closed-loop system. As shown in section 1.3, BCI paradigms vary widely based on different brain systems.

[2]The BCI Society [254] has recently formalized this into the following definition: *"A brain-computer interface is a system that measures brain activity and converts it in (nearly) real-time into functionally useful outputs to replace, restore, enhance, supplement, and/or improve the natural outputs of the brain, thereby changing the ongoing interactions between the brain and its external or internal environments. It may additionally modify brain activity using targeted delivery of stimuli to create functionally useful inputs to the brain."*

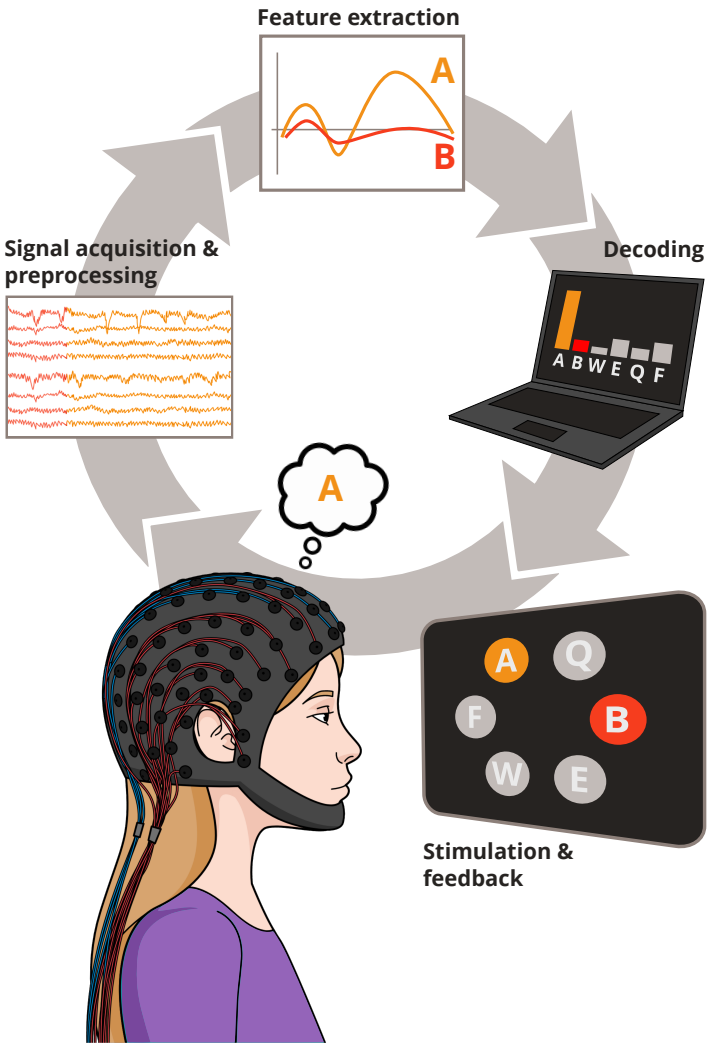


Figure 1.1: The BCI loop. The user interacts with the BCI via a specific paradigm, in this case involving visual stimulation. Electrical neural signals are recorded, and neurophysiological features related to the paradigm are extracted. Using machine learning, a decision can then be made based on these features, which can be presented back to the user. In a closed-loop BCI, this feedback allows the user to adapt to the BCI.

Another critical component of a successful BCI is identifying brain signatures related to the paradigm within the recorded activity. The brain's electrical activity is weak compared to environmental noise from electronic devices and interference. Additionally, the brain continuously processes information and carries out 'background tasks', which generate more brain activity. The desired activity often originates from specific brain regions or networks. This signal can be difficult to isolate from all other brain and environmental activity. It is similar to trying to hear a single speaker at a noisy party where everyone is shouting. Some interference can be filtered out through *preprocessing* (see section 1.4) and by selecting the correct signal representation (i.e., *feature extraction*), but retrieving or *decoding* the relevant signal remains challenging. Supervised machine learning can help solve this. The information encoded in the brain signal can be retrieved by the decoder and used to determine the desired BCI output. Section 1.5 reviews some common decoding techniques.

Finally, the loop can be closed by coupling this to a device or actuator. There are two aspects to this. On the one hand, the BCI gains its function by allowing the user to interact with their environment. For a communication system, this means coupling the decoded information to, e.g., a virtual keyboard or speech synthesizer. On the other hand, the actions performed by the BCI can themselves alter the user's cognitive state or brain activity, creating an adaptive system. This happens either directly through electrical neurostimulation, as is the case for adaptive deep brain stimulation in Parkinson's disease or for the cochlear implant as a hearing aide, or indirectly through sensory stimulation. The latter involves specific, paradigm-related, sensory input, like tactile feedback for movement BCI [75], or simply presenting selected actions back to the user. The user's brain will then adapt through learning, causing changes in behavior or strategy. Neuroplasticity, the brain's ability to adapt and reinforce specific neural pathways, can have a positive impact on BCI performance and gives rise to rehabilitation applications.

To summarize, a BCI can facilitate direct interaction with the central nervous system. To do this, information is modulated in the brain signal according to the BCI paradigm by stimulating the user or having them perform a task. Their brain signal is subsequently recorded and preprocessed to clean it. Relevant features that contain aspects of the modulated signal are extracted, and the modulated information is retrieved by the decoder, a machine learning model. This information can then be acted on, and the loop is closed if this action affects the user. Let us take a look at each of these elements separately in the next section.

1.2 Recording technologies

The brain's activity can be recorded using various neuroimaging technologies. These can range from brain scans [285] using functional magnetic resonance imaging (fMRI) to more portable technologies like the acoustic signals obtained by functional ultrasound imaging (fUS) [316], and functional near-infrared spec-

troscopy (fNIRS) [31]. These technologies all measure blood oxygen levels or blood flow, resulting in the *hemodynamic* signal. This signal indirectly reflects brain activity through metabolic reactions to an increas or decreas in activity. Usually, it reacts too slowly to reflect the real-time activity that is of interest in BCI for communication, and it carries little information other than the brain region from where it is originating.

A better signal candidate is the neuronal electrical activity. The brain contains 86 billion neurons, which are highly interconnected cells that are the smallest units of information processing. The *action potential* and its related *postsynaptic potential* are electrical pulses occurring when a neuron receives input from other neurons and is activated. The firing of a neuron, or the combined firing of groups of synchronized neurons, thus generates an electrical field in and near the brain, which can be measured using electrodes. The related discipline is *electrophysiology*.

Because of the desirable high temporal resolution [63] and practicality of electrophysiology, we will focus only on methods to record electromagnetic activity originating from neuronal action potentials and postsynaptic potentials. The magnetic field of the brain can be measured using magnetoencephalography (MEG) [164]. While MEG is non-invasive and results in high-quality signals, the necessary equipment is rather expensive and impractical. Recent advances have been made using optically pumped magnetometers (OPM-MEG) [295], but this technology still falters outside of lab conditions. As a consequence, the BCI field relies heavily on the electrical activity of the brain.

1.2.1 Invasive electrophysiology

The *invasiveness* of the technology forms an important concept in determining its suitability for a given application and user. As illustrated in fig. 1.2, invasive electrodes can often measure a very specific brain region and can result in better signal quality, at the cost of the risks involved with brain surgery.

Specifically, invasive methods are less noisy and come with an increased *spatial resolution*, meaning they can extract a signal from a specific brain region or even a set of neurons of interest. This makes it easier to retrieve a signal with a known origin from the brain, which might be leveraged by the paradigm to modulate information. Generally, microelectrodes penetrating the cortex are considered to have the highest spatial resolution. These are needles with a diameter smaller than 10 μm and one or multiple electrodes that extend several millimeters into the cortex. Microelectrodes can measure the Local Field Potential (LFP), the extracellular potential of a small group of neurons, or even intracellular single neuron action potential spikes. They come in single-electrode form, or, more popularly, in microelectrode arrays. Well-known examples are the Utah array [161], the more recent NeuraLink implant [182], and IMEC's Neuropixels 2.0 [248].

Larger implanted electrodes (usually a few millimeters in size) are referred to as intracranial electroencephalography (EEG) (iEEG). Depth electrodes, like those found in in stereo-EEG, can be used for BCI [302] and other applications,

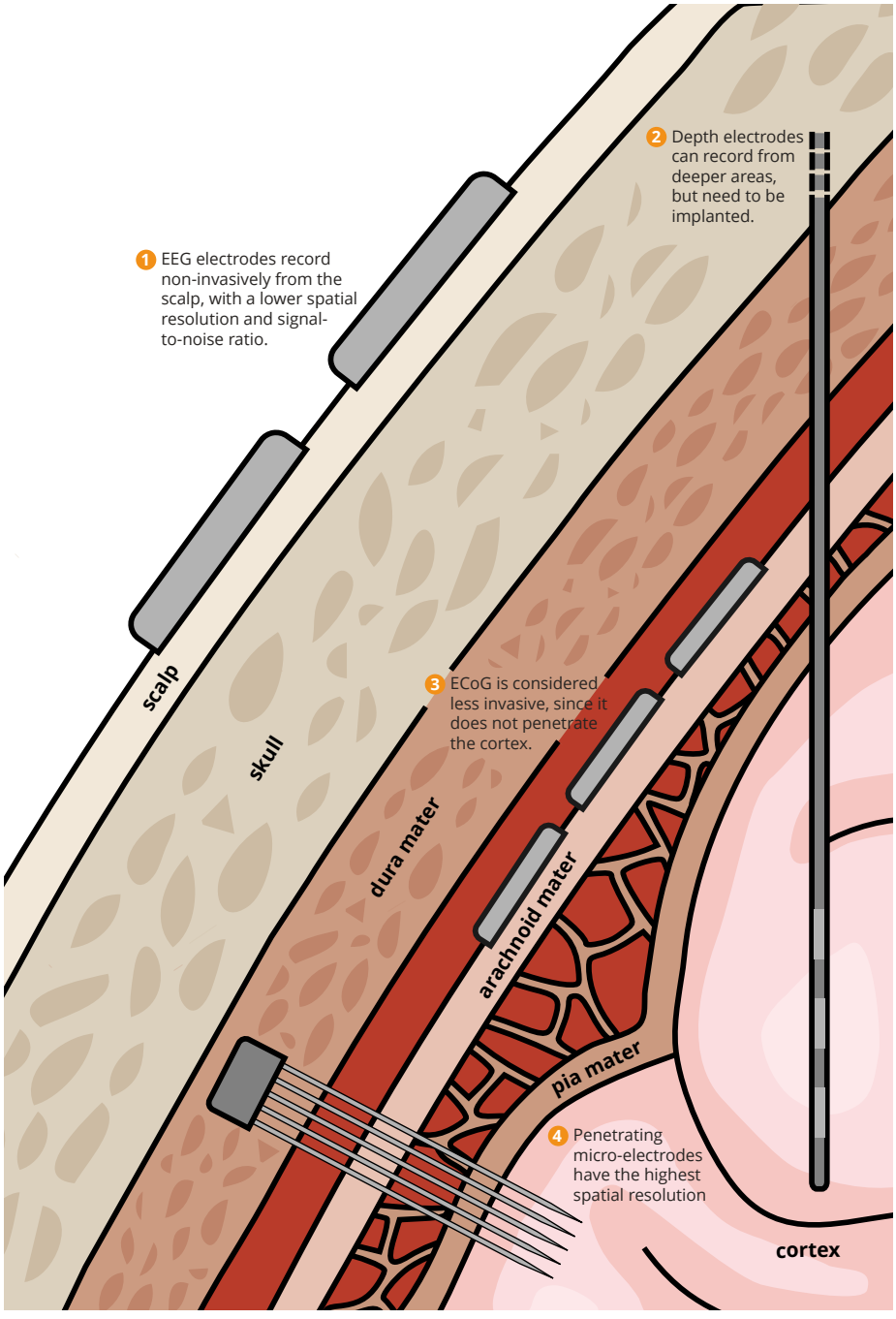


Figure 1.2: Various electrophysiology recording technologies and their respective position in a cross-section of
the skull. There is a trade-off between invasiveness
and spatial resolution.

like closed-loop adaptive control for Deep Brain Stimulation to mitigate Parkinson's disease symptoms [15]. Depth electrodes are longer rods with multiple electrodes that can extend deeper into the brain. Since these still penetrate the cortex, electrodes that sit on top of the cortex are more popular, as they are considered less invasive. Electrocorticography (ECoG) is a powerful method with many applications in BCI [234] and epilepsy diagnosis. Newer developments like micro-ECoG (μECoG), with a higher number of smaller electrodes per surface area (recording from sites with diameter of 10s of micrometers instead of $\pm$ 1 cm) [240], allow for more precise measurements.

Some impressive recent breakthroughs in speech and motor BCI for communication have been realized using invasive microelectrode arrays [289] and ţECoG [165]. Furthermore, recent advances in recording technology focus on improving implant durability and resolution [248], and on balancing the invasiveness tradeoff by finding new, minimally invasive ways to record from closer to the cortex. The Synchron Stentrode [169], for instance, can be delivered through the bloodstream via a catheter, removing the need for open brain surgery. Nevertheless, surveys [113, 114, 33] in different potential communication BCI user groups consistently show that non-invasive technology is preferred over implanted electrodes, unless the added value of invasive BCI is sufficiently large [125]. This provides justification for focusing part of the current research effort on non-invasive BCI.

1.2.2 Electroencephalography

EEG is the de facto non-invasive electrophysiology standard in clinical neurology practice as well as in BCI research. Developed since 1924 [25], it is cheap and relatively practical and offers great temporal sampling resolution up to thousands of Hz. EEG measures the electrical potential over time at a given electrode relative to a reference electrode in volts (V), often on the scale of microvolts (ţV). As it is easily applicable on the outside of the scalp, many practical consumer-grade EEG headsets exist in the form of (often wireless) electrode caps or helmets. Current consumer systems sometimes feature dry electrodes, but clinical and research systems often use active electrodes with preamplifiers and electrolyte gel to reduce impedance for improved signal quality, which slightly decreases practicality, since they require a trained individual to properly administer. Nevertheless, it is the most accessible BCI technology for both users and researchers.

The major drawback of EEG is its poor spatial resolution [73]. Since they sit on the outside of the scalp, electrodes are large and distant relative to the neurons in the cortex they ought to measure. Furthermore, the layers of skull and cerebrospinal fluid in between can attenuate the signal and cause volume conduction [36], which results in electrodes measuring a mixture of activity from sources elsewhere in the brain. This also has a negative effect on spectral resolution. The noise problem mentioned earlier is also prominently present in EEG recordings. In addition to picking up brain activity other than that of the region of interest, EEG also records other artifacts from the environment, the

power supply, nearby electrical equipment, or the user's muscle movements [269]. Combined, this means EEG is inherently noisy.

Despite its noisy nature, EEG is the recording methodology of choice for our BCI because of its wide acceptance by users. To counteract the noise present in our recording, we must on the one hand evoke strong, informative brain activity using a suited BCI paradigm, and on the other hand pick a suited decoder which can isolate the signal of interest and filter out the noise.

1.3 Paradigms

1.3.1 Active & passive BCIs

The BCI paradigm [304, 186] is the key to translating brain activity into useful output actions, since it defines which neural responses will be elicited and captured as features. A paradigm can be loosely seen as *a way of interacting with the BCI*. According to Zander and Kothe [310], they can be arranged by their reliance on external stimulation and the engagement of the user as shown in fig. 1.3.

In contrast to paradigms that require the user's active participation, for instance through redirecting attention or initiating imagined actions, paradigms free of this requirement are classified as *passive BCI*. From the user's point of view, this concept is probably the most attractive, since their cognitive state would be inferred from just their brain activity. However, without active user participation, it can be hard to establish reliable communication or control. Therefore, passive BCI are currently used for control tasks that require input, like emotional affect [261, 146, 179], workload [311] or fatigue [5] detection, or, more dependent on stimulation, error detection [172]. Aside from control, clinical applications are found in the diagnosis of a multitude of neurological and psychological disorders through various forms of neuroimaging or monitoring, in resting state or while performing a specific task. Nevertheless, active participation facilitates training data collection for supervised machine learning, since the conditions can be more easily controlled by the user or the stimulation.

Paradigms that do require high active participation can be split up into *active BCI* and *reactive BCI*. Active BCI paradigms encode endogenous activity initiated by the user, such as imagined movement or imagined speech. These tasks can encode complex information, but current non-invasive communication methods often limit the considered domain to a few movement directions to control a cursor, or a few words [199]. Invasive recordings are necessary for decoding more complex encoded activities like imagined natural speech [165] or meaningful motor trajectories, like handwritten symbols [289] or sign language [34]. Active BCI also has clinical applications in the form of neurofeedback [96], where the user adapts mental strategies to perform active self-regulation of their brain activity which is presented to them, causing a feedback loop.

Active paradigms are subject to large inter-subject variability due to the complexity of the performed tasks. They might require extensive user training

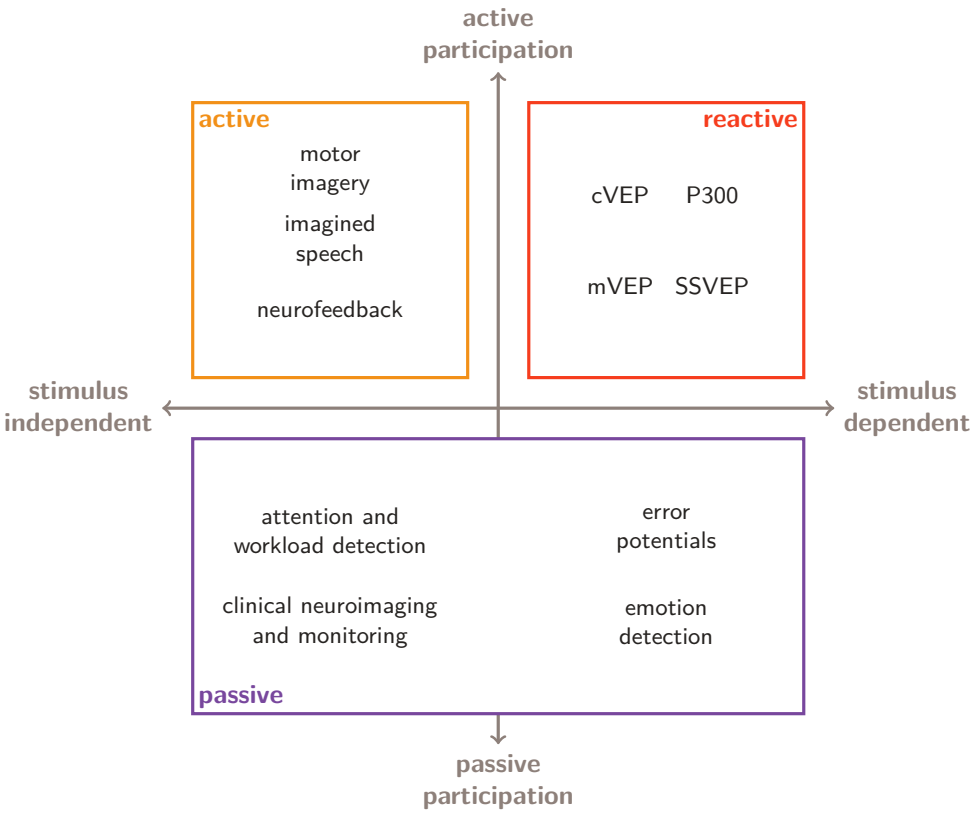


Figure 1.3: A BCI paradigm defines which brain activity will be used for control. They can be categorized in active, reactive and passive paradigms, depending on their reliance on external stimulation and active participation required from the user. Adapted from Mühl et al. [179].

and are sensitive to correct task instruction. Imagined speech or movements, or mental strategies in general, can be performed in a multitude of ways, which can themselves have different neural representations for separate individuals. This can make it hard to adapt the BCI to specific individuals, causing poor performance and giving rise to the concept of BCI illiteracy [7][3].

1.3.2 Reactive BCIs

Reactive BCI take another approach by providing a discrete set of sensory stimulations towards which the user's attention can be directed. Compared to active BCI, reactive paradigms can more easily induce fatigue due to the constant stimulation. These paradigms are also somewhat less intuitive compared to, e.g., speech, and their expressive power is limited by how many different stimuli can effectively be presented and attended to within a reasonable timeframe. Nevertheless, reactive BCI work well with EEG recordings, and, most importantly, they work for most people [6, 64].

Reactive BCI can be realized in multiple ways, depending on the stimulation used. Some examples include the following: In the steady-state somatosensory evoked potentials paradigm [211], the user can attend to one of multiple vibrotactile stimulations in different limbs, which encodes the information of the attended limb in the brain signal. In auditory paradigms, information can be modulated by the volume, tone, pitch, or spatial origin of presented sounds, to which the user can attend [127]. Most common are the visual paradigms. These can be more performant since they allow the user to use their visual system, one of the most evolved sensory systems in humans.

The major visual paradigms rely on steady-state visually evoked potential (SSVEP), code-modulated visually evoked potential (cVEP), oddball, and motion-onset visually evoked potential (mVEP) brain responses. In SSVEP, information is modulated by the frequency of a stimulus attended among many that each flicker continuously at different frequencies [49]. In cVEP, the stimuli flicker instead with distinct on-off patterns, which can be correlated with the brain activity to retrieve the attended stimulus [250]. Instead of stimulating all possible selections at once, the oddball paradigm stimulates them one by one with single flashes [198]. Since the times of stimulation of each target are known, and time-points at which an attention signature is detected can be correlated to selected targets. MVEP is usually similarly time-modulated, but stimuli make

[3]Recently, the concept of BCI illiteracy is increasingly being seen as outdated. Instead of attributing poor performance to the user's inherent limitations, critiques point out that the issue lies more with the design of the systems themselves. BCI systems often fail to account for individual variability in brain anatomy and cognitive strategies. As a result, what has been labeled as 'illiteracy' may, in fact, be a reflection of systems that are not flexible enough to adapt to different users. This perspective shifts the responsibility to the individual, rather than recognizing that better, more adaptive systems could overcome many of these performance challenges, thus standing in the way of progress. Moreover, the term 'illiteracy' itself carries negative connotations, implying a user deficiency that reinforces normative assumptions, when in reality, a more inclusive approach to design could enhance BCI accessibility for a wider range of individuals. This evolving view calls for a reframing of the problem, focusing on improving the adaptability of BCIs rather than labeling users as illiterate [21, 257].

sudden movements in specific directions instead of flashing. Information is then also carried by movement direction [149, 148].

The BCI paradigms mentioned above can also be combined to gain performance. Straightforward examples are activating or deactivating SSVEP stimulation with a motor command [186], or combining multiple visual paradigms by stimulating along multiple dimensions at once. Han et al. [97] recently used this strategy, combining frequency and phase coding in SSVEP with the mVEP and oddball paradigms to develop an efficient BCI where one of 200 targets can be accurately selected with only 800 ms of stimulation.

1.4 Preprocessing

EEG preprocessing is a critical step in brain-computer interface (BCI) systems, as it significantly improves the quality of the recorded data for subsequent analysis and classification. Common preprocessing techniques in BCI systems include rereferencing, band-pass and notch filtering, and independent component analysis (ICA).

Rereferencing is an essential first step, where the potential of each EEG channel is recalculated relative to a common reference point, such as the average of all electrodes (common average reference) or a selected neutral set of electrodes. This technique helps reduce noise common across all channels, improving signal clarity and enhancing the accuracy of subsequent processing steps.

Band-pass filtering is then usually applied to isolate the frequency ranges of interest by removing frequencies that are too high or too low to contain relevant neural signals. Band-pass filters eliminate slow drifts and high-frequency noise, such as muscle artifacts or environmental electrical interference, improving the signal-to-noise ratio (SNR). Additionally, notch filtering can be used to remove strong power line interference[4]. This interference can heavily contaminate EEG recordings and even leak through when outside the passband of a previous filter. Notch filtering efficiently attenuates this narrowband noise without affecting the surrounding frequencies.

Finally, to address artifacts like eye blinks and saccades that can significantly distort the EEG signal, ICA is performed [61]. ICA is a blind source separation technique that decomposes the EEG data into statistically independent components. Components corresponding to eye movements, eye blinks, and other bodily or environmental artifacts can be identified, either visually based on their characteristic topographies and time courses, or statistically, and then removed. This leads to cleaner data for subsequent processing, especially in the case of visual BCI paradigms. Since eye movement plays a role in the paradigm, it can strongly be present in the signal and correlate with the task, forming a confounding factor for studying control based solely on brain activity.

Together, these preprocessing steps ensure that only the relevant neural signals are passed on to the feature extraction and classification stages in BCI applications, improving the accuracy and reliability of the system.

[4]Typically at 50 Hz in Europe.

1.5 Decoders

1.5.1 Design principles

After preprocessing and extracting features, machine learning algorithms can perform BCI decoding. Machine learning decoders use some training data for which the performed tasks are known, and learn to recognize the activity elicited by the task in recorded data. The training data can be either obtained from the BCI user themselves or from other users. In the first case, the user would perform a short calibration session before using the BCI, where they are instructed to perform a known task. This calibration session can be eliminated by training the decoder on preexisting labeled data from other sessions and users, but this is harder due to the variability between subjects and sessions.

If the goal is to eliminate the per-session calibration, one can also use a decoder pre-trained on an existing dataset of the same task. This is complicated by large variability in measured brain responses between subjects and even between different sessions within the same subject [92, 230]. Therefore, pre-trained decoders must either be trained on a very large, diverse sample of subjects, or some form of transfer learning must be applied. Nevertheless, in practice, pre-trained models or models using transfer learning often still require some per-session fine-tuning, which necessitates some calibration. Instead, it is better to opt for keeping the calibration time as short as possible by using an algorithm that can learn from very few training samples. This can work well but requires strong regularization [271].

Decoder choice is heavily dependent on constraints set by the paradigm and the application. The number of output degrees of freedom, the serial or parallel nature of performed tasks, and whether the paradigm is active or reactive, all need to be taken into account. One of these choices is whether to solve a *regression* or *classification* problem. In regression, a continuous output variable is predicted for a new sample, while classification attempts to sort it into one of a set of discrete classes. Regression can be useful in BCIs for applications like mapping imagined movement to an external robot arm. For communication BCIs, however, conveying information in a symbolic manner is often of interest. Many applications resemble virtual keyboards or map a user's actions to discrete words, predefined actions, or letters for full control. Hence, such problems often present themselves as classification tasks.

1.5.2 Implementation & evaluation

Lotte et al. [155] and Xu et al. [304] present a relatively recent overview of state-of-the-art classification algorithms for decoding. Classic linear methods, such as common spatial patterns (CSP) feature extraction for motor imagery [201] and canonical correlation analysis (CCA) for SSVEP [184], and linear discriminant analysis (LDA) for event-related potential (ERP) classification [244] still perform relatively well, given some regularizing constraints or extensions. Multilinear techniques exploiting the tensor structure of neural signals are also promis-

ing [155]. Riemannian Geometry [19] is a popular and robust new strategy. Furthermore, they lend themselves well to applications like adaptive learning [23] or transfer learning [312]. Riemannian Geometry classifiers are often considered the current state-of-the-art. Finally, deep learning [26] is also sometimes considered, albeit only when sufficiently large training datasets are available. If decoders tailored to a specific user that keep the calibration session as short as possible are of interest, not enough training data is available to properly train a deep learning model.

Usually, a decoder makes a prediction for a given *trial* while operating the BCI. A trial is the smallest unit on which a selection decision can be made, for instance, one imagined movement or one repetition of flashing all targets in a visual ERP BCI. Accuracy is a valuable metric to assess decoder performance in the classification case. Accuracy is calculated as the proportion of correct selections to all selections made. It should be carefully interpreted in the presence of imbalanced data and always compared to the random chance accuracy level in the presence of more than two possible selections per trial. Alternatively, area under the receiver-operator characteristic curve (ROC-AUC) is also a measure of classifier predictive power but is more suited for evaluation of classification of single epochs of data and in the presence of imbalanced data, e.g., when comparing single ERPs. Higher ROC-AUC usually translates to a higher target selection accuracy.

Finally, an important concept in the evaluation of a BCI decoder, and of BCIs in general, is information transfer rate (ITR). ITR reflects how fluently a BCI can be used for communication and can be calculated as

$$\mathrm{ITR} = Q\left(\log_2 N + P\log_2 P + (1-P)\log_2 \frac{1-P}{N-1}\right) \tag{1.1}$$

ITR is expressed in bits/s and is dependent on N, the number of different options that can be selected per trial[5], P, the selection accuracy of the decoder, and Q, the number of trials per second. The parameters N, P, and Q of this formula give us some insight into the building blocks of a successful, high-ITR BCI. To improve performance, we can aim to 1) increase N by selecting a paradigm and interface that offer a broad range of informative selections per trial, 2) increase P by engineering more performant machine learning methods for decoding, and 3) increase Q by selecting a paradigm that allows fast stimulation or responses.

1.6 A case study: the visual oddball BCI

The visual oddball paradigm is a *reactive*, *stimulus-dependent*, BCI paradigm with all the benefits and drawbacks. Nevertheless, it can score relatively high on the ITR parameters established above. The brain response of interest is the visual ERP, which can be accurately measured and decoded from non-invasive EEG signals with a relatively low computational effort and short calibration

[5]Given that selections are independent of each other. This formula requires some adaptations in the case of sequential or hierarchical selection.

time. First established by Farwell and Donchin [71] in 1988, it is well supported by literature and has been proven to work for those with severe speech and physical impairment in day-to-day home use [298].

1.6.1 Stimulation paradigm

One by one, visual elements on a computer screen are *intensified* for a short period of time by changing in color, brightness, or size. Intensifications of different targets are usually 100-200 ms apart and last for about 100 ms [237]. On each flash observed by the user, specific brain activity is elicited. If one of these intensified elements is rare with respect to the others, i.e., it is the *odd-one out* or *oddball*, this brain activity is altered. A stimulus can be rare either due to its inherent properties like color or location. Crucially, this is also the case if it is 'marked' as rare by the user, e.g., by paying explicit attention to only a given stimulus and not the others or counting how many times it occurs. If these visual elements represent letters, and we know the timing of the intensifications of each letter, we can now establish communication by detecting for which letter an oddball response was present in the brain signal at its time of intensification.

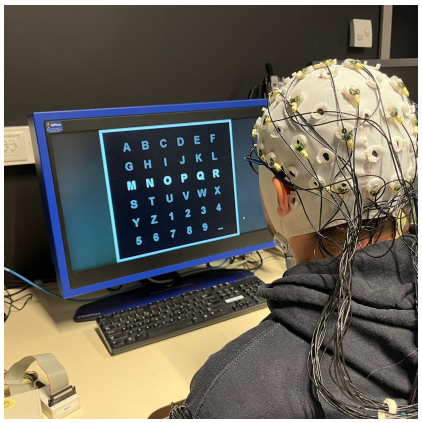
ITR in the oddball paradigm is optimized by intensifying multiple targets at once in a row-column strategy and using a sequence of selections, giving rise to the classic matrix speller of Farwell and Donchin [71]. Other optimizations increase response SNR, like using flashing face images as intensifications [122] or adding distinct colors and shapes to the stimuli [263]. The number of targets in a visual oddball BCI is limited by the crowding effect, which imposes a limit on how close targets can be to each other while not distracting the user's attention [237, 145]. This may be overcome by making hierarchical selections of stimuli representing groups of selections [264].

Together with SSVEP, oddball paradigms are most frequently used in visual BCI. There are, however, indications that users prefer oddball stimulation over SSVEP, which can cause eye strain and fatigue due to the continuous oscillatory stimulation of all targets at once [304][6]. Furthermore, SSVEP relies more on directing the gaze at the intended target than oddball, detracting from the concept of control independent from muscle movement.

1.6.2 The event-related potential

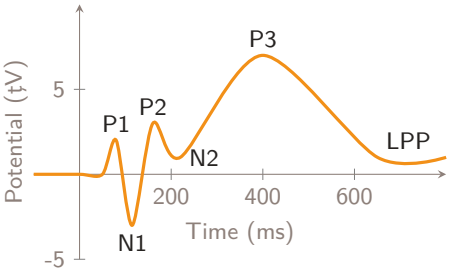
The time-domain response elicited immediately after the intensification of a stimulus is known as the visual ERP [157]. An ERP is a waveform consisting of multiple components. Some of these components are modulated by whether the stimulus was an oddball or not. These components appear as positive peaks and negative troughs in the ERP waveform. They follow a naming scheme based on polarity (Negative or Positive) and latency (e.g., N2 or N200 for a negative component after ±200 ms). The component most prominently modulated

[6] Recent work mitigates this by making SSVEP stimulation more comfortable without compromising selection accuracy through high-frequency low-contrast stimulation [136].

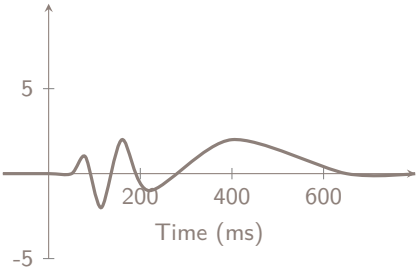


(a) Stimuli are flashed while the EEG is recorded.

(b) The P3 matrix speller interface. Entire rows or columns flash at once.



(c) The idealized ERP evoked by an attended stimulation in a single channel.



(d) The idealized ERP evoked by an unattended stimulation in a single channel.

by attention is called the P3 (positive deflection after ±300 ms). Some BCI[7]. Therefore, oddball paradigms are frequently referred to as P3(00) paradigms. However, due to the equally important contributions of other ERP components in decoding, as we will show later, this could be considered a misnomer. Thus, we will adhere to the oddball paradigm naming.

The visual ERP components primarily include P1, N1, P2, N2, and P3. Each component is characterized by specific latency, amplitude, and neural origins. These factors influence the perception of visual stimuli and attentional

[7]In ERP analysis, components are sometimes referred to with the timing around which they occur (e.g., N170, negative after 170 ms), or by their rank in order of occurrence (N1, the first negative component). The timing nomenclature is based on average latencies for neurotypical individuals, but is seldom correct for specific BCI users due to the large variability across ERP stimulation paradigms and subjects, and can also depend on the user's pathology. Furthermore, as we will see later, this work focuses specifically on the intra-session variability in ERP latency, explicitly assuming the latencies are not fixed. Therefore, we adhere to the ranking nomenclature. Fortunately, for the visual oddball paradigm, the rankings do roughly correspond to their expected timings, i.e., P1=P100, N1=N100, etc.

mechanisms [158].

The P1 component occurs approximately 100 ms after stimulus presentation. This component reflects initial visual processing and is primarily linked to activity in the primary visual cortex. P1 is particularly sensitive to gaze fixation, where the gaze is directed toward a stimulus. When a stimulus is fixated on, the amplitude of P1 increases, reflecting the enhanced processing of that visual input. In contrast, P1 shows less modulation when stimuli are attended to via attention without direct gaze.

The N1 component peaks around 150 to 200 ms post-stimulus. It represents an extension of sensory processing related to both gaze fixation and attention. For attended stimuli, the amplitude is greater, indicating prioritization for further cognitive processing. This highlights a more selective form of attention allocation, demonstrating the impact of both types of attention on visual processing.

The P2 component occurs between 200 to 300 ms. It reflects higher-order processing, particularly in contexts demanding stimulus evaluation and feature detection. P2 is sensitive to attention. When participants actively engage with specific features or categories of stimuli, P2 shows increased amplitude. Conversely, P2 may be less responsive when gaze is not directed toward the stimulus but attention is still maintained.

The N2 component peaking around 200 to 350 ms is associated with cognitive control processes such as conflict monitoring and inhibition. It is particularly pronounced in tasks requiring differentiation between competing stimuli. N2 reflects the allocation of attentional resources required to resolve conflicts, irrespective of whether the gaze is directly on the competing stimuli or not. Higher amplitudes in N2 are seen when cognitive demands increase, indicating the influence of attention strategies.

The P3 component is elicited in oddball paradigms and peaks around 300 ms post-stimulus. It reflects attentional engagement and the processing of rare or unexpected events. The P3 is typically subdivided into two subcomponents: P3a and P3b. The P3a is associated with the allocation of attention to novel or unexpected stimuli, indicating the initial detection of a change in the environment. It is particularly responsive to stimuli that capture attention, whether through gaze fixation or attention. P3a typically exhibits a frontal distribution and peaks earlier than P3b.

Conversely, the P3b relates to the evaluation of the stimulus. It involves the updating of cognitive resources in response to task relevance. This subcomponent is typically observed over parietal regions. P3b is elicited when participants must process the stimulus meaningfully, often requiring a decision or response. Both P3a and P3b amplitudes are sensitive to the probability of occurrence and

the participant's attentional focus, showing how attentional strategies affect cognitive processing.

The late positive potential (LPP) is a later component occurring beyond 400 ms. Sustained attention and emotional processing of stimuli are reflected in the LPP. Typically enhanced for emotionally salient or motivationally relevant information, the LPP serves as an indicator of cognitive appraisal beyond mere recognition tasks. This component can be influenced by attention when the participant evaluates the emotional significance of stimuli, even if gaze is not directed at them.

In summary, ERP components such as P1, N1, P2, N2, P3, and LPP collectively elucidate the neural underpinnings of visual processing and attentional mechanisms. These components not only inform the design of an effective oddball BCI system but also enhance our understanding of the cognitive processes involved in visual attention. This includes both gaze fixation and attention. Understanding these components is essential for optimizing BCI protocols and improving user interaction, particularly in contexts involving visual oddball tasks.

1.6.3 Preprocessing & feature extraction

For ERP analysis, the EEG is usually band-pass filtered between 0.5 Hz and 16 Hz. Re-referencing can be done to average mastoids to highlight the contribution of the P3 component. In addition to the steps outlined in section 1.4, the signal is usually cut into *epochs* for ERP analysis. These are time windows surrounding the stimulus event. They usually include some baseline interval before the stimulus, which can be used to normalize the response.

The epochs contain values representing the voltage of a specific channel relative to the reference at a specific time relative to the stimulus. In ERP classification, they are usually directly used as features, by concatenating all C channels of length S time samples in the matrix format epoch $\mathbf{X} \in \mathbb{R}^{C \times S}$ into one feature vector $\mathbf{x} \in \mathbb{R}^{CS}$ such that $\mathbf{x} = \text{vec}(\mathbf{X})$.

1.6.4 Decoders

Due to the unfavorable SNR of ERPs, it is often not feasible to accurately decode the encoded information based on just one example. Therefore, classification algorithms are trained on a set of epochs but are often evaluated on an average of multiple epochs. These averages are constructed over multiple stimulations of the stimulus representing a given selection, enhancing the SNR. The drawback, however, is that performing these multiple stimulations takes more time. To optimize the ITR, the trade-off between the number of repetitions, stimulation speed, and decoding performance must be carefully balanced. Performant classifiers can shift this balance to the benefit of ITR. Several techniques are suited to decode ERPs. Below are some methods that can achieve state-of-the-art performance and are relevant to this work.

Linearily constrained minimum-variance (LCMV) beamforming

Linear decoding techniques such as the LCMV beamformer [292] generally work well for classifying ERP epochs. LCMV beamforming learns a weight vector $\mathbf{w} \in \mathbb{R}^{CS}$ from a set of epochs that can be multiplied with the feature vector $\mathbf{x}$ of a new epoch to obtain a prediction score. The weights are obtained by minimizing the variance of the output

$$\arg\min_{\mathbf{w}} \mathbf{w}^\intercal \mathbf{C} \mathbf{w}^\intercal \tag{1.2}$$

under the linear constraint

$$\mathbf{a}^\intercal \mathbf{w} = 1 \tag{1.3}$$

with *activation pattern* $\mathbf{a} \in \mathbb{R}^{CS}$ the difference between the average of feature vectors of the attended and non-attended epochs in the training data, and $\mathbf{C} \in \mathbb{R}^{CS \times CS}$ the background covariance matrix which can be estimated empirically from the training data. This is then solved by the method of Lagrange multipliers as

$$\mathbf{w} = \frac{\mathbf{C}^{-1}\mathbf{a}^\intercal}{\mathbf{a}\mathbf{C}^{-1}\mathbf{a}^\intercal} \tag{1.4}$$

Linear discriminant analysis (LDA)

LDA is a more well-known method that is, in fact, nearly equivalent to LCMV beamforming [265] under certain conditions. It assumes normally distributed data with a different mean per class but a common covariance. Similar to LCMV beamforming, LDA learns a weight vector $\mathbf{w} \in \mathbb{R}^{CS}$ which satisfies

$$\arg\max_{\mathbf{w}} \frac{\mathbf{w}^\intercal \mathbf{S}_\mathrm{b} \mathbf{w}}{\mathbf{w}^\intercal \mathbf{S}_\mathrm{w} \mathbf{w}} \tag{1.5}$$

The between-class scatter matrix $\mathbf{S}_\mathrm{b} \in \mathbb{R}^{CS \times CS}$ models the variability between the expected responses, while the within-class scatter matrix $\mathbf{S}_\mathrm{w} \in \mathbb{R}^{CS \times CS}$ models the expected noise. The optimization criterion corresponds to finding a projection that maximizes the spread between classes (signal) while minimizing the spread within classes (noise). $\mathbf{w}$ can be obtained as the eigenvector of $\mathbf{S}_\mathrm{w}^{-1}\mathbf{S}_\mathrm{b}$ corresponding to the largest eigenvalue.

There is one caveat with these linear methods, however: the number of values ($= C \cdot S$) is usually relatively high compared to the size of the training dataset, since calibration time should be kept minimal. Therefore, strong regularization, such as covariance/scatter shrinkage, should be applied. Regularization can also be performed by retaining structure in the data, such as in the case for the multilinear methods presented by Lotte et al. [155].

Block-Toeplitz linear discriminant analysis (tLDA)

Another example of this is tLDA [244], which assumes temporal stationarity in the background noise to regularize the problem. It does this by imposing

a block-Toeplitz structure to the within-class scatter $\mathbf{S}_{\mathrm{w}}$ matrix of LDA. This means

$$\mathbf{S}_{\mathrm{w}} = \begin{bmatrix} \mathbf{S}_{\mathrm{w}}^{(1,1)} & \mathbf{S}_{\mathrm{w}}^{(1,2)} & \dots & \mathbf{S}_{\mathrm{w}}^{(1,C)} \\ \mathbf{S}_{\mathrm{w}}^{(2,1)} & \mathbf{S}_{\mathrm{w}}^{(2,2)} & \dots & \mathbf{S}_{\mathrm{w}}^{(2,C)} \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{S}_{\mathrm{w}}^{(C,1)} & \mathbf{S}_{\mathrm{w}}^{(C,2)} & \dots & \mathbf{S}_{\mathrm{w}}^{(C,C)} \end{bmatrix} \tag{1.6}$$

consists of $C \times C$ temporal scatter matrices $\mathbf{S}_{\mathrm{w}}^{(c_1,c_2)} \in \mathbb{R}^{S \times S}$ corresponding to a pair of channels (c_1, c_2), such that every temporal scatter matrix has the Toeplitz structure

$$\mathbf{S}_{\mathrm{w}}^{(c_1,c_2)} = \begin{bmatrix} s_1^{(c_1,c_2)} & s_2^{(c_1,c_2)} & \dots & s_S^{(c_1,c_2)} \\ s_2^{(c_1,c_2)} & s_1^{(c_1,c_2)} & \dots & s_{S-1}^{(c_1,c_2)} \\ \vdots & \vdots & \ddots & \vdots \\ s_S^{(c_1,c_2)} & s_{S-1}^{(c_1,c_2)} & \dots & s_1^{(c_1,c_2)} \end{bmatrix} \tag{1.7}$$

The structuring is performed by setting all elements to the average of their corresponding block-subdiagonal and subsequently applying shrinkage regularization. This method is simple yet effective, reaching state-of-the-art decoding performance.

Riemannian geometry

Alternative methods that currently work well include non-linear methods using Riemannian geometry [19]. Methods leveraging Riemannian geometry use a different feature extraction method. They do not use the raw epoch values but rather operate on the spatial covariance of an epoch. These covariance matrices C_i are symmetric and should have only positive eigenvalues; hence they lie on the Riemannian manifold of symmetric positive definite matrices, which defines the distance metric

$$\delta_{\mathrm{R}}\left(\mathbf{C}_1, \mathbf{C}_2\right) = \left\| \log\left(\mathbf{C}_1^{-\frac{1}{2}} \mathbf{C}_2 \mathbf{C}_1^{-\frac{1}{2}} \right) \right\| \tag{1.8}$$

This geodesic distance metric is naturally suited for this type of data and has some desirable properties, like invariance to linear transformations, which result in a robustness particularly useful for BCI applications [20].

With this distance defined, the epochs can either be directly classified on the Riemannian manifold using a minimum distance to the mean classifier, or they can be projected to the tangent space. This is a Euclidean space that locally approximates the Riemannian distance, which allows us to apply regular machine learning methods like an SVM or LDA. To enhance classifications of ERPs which rely on time-domain information, which would otherwise be lost in the spatial representation of the extracted feature, the class averages can be concatenated as additional channels to the epochs, embedding the temporal structure in the feature [19]. To improve performance even further, the ERPs

can be decomposed by a method like XDAWN before calculating the extended covariances [143].

These decoders define the last building blocks of our visual oddball BCI. Yet, the overview until now has been rather theoretical. Interesting and complex challenges arise when applying these concepts outside of the lab in augmented and assistive communication technologies for users with severe speech and physical impairment.

1.7 The gaze-dependence problem

In a nutshell, we can state that BCIs decode brain activity with the aim to establish a direct communication pathway bypassing speech and other forms of muscular activity [183, 47]. They have raised great hopes for individuals devoid of these abilities, for whom BCIs can provide a means to communicate or to control devices.

Furthermore, the EEG-based visual event-related potential (ERP)-based oddball interface is an effective and proven method [298, 239]. Targets are shown in short flashes on a computer screen that evoke an ERP when observed, which can be detected in the EEG signal. The ERP consists of multiple components, some of which are modulated by perception and the attention of the user. Decoding this modulation allows for information transfer controlled by the user's brain activity, as sketched in fig. 1.1.

However, when carefully examining the properties of the ERP components, we notice that while attention plays a role in the modulation of specific components, some part is also related to visual perception. How a stimulus is perceived, and, as a consequence, which ERP components are modulated, is directly related to whether the user gazes directly at it (i.e., *fixates* their gaze on the target) or not. Indeed, research has shown that visual oddball BCIs cannot operate efficiently when the user does not direct their gaze onto the desired target [38, 78].

This leads to a problem that is not often addressed: a visual BCI is dependent on the gaze of the user, and hence on muscle control. This is exactly what BCIs for augmented and assistive technology try to avoid. The BCI target population consists of individuals with severe physical impairment due to with various degrees of paralysis, or even with Locked-in Syndrome (LiS), the complete loss of muscle control with preserved consciousness. While BCIs are most attractive as a solution for individuals with SSPI, studies paradoxically often report poor performance in exactly this group. This is caused by a multitude of factors, including diminished electrophysiological responses and cognitive defects linked to the pathophysiology, psychological factors such as the "extinction of goal directed thinking" phenomenon in complete impairment, and, importantly, sensory impairments such as vision and oculomotor defects [236]. While visual BCIs often rely on visual sensory functions, it is precisely this group of patients that could benefit from them, that lacks adequate eye motor control. Equally often, participants with gaze impairments were excluded from the study.

Different degrees of eye motor impairment can render visual BCIs uncomfortable or outright impossible to use. When operating a visual BCI, the usual rapid series of forced saccades followed by fixation is tiring over time, even for healthy control subjects. A suitable alternative would allow the user to keep their eyes in an at-rest position of their choice while operating the BCI. This points to the need for gaze-independent solutions, which we will explore further in chapter 2 and the rest of this thesis.

Chapter 2

Gaze-independent visual BCIs

The last paragraph of section 2.2, section 2.1.2, and the first paragraph of section 2.3 were adapted from Van Den Kerchove et al. [272].

2.1 Eye motor impairment in BCI users

One of the goals of brain-computer interfacing is establishing a communication channel that does not rely on speech or muscular activity, which in turn can provide solutions to individuals with severe speech and physical impairment (SSPI). In the strictest interpretation, this means that an interface should not rely on the control of eye muscles used to redirect the gaze or for blinking. It is exactly this potential of BCIs that makes them suitable as assistive communication devices.

2.1.1 Incidence of eye motor impairment

SSPI is often caused by damage to the central or peripheral nervous system, either through congenital diseases (cerebral palsy (CP), Friedreich's Ataxia (FRDA), ...), neurodegenerative (Amyotrophic Lateral Sclerosis (ALS), Multiple Sclerosis (MS), ...) or acquired (stroke and traumatic brain injury (TBI)). Many individuals in these groups are unfortunately also afflicted by some form of eye motor impairment, requiring BCIs adapted to their condition. Table 2.1 reports the relatively high frequencies of eye motor impairment, which can range from minor (nystagmus, [1], other eye tremors [2], gaze fixation fatigue or discom-

[1] Involuntary, rhythmic, and repetitive eye movements recognizable by their consistent directionality (horizontal, vertical, or rotational).

[2] These can include square-wave jerks, saccadic intrusions, microtremors, or microsaccades while resting or fixating the gaze.

	ALS	MS	Stroke	CP	FRDA	DMD	SMA	LiS
Minor	50%	31%	40-70%	+	100%	+	-	
Severe	33%	3%	+	60-100%	>5%	-	-	98%
Complete	17%[5]	-	+	-	-	-	-	2%

Table 2.1: Incidence of eye motor impairment in selected BCI user target populations. ALS: Amyotrophic Lateral Sclerosis, MS: Multiple Sclerosis, DMD: Duchenne's Muscular Dystrophy, SMA: Spinal Muscular Atrophy, CP: cerebral palsy, LiS: Locked-in Syndrome. +: frequent, −: infrequent.

fort, ...), to severe (partial ophthalmoplegia [3], involuntary movements, impaired pursuit, ...), and even complete ophthalmoplegia[4] or eye motor paresis. This not only affects vision and coordination but also their ability to operate a visual BCI [79].

Among the most affected are those recovering from stroke [216, 228], and most of all those with due to brainstem or cerebellar stroke [173, 30]. Stroke can lead to the most severe severe or even complete eye motor impairment from the onset of their condition, resulting in the Locked-in Syndrome (LiS). However, case study reports [204, 87] show that even in individuals with LiS due to stroke, the group with a complete lack of eye motor control is very small.

ALS is progressive disease affecting motor neurons. This initially results in general weakness and loss of muscle tone, but eventually leads to full body paralysis. Although eye movement is often cited as one of the longest preserved capabilities in ALS, studies show that minor issues are still fairly common [126, 93, 175]. The bulbar-onset variant is characterized by an early loss of speech and increased involvement of eye motor symptoms [93]. Furthermore, Hayashi, Kato, and Kawada [105] show that as ALS progresses past the point of independent breathing, symptoms will eventually also involve eye muscle paralysis. One of the goals of BCI has always been to support these individuals to ensure quality of life.

Various forms of eye movement abnormalities also occur often in MS. MS is a neurodegenerative disease involving demyelination of nerves. Eye motor abnormalities are especially well studied [181, 217, 45, 238, 214] and are often used as diagnostic tools. These abnormalities can be minor or severe, seldom progressing to complete paralysis. However, MS often comes with vision loss, further complicating interaction with visual BCIs.

FRDA is a neurodegenerative disease affecting the spinal cord, peripheral nervous system, and cerebellum, resulting in an impairing loss of muscle coordination. This almost always heavily affects eye movements [70, 110, 81, 52], with various forms of involuntary movements and trouble pursuing or fixating on targets. They also gradually have more trouble speaking but often retain

[3] Weakness or limited paralysis of one or more of the muscles that control eye movement, leading to restricted eye motion, but not complete paralysis. Often, a specific movement direction (up-down, left-right) is preserved.

[4] Full eye movement paralysis.

some muscular control.

Another group that is heavily affected, is CP [72]. Additionally, individuals with other neurodegenerative diseases like Spinal Muscular Atrophy (SMA) [10] and Duchenne's Muscular Dystrophy (DMD) [159] are sometimes also interested in BCI use, but their eye motor capabilities are mostly preserved.

2.1.2 Gaze impairment and Locked-in Syndrome

We believe it is generally not opportune to delve too deeply into the underlying ophthalmological and neurological mechanisms for each of the specific conditions, or even the exact symptoms. In clinical reality, every BCI user has different symptoms, which lead to their own unique set of preserved capabilities and visual skills. From a solution-oriented BCI engineering point of view, the etiology of the symptoms can be abstracted away. Instead of categorizing users by their etiology, we will refer to those who might benefit from BCI assistive communication technology as having severe speech and physical impairment (SSPI), in line with the terminology used by Fried-Oken et al. [79]. If they have eye motor impairment that affects their use of visual BCI or eye-tracking solutions, we use the term severe speech, physical and gaze impairment (SSPGI) (see fig. 2.1).

The most severely impaired of the user groups above constitute the LiS group, which forms one of the main BCI interest groups. In this work, we use the term LiS for a situation of (near) complete paralysis and difficulties or the inability to communicate independently, even with assistive technology. This corresponds to classes one and two defined by Wolpaw et al. [297][6]. Hence, some degree of severe or complete eye motor impairment is usually necessary to qualify as locked-in[7].

In general, the group of locked-in individuals with complete loss of eye motor function is very small. Hence, it would be interesting to focus BCI development efforts on the larger population of individuals with SSPGI who currently slip through the cracks of the assistive technology offerings. These are individuals whose severe eye motor impairment prevents them from using eye-tracking-based solutions. They currently use their remaining motor control to communicate by

[6] "The first class consists of people who are truly totally locked-in (e.g., due to end-stage ALS or severe cerebral palsy), who have no remaining useful neuromuscular control of any sort, including no eye movement. [...] This class is very small. [...] The second class of potential BCI users comprises those who retain a very limited capacity for neuromuscular control. This group includes people who retain some useful eye movement or enough limb muscle function to operate a single-switch system. Such control is often slow, unreliable, or easily fatigued. This group is much larger than the first." [297]

[7] Multiple definitions of LiS are encountered in BCI and neurological literature. Some definitions include only those with tetraplegy without eye movements used for communications. Others distinguish Complete Locked-in Syndrome (CLIS) with full body paralysis, including no eye motor control at all, from a LiS state with some preserved eye movements or minor motor output. While some definitions only include stroke or TBI with damage to specific regions in the brain (midbrain, brainstem, or cerebellum) [242], it can also generally refer to the state of full body paralysis or loss of muscle tone incurred in neurodegenerative diseases, combined with the inability to speak, such as occurs in late-stage ALS.

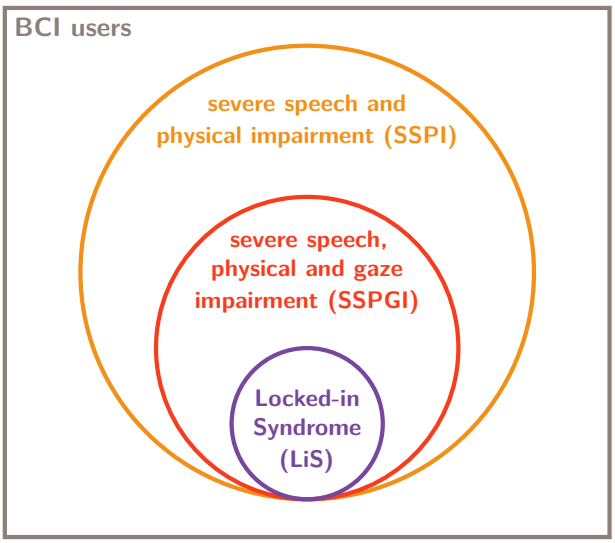


Figure 2.1: Venn diagram of potential users of assistive technology. Patients with severe speech and physical impairment (SSPI) are the general target population. If eye gaze-based solutions are not properly suited for an individual, they have severe speech, physical and gaze impairment (SSPGI). Individuals with LiS have no muscular control left, or it is very limited, such as the ability to communicate through a binary switching system. BCIs are especially relevant for the last two groups, as they are not able to use eye tracking solutions.

indicating symbols on a letterboard with great effort, or signal with upward eye movements or blinks to confirm prompted letters. They require a caregiver or relative to interpret their signals and crave the ability to communicate independently, which is crucial to retaining an acceptable quality of life.

This brings us back to the gaze-dependence problem in visual BCI: Traditional visual BCI scenarios require the user to *overtly* direct both their visuospatial attention (VSA) and gaze toward the screen target they intend to select. However, a critical challenge arises when users rely solely or in part on *covert* VSA, which involves directing VSA without corresponding eye gaze. In these cases, classical solutions often fall short of the widely accepted 80% target selection accuracy threshold [38, 78, 264, 226] deemed necessary for a comfortable user experience [186], calling for alternative, gaze-independent solutions.

In this work, we will use the term *gaze-independent* to mean *'dealing explicitly with the fact that a user cannot control their gaze.'* In the context of a visual BCI, this means that the user's visuospatial attention and their gaze do not necessarily coincide. There are multiple approaches to implement gaze-independence in a BCI. The next section highlights some noteworthy examples.

2.2 State-of-the-art

2.2.1 Gaze-independent modalities

Gaze-independent event-related potential (ERP)-based BCIs [223, 8] can be realized in three ways. Firstly, active BCI communication paradigms relying on endogenous activation from the user do not rely on sensory stimulation. Examples of this are imagined movement or imagined speech paradigms. Active paradigms can yield very high information transfer rates [289, 165], both due to their intuitiveness and the complexity that can be captured in commands, but they often do so only when paired with invasive recording. Non-visual reactive paradigms that use, for example, auditory [95] and somatosensory [37] stimulation do not rely on gaze redirection but also result in lower information transfer rates (ITRs) compared to visual paradigms. A survey by Riccio et al. [223] from 2012 reported that ITRs ranged between 1 and 10 bits/min for auditory paradigms and 0.5 and 5 bits/min for tactile [223].

On the tactile front, recent studies use steady-state somatosensory evoked potentials [211] or tactile oddball paradigms [280] These are analogous to the steady-state visually evoked potential (SSVEP) and visuall oddball paradigms. Instead of visual stimulation, however, they use physical vibrotactile pulses stimulation at different spatial locations and/or frequencies [98]. Some noteworthy advancements have achieved an ITR of 14.77 bits/min [151] and 20.73 bits/min [107] with a tactile oddball paradigm through highly optimized stimulation. Still, other recent tactile BCI studies [106, 98, 121, 144, 66] fail to exceed 5 bits/min.

Auditory BCIs are similarly implemented as paradigms oddball or steady-state stimulation with spatial or tone modulation. Zhang, Zhou, and Jiang [313] report that recent decoding methods can reach 16.99 bits/min ITR, but, in general, auditory BCIs are also limited in their speed and usually outperformed by visual paradigms [222].

While auditory and somatosensory stimulation on their own might yield poor ITR, Yin et al. [308] showed that these modalities can provide added value in gaze-independent settings when coupled with visual oddball stimulation in a hybrid paradigm. Their solution using the P3 components of bimodal stimulation reached an ITR of 14.94 bits/min. In general, their approach is interesting as it follows a philosophy that relies on establishing as many communication pathways as possible. Auditory and tactile stimulation can even be paired to create a fully non-visual, hybrid BCI, with a reported ITR of 11.66 bits/min [313].

Yet, auditory and somatosensory BCIs also suffer from increased mental effort in operation and from user-dependent variability [239, 221]. Severens et al. [239] showed that the visual Hex-o-Spell [264] outperformed a somatosensory alternative in participants with ALS whose eye motor capabilities were effectively impaired. However, non-visual stimulation modalities are still valuable for developing BCIs for individuals affected with severe vision loss or blindness [187, 229], which can occur in some cases of LiS [204].

2.2.2 Gaze-independent visual stimulation

This leads us to the second approach: current visual stimulation paradigms can be optimized so that the stimuli are always present in the field of view, either overtly [3, 300, 150] or covertly [213, 141]. Some noteworthy examples include the GIBS Block Speller [213], the GeoSpell interface [8], and the rapid serial visual presentation (RSVP) speller [2]. The RSVP paradigm, in particular, is a prime contender for a performant gaze-independent BCI. Lin et al. [150] reported an average online ITR of 20.259 bits/min using a paradigm with subsequently flashing sets of three characters filling the user's field of view.

We make a distinction here between spatially organized interfaces, where multiple targets are displayed at the same time at different spatial locations, and temporally organized interfaces, like RSVP, where targets or small sets of targets (e.g. 3) are shown consecutively. The Hex-o-Spell, GeoSpell, and GIBS Block Speller are spatially organized, while variants of the RSVP paradigm are temporally organized. Lin et al. [150] follow the philosophy that spatially organized interfaces have historically not performed well in the presence of gaze impairment. However, we argue that spatially organized interfaces generally have a much higher ITR in healthy controls. For instance, a regular matrix-based speller can achieve an ITR of up to 30 bits/min. With the right set of adaptations, spatial attention could potentially also bring added value to individuals with SSPGI. While spatial interfaces are indeed more gaze-dependent than temporal interfaces, they also provide an extra channel that can be used to transfer information, i.e., the location of the stimulus.

As an alternative to spatial attention, non-spatial visual attention (feature attention) can also be exploited, such as attention to stimulus color, shape, or symbol [314, 263, 117]. The RSVP speller is already an example of this, relying on the user to attend to the appearance of the intended character to type. Visual stimulation paradigms relying on alternative types of attention can modulate specific extra ERP components that either improve performance because they embed extra information in the brain signal [306] or because they are more sensitive to stimulation in the visual periphery [232]. Nevertheless, solely relying on alternative types of attention can also suffer from reduced information transfer rates [50]. Furthermore, the previously mentioned systems were typically tested only in settings where the user was required to focus on a central fixation point while selecting peripheral targets. This entails that they still rely to some extent on eye motor control, often necessitating central gaze fixation.

2.2.3 Gaze-independent decoding

Thirdly, stimuli can be presented in a standard BCI paradigm, but visuospatial attention can be decoded separately from gaze direction. Aloise et al. [8] aimed to bridge the performance gap between covert and overt VSA decoding performance. They compared classical linear and non-linear ERP classifiers on a covert attention oddball ERP paradigm dataset. The results revealed no

significant performance improvement in covert VSA decoding for any of the investigated decoders.

More recent work has made advances in decoding lateralized covert VSA by harnessing the N2pc ERP component [255, 221, 283]. The amplitude of the N2pc component in the contralateral hemisphere is directly modulated by the location of covert VSA. Another approach is directly decoding covert shifts in VSA from spectral content [260]. These methods are promising, but to date often require a slower stimulation pace or cannot display as many targets as the classical ERP BCI paradigms. As with the alternative modality approach, they are currently only helpful as extra communication channels in a hybrid paradigm. Xu et al. [305] reached an ITR of 23.56 bits/min by detecting the N2pc component while performing SSVEP stimulation. Egan et al. [65] also showed an increase in SSVEP performance by concurrently decoding the covert VSA shift from spectral content, but their study only used two stimulation targets. In general, we conclude that gaze-independent decoding in the fast-paced reactive ERP paradigm leveraging spatial attention remains underexplored.

2.3 Research objectives

To summarize, for individuals with SSPGI, gazing directly at a screen target may be uncomfortable, impractical, or even impossible. Hence, assistive devices that rely on eye tracking are often inefficient for them. Consequently, while visual BCIs hold great promise for these individuals, if they do not suffer from severe vision loss, conventional gaze-dependent BCI solutions do not meet their needs due to the absence of gaze control. Therefore, the development of decoding strategies that account for covert VSA becomes crucial in the pursuit of high-performance gaze-independent BCIs.

The general goal of this PhD is to tackle the gaze-dependence problem in the visual oddball BCI. Section 2.2 showed that visual BCI paradigms are a good candidate for a performant gaze-independent BCI. While the central gaze fixation of most visual gaze-independent paradigms still relies to some extent on eye motor control, we aim to circumvent this. In such an interface, stimuli can be presented in a standard BCI paradigm, but visuospatial attention can be decoded separately from gaze position, which do not necessarily need to coincide.

This leads us to adopt the following central hypothesis: *it is possible to improve the (gaze-independent) accuracy of a spatially organized visual oddball ERP-based brain computer interface by using a suited decoding strategy.*

To test this hypothesis, we chose a suitable interface, innovate on decoder development and test if these innovations improve BCI performance in a meaningful way. Accordingly, we collected data from healthy control subjects to test proposed decoders, and the findings have been verified in experiments with individuals with SSPGI, including individuals with LiS, focusing both accuracy and comfort.

The specific target of improving BCI performance under these conditions is

embedded within a broader goal of enabling effective communication for those with SSPGI. Ultimately, we wish to design a comfortable interface that allows them to maximally exploit their residual gaze capabilities. We believe this can be done with a non-invasive, spatially organized visual BCIs with high ITR. Finally, our efforts also aim to improve ERP decoding performance in general, which will also contribute to their effectiveness in both gaze-dependent and gaze-independent settings.

2.4 Approach

2.4.1 Decoder design

During this PhD, we explored different lines of decoding strategies, trying to tackle several challenges, be they general problems in ERP decoding or those that arise specifically from gaze-independence. Examples of these challenges include the lack or decrease in the amplitude of specific ERP components, and the increased non-stationarity of the signal. As mentioned, state-of-the-art decoders have poor performance in covert attention settings. The general goal is thus to design a machine learning classifier that represents the ERP signal in a way that makes it more robust to the problems occurring in covert attention conditions.

Regularized spatiotemporal beamforming

Due to the decreased amplitude of the N1 and P3 components in covert attention settings [264], the signal-to-noise ratio (SNR) of the ERP is lower than in overt attention settings. Therefore, a straightforward way to reach satisfactory gaze-independent decoding performance may be to increase overall ERP decoding performance.

To address this, we improved upon an in-house developed, state-of-the-art ERP decoder, the spatiotemporal beamformer (STBF) [292], by reformulating this classifier as a linear discrimination problem and imposing regularizing constraints by structuring the noise covariance matrix (STBF-struct). Furthermore, these regularizing constraints impose temporal stationarity on the background noise, yielding insights for our next efforts to cope with the non-stationarity of the P3 signal component. This approach has been published by Van Den Kerchove et al. [271] and will be described in chapter 3.

Tensor discriminant analysis

In the context of BCI decoding, extracting robust and discriminative features from multidimensional neural data is critical. Tensor decoding methods offer a powerful approach by preserving the multiway structure of electroencephalography (EEG) data while optimizing class separability. Unlike traditional methods that flatten data, TDA operates directly on tensors, making it partic-

ularly suited for BCI applications, where signals are structured across multiple channels and time points.

A relatively well-known example is Higher Order Discriminant Analysis (HODA) [212] It projects tensors onto subspaces optimized for class separation, similar to linear discriminant analysis (LDA) but extended to multidimensional data. HODA has been applied to various BCI tasks, including ERP and motor imagery (MI) decoding, enabling the extraction of class-relevant features while reducing the data's dimensionality.

While HODA is effective, it is limited by its reliance on a rigid core structure, which may not fully capture the complexity of neural data in some scenarios. To address this, Block-Term Tensor Discriminant Analysis (BTTDA) introduces a more flexible block-term structure. BTTDA iteratively extracts multiple blocks of discriminative information, which can help in capturing more complex patterns in neural data. While BTTDA is not specifically tailored to gaze-independent BCI decoding, specific tensorizations of neural data can offer advantages in coping with several obstacles arising from gaze-independence. The proposed BTTDA method will be discussed in chapter 4.

Classifier-based Latency Estimation with Woody iterations

The previous two methods did not yet yield specific results in gaze-independent decoding. The following approach is targeted at improving decoding performance in gaze-independent settings by accounting for a known property of the covert attention ERP response.

P3 latency generally falls between 350 ms and 600 ms [157], but this value is heavily dependent on the subject and the task and can vary from trial to trial [196]. The work of [12] illustrates that the variation in single-trial P3 latencies is important in gaze-independent decoding and has been hampering covert VSA decoding performance. We reprised their hypothesis, stating that jitter compensation through latency estimation and alignment improves covert VSA performance and extended it by developing a decoder.

Existing latency estimation methods are either not applicable to the classification problem of labeling unseen data, or are not robust enough to deal with the low SNR of the ERP. Classifier-based Latency Estimation (CBLE) as described by Mowla, Huggins, and Thompson [177] is a technique that can leverage ERP latency estimation in a decoding setting, but our results show that it yields no improvement in gaze-independent settings. We improved upon this technique and extended it to a probabilistic and iterative method, named Classifier-based Latency Estimation with Woody iterations (WCBLE). This approach has been published by Van Den Kerchove et al. [272] and is presented and evaluated on simulated data in chapter 5. We have also collected datasets to evaluate this approach on real ERP data in gaze-independent settings.

2.4.2 Data collection

The CVSA-ERP dataset

In the first series of experiments, we recorded data using this interface from healthy participants. The goal of these experiments was to benchmark gaze-independent ERP decoding algorithms. We denote this dataset as the Covert Visuospatial Attention ERP dataset (CVSA-ERP). This study was approved by the Ethics Commission of University Hospital Leuven (S62547). To simulate the dissociation between the eye gaze and the intended target (VSA), which could occur in individuals with gaze impairment, healthy participants were cued to operate the BCI in specific VSA conditions.

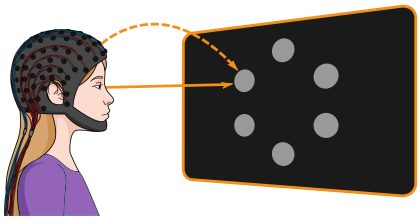
We chose a visual oddball interface to study and adapt to the effects of eye motor impairment on the ERP, using both existing state-of-the-art decoders and our own proposed decoders. Using a hexagonal layout interface, similar to the visual Hex-o-Spell proposed by Treder and Blankertz [264], we presented six flashing circular targets (without letters or symbols) to the participant while the EEG and electro-oculogram (EOG) were recorded, along with eye gaze using eye tracking. The Hex-o-Spell was chosen since it is optimized for gaze-independent performance and because it has already been tested in individuals with SSPI by Severens et al. [239].

The interface can be operated in different VSA conditions, as illustrated in fig. 2.2. In the overt case, users gazed at the cued target they were also mentally attending; in the covert case, users gazed at the center of the screen while mentally attending to the cued target. We introduced a VSA condition that is understudied in the context of gaze-independent BCI development: split VSA. In split VSA, the user mentally focuses on one cued target while gazing at another (the distractor). This last option has been scarcely studied but completes the options to dissociate gaze and visuospatial attention, allowing us to investigate the effect of (the lack of) gaze control on BCI performance. This dataset and the performance of our proposed WCBLE decoder have also been published by Van Den Kerchove et al. [272] and will be presented in chapter 6. Performance is also evaluated on a publicly available dataset.

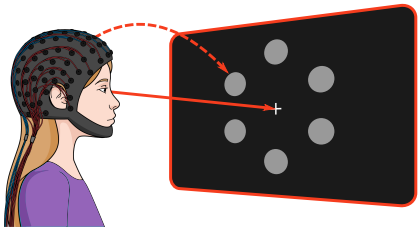
Case studies with gaze-impaired individuals

After collecting data from healthy participants and developing initial decoders, we conducted a study with seven individuals with SSPGI to evaluate the impact of eye motor deficits on BCI performance. This study recruited participants from neurorehabilitation centers and specialized care homes, including participants with conditions such as ALS, FRDA, and stroke.

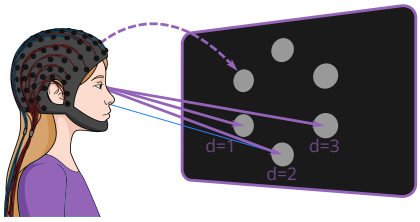
The primary aim was to assess whether gaze-independent decoding strategies could improve BCI performance in this group. We used the visual Hex-o-Spell interface [264], adapting the task by adding a free VSA setting, where participants could operate the interface without being required to fixate their gaze. This setting allowed us to investigate how participants with varying eye motor



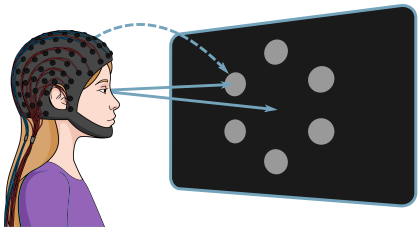
(a) *Overt* VSA. Gaze and VSA coincide on a target.



(b) *Covert* VSA. Gaze rests on the center of the screen, while VSA is directed towards a target.



(c) *Split* VSA. VSA is directed towards a target, while the gaze rests on another.



(d) *Free* VSA. The user is free to direct their gaze as they deem most comfortable. All previous VSA conditions are possible.

Figure 2.2: Visuospatial attention (VSA) conditions defined in our hexagonal spatial ERP paradigm interface. Full lines indicated gaze fixation, dashed lines VSA.

capabilities naturally interacted with a BCI, compared to standard overt and covert VSA tasks.

The hypothesis was that gaze-independent decoders, such as WCBLE, would improve performance in covert and free VSA conditions compared to standard decoders like block-Toeplitz linear discriminant analysis (tLDA). We aimed to determine if calibration in an overt VSA setting could also enhance performance in free and covert conditions, leveraging residual eye motor control during calibration to improve overall BCI accuracy for gaze-impaired users. The results of this study are presented in chapter 7.

Chapter 3

Kronecker-structured linear decoding

This chapter, except tables 3.1 to 3.4, was published as Van Den Kerchove et al. [271].

3.1 Introduction

There are multiple state-of-the-art P3 classification methods, like support vector machines (SVMs) [253], deep learning models [276, 32], and Riemannian Geometry classifiers [19]. While these models often return a high classification accuracy, there is a need for lightweight models – lightweight models lead to fast off-line analyses and can be transferred to consumer-grade hardware. When moving towards plug-and-play solutions, brain-computer interface (BCI) calibration sessions should be short and model training times low. The spatiotemporal beamformer (STBF) [274, 292] belongs to this class of event-related potential (ERP) decoding models as it achieves state-of-the-art performance and is fast to train. Earlier work has shown that it is possible to apply the spatiotemporal beamformer to multiple time-locked visual BCI paradigms, including the P3 oddball paradigm, steady-state visually evoked potential (SSVEP), code-modulated visually evoked potential (cVEP) [294], and the motion-onset visually evoked potential (mVEP) [147].

This work shows that the original spatiotemporal beamformer [292] can fall short in performance when BCI calibration data are restricted. We also show that the spatiotemporal beamformer does not scale well for higher spatial and temporal resolution cases. As a response to these issues, we introduce a regularization method that exploits prior knowledge about the spatiotemporal nature of the electroencephalography (EEG) signal to improve the accuracy for low data availability settings and speed up the classifier training time, thereby considerably reducing memory usage. Similarly structured regularization approaches

have been applied to other linear ERP classifiers [86, 279] and have shown significant increases in performance.

3.2 Materials & methods

3.2.1 Notation

We represent matrices with bold capital letters, vectors with bold lowercase letters, fixed scalars with uppercase cursive letters and variable scalars with cursive lowercase letters. The epoched EEG data with N epochs, C channels, and S samples are represented in epoch format as $\{\mathbf{X}_n \in \mathbb{R}^{C\times S}\}_{n=1}^{N}$ or flattened vector format by concatenating all channels for each epoch. Flattening results in $\{\mathbf{x}_n \in \mathbb{R}^{CS}\}_{n=1}^{N}$ such that $\mathbf{x}_n = \mathrm{vec}(\mathbf{X}_n)$. The real covariance matrix of the epochs in vector format is denoted by $\mathbf{C}$, estimators thereof as $\hat{\mathbf{C}}$.

3.2.2 Spatiotemporal beamforming

linearily constrained minimum-variance (LCMV)-beamforming was initially introduced to EEG analysis as a filter for source localization [273] to enhance the signal-to-noise ratio (SNR). Van Vliet et al. [274] first applied the spatiotemporal LCMV-beamformer as a method for the analysis of ERPs. The extension to the combined spatiotemporal domain [274] and the data-driven approaches proposed by Treder et al. [265] and Wittevrongel and Van Hulle [292] allow for its application to classification problems.

For the following analyses, we assume that all EEG channels are normalized with zero mean and unit variance without loss of generality. Solving eq. (3.1) under the linear constraint given by eq. (3.2) returns the filter weights $\mathbf{w}$ defining the spatiotemporal LCMV-beamformer.

$$\arg\min_{\mathbf{w}} \mathbf{w}^\intercal \mathbf{C} \mathbf{w}^\intercal \tag{3.1}$$

$$\mathbf{a}^\intercal \mathbf{w} = 1 \tag{3.2}$$

These weights minimize the variance of the output of the filter while enhancing the signal characterized by the constraint. $\mathbf{a} = \mathrm{vec}(\mathbf{A})$ is the data-driven activation pattern, a template of the signal of interest maximizing the difference between two classes of epochs, here denoted as the 'target' (T) and 'non-target' (NT) class. The activation pattern is then determined as follows:

$$\mathbf{a} = \frac{1}{N_\mathrm{T}} \sum_{\mathrm{T}} \mathbf{x_n} - \frac{1}{N_\mathrm{NT}} \sum_{\mathrm{NT}} \mathbf{x_n} \tag{3.3}$$

The method of Lagrange multipliers then gives the closed-form solution to the minimization problem posed by eq. (3.1) and eq. (3.2) as:

$$\mathbf{w} = \frac{\mathbf{C}^{-1}\mathbf{a}^\intercal}{\mathbf{a}\mathbf{C}^{-1}\mathbf{a}^\intercal} \tag{3.4}$$

The beamformer can be applied to epochs (unseen or not) as:

$$y_n = \mathbf{w}\mathbf{x}_n \tag{3.5}$$

resulting in a scalar output per epoch. The linear constraint in eq. (3.2) ensures that the beamformer maps epochs containing a target response to a score close to one and, conversely, epochs not containing a target response to a score close to zero.

3.2.3 Covariance matrix regularization

While the spatiotemporal beamformer, in theory, achieves optimal separation between target and non-target classes, in analogy to linear discriminant analysis [265], it does not always perform well on unseen data. The main challenge is to find a good estimator for the inverse covariance matrix $\mathbf{C}^{-1}$ since the real underlying covariance matrix generating the data is, in principle, unknown.

Empirical covariance estimation

Earlier spatiotemporal beamformer studies [292, 291, 293, 294] use the empirical covariance and inverse covariance calculated as follows:

$$\hat{\mathbf{C}}_{\text{emp}} = \frac{1}{N-1}\sum_{n=1}^{N}\mathbf{x}_n\mathbf{x}_i^{\intercal} \tag{3.6}$$

$$\widehat{\mathbf{C}^{-1}}_{\text{emp}} = \hat{\mathbf{C}}_{\text{emp}}^{+} \tag{3.7}$$

The Moore-Penrose pseudoinverse, $^{+}$, ensures a solution exists when $\hat{\mathbf{C}}_{\text{emp}}$ is singular. fig. 3.1a and fig. 3.1b respectively show examples of the empirical estimators of the covariance and the inverse covariance matrices. The empirical estimator suffers from performance and stability issues if the number of epochs N used or estimation is not much larger than the number of features CS [247, 129].

Shrunk covariance estimation

The shrinkage covariance estimator creates a better conditioned inversion matrix problem and generally performs better when applied to unseen data. The estimators for the covariance and inverse covariance are given by:

$$\hat{\mathbf{C}}_{\alpha} = (1-\alpha)\hat{\mathbf{C}}_{\text{emp}} + \alpha\frac{\text{Tr}(\hat{\mathbf{C}}_{\text{emp}})}{CS}\mathbb{I} \tag{3.8}$$

$$\widehat{\mathbf{C}^{-1}}_{\alpha} = \hat{\mathbf{C}}_{\alpha}^{+} \tag{3.9}$$

with $0 < \alpha < 1$. Analogous to L_2 regularization of the beamforming problem, shrinkage reduces the ratio between the smallest and largest eigenvalues of the

covariance matrix by strengthening the diagonal. fig. 3.1c and fig. 3.1d respectively show examples of the shrunk estimator of the covariance and the inverse covariance matrices.

Earlier work [147] applied shrinkage regularization to ERP decoding with the spatiotemporal beamformer and showed competitive performance compared to other state-of-the-art decoding techniques like stepwise LDA or SVM. The abovementioned work chooses the shrinkage coefficient α as a fixed hyperparameter. However, its optimal value depends on the number of training epochs, the covariance matrix's dimensionality, and the independence and variance of the data, which can vary across evaluation settings and per session. The optimal value for α can be found with a line search using cross-validation, but this can be a costly procedure.

Methods exist to estimate an optimal shrinkage value from the data directly. Most notable among these are the Ledoit-Wolf procedure [138], Rao-Blackwell Ledoit-Wolf [48], and Oracle Approximating Shrinkage [48]. A more recent estimation method [259] emulates a leave-one-out cross-validation (LOOCV) scheme expressed by the data-driven closed-form estimate:

$$\alpha = 1 - \frac{\begin{array}{r}\frac{N}{N-1}\operatorname{Tr}(\hat{\mathbf{C}}^2_{\text{emp}}) - \frac{2}{CS}\left[\operatorname{Tr}(\hat{\mathbf{C}}_{\text{emp}})\right]^2 + \frac{1}{CS}\operatorname{Tr}(\hat{\mathbf{C}}^2_{\text{emp}}) \\ - \frac{1}{N(N-1)}\sum_{n=1}^{N}||\mathbf{x}_n||_2^4\end{array}}{\begin{array}{r}\frac{N^2-2N}{(N-1)^2}\operatorname{Tr}(\hat{\mathbf{C}}^2_{\text{emp}}) - \frac{2}{CS}\left[\operatorname{Tr}(\hat{\mathbf{C}}_{\text{emp}})\right]^2 + \frac{1}{CS}\operatorname{Tr}(\hat{\mathbf{C}}^2_{\text{emp}}) \\ + \frac{1}{N(N-1)^2}\sum_{n=1}^{N}||\mathbf{x}_n||_2^4\end{array}} \tag{3.10}$$

We opt for the LOOCV shrinkage estimator because it avoids some of the assumptions made by Ledoit and Wolf [138] and Chen et al. [48] and because it generalizes to structured covariance estimation as described in section 3.2.3.

Spatiotemporal beamforming with Kronecker-Toeplitz structured covariance

Exploiting prior knowledge about the spatiotemporal structure of the EEG signal leads to a more regularized estimator of the covariance. When viewing the example of empirical spatiotemporal EEG covariance in fig. 3.1a, it becomes clear that this matrix consists of a block pattern of repeated, similar matrices. Due to the multi-channel nature of the signal, we assume that the covariance of spatiotemporal EEG epochs is a Kronecker product of two smaller matrices [180, 59, 115], as expressed by:

$$\hat{\mathbf{C}}_{\text{struct}} = \hat{\mathbf{S}} \otimes \hat{\mathbf{T}} \tag{3.11}$$

with $\otimes$ the Kronecker product operator. $\hat{\mathbf{S}} \in \mathbb{R}^{C\times C}$ and $\hat{\mathbf{T}} \in \mathbb{R}^{S\times S}$ respectively correspond to estimators of the spatial and temporal covariance of the data. Furthermore, because the temporal covariance of the EEG-signal is stationary (i.e., it is only dependent on interval length between covarying time samples) [28], it is assumed to have a Toeplitz-matrix structure:

$$\hat{\mathbf{T}}_{i,j} = \hat{\mathbf{T}}_{i+1,j+1} \tag{3.12}$$

Property 1 then leads to eq. (3.13) to estimate the inverse covariance.

Property 1 $(\mathbf{U} \otimes \mathbf{V})^+ = \mathbf{U}^+ \otimes \mathbf{V}^+$ *for any non-singular matrices* $\mathbf{U}$ *and* $\mathbf{V}$ *[137].*

$$\widehat{\mathbf{C}^{-1}}_{\text{struct}} = \hat{\mathbf{S}}^+ \otimes \hat{\mathbf{T}}^+ \tag{3.13}$$

Finally, based on property 2, eq. (3.4) can be reformulated more efficiently as eq. (3.14).

Property 2 $(\mathbf{U} \otimes \mathbf{V}) \cdot \text{vec}(\mathbf{W}) = \text{vec}(\mathbf{V}\mathbf{W}\mathbf{U}^\intercal)$ *for any matrices* $\mathbf{U} \in \mathbb{R}^{P \times P}$, $\mathbf{V} \in \mathbb{R}^{Q \times Q}$ *and* $\mathbf{W} \in \mathbb{R}^{P \times Q}$ *[153].*

$$\hat{\mathbf{w}}_{\text{struct}} = \frac{\hat{\mathbf{S}}^+ \mathbf{A}^\intercal \hat{\mathbf{T}}^+}{\mathbf{a} \cdot \text{vec}\left(\hat{\mathbf{S}}^+ A^\intercal \hat{\mathbf{T}}^+\right)} \tag{3.14}$$

Using eq. (3.14) removes the need to calculate the full, high dimensional Kronecker product $\hat{\mathbf{S}}^+ \otimes \hat{\mathbf{T}}^+$. Figure 3.1e and fig. 3.1f respectively show examples of the structured covariance and inverse covariance estimators, consisting of a spatial Kronecker factor (fig. 3.1g and fig. 3.1h) and a temporal component (fig. 3.1i and fig. 3.1j).

The Kronecker approach has shown significant performance yields in different linear spatiotemporal EEG and MEG applications [60, 115, 22, 85, 86]. Vliet and Salmelin [279] have applied a Kronecker-structured covariance estimator to ERP classification with linear models in a post-hoc fashion. Our work goes further by embedding the Kronecker structure in the spatiotemporal beamformer training process, using a data-adaptive shrinkage method, and regularizing the covariance further by imposing a Toeplitz structure on the temporal covariance.

Kronecker-Toeplitz structured covariance estimation

The question remains how to estimate $\hat{\mathbf{S}}$ and $\hat{\mathbf{T}}$. While the Flip-Flop and Non-iterative Flip-Flop algorithms [156, 286, 290] can estimate Kronecker or Kronecker-Toeplitz structured covariances, new results show that a fixed point iteration is more efficient [288, 287]. After each iteration, the spatial and temporal covariances matrices are scaled to unit variance to ensure the fixed point iteration converges. Finally, shrinkage can also be introduced in the Fixed Point Iteration to improve stability and achieve more robust regularization [287, 91, 22, 35]. The spatial and temporal covariance matrices are shrunk at every fixed-point iteration with shrinkage factors β_k and γ_k before matrix inversion in the next iteration.

Combined, this leads to the iterative estimation algorithm described by the following equations:

$$\tilde{\mathbf{S}}_{k+1} = \frac{1}{N} \sum_{n=1}^{N} \mathbf{X}_n^\intercal \hat{\mathbf{T}}_k^+ \mathbf{X}_n \tag{3.15a}$$

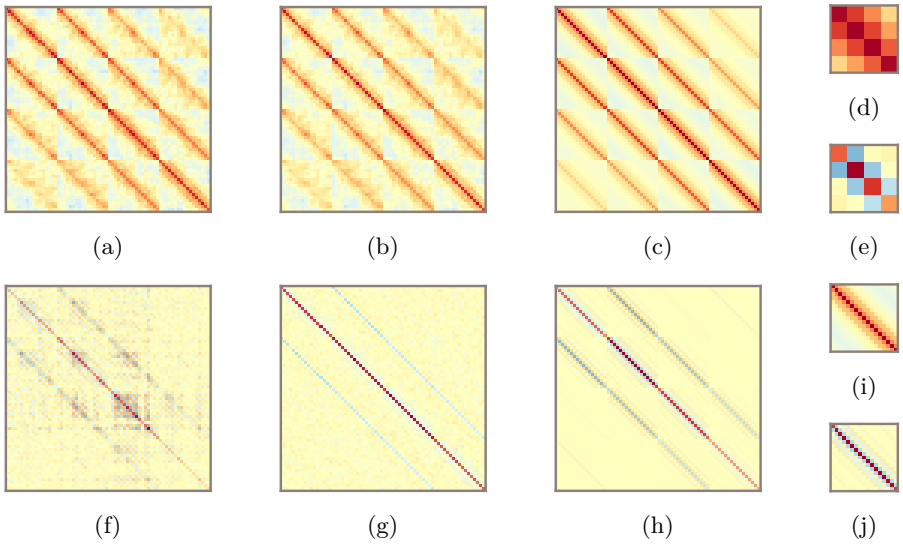


Figure 3.1: Different estimators of the covariance and inverse covariance of 100 epochs of data from *Subject 01* for channels *Fz*, *Cz*, *Pz*, and *Oz* and time samples between 0.1s and 0.6s. Regularized estimators of the inverse covariance exhibit less extreme values and have a sparser structure. (**a,f**) Empirical covariance and inverse covariance matrices. (**b,g**) Shrunk covariance and inverse covariance matrices with $\alpha = 0.14$ as determined by the closed-form LOOCV method. (**c,h**) Kronecker-Toeplitz structured covariance and inverse covariance matrices. (**d,e**) Spatial Kronecker factor of the Kronecker-Toeplitz structured shrunk estimator and its inverse. (**i,j**) Temporal Kronecker factor of the Kronecker-Toeplitz structured shrunk estimator and its inverse.

$$\tilde{\mathbf{T}}_{k+1} = \frac{1}{N}\sum_{n=1}^{N} \mathbf{X}_n \hat{\mathbf{S}}_k^{+} \mathbf{X}_n^{\intercal} \tag{3.15b}$$

$$\tilde{\mathbf{S}}_{k+1}^{(\beta)} = (1-\beta_{k+1})\tilde{\mathbf{S}}_{k+1} + \beta_{k+1}\frac{\mathrm{Tr}(\tilde{\mathbf{S}}_{k+1})}{C}\mathbb{I} \tag{3.16a}$$

$$\tilde{\mathbf{T}}_{k+1}^{(\gamma)} = (1-\gamma_{k+1})\tilde{\mathbf{T}}_{k+1} + \gamma_{k+1}\frac{\mathrm{Tr}(\tilde{\mathbf{T}}_{k+1})}{S}\mathbb{I} \tag{3.16b}$$

$$\hat{\mathbf{S}}_{k+1} = \frac{C}{\mathrm{Tr}\left[\tilde{\mathbf{S}}_{k+1}^{(\beta)}\right]}\tilde{\mathbf{S}}_{k+1}^{(\beta)} \tag{3.17a}$$

$$\hat{\mathbf{T}}_{k+1} = \frac{S}{\mathrm{Tr}\left[\tilde{\mathbf{T}}_{k+1}^{(\gamma)}\right]}\tilde{\mathbf{T}}_{k+1}^{(\gamma)} \tag{3.17b}$$

$\hat{\mathbf{S}}_0$ and $\hat{\mathbf{T}}_0$ can be initialized to any positive definite matrix. We choose to use the identity matrices $\mathbb{I}^{C\times C}$ and $\mathbb{I}^{S\times S}$. After each iteration, all diagonals of

$\hat{\mathbf{R}}_{k+1}$ are set to their mean values to ensure that $\hat{\mathbf{R}}_{k+1}$ and $\hat{\mathbf{T}}_{k+1}$ are Toeplitz structured.

Xie et al. [303] show that the LOOCV estimates for the optimal values of β_{k+1} and γ_{k+1} also yield a closed-form solution for the Kronecker fixed-point-iteration algorithm:

$$\beta_{k+1} = 1 - \frac{\frac{N}{N-1}\operatorname{Tr}(\tilde{\hat{\S}}^2_{k+1}) - \frac{2}{C}\left[\operatorname{Tr}(\tilde{\mathbf{S}}_{k+1})\right]^2 + \frac{1}{C}\operatorname{Tr}(\tilde{\mathbf{S}}^2_{k+1}) - \frac{1}{N(N-1)}\sum_{i=1}^{N}\left[\operatorname{Tr}(\mathbf{X}_i\hat{\mathbf{T}}_k^{+}\mathbf{X}_i^{\intercal})^2\right]}{\frac{N^2-2N}{(N-1)^2}\operatorname{Tr}(\tilde{\mathbf{S}}^2_{k+1}) - \frac{2}{C}\left[\operatorname{Tr}(\tilde{\mathbf{S}}_{k+1})\right]^2 + \frac{1}{C}\operatorname{Tr}(\tilde{\mathbf{S}}^2_{k+1}) + \frac{1}{N(N-1)^2}\sum_{i=1}^{N}\left[\operatorname{Tr}(\mathbf{X}_i\hat{\mathbf{T}}_k^{+}\mathbf{X}_i^{\intercal})^2\right]} \tag{3.18a}$$

$$\gamma_{k+1} = 1 - \frac{\frac{N}{N-1}\operatorname{Tr}(\tilde{\mathbf{T}}^2_{k+1}) - \frac{2}{S}\left[\operatorname{Tr}(\tilde{\mathbf{T}}_{k+1})\right]^2 + \frac{1}{S}\operatorname{Tr}(\tilde{\mathbf{T}}^2_{k+1}) - \frac{1}{N(N-1)}\sum_{n=1}^{N}\left[\operatorname{Tr}(\mathbf{X}_n^{\intercal}\mathbf{S}_k^{+}\mathbf{X}_n)^2\right]}{\frac{N^2-2N}{(N-1)^2}\operatorname{Tr}(\tilde{\mathbf{T}}^2_{k+1}) - \frac{2}{S}\left[\operatorname{Tr}(\tilde{\mathbf{T}}_{k+1})\right]^2 + \frac{1}{S}\operatorname{Tr}(\tilde{\mathbf{T}}^2_{k+1}) + \frac{1}{N(N-1)^2}\sum_{n=1}^{N}\left[\operatorname{Tr}(\mathbf{X}_n^{\intercal}\hat{\mathbf{S}}_k^{+}\mathbf{X}_n)^2\right]} \tag{3.18b}$$

The shrinkage parameters $0 < \beta_{k+1} < 1$ and $0 < \gamma_{k+1} < 1$ should be re-determined after each iteration. The Oracle Approximation Shrinkage method can also be used to determine β_{k+1} and γ_{k+1} [48, 303] but performs worse for spatiotemporal EEG data since not all assumptions are met.

3.2.4 Dataset

We use the dataset from Wittevrongel and Van Hulle [292], containing P3 oddball EEG recordings of 21 healthy subjects since it is a high-quality dataset with a high number (32) of electrodes and concurrently recorded EOG responses for ocular artifact rejection. Nine targets were arranged on a monitor before the subject during an experimental session. The subject was asked to pay attention to a cued target for a block of stimulations. The stimulations in a block are organized in 15 separate subsequent trials. A trial is defined as 9 stimulations in which each target is flashed precisely once per trial. Each target was cued four times, resulting in a dataset consisting of 36 blocks (4860 stimulations) per subject. Each stimulation will correspond to a single epoch in the preprocessed dataset. See Wittevrongel and Van Hulle [292] for a complete description of the dataset and the recording procedure.

3.2.5 Software and preprocessing

Data processing and classifier analysis were performed in Python using Scikit-Learn (v1.0.1) [205] and SciPy (version 1.7.1) [278]. The preprocessing pipeline was implemented using the MNE-Python toolbox (v0.24.0) [88]. The dataset

was converted to BIDS-EEG format [208] and managed and loaded with MNE-BIDS (version 0.9) [11]. The Riemannian classifier from section 3.2.6 was implemented using pyRiemann (v0.2.7). Statistical tests were performed in R (v4.1.2).

The EEG recorded at 2048 Hz was re-referenced off-line to the average of the mastoids. The reference electrodes were dropped from the analysis. Data were subsequently filtered between 0.5 Hz and 16 Hz using forward-backward filtering with a fourth-order Butterworth IIR filter. The EEG signal was corrected for ocular artifacts using Independent Component Analysis (ICA) by rejecting components that correlated with the bipolar EOG channels vEOG and hEOG using iterative Z-score thresholding. Components with a Z-scored pearson correlation coefficient exceeding 3 times the standard deviation of the Z-scored correlation coefficients of other components, were iteratively rejected until none that exceed the threshold remained. Finally, epochs were cut from 0.1s to 0.6s after stimulus onset. No baseline correction was performed since this affects the temporal covariance of the data, violating the Toeplitz structure assumption [28].

3.2.6 Classification

Cross-validation scheme per subject

We use a variation of grouped fold cross-validation per subject to evaluate the classifiers. We apply 4-fold cross-validation by splitting the blocks of each subject into four continuous folds. Unlike regular cross-validation, we only use a single fold to train the classifiers while using the other three folds for validation. This scheme results in a training set of 9 blocks of 135 epochs each. We chose this approach since we are primarily interested in the performance of the classifiers in the case of low data availability. The classification task is to determine the cued target for each block. The fraction of correctly predicted cues provides the accuracy of a classifier. Data from all trials are used in the training fold, while classifier validation is performed multiple times per fold, each time using an increasing amount of trials (i.e., using the first trial, using the first two trials, etc. until all 15 trials are used). For each of the 9 stimulated targets, the averages over the corresponding epochs across the utilized trials are used to predict the cued target in that block. The target with the maximal classifier score was then chosen as the predicted cued target. Before training the classifiers, a Z-score normalization transformation was developed on the training data to scale all EEG channels to unit variance. This transformation was then applied to the validation data.

Spatiotemporal beamformer classifier

Before calculating the STBF, the signal was downsampled to 32 Hz or twice the low-pass frequency 16 Hz, resulting in 17 time samples between 0.1 and 0.6s. According to the Nyquist Theorem, more samples would not contain more information hence the minimum temporal resolution is chosen to reduce the

dimensionality of the covariance and improve its condition number. The activation pattern is the difference between the averages of epochs in response to cued targets and the averages of those in response to non-cued targets. We constructed three variations of the spatiotemporal beamformer the STBF with empirical covariance estimation (STBF-emp) as in section 3.2.3, the STBF with shrunk covariance estimation (STBF-shrunk) as in section 3.2.3, and the structured spatiotemporal beamformer (STBF-struct) with LOOCV shrinkage for the Kronecker factors as in section 3.2.3.

Riemannian geometry classifier

We opted for a Riemannian geometry-based classifier to compare our results. The Riemannian model (XDAWN+RG) uses the XDAWN spatial filter combined with Riemannian geometry in tangent space as implemented by Barachant [18]. This classifier uses four XDAWN spatial filters and each epoch's empirical spatial covariance matrix. The target with the maximum score is the prediction of the cued target. XDAWN+RG was trained and validated without downsampling using epochs at the original sample rate of 2048 Hz. Preliminary experimentation revealed that best performance in low and high data availability setting was achieved when no shrinkage was applied to the estimation of XDAWN filters or the subsequent covariance matrices for this specific dataset. Hence, no shrinkage regularization will be applied to XDAWN+RG in further experiments.

3.3 Results

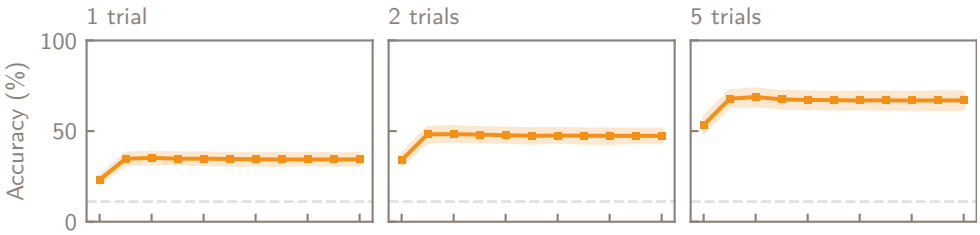
3.3.1 Minimum required fixed-point iterations

The fixed point iteration algorithm described in eqs. (3.15a) to (3.17b) estimates the Kronecker-Toeplitz structured covariance for the STBF-struct classifier. Fixed-point iteration is an iterative procedure starting from (in our case) non-informed initial guesses for the spatial and temporal covariance matrices. As a stopping criterion, one could impose a threshold on the difference in outcome of successive steps, e.g., based on the covariance norm or the classifier accuracy. However, few iterations or even just one [44] suffice to achieve satisfactory performance in practice.

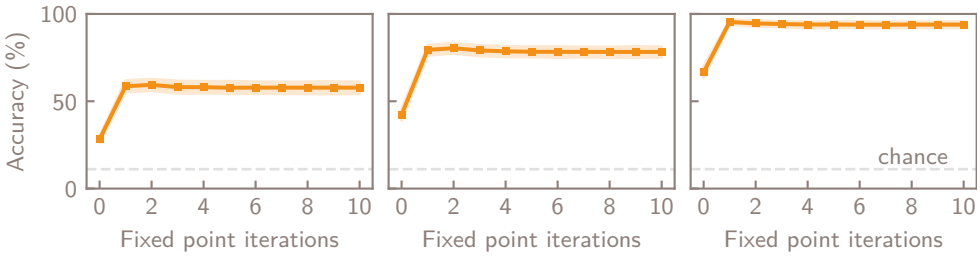
fig. 3.2 confirms these results for the STBF-struct classifier. Using more than one fixed-point iteration does not significantly improve the accuracy across the amounts of training data and the number of trials used for evaluation. Hence, only one iteration is used for the STBF-struct classifier, leading to a drastic speed-up of the training process.

3.3.2 Classifier accuracy for limited training data

It is of interest to keep the calibration time before BCI operation as short as possible. We mimic this problem by training the classifier with as few training



(a) Results for 1, 2, and 5 trials using only the first block in each training fold for training.



(b) Results for 1, 2, and 5 trials using all nine training blocks in the training folds.

Figure 3.2: Average cross-validated STBF-struct accuracy using one trial per block over all 21 subjects relative to the number of iterations used to estimate the Kronecker-Toeplitz structured shrunk covariance. Shaded area represents a 95% confidenced interval obtained through 1000 bootstrapping iterations. Accuracy does not improve when using more than one iteration.

epochs as possible. We evaluate the performance of all classifiers for different levels of available training data and apply the cross-validation procedure nine times (the number of blocks in the training fold) for all subjects, keeping the corresponding number of blocks in the training folds and dropping the rest. Figure 3.3, fig. 3.5 and tables 3.1 to 3.4 shows each classifier's accuracy relative to the data availability and trials used for agveraging. We statistically compare the two newly proposed classifiers, STBF-struct and STBF-shrunk with different levels of training data availability using a one-sided Wilcoxon signed-rank test with Holm correction for the multiple pairwise comparisons between classifiers. Statistical comparisons were performed for trial table 3.5a, two table 3.5b and five trials table 3.5c.

The tables show that STBF-struct has a significant advantage over STBF-shrunk when the number of training blocks is low, with significance $p = 0.005$ and an accuracy increase of 4,14%. for a single training block and testing trial. This effect is present for 1-, 2- and 5-trial evaluation, up to 3, 4 and 3 blocks respectively, but the advantage decreases when adding more training blocks. Both STBF-struct and STBF-shrunk perform significantly better than STBF-emp for all evaluated settings, with $p < 0.001$ for all number of training blocks

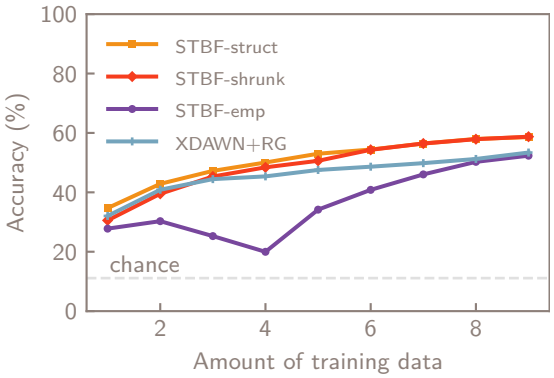


Figure 3.3: Accuracy of the different classifiers for all 21 subjects relative to the number of blocks available for training. One block consists of 135 epochs and corresponds to 27 seconds of stimulation. Accuracies are shown for the evaluation settings averaging over 1, 2, and 3 trials of testing stimuli. fig. 3.5 contains results for all numbers of trials. While STBF-emp is unstable when little training data are available, regularization of the covariance matrix (STBF-shrunk and STBF-struct) drastically improves performance.

and testing trials, with accuracy increase of 6.92 and 2.78%. respectively for a single training block and testing trial. Compared to XDAWN+RG, STBF-struct also has significantly higher accuracy in almost all evaluated settings, except when using only one training block, with significance $p < 0.001$ and an accuracy increase of 5.20%. when using all 9 training blocks and 1 testing trial. STBF-shrunk does not outperform XDAWN+RG when training data is low (up to 3,3 and 2 blocks respectively for 1, 2 and 5 testing trials), but gains a significant advantage when using more training data, with significange $p = 0.001$ and an accuracy increase of 5.33%. when using all 9 training blocks and 1 testing trial.

3.3.3 Classifier training time

In order to evaluate the training time of the investigated classifiers, the cross-validation scheme is run for each subject. For this analysis, downsampling to 32 Hz was also performed for the XDAWN+RG classifier for fair comparison. fig. 3.4 shows the measured training times. These results were obtained using a laptop with an Intelő Core i7-8750H CPU and 16GB of RAM.

Figure 3.4 shows that the training time of STBF-struct has a lower median training time (0.06 s) than STBF-shrunk (0.33 s), STBF-emp (0.30 s) and XDAWN+RG (0.20 s) when training on all available data and using 32 channels.

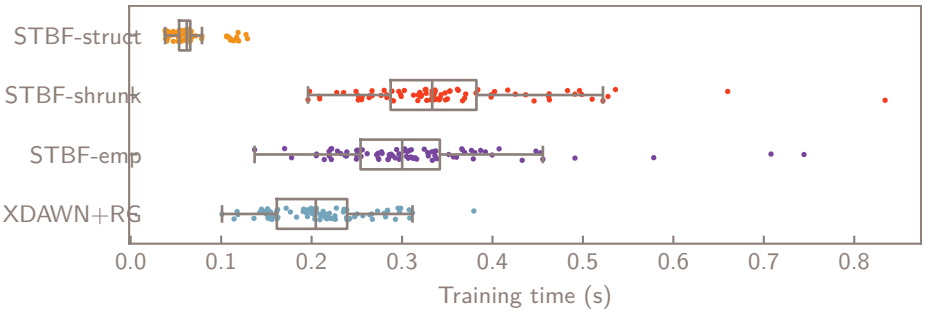


Figure 3.4: Training time boxplots of different classifiers and individual results for all training folds for all subjects using 32 channels and a sampling rate of 32 Hz. Using structured covariance estimation greatly reduces training time.

3.4 Discussion

3.4.1 Classification accuracy

As evidenced by fig. 3.3 and table 3.5, the regularized classifiers STBF-shrunk and STBF-struct significantly improve the classification accuracy compared to the original STBF-emp for all numbers of training blocks indicated. We believe there are three reasons for this. First and foremost, the empirical covariance matrix in STBF-emp becomes ill-conditioned when the number of available training epochs is smaller than the number of features ($N < CS$), rendering its inversion with the Moore-Penrose pseudoinverse unstable. This is the case STBF-emp when $N = CS = 32 * 17 = 544$, after which the accuracy of STBF-emp starts to increase. This effect is visible in fig. 3.3, where the accuracy starts increasing when using more than four training blocks, amounting to 540 epochs. The noticeable dip in accuracy when using around 540 epochs can be explained by numerical effects in the pseudoinverse for very small eigenvalues [29, 220, 233, 132]. Regularization of the covariance matrix with shrinkage ensures that the covariance matrix is non-singular and better conditioned so it can stably be inverted. Second, covariance regularization introduces a trade-off between variance and bias of the model [138]. Better performance on unseen data can be achieved when some model variance is traded for extra bias. Regularization reduces extreme values present, as shown in fig. 3.1, resulting in a classifier with better generalization. Third, the true spatiotemporal covariance matrix may vary throughout BCI sessions, e.g., due to movement of the EEG-cap, changing impedances of electrodes, subject fatigue, the introduction of new spatiotemporal noise sources, and other possible confounds. A regularized covariance matrix should better account for changes in true covariance. Note that the LOOCV method in principle assumes that the covariances of the training data and unseen data are the same. Because the covariance might have changed for unseen data, the shrinkage estimate obtained with LOOCV is probably still an underestimation of the optimal – but unknown – shrinkage coefficient that would yield

the best classification accuracy for the unseen data.

Another observation is the significantly better accuracy score of STBF-struct over STBF-shrunk when the amount of available training data is small (< 540 epochs). This property is an attractive advantage in a BCI setting since it is desirable to keep the calibration (training) phase as short as possible without losing accuracy. The accuracy advantage of the structured estimator is a consequence of the Kronecker-Toeplitz covariance structure, which is informative for the underlying process generating the epochs, if it is assumed that the EEG signal is a linear combination of stationary activity generated by random dipoles in the brain with added noise [180, 60, 86]. Hence, STBF-struct can utilize this prior information to better estimate the inverse covariance. The increase in accuracy for small training set sizes can also be explained by the smaller number of parameters necessary to estimate the inverse covariance (see section 3.4.2), increasing the stability of matrix inversions.

When compared to the state-of-the-art XDAWN+RG classifier, we conclude that STBF-struct reaches similar accuracy when using only one block of training data, with accuracies of and 34.70% 32.5% respectively and $p = 0.086$. The authors suspect this is due to both classifiers having insufficient training information to reach satisfactory classification accuracy. When more data are available, STBF-struct reaches significantly better accuracies. Combined with the benefits laid out in section 3.4.2, this makes it an attractive option for ERP classification. STBF-shrunk does not show decisive accuracy improvements over XDAWN+RG using a few training blocks, but this improves as the training data increases.

3.4.2 Time and memory complexity

As mentioned above, inverting the full $CS \times CS$ dimensional covariance matrix to construct STBF-emp and STBF-shrunk can be costly and unstable, in particular in high-resolution settings with many EEG channels or time samples. Constructing the full covariance and inverse covariance matrices also requires a considerable amount of memory. The structured covariance estimator of STBF-struct has two advantages here.

First, because of property 1 and property 2 there is no need to calculate or keep in memory the full $cs \times cs$ symmetric covariance and inverse covariance matrices for STBF-struct; they can instead be replaced by two smaller symmetric matrices respectively of dimensions $c \times c$ and $s \times s$. Furthermore, since the temporal component of the Kronecker product is Toeplitz-structured, it only requires s parameters to estimate. While the inverse covariance of STBF-emp and STBF-shrunk is defined by $\frac{CS(CS+1)}{2} = \frac{32 \cdot 17(32 \cdot 17+1)}{2} = 122\,128$ parameters accounting for the symmetric nature of covariance, the structured estimator only requires $\frac{C(C+1)}{2} + S = \frac{32(32+1)}{2} + 17 = 545$ unique parameters. This reduction in parameters to estimate reduces memory usage and contributes to the regularization effect for low data availability settings. The inverse covari-

ances of STBF-emp and STBF-struct, represented as $32*17 \times 32*17$ symmetric matrices of single-precision real floating point numbers for weight calculation, use 9.03MiB of memory. The 32×32 and 17×17 matrices of STBF-struct only require 5.12KiB.

Second, structured estimation has better time complexity. Covariance estimation and inversion occupy the largest part of the STBF training time. For STBF-emp and STBF-shrunk, the time complexity of this process is described by $\mathcal{O}(NC^2S^2+C^3S^3)$. Thanks to Property 1, the complexity can be reduced to $\mathcal{O}(NS^2S^2+C^3+S^3)$ for the structured estimator of STBF-struct. The results presented in fig. 3.4 confirm that STBF-struct is much faster in operation than other classifiers.

Since the training times of all STBF-based classifiers are already in the order of tenths of seconds, the question arises whether the improvements achieved by using the structured estimator would be relevant in practice. However, the authors believe that these results could significantly impact some use cases of the spatiotemporal beamformer, like high spatial or temporal resolution ERP analyses (e.g., $C >> 32, S >> 256$). One example is single-trial ERP analysis with a high-temporal resolution to extract ERP time features. Such higher-resolution analyses can later be incorporated into an ERP classification framework. In addition, the speed-up provided by structured estimation yields a faster off-line evaluation of the STBF ERP classifier, where often multiple cross-validation folds, subjects, and hyperparameter settings need to be explored, which can quickly increase runtime. Improvements in computation speed and memory usage can remove the need for dedicated computation hardware and enable running group analyses on a personal computer.

3.5 Conclusion

While it is possible to regularize the spatiotemporal LCMV beamformer classifier for ERP detection with other methods such as by employing feature selection, by adding regularizing penalties to the cost function beamforming problem, or by crafting a cleaner activation pattern, our work focused on estimation methods for the spatiotemporal covariance. We introduced a covariance estimator using adaptive shrinkage and an estimator exploiting prior knowledge about the spatiotemporal nature of the EEG signal. We compared these estimators with the original spatiotemporal beamformer and a state-of-the-art method in an off-line P3 detection task. Our results show that the structured estimator performs better when training data are sparsely available and that it can be computed faster and with substantially less memory usage. Since these algorithms are not paradigm-specific, the conclusions can be generalized to other ERP-based BCI settings.

Future work should focus on introducing more robust regularization strategies using prior knowledge, such as shrinking the spatial covariance to a population mean or a priorly known matrix based on sensor geometry or characterizing the temporal covariance as a wavelet or autoregressive model. More accurate

results could be obtained by expressing the covariance as the sum of multiple Kronecker products to account for spatial variation in temporal covariance. It could also be interesting to explore the impact of covariance regularization on transfer learning of the STBF between subjects to alleviate calibration entirely. Finally, it could be insightful to evaluate the proposed algorithms in a real-world on-line BCI setting.

3.A STBF-struct selection accuracies

blocks trials	1	2	3	4	5	6	7	8	9
1	34.70	42.86	47.22	50.04	53.00	54.41	56.44	57.98	58.60
2	48.28	61.51	66.89	71.03	74.87	76.32	77.25	78.62	79.45
3	56.48	71.21	76.81	80.64	83.55	85.85	87.26	87.96	88.71
4	62.83	78.22	82.80	84.74	88.01	89.73	91.18	92.28	92.99
5	67.86	82.10	86.38	88.32	91.27	92.77	93.65	94.66	95.41
6	72.18	85.63	88.62	91.09	93.17	94.22	95.06	95.94	96.52
7	75.62	86.82	91.45	93.03	94.40	95.46	95.94	96.43	96.96
8	76.76	87.39	91.89	93.92	95.06	96.34	96.56	96.91	97.18
9	77.95	89.51	93.08	94.66	95.37	96.65	97.09	97.40	97.66
10	79.81	90.17	93.52	95.11	95.81	96.69	97.00	97.27	97.62
11	80.95	91.45	94.22	95.81	96.47	97.18	97.40	98.02	97.75
12	82.28	92.33	94.49	95.94	96.52	97.27	97.53	97.62	97.75
13	83.82	92.59	95.24	96.12	96.52	97.05	97.40	97.62	97.88
14	84.26	92.99	95.50	96.25	96.83	97.44	97.80	97.97	97.93
15	85.76	93.61	95.24	96.12	96.83	97.71	97.80	98.10	98.06

Table 3.1: Accuracies (%) for STBF-struct.

3.B STBF-shrunk selection accuracies

blocks trials	1	2	3	4	5	6	7	8	9
1	30.56	39.51	45.33	48.41	50.62	54.32	56.44	57.89	58.73
2	44.36	57.45	63.80	68.30	71.83	75.62	76.68	79.54	80.34
3	53.17	68.17	74.56	78.31	81.70	85.23	87.21	88.14	88.93
4	59.83	74.96	80.25	84.44	87.87	89.68	91.27	92.37	93.25
5	64.42	78.92	84.13	87.87	90.74	92.20	93.65	94.58	95.02
6	68.52	82.36	86.60	90.43	92.50	93.78	95.37	96.16	96.47
7	71.47	84.70	88.93	92.06	93.83	95.19	95.99	96.87	97.13
8	73.46	87.21	90.52	93.12	94.80	96.34	97.00	97.27	97.31
9	76.63	87.65	91.45	94.00	95.55	96.65	97.31	97.40	97.44
10	78.09	88.49	92.15	94.14	95.81	96.74	97.31	97.49	97.53
11	78.53	90.78	93.21	95.41	96.25	97.00	97.57	97.62	97.71
12	79.32	91.40	93.25	95.41	96.25	97.40	97.44	97.66	97.62
13	80.47	91.89	93.56	95.37	96.30	97.22	97.57	97.62	97.75
14	81.35	92.72	94.14	95.68	96.16	97.35	97.75	97.75	97.88
15	82.50	92.77	94.22	95.90	96.38	97.53	97.84	98.06	98.24

Table 3.2: Accuracies (%) for STBF-shrunk.

3.C STBF-emp selection accuracies

blocks trials	1	2	3	4	5	6	7	8	9
1	27.78	30.29	25.26	19.97	34.17	40.83	46.03	50.26	52.34
2	39.90	42.15	33.99	26.37	46.52	56.48	63.49	68.30	71.74
3	47.71	51.28	41.18	28.84	56.61	67.11	74.60	79.37	82.19
4	54.76	58.16	45.90	34.44	62.65	74.38	81.00	84.92	87.83
5	59.17	63.10	51.10	36.11	67.50	79.54	85.71	89.02	91.71
6	63.23	67.90	54.63	38.80	71.38	83.20	88.14	91.36	93.47
7	64.81	71.08	58.73	41.49	74.29	86.60	91.49	93.78	94.71
8	67.77	73.46	60.98	42.55	77.69	88.49	92.42	94.71	95.37
9	70.28	75.44	64.02	44.80	78.92	89.15	94.09	95.41	96.30
10	72.49	77.47	65.34	46.87	80.91	91.31	94.53	96.21	96.65
11	73.24	79.28	66.84	49.12	81.83	91.45	94.93	96.43	96.83
12	74.34	80.47	69.49	50.00	83.86	92.11	95.02	96.56	96.91
13	75.04	82.32	70.77	51.28	84.70	93.25	95.55	96.69	97.13
14	76.37	83.16	71.43	51.63	86.99	93.43	95.94	97.05	97.18
15	77.69	83.86	72.57	53.75	87.39	93.87	96.25	97.31	97.62

Table 3.3: Accuracies (%) for STBF-emp.

3.D XDAWN+RG selection accuracies

blocks trials	1	2	3	4	5	6	7	8	9
1	32.05	40.87	44.44	45.37	47.53	48.63	49.82	51.23	53.40
2	45.81	57.19	61.20	63.67	66.71	69.62	71.43	73.06	75.00
3	54.94	66.45	72.09	73.81	76.54	78.62	80.91	82.76	84.04
4	61.07	73.02	77.91	82.14	83.20	85.32	86.77	88.67	90.04
5	65.30	78.26	81.88	83.55	86.64	88.45	89.73	92.28	93.08
6	68.69	80.51	84.08	87.35	88.84	91.49	92.37	93.61	94.49
7	71.78	83.38	87.08	89.15	90.56	92.90	93.69	94.44	95.77
8	74.34	84.79	88.84	90.78	91.89	93.43	94.58	95.50	96.03
9	77.16	86.07	89.81	92.28	93.25	94.49	95.37	95.81	96.65
10	78.44	87.08	90.34	92.77	93.43	94.80	95.68	95.99	96.56
11	79.45	89.02	90.96	93.34	94.58	95.59	96.47	96.56	97.31
12	80.78	89.95	91.40	93.47	94.49	95.90	96.69	96.91	97.09
13	82.23	90.65	91.75	94.00	95.06	95.86	96.56	97.00	97.31
14	83.77	91.05	92.06	93.96	95.28	96.47	96.69	97.66	97.62
15	84.30	91.18	92.72	94.62	95.63	96.03	96.74	97.44	97.44

Table 3.4: Accuracies (%) for XDAWN+RG.

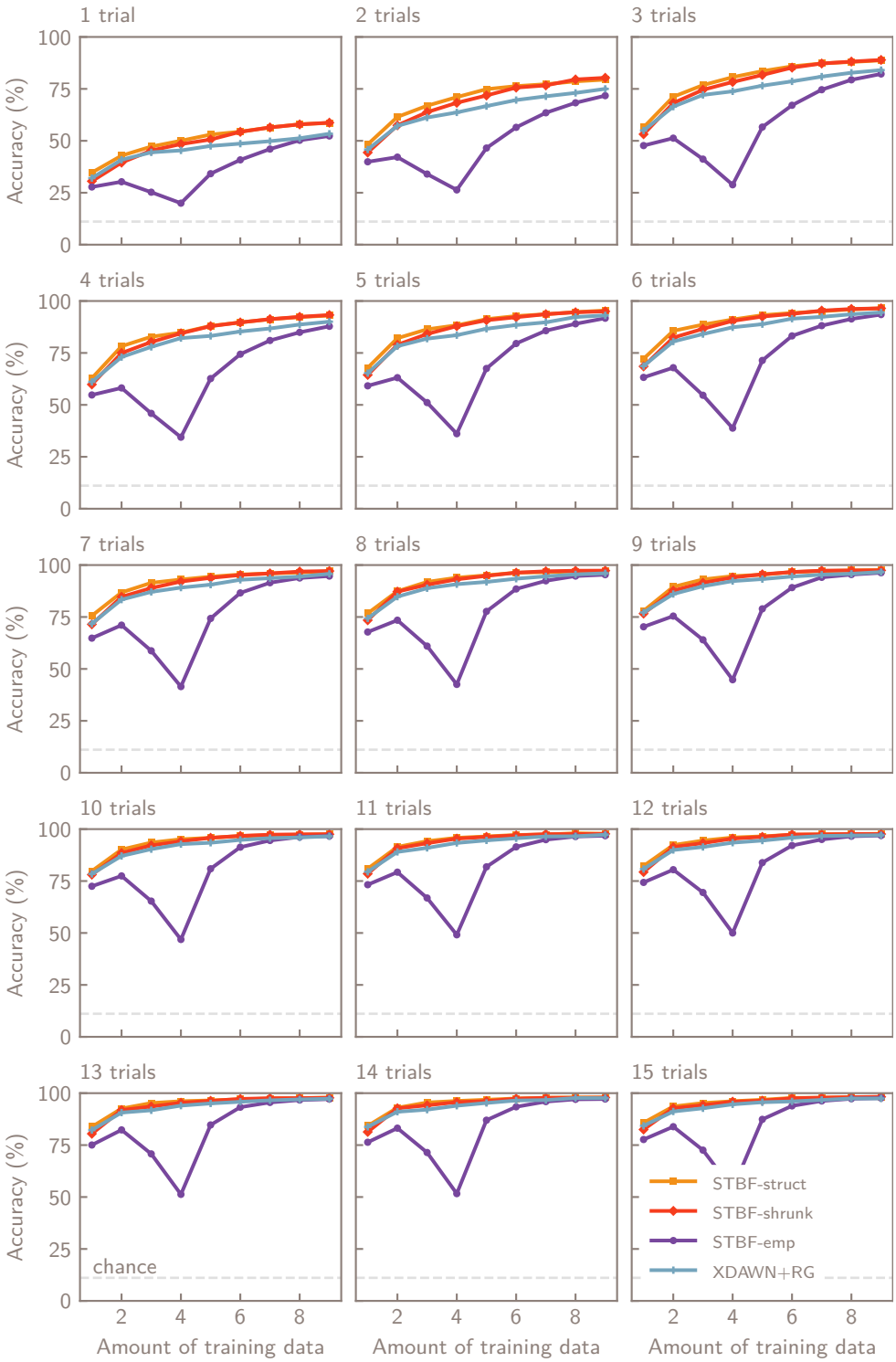


Figure 3.5: Accuracy of the different classifiers for all 21 subjects relative to the number of blocks available for training. One block consists of 135 epochs and corresponds to 27 seconds of stimulation. Accuracies are shown for the averaging over different numbers of trials, ranging from 1 to 15.

3.E Statistical comparisons

# train blocks	1	2	3	4	5	6	7	8	9
STBF-struct > STBF-shrunk	0.005	0.030	0.015	0.543	0.284	0.159	1.000	1.000	0.952
STBF-struct > STBF-emp	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
STBF-struct > XDAWN+RG	0.086	0.002	0.000	0.000	0.000	0.000	0.000	0.000	0.000
STBF-shrunk > STBF-emp	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
STBF-shrunk > XDAWN+RG	1.000	0.499	0.071	0.000	0.000	0.000	0.000	0.001	0.001

(a) Evaluation on one testing trial.

# train blocks	1	2	3	4	5	6	7	8	9
STBF-struct > STBF-shrunk	0.014	0.006	0.040	0.040	0.004	0.846	0.888	1.000	1.000
STBF-struct > STBF-emp	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
STBF-struct > XDAWN+RG	0.103	0.004	0.000	0.000	0.000	0.000	0.000	0.000	0.000
STBF-shrunk > STBF-emp	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
STBF-shrunk > XDAWN+RG	1.000	1.000	0.163	0.001	0.001	0.000	0.000	0.000	0.000

(b) Evaluation averaging over two testing trials.

# train blocks	1	2	3	4	5	6	7	8	9
STBF-struct > STBF-shrunk	0.005	0.030	0.015	0.543	0.284	0.159	1.000	1.000	0.952
STBF-struct > STBF-emp	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
STBF-struct > XDAWN+RG	0.086	0.002	0.000	0.000	0.000	0.004	0.006	0.000	0.000
STBF-shrunk > STBF-emp	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
STBF-shrunk > XDAWN+RG	1.000	0.499	0.000	0.000	0.000	0.000	0.000	0.001	0.001

(c) Evaluation averaging over five testing trials.

Table 3.5: Statistical significance of differences in classifier performance. p-values calculated by one-sided Wilcoxon signed-rank test with Holm correction using one testing trial for different classifiers and levels of data availability. p-values < 0.05 are considered significant.

Chapter 4

Block-term tensor discriminant analysis

This chapter was submitted for publication as Van Den Kerchove et al. [270].

4.1 Introduction

Brain-computer interfaces (BCIs) have the potential to bypass defective neural pathways by providing an alternative communication channel between the brain and an external device. These interfaces find applications in the development of neuroprosthetics and assistive technologies, among other applications [296]. To achieve their functionality, BCIs record and process neural data obtained through a neuroimaging technique, with electroencephalography (EEG) being the most popular.

4.1.1 Tensors & tensor methods

Due to its multichannel time series structure, EEG data, like most neural signal acquisition modalities used for BCIs, naturally exist as multiway data, capturing information in both spatial and temporal domains. Common preprocessing transformations, such as time-frequency transformation, time-binning, or integrating information across multiple subjects or conditions, can further expand the data into additional analytic domains. This can result in high-dimensional datasets which are usually flattened into a set of sample vectors, stripping the original data from its structure. Yet, the intrinsic multiway structure of neural data [67] is well-suited for representation as *tensors*, or multiway arrays, where each domain is represented as a tensor *mode*. Tensors provide a structured data representation for this highly dimensional multiway data. This in turn paves the way to the development of tensor methods which can counteract some of the drawbacks of the dimensionality problem. Tensor methods are machine learning

techniques that consider each tensor mode separately, reducing a given problem into partial, per-mode problems.

These tensor methods decompose a tensor into a lower dimensional structure of a core tensor and factor tensors. The most common approaches adhere to either the Tucker structure or the PARAFAC structure. A Tucker decomposition reduces an input tensor of size $(D_1, D_2, \ldots, D_K)$ to a dense tensor of size $(r_1, r_2, \ldots, r_K)$ with $r_k \leq D_k$ using a set of per-mode factor matrices. Effective unsupervised tensor decomposition and approximation in the Tucker format can be achieved using the Higher Order Singular Value Decomposition (HOSVD) [57, 243]. Alternatively, the PARAFAC structure can be used. Here, the tensor is decomposed into a sum of rank-1 tensors, each the product of a scalar and a vector per mode. This is equivalent to a Tucker structured decomposition with all core elements off the hyperdiagonal set to 0. One way of obtaining an unsupervised PARAFAC decomposition is through the Canonical Polyadic Decomposition [109, 185].

While commonly used, these Tucker or PARAFAC structures might still not be able to efficiently represent relevant neural information in a compressed format. The block-term tensor structure is a generalization of the Tucker and PARAFAC structures. It represents the tensor as a sum of Tucker structured terms. If the number of terms is equal to 1, it is equivalent to the Tucker structure; if the rank of each term is equal to 1, it is equivalent to the PARAFAC structure. The generalized structure can be calculated in an unsupervised way using the Block-Term Decomposition [55, 56, 58, 227]. Performance of methods leveraging either the Tucker and PARAFAC structures are heavily dependent on the prior choice of hyperparameters describing the multilinear rank or the number of rank-1 terms. The block-term structure is more flexible than either the Tucker or the PARAFAC structures, since it is not constrained to problems that cannot be expressed by one of these structures and the chosen hyperparameters. However, this increased flexibility also increases the number of hyperparameters to both the number of terms and the multilinear rank of each term.

4.1.2 Supervised tensor decomposition

If the decompositions are not full rank, the Tucker, PARAFAC and block-term structures are not unique and can be obtained by optimizing different criteria. Given the low signal-to-noise ratio and specific, task-related output expected in a BCI application, supervised feature extraction and machine learning techniques are favored [155] over the unsupervised decomposition methods presented above. A decomposition that is helpful for classification purposes should ideally optimize the discriminability between classes in the resulting core tensors, which can be considered as extracted features. In this philosophy, the Tucker decomposition can also be obtained using Higher Order Discriminant Analysis (HODA) [307, 212, 80], which optimizes class separability in the Fisher sense, analogous to linear discriminant analysis. Extracted features can subsequently be further classified, most commonly using linear discriminant analysis (LDA) or a support vector machine (SVM) to predict class labels.

Variants of HODA have been applied to BCI problems such as event-related potential (ERP) [191, 108] and motor imagery (MI) [152, 40] decoding with positive results [155]. Recent work proposes adaptations such as suited objective functions and regularization [120, 123, 4]. Discriminant tensor features have also been extracted in the PARAFAC structure through manifold optimization [80]. However, it is not immediately obvious if either the Tucker or PARAFAC structure are most suited to represent the neural data of interest for the BCI paradigm and for decoding.

Recent research has shown that supervised decoders adopting a more flexible structure can improve BCI performance. Promising results have been achieved for regression tasks using Higher Order Partial Least Squares [41] and Block-Term Tensor Regression (BTTR) [68, 69]. BTTR has also been adapted into a classification variant [42] but this leaves room for improvement. Instead of optimizing features directly for class separability, a dummy independent variable was regressed towards, and the method cannot be extended to a multi-class setting. Furthermore, structures employed in these regression approaches are still more constrained than what could be achieved with a full block-term tensor structured decomposition optimized for discriminability, since they rely on a low-rank common subspace between the input and classification labels. Huang et al. [112] propose a supervised approach for finding multiple discriminant multilinear spectral filter terms and apply it to motor imagery BCI, but their decomposition is also restricted in flexibility, since the solution can only extract rank $(r_1, r_2, 1)$ with mode 3 corresponding to the frequency domain.

4.1.3 A block-term structured model for classification

A block-term decomposition that is directly optimized for discriminability and with a good choice of ranks and number of terms might better represent the behavior of generators of neural activity through increased flexibility, or might achieve better regularization through increased sparsity. An alternative view on the same approach goes as follows: If HODA with a well-chosen multilinear rank extracts some discriminant features from the input tensor, it is likely that it does not yet retrieve all useful information due to the restriction following from its Tucker structure. Could we therefore not keep using HODA to extract discriminant Tucker structured terms as long as decoding performance increases?

We propose to implement this idea as a new supervised feature extraction tensor method that is a generalization of the aforementioned HODA algorithm and is more suited for the extraction of discriminant features while adhering to a flexible and efficient block-term tensor structure. This work features the following contributions: 1) We develop a forward model for HODA. It can reconstruct a given input tensor from the extracted features. 2) This allows us to introduce a state-of-the-art BCI feature extraction method based on the block-term tensor structure, named Block-Term Tensor Discriminant Analysis (BTTDA). 3) We evaluate a BCI decoder based on BTTDA and its special PARAFAC-structured case on decoding tasks for both ERP and MI paradigm BCI datasets.

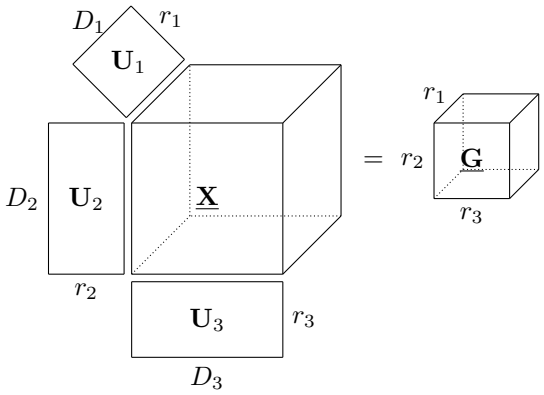


Figure 4.1: A visualization of the multilinear projection obtained by Higher Order Discriminant Analysis (HODA) applied to a third-order tensor sample $\underline{\mathbf{X}}$ with shape (D_1, D_2, D_3). HODA finds projection matrices $\mathbf{U}_k$ such that maximal discriminability between classes is achieved in the projected latent tensors $\underline{\mathbf{G}}$ with reduced dimensionality (r_1, r_2, r_3).

4.2 Methods

4.2.1 Notation

Tensors are indicated by bold underlined letters $\underline{\mathbf{X}}$, matrices by bold letters $\mathbf{U}$, fixed scalars by uppercase letters K, and variable scalars as lowercase letters k. The n^{th} sample of a tensor dataset with N samples is written as $\underline{\mathbf{X}}(n)$. A tensor $\underline{\mathbf{X}} \in \mathbb{R}^{D_1 \times D_2 \times \cdots \times D_K}$ can be unfolded in mode k to a matrix $\mathbf{X}_k \in \mathbb{R}^{(D_k \times \prod_{j \neq k}^{K} D_j)}$, by concatenating all mode $j \neq k$ fibers. The tensor-matrix product of tensor $\underline{\mathbf{X}}$ with matrix $\mathbf{U}$ along a given mode k is written as $\underline{\mathbf{X}} \times_k \mathbf{U}_k$. For ease of notation, let $\underline{\mathbf{X}} \times \{\mathbf{U}\} = \underline{\mathbf{X}} \times_1 \mathbf{U}_1 \times_2 \mathbf{U}_2 \cdots \times_K \mathbf{U}_K$. When skipping one of the modes k, this is written as $\underline{\mathbf{X}} \times_{-k} \{\mathbf{U}\} = \underline{\mathbf{X}} \times_1 \mathbf{U}_1 \times_2 \mathbf{U}_2 \cdots \times_{k-1} \mathbf{U}_{k-1} \times_{k+1} \mathbf{U}_{k+1} \ldots \times_K \mathbf{U}_K$.

4.2.2 Higher Order Discriminant Analysis

Higher Order Discriminant Analysis [212] is a supervised tensor-based feature extraction technique. Let $\left\{\underline{\mathbf{X}}(n) \in \mathbb{R}^{D_1 \times D_2 \times \cdots \times D_K}\right\}_n^N$ be a set of N tensors of order$K$. HODA finds projection matrices $\mathbf{U_k}$ for each mode k that project a given $\underline{\mathbf{X}}$ to a latent tensor $\underline{\mathbf{G}} \in \mathbb{R}^{r_1 \times r_2 \times \cdots \times r_K}$, usually with lower dimensionality $(r_1 \leq D_1, r_2 \leq D_2, \ldots, r_K \leq D_K)$ using tensor-matrix mode products:

$$\underline{\mathbf{G}} = \underline{\mathbf{X}} \times \{\mathbf{U}\} \tag{4.1}$$

as visualized in fig. 4.1.

Analogous to the HOSVD, HODA is a dimensionality reduction decomposition that results in a dense latent tensor $\underline{\mathbf{G}}$, and imposes an orthogonality

constraint on $\mathbf{U}_k$ to ensure uniqueness. However, while for the HOSVD decomposition the projection matrices are chosen to minimize the reconstruction error, the projection matrices $\mathbf{U}_k$ of HODA are optimized for maximal discriminability of tensors $\underline{\mathbf{G}}(n)$ belonging to classes with labels c_n. This is a desirable property in a classification setting where samples are high-dimensional tensors.

HODA optimizes discriminability in the Fisher sense, by optimizing the Fisher ratio ϕ between the latent tensors $\underline{\mathbf{G}}(n)$:

$$\phi = \frac{\sum_c^C N_c \left\| \bar{\underline{\mathbf{G}}}(c) - \bar{\bar{\underline{\mathbf{G}}}} \right\|_F^2}{\sum_n^N \left\| \underline{\mathbf{G}}(n) - \bar{\underline{\mathbf{G}}}(c_n) \right\|_F^2} \tag{4.2}$$

for C classes with each N_c samples. $\bar{\underline{\mathbf{G}}}(c)$ is the mean of latent tensors of class c, and $\bar{\bar{\underline{\mathbf{G}}}}$ the mean of these class mean latent tensors. If the ranks $(r_1, r_2, \ldots, r_k)$ are set a priori, the goal is now to find the optimal projection matrices:

$$\{\mathbf{U}^*\} = \arg\max_{\{\mathbf{U}\}} \phi \tag{4.3}$$

which is solved through the HODA algorithm. To start, $\mathbf{U}_k$ are initialized to orthogonal matrices, e.g. as random orthonormal matrices, by a per-mode Singular Value Decomposition (SVD), or as the partial HOSVD of all stacked tensors in the dataset. At each iteration, the algorithm loops through the modes and fixes all projections but $\mathbf{U}_k$ corresponding to mode k. It then finds a partial latent tensor:

$$\underline{\mathbf{G}}_{-k} = \underline{\mathbf{X}} \times_{-k} \{\mathbf{U}\} \tag{4.4}$$

Subsequently, a new projection matrix $\mathbf{V}_k$ can be found analogous to Linear Discriminant Analysis by constructing the within-class scatter matrix:

$$\mathbf{S}_{-k,\mathrm{w}} = \sum_n^N \tilde{\mathbf{G}}_{-k,k}(n) \cdot \tilde{\mathbf{G}}_{-k,k}^{\intercal}(n) \tag{4.5}$$

with $\tilde{\underline{\mathbf{G}}}_{-k}(n) = \underline{\mathbf{G}}_{-k}(n) - \bar{\underline{\mathbf{G}}}_{-k}(c_n)$, and the between-class scatter matrix:

$$\mathbf{S}_{-k,\mathrm{b}} = \sum_c^C N_c \tilde{\bar{\mathbf{G}}}_{-k,k}(c) \cdot \tilde{\bar{\mathbf{G}}}_{-k,k}^{\intercal}(c) \tag{4.6}$$

with $\tilde{\bar{\underline{\mathbf{G}}}}_{-k}(c) = \bar{\underline{\mathbf{G}}}_{-k}(c) - \bar{\bar{\underline{\mathbf{G}}}}_{-k}$, and solving for the r_k leading eigenvectors in the eigenvalue problem:

$$\mathbf{S}_{-k,\mathrm{b}} - \varphi_k \mathbf{S}_{-k,\mathrm{w}} = \mathbf{V}_k \mathbf{\Lambda} \mathbf{V}_k^{\intercal} \tag{4.7}$$

with $\varphi_k = \mathrm{Tr}\left(\mathbf{U}_k^{\intercal} \mathbf{S}_{-k,\mathrm{b}} \mathbf{U}_k\right) / \mathrm{Tr}\left(\mathbf{U}_k^{\intercal} \mathbf{S}_{-k,\mathrm{w}} \mathbf{U}_k\right)$ using the $\mathbf{U}_k$ obtained in the previous iteration. Finally, the orthogonal transformation invariant projections $\mathbf{U}_k$ are obtained by calculating the per-mode total scatter matrices:

$$\mathbf{S}_{k,\mathrm{t}} = \sum_n^N \mathbf{X}_k(n) \cdot \mathbf{X}_k^{\intercal}(n) \tag{4.8}$$

Algorithm 1 The HODA backward solution.

Require: $\{\underline{\mathbf{X}}(n)\}_n^N$, $\{c_n\}_n^N$, $(r_1, r_2, \ldots, r_K)$, I_{max}, ϵ

$\mathbf{U}_k \leftarrow$ orthonormal matrix $\in \mathbb{R}^{D_k \times r_k} \quad \forall k$

$\mathbf{S}_{k,\mathsf{t}} \leftarrow \sum_n^N \mathbf{X}_k(n) \cdot \mathbf{X}_k^\intercal(n) \quad \forall k$

$i \leftarrow 1$

repeat

  for $k = 1, 2 \ldots, K$ **do**

    $\underline{\mathbf{G}}(n)_{-k} \leftarrow \underline{\mathbf{X}}(n) \times_{-k} \{\mathbf{U}\} \quad \forall n$

    $\mathbf{S}_{-k,\mathsf{w}} \leftarrow \sum_n^N \tilde{\mathbf{G}}_{-k,k}(n) \cdot \tilde{\mathbf{G}}_{-k,k}^\intercal(n)$

    $\mathbf{S}_{-k,\mathsf{b}} \leftarrow \sum_c^C N_c \tilde{\bar{\mathbf{G}}}_{-k,k}(c) \cdot \tilde{\bar{\mathbf{G}}}_{-k,k}^\intercal(c)$

    $\mathbf{V}_k \leftarrow$ largest r_k eigenvectors of $\mathbf{S}_{-k,\mathsf{b}} - \varphi_k \mathbf{S}_{-k,\mathsf{w}}$

    $\mathbf{U}_k \leftarrow$ largest r_k eigenvectors of $\mathbf{V}_k \mathbf{V}_k^\intercal \mathbf{S}_{k,\mathsf{t}} \mathbf{V}_k \mathbf{V}_k^\intercal$

  end for

  $i \leftarrow i + 1$

until $i = I_{\mathsf{max}}$ or $||\mathbf{U}_k^{(i)} - \mathbf{U}_k^{(i-1)}|| < \epsilon \quad \forall k$

and finding the r_k leading eigenvectors of:

$$\mathbf{V}_k \mathbf{V}_k^\intercal \mathbf{S}_{k,\mathrm{t}} \mathbf{V}_k \mathbf{V}_k^\intercal = \mathbf{U}_k \mathbf{\Lambda} \mathbf{U}_k^\intercal \tag{4.9}$$

at each iteration [284]. The iterative process halts after a fixed number of iterations, or when the update of each $\mathbf{U}_k$ is lower than a predetermined threshold. The full HODA procedure is summarized in algorithm 1.

To apply HODA in a classification setting, the projections are first learned on a training dataset with known class labels. Next, these projections are used to extract latent tensors from the tensors in the training dataset. These latent training tensors are then reshaped (*vectorized*) into feature vectors $\mathbf{g} = \mathrm{vec}(\underline{\mathbf{G}})$ and used to train a decision classifier with the corresponding class labels. At the evaluation stage, the projections learned from the training dataset are used to extract latent tensors from an unseen test dataset with unknown class labels, which can also be vectorized and passed on to the trained decision classifier.

To avoid overfitting and improve performance in low sample size settings, the HODA problem can be regularized by shrinking the partial within-class scatter matrices [212] with a shrinkage factor α_k at each step such that the eigenvalue problem becomes:

$$\mathbf{S}_b^{(-k)} - \varphi \left[(1 - \alpha_k) \, \mathbf{S}_{-k,\mathrm{w}} + \alpha_k \mathbf{I} \right] = \mathbf{V}_k \mathbf{\Lambda} \mathbf{V}_k^\intercal \tag{4.10}$$

As in Linear Discriminant Analysis, the shrinkage parameter α_k can also be estimated in a data-driven way in HODA [123], e.g., using the Ledoit-Wolf procedure [139] at every iteration.

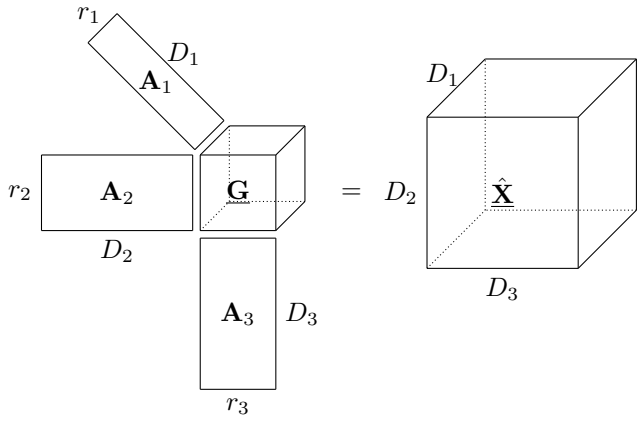


Figure 4.2: The forward projection for HODA. By calculating activation patterns $\mathbf{A}_k$, the original tensor $\underline{\mathbf{X}}$ can approximately be reconstructed as $\hat{\underline{\mathbf{X}}}$ from projected latent tensor $\underline{\mathbf{G}}$. The reconstruction is accurate up to an error term $\underline{\boldsymbol{\mathcal{E}}}$, such that $\underline{\mathbf{X}} = \hat{\underline{\mathbf{X}}} + \underline{\boldsymbol{\mathcal{E}}}$. $\mathbf{A}_k$ are chosen such that the variability captured in the latent tensor is maximally expressed in the reconstructed tensor $\hat{\underline{\mathbf{X}}}$ and not in the error term $\underline{\boldsymbol{\mathcal{E}}}$.

4.2.3 A forward model for HODA

As a prerequisite to the proposed BTTDA model, we must find a way to reconstruct the original data tensor $\underline{\mathbf{X}}$ as accurately as possible from $\underline{\mathbf{G}}$ after dimensionality reduction. This is usually referred to as finding a corresponding *forward model*. While a *backward model* extracts latent sources or properties from the observed data based on some task-related criterion or on prior domain knowledge, a forward model is a generative model that expresses the observed data in terms of these latent properties or sources that are given. Forward models are useful for, among other things, interpretability and data compression.

The HODA projection in eq. (4.1) is an example of a backward model. A straightforward and computationally efficient candidate for a corresponding forward model, visualized in fig. 4.2, is given by:

$$\underline{\mathbf{X}} = \underline{\mathbf{G}} \times \{\mathbf{A}^\intercal\} + \underline{\boldsymbol{\mathcal{E}}} = \hat{\underline{\mathbf{X}}} + \underline{\boldsymbol{\mathcal{E}}} \tag{4.11}$$

with *activation patterns* $\mathbf{A}_k \in \mathbb{R}^{D_k \times r_k}$, reconstructed tensor $\hat{\underline{\mathbf{X}}}$, and error term $\underline{\boldsymbol{\mathcal{E}}}$.

A forward model should ensure that reconstruction error is minimized. In other words, variation captured in the latent tensor should be maximally captured by the reconstruction term $\hat{\underline{\mathbf{X}}} = \underline{\mathbf{G}} \times \{\mathbf{A}^\intercal\}$, and not by the error term $\underline{\boldsymbol{\mathcal{E}}}$ [104]. Hence, we aim to minimize the expected value of the cross-covariance between the noise term and the extracted latent tensors:

$$\{\mathbf{A}^*\} = \underset{\{\mathbf{A}\}}{\arg\min}\, \mathrm{E}\left[\mathrm{vec}\left(\underline{\boldsymbol{\mathcal{E}}}(n)\right) \mathrm{vec}\left(\underline{\mathbf{G}}(n)\right)\right]_n \tag{4.12}$$

Algorithm 2 The HODA forward solution.

Require: $\{\underline{\mathbf{G}}(n)\}_n^N$, $\{\underline{\mathbf{X}}(n)\}_n^N$, $I_{\max}$, ϵ
1: $\mathbf{A}_k \leftarrow \mathbf{U}_k \quad \forall k$
2: $i \leftarrow 1$
3: **repeat**
4: **for** $k = 1, 2 \ldots, K$ **do**
5: $\mathbf{X}_{-k}(n) \leftarrow [\underline{\mathbf{G}}(n) \times_{-k} \{\mathbf{A}\}]_k \quad \forall n$
6: $\mathbf{A}_k \leftarrow \arg\min_{\mathbf{A}_k} \sum_n^N [\mathbf{X}_k(n) - \mathbf{A}_k \mathbf{X}_{-k}(n)]^2$
7: **end for**
8: $i \leftarrow i + 1$
9: **until** $i = I_{\max}$ or $||\mathbf{A}_k^{(i)} - \mathbf{A}_k^{(i-1)}|| < \epsilon \quad \forall k$

or, equivalently [202, 104],

$$\{\mathbf{A}^*\} = \arg\min_{\{\mathbf{A}\}} \sum_n^N \left[\underline{\mathbf{X}}(n) - \hat{\underline{\mathbf{X}}}(n)\right]^2 \tag{4.13}$$

$$= \arg\min_{\{\mathbf{A}\}} \sum_n^N [\underline{\mathbf{X}}(n) - \underline{\mathbf{G}}(n) \times \{\mathbf{A}\}]^2 \tag{4.14}$$

This least squares tensor approximation problem can be solved efficiently using the alternating least squares algorithm [24], iteratively fixing all but one of the activation patterns such that:

$$\mathbf{A}_k = \arg\min_{\mathbf{A}_k} \sum_n^N [\mathbf{X}_k(n) - \mathbf{A}_k \left(\underline{\mathbf{G}}(n) \times_{-k} \{\mathbf{A}\}\right)_k]^2 \tag{4.15}$$

at every iteration, which can be solved directly by ordinary least squares. The activation patterns are initialized to the weights $\{\mathbf{U}\}$ of the backward model. Similar to fitting the backward model, the iterative process for the forward model halts after a fixed number of iterations or when the update of each $\mathbf{A}_k$ is lower than a predetermined threshold. The full algorithm to determine the HODA forward projection is listed in algorithm 2.

The forward model could alternatively be specified in the classical manner by linear regression or with the solution of Haufe et al. [104], but this would require vectorizing the tensor representation. While the reconstruction error might be reduced, we would lose the computational efficiency and regularizing constraints of the tensor form. More efficiently, the forward model could be obtained through other tensor regression methods, like Higher Order Partial Least Squares. This method still requires fitting more parameters than our proposed model, on top of additional hyperparameters defining the dimensionality of the common subspace. An attractive property of our multilinear forward model for HODA is that it estimates exactly as many parameters as the backward model, with the intuition that it should not disproportionally contribute to overfitting.

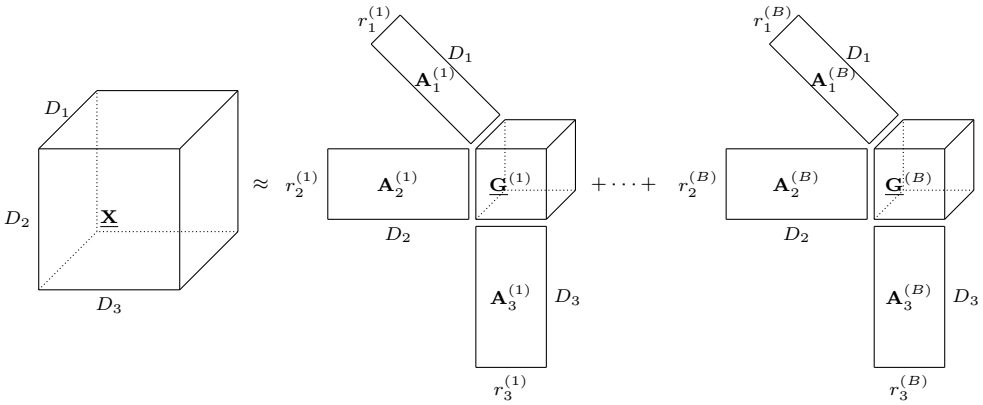


Figure 4.3: A forward model for Block-Term Tensor Discriminant Analysis (BTTDA). BTTDA can extract more features than HODA by iteratively finding a latent tensor $\underline{\mathbf{G}}^{(b)}$ in a deflation scheme. The HODA backward projection is first applied. Next, the input data is reconstructed via the HODA forward model and the difference between the two is found. Finally, this process is repeated with this difference as input data, until a desired number of blocks B has been found.

4.2.4 Block-Term Tensor Discriminant Analysis

After defining the forward model, we can construct our proposed block-term tensor model. Assuming the latent tensors $\underline{\mathbf{G}}$ obtained by the backward projection of HODA do not achieve perfect class separation, the error term $\underline{\boldsymbol{\mathcal{E}}}$ in eq. (4.11) should still contain some discriminative information, which can be exploited to improve classifier performance. Useful features can still be extracted from $\underline{\boldsymbol{\mathcal{E}}} = \underline{\mathbf{X}} - \hat{\underline{\mathbf{X}}}$, so it is further projected onto another core tensor $\underline{\mathbf{G}}^{(2)}$ (assuming $\underline{\mathbf{G}}$ as $\underline{\mathbf{G}}^{(1)}$).

We thus extend the HODA feature extraction scheme to Block-Term Tensor Discriminant Analysis (BTTDA). BTTDA finds multiple discriminative blocks, such that its forward model adheres to the block-term tensor structure:

$$\underline{\mathbf{X}} = \sum_{b}^{B} \underline{\mathbf{G}}^{(b)} \times \left\{ \mathbf{A}^{(b)} \right\} + \underline{\boldsymbol{\mathcal{E}}} \tag{4.16}$$

for B extracted latent tensors $\underline{\mathbf{G}}^{(b)}$ and residual error term $\underline{\boldsymbol{\mathcal{E}}}$. The BTTDA model is further illustrated by fig. 4.3. The block-term structure of this model implies that it is a generalization of both the Tucker-structured HODA and PARAFAC-structured discriminant feature extraction. If B in eq. (4.16) is set to one, BTTDA is equivalent to HODA; if at each term b the rank of the core tensor ($r_1^{(b)} = r_2^{(b)} = \ldots = r_k^{(b)} = 1$), a PARAFAC structure is assumed.

Since BTTDA is specified above as a forward model, a backward procedure is required which finds the latent tensors $\underline{\mathbf{G}}^{(b)}$ given $\underline{\mathbf{X}}$ for BTTDA to be useful as a feature extraction method. The extracted features represented by the

Algorithm 3 BTTDA.

Require: $\{\underline{\mathbf{X}}(n)\}_n^N$, $\{c_n\}_n^N$, $\{(r_1^{(b)}, r_2^{(b)}, \ldots, r_K^{(b)})\}_b^B$
1: $\underline{\boldsymbol{\mathcal{E}}}(n) \leftarrow \underline{\mathbf{X}}(n)\ \forall n$
2: **for** $b = 1, 2, \ldots, B$ **do**
3: $\{\mathbf{U}^{(b)}\} \leftarrow$ HODA on $\{\underline{\boldsymbol{\mathcal{E}}}(n)\}_n^N$ and $\{c_n\}_n^N$ with rank $(r_1^{(b)}, r_2^{(b)}, \ldots, r_K^{(b)})$
4: $\underline{\mathbf{G}}^{(b)}(n) \leftarrow \underline{\boldsymbol{\mathcal{E}}}(n) \times \{\mathbf{U}^{(b)}\}\ \forall n$
5: $\{\mathbf{A}^{(b)}\} \leftarrow$ Forward HODA on $\{\underline{\mathbf{G}}^{(b)}(n)\}_n^N$ and $\underline{\boldsymbol{\mathcal{E}}}$
6: $\hat{\underline{\boldsymbol{\mathcal{E}}}}(n) \leftarrow \underline{\mathbf{G}}^{(b)}(n) \times \{\mathbf{A}^{\intercal(b)}\}\ \forall n$
7: $\underline{\boldsymbol{\mathcal{E}}}(n) \leftarrow \underline{\boldsymbol{\mathcal{E}}}(n) - \hat{\underline{\boldsymbol{\mathcal{E}}}}(n) \forall n$
8: **end for**

latent tensors $\underline{\mathbf{G}}^{(b)}$ can be computed through a deflation scheme summarized in algorithm 3. For each block b, the latent tensor is extracted using the HODA backward projection in eq. (4.1) from the residual error term of the previous block $\underline{\boldsymbol{\mathcal{E}}}^{(b-1)}$:

$$\underline{\mathbf{G}}^{(b)} = \underline{\boldsymbol{\mathcal{E}}}^{(b-1)} \times \left\{\mathbf{U}^{(b)}\right\} \tag{4.17}$$

This residual error term is calculated by finding the difference between the previous error and its reconstruction after backward and forward HODA projection:

$$\underline{\boldsymbol{\mathcal{E}}}^{(b)} = \underline{\boldsymbol{\mathcal{E}}}^{(b-1)} - \hat{\underline{\boldsymbol{\mathcal{E}}}}^{(b-1)} = \underline{\boldsymbol{\mathcal{E}}}^{(b-1)} - \underline{\mathbf{G}}^{(b)} \times \left\{\mathbf{A}^{\intercal(b)}\right\} \tag{4.18}$$

with $\underline{\boldsymbol{\mathcal{E}}}^{(0)} = \underline{\mathbf{X}}$.

The resulting latent tensors can be vectorized and concatenated into one single feature vector per input tensor:

$$\mathbf{g} = \left[\text{vec}\left(\underline{\mathbf{G}}^{(1)}\right)\ \text{vec}\left(\underline{\mathbf{G}}^{(2)}\right)\ \cdots\ \text{vec}\left(\underline{\mathbf{G}}^{(B)}\right)\right] \tag{4.19}$$

so that they can be classified in a similar manner to HODA.

4.2.5 Model and feature selection

Similar to the unsupervised Block-Term Decomposition, the performance of BTTDA is heavily dependent on the rank $(r_1^{(b)}, r_2^{(b)}, \ldots, r_K^{(b)})$ and on the number of blocks B. If these are not known a priori, i.e., if they cannot be set based on insights into the data generation process, a model selection step is necessary in order to determine the optimal values for $r_k^{(b)}$ and B.

Furthermore, HODA, and by extension BTTDA, can extract a substantial amount of redundant features, which can be dropped after projection and before proceeding to the classification step [212]. Specifically, in BTTDA redundant features can accumulate over the number of blocks, hampering performance. Relevant features can be retained by calculating the univariate Fisher score for

Algorithm 4 Greedy model selection.

Require: $\{\underline{\mathbf{X}}(n)\}_n^N$, $\{c_n\}_n^N$,$B_{\max}, \alpha$
for $b = 1, 2, \ldots, B_{\max}$ **do**
  for $r = 1, 2, 4, \ldots, \max_k D_k$ **do**
    $r_k \leftarrow \min(r, D_k) \quad \forall k$
    $\sigma_r^{(b)} \leftarrow$ Cross-validated BTTDA+LDA classification score for $\{\underline{\mathbf{X}}(n)\}_n^N$ and $\{c_n\}_n^N$ with ranks $\{(r_1^{(c)}, r_2^{(c)}, \ldots, r_K^{(c)}) \forall \quad c \leq b\}$ and features where $p(i) < \alpha$
  end for
  $\sigma_*^{(b)} \leftarrow \max_r \sigma_r^{(b)}$
  $r_*^{(b)} \leftarrow \arg\max_r \sigma_r^{(b)}$
  $r_k^{(b)} \leftarrow r_*^{(b)} \quad \forall k$
end for
$B \leftarrow \arg\max_b \sigma_*^{(b)}$

all features. The Fisher score $\phi(i)$ for feature i is obtained as:

$$\phi(i) = \frac{\sum_c^C N_c \left(\bar{g}_i(c) - \bar{\bar{g}}_i\right)^2}{\sum_n^N \left(g_i(n) - \bar{g}_i(c_n)\right)^2} \tag{4.20}$$

The features can then either be sorted by $\phi(i)$, and a given number of features retained for classification, or the decision to retain a feature can be made by a threshold on the statistical p-values corresponding to the $\phi(i)$. We opted for the latter strategy since it does not require redetermining the optimal number of features to retain at each block.

Combined, this results in hyperparameters B for the number of blocks, the threshold value for feature retention, and the ranks of the individual blocks. While we can reasonably set the feature selection p-value threshold to 0.05, B and the block ranks must be tuned to select a performant feature extraction model. While these hyperparameters can be set through cross-validation, this can be computationally expensive. To reduce the computational cost of model selection, algorithm 4 proposes a heuristic model selection algorithm that leverages cross-validation in a greedy way per block, to iteratively find the optimal rank for the next block given the ranks of the previous block. The area under the receiver-operator characteristic curve (ROC-AUC), or the accuracy in the case of a balanced multi-class problem, of classification of extracted feature vectors after feature selection is used as a cross-validation score. Finally, the series of blocks can be truncated to the point with the highest validation score to determine B.

Alternatively, a special case of BTTDA can be constructed using only rank-one blocks such that the resulting forward model adheres to the PARAFAC structure. The optimal number of rank-one blocks can be found by truncating as above. We refer to this strategy as PARAFACDA.

Dataset	Sub.	Chan.	Cls.	Trials	Epoch len. (s)	S. freq. (Hz)	Sess. / Runs
ERP							
BNCI2014-008 [224]	8	8	2	3500/700	1	256	1
BNCI2015-003 [92]	10	8	2	1500/300	0.8	256	1
MI							
BNCI2014-001 [252]	9	22	4	144	4	250	2/6
BNCI2014-004 [140]	9	3	2	180	5	250	5

Table 4.1: MOABB datasets used for evaluation, with the number of subjects (Sub.), the number of EEG channels (Chan.), the number of classes (Cls.), the number of trials or trials per class for ERP datasets (Trials), the epoch length (Epoch len.), the sampling frequency (S. freq.), the number of sessions per subject (Sess.) and the number of runs. Adapted from Aristimunha et al. [14] and Chevallier et al. [51].

4.3 Experiments

4.3.1 Datasets and decoders

We evaluated our proposed model in two offline EEG-based BCI decoding problems: the event-related potential (ERP) and motor imagery (MI) paradigms using a selection of the openly available MOABB benchmarking datasets [14]. Two ERP and two MI datasets were retained to reduce computational demand. Details about these datasets are found in table 4.1. For the ERP datasets, the task is to distinguish target from non-target ERPs, while the MI datasets consist of distinguishing different imagined limb movements. Within-session classification performance was assessed using stratified 5-fold cross-validation to calculate the area under the receiver operating characteristic curve (ROC-AUC).

To use HODA as a decoder, it is paired with LDA to classify the extracted feature (HODA+LDA), with hyperparameters r_k. Similarly, we implemented BTTDA+LDA with the proposed BTTDA feature extraction with hyperparameters $r_k^{(b)}$ for each block b and the number of blocks $1 \leq B \leq 16$. Additionally, we also introduce PARAFACDA+LDA, which is the special case of BTTDA+LDA where each $r_k^{(b)} = 1$, with B and the feature selection threshold as only hyperparameters. Hyperparameters were determined separately for each fold using nested stratified 5-fold cross-validation, and, for BTTDA+LDA, in conjunction with the greedy model selection algorithm in algorithm 4. For HODA+LDA, as well as for the HODA blocks in BTTDA+LDA, we chose $r = r_1 = r_2 = \ldots = r_K$ with possible values $1, 2, 4, 8, \ldots, \max_k D_k$ to reduce computational cost. Other HODA and BTTDA hyperparameters were set to $B_{\max} = 10$, $\varepsilon = 1 \times 10^{-8}$ and $I_{\max} = 128$.

Furthermore, as additional comparison methods, we used the methods evaluated by Chevallier et al. [51]. For the ERP datasets, these were the Riemannian Geometry-based methods using augmented ERP covariance matrices with

and without SVD dimensionality reduction features for a Minimum Riemannian Distance to Mean classifier (ERPCov+MDM and ERPCovSVD+MDM), augmented ERP covariance matrices after XDAWN [225] filtering paired with a Riemannian Minimum Distance to Mean classifier or a projection to tangent space and a support vector machine (XDAWNCov+MDM and XDAWNCov+TS+SVM), and LDA applied after XDAWN filtering (XDAWN+LDA). For the MI datasets, the comparison methods were selected from Riemannian methods. These include projection onto the Riemannian tangent space combined with an ElasticNet classifier (TS+EL), a Fisher Geodesic Minimum Riemannian Distance to Mean classifier and the classification of augmented covariance matrices projected onto the Riemannian tangent space with a support vector machine (ACM+TS+SVM). Additionally, we report the performance of the Common Spatial Patterns filter and a frequency filterbank combined with a support vector machine (CSP+SVM and FilterBank+SVM, respectively). We refer to Chevallier et al. [51] for implementation details of these comparison methods.

4.3.2 Event-Related Potentials

ERPs are spatiotemporal features, with each sample forming a 2^{nd} order tensor with $K = 2$ modes (a matrix), representing EEG channels and time samples per epoch. EEG signals for the evaluated datasets were recorded at the sample rate given by table 4.1 and band-pass filtered between 1 Hz and 24 Hz. The signals were cut into epochs starting from stimulus onset with a dataset-specific length given by table 4.1. For HODA+LDA, PARAFACDA+LDA, and BTTDA+LDA decoders, epochs were downsampled to 48 Hz. For the BNCI2014-008 and BNCI2015-003 datasets, this resulted in matrices of dimensionality $(8, 48)$ and $(8, 38)$, respectively.

Table 4.2 lists the cross-validated ROC-AUC for all evaluated decoders. The highest performance is achieved with the proposed BTTDA+LDA. One-sided Wilcoxon signed-rank tests with a significance level of $\alpha = 0.05$ reveal that both PARAFACDA+LDA and BTTDA+LDA significantly outperform HODA+LDA in both the BNCI2014-008 (PARAFACDA+LDA: $p = 0.0039$, BTTDA+LDA: $p = 0.0039$) and the BNCI2015-003 (PARAFACDA+LDA: $p = 0.0001$, BTTDA+LDA: $p = 0.0049$) datasets. No significant difference was found between BTTDA+LDA and PARAFACDA+LDA.

4.3.3 Motor Imagery

For MI, discriminatory information is represented in the EEG data as Event-Related Synchronizations/Desynchronizations (ERS/Ds). Contrary to the time-domain analyses performed on ERPs, ERS/Ds are usually well discerned in the time-frequency domain. Hence, for the MI task, we transform the EEG signal into the time-frequency domain, forming 3^{rd} order tensors, with $K = 3$ modes representing the channels, frequencies, and time bins.

Pipelines	BNCI2014-008	BNCI2015-003
ERPCov+MDM	74.30±9.77	76.79±10.95
ERPCovSVD+MDM	75.42±9.91	76.93±11.26
XDAWNCov+MDM	77.62±9.81	83.08±7.55
XDAWN+LDA	82.24±5.26	78.62±7.19
XDAWNCov+TS+SVM	85.61±4.43	82.95±8.57
HODA+LDA	83.25±6.25	82.57±7.52
PARAFACDA+LDA	86.19±4.62	84.85±7.93
BTTDA+LDA	**86.43±4.51**	**85.08±7.36**

Table 4.2: Area under the receiver operating characteristic curve for cross-validated within-session evaluation for HODA and our proposed decoders PARAFACDA and BTTDA evaluated on 2 event-related potential datasets. Scores for other decoders were taken from Chevallier et al. [51]. BTTDA reaches the highest performance for both evaluated datasets, closely followed by PARAFACDA.

The MI datasets listed in table 4.1 were first band-pass filtered between 8 and 30 Hz and cut into epochs with time windows as specified by table 4.1. Next, time-frequency transformation was performed using a complex Morlet wavelet convolution, with 16 wavelet frequencies logarithmically spaced between 8 and 32 Hz. The number of wavelet cycles c varied with wavelet frequency f as $c = 0.7 * f$. Features were extracted by taking the log-transformed envelope of the complex wavelet transformation and averaging each epoch along the time axis into time bins of length 1/4s. For the BNCI2014-001 and BNCI2014-004 datasets, this resulted in tensors of dimensionality $(22, 4, 16)$ and $(3, 4, 18)$, respectively.

table 4.3 lists the cross-validated classification scores for the evaluated motor imagery datasets. Note that, in line with Chevallier et al. [51], accuracy is reported for the multi-class classification problem in BNCI2014-001, while ROC-AUC was reported for the binary classification problem in BNCI2014-004. For the BNCI2014-001 dataset, HODA+LDA and our proposed decoders do not reach satisfactory performance compared to the comparison methods, yet both PARAFACDA+LDA and BTTDA+LDA improve upon HODA+LDA, and for BTTDA+LDA this difference is significant ($p = 0.0269$). For the BNCI2014-004 dataset, the performance gap with comparison methods is smaller, but no significant differences were found between HODA+LDA, PARAFACDA+LDA, and BTTDA+LDA.

4.3.4 Block contribution

To analyze the contribution of extra feature blocks extracted by BTTDA over the first one found by HODA, we pick an ERP ($K = 2$) dataset (BNCI2014-008) and an MI ($K = 3$) dataset (BNCI2014-001). We report within-session ROC-AUC scores for training, validation, and test data as a function of the number

Pipelines	BNCI2014-001	BNCI2014-004
CSP+SVM	66.88±15.22	79.27±15.68
FilterBank+SVM	66.53±12.05	80.39±16.05
FgMDM	70.14±15.13	79.28±15.25
TS+EL	72.38±14.85	79.75±15.44
ACM+TS+SVM	**77.82±12.23**	**82.67±15.33**
HODA+LDA	53.51±15.32	80.88±16.78
PARAFACDA+LDA	54.34±15.62	81.18±17.01
BTTDA+LDA	55.13±15.22	80.60±16.71

Table 4.3: Classification score for cross-validated within-session evaluation for HODA+LDA and our proposed decoders PARAFACDA+LDA and BTTDA+LDA evaluated on 2 motor imagery datasets. Scores for other decoders were taken from Chevallier et al. [51]. Accuracy is listed for BNCI2014-001, and ROC-AUC for BNCI2014-004. BTTDA outperforms HODA and PARAFACDA for BNCI2014-001 but does not reach performance comparable to the current state-of-the-art. For BNCI2014-004, PARAFACDA and BTTDA perform approximately on par with HODA and the current state-of-the art.

of blocks, shown in fig. 4.4. Training and validation folds were taken from the model selection procedure. Additionally, the Normalized Mean Squared Error (NMSE) is reported for the reconstructed from the truncated BTTDA decomposition $\hat{\underline{\mathbf{X}}}^{(B)} = \sum_b^B \underline{\mathbf{G}}^{(b)} \times \left\{\mathbf{U}^{(b)}\right\}$. NMSE is calculated as:

$$\text{NMSE}\left(\underline{\mathbf{X}}, \hat{\underline{\mathbf{X}}}^{(B)}\right) = \frac{\sum_n^N \left\|\underline{\mathbf{X}}(n) - \underline{\mathbf{X}}_{\text{rec}}^{(B)}(n)\right\|_F^2}{\sum_n^N \|\underline{\mathbf{X}}(n)\|_F^2} \tag{4.21}$$

4.4 Discussion

4.4.1 Contribution

The results listed in table 4.2 and Table 4.3 show that, for the tested ERP datasets, BTTDA+LDA and PARAFACDA+LDA reach state-of-the-art decoding performance by exceeding all comparison methods for the two evaluated datasets, but for MI datasets, results fall short of those of comparison methods. Performance on the MI classification task could, however, be greatly influenced by the exact tensorization method used, i.e., the time-frequency transform in this case, which is not the main focus of this work.

Nevertheless, when considering the relative improvement over the original HODA model, our results show that BTTDA+LDA performs consistently on par or higher than HODA+LDA, with performances of respectively $86.43 > 83.25$ and $85.08 > 82.57$ for ERP datasets and $55.13 > 53.51$ and $80.60 < 80.88$

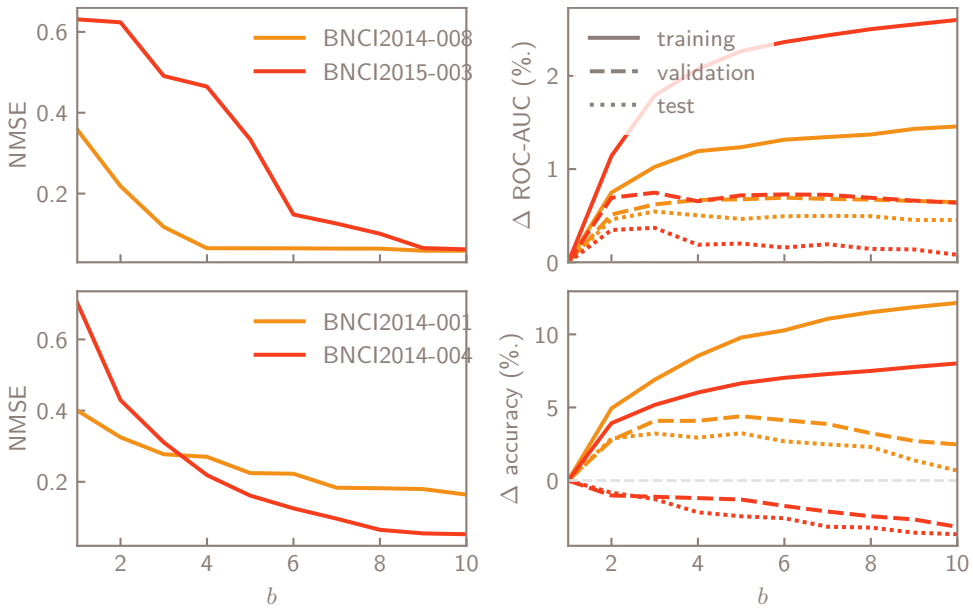


Figure 4.4: Normalized Mean Square Error (NMSE) (left), and difference in classification score for the training and validation for the ERP datasets (top row) and the MI dataset (bottom row) of the greedy model selection procedure (right), as a function of the number of BTTDA blocks b. While NMSE monotonically decreases for the evaluated datasets, better class separation will be achieved, but eventually overfitting occurs and validation and test scores will drop, or plateau due to feature selection.

for MI datasets. The classification score obtained by BTTDA+LDA was also slightly higher than that of the restricted PARAFACDA+LDA for the ERP datasets ($86.43 > 86.19$ and $85.06 > 84.85$ respectively) and one of the MI datasets ($55.13 > 54.34$ and $80.60 < 80.88$ respectively), but these results were not significant, and further studies with more datasets and subjects should show whether this holds. Figure 4.4 shows that there is added value in finding extra blocks over the first HODA block. While no proof is given, we notice that NMSE monotonically decreases, indicating that since the training score keeps increasing, this suggests that eventually all the variation in the signal will be explained by the model while still extracting features that are maximally discriminant. Eventually, the number of blocks will reach a point of diminishing validation score returns, when adding extra features to the decision classifier increases its risk of overfitting instead of adding extra useful discriminatory information. Together with the improved classification scores presented, these results point to the potential of our more flexible model block-term or its special PARAFAC-structured case over a Tucker-structured model given proper block and rank selection. Since the optimal ranks for HODA+LDA were also

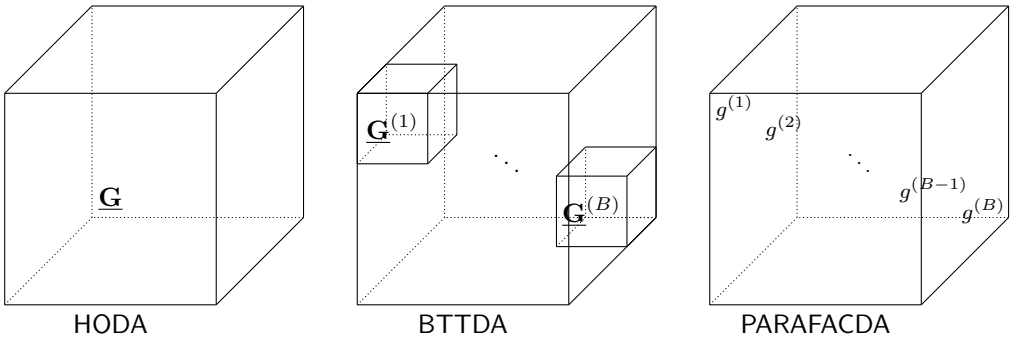


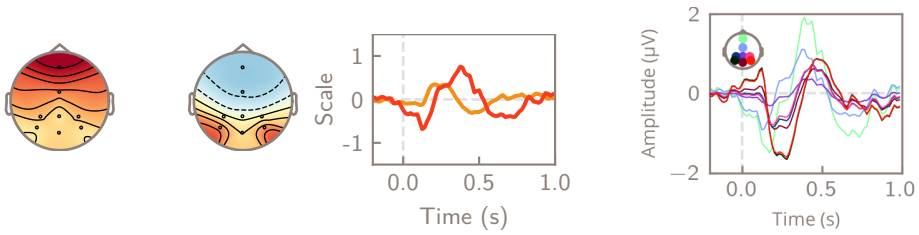
Figure 4.5: BTTDA and PARAFACDA can find a sparser expression of the discriminant information captured by the dense, Tucker-structured HODA algorithm.

determined through cross-validation, BTTDA+LDA can improve over the first HODA block. If validation shows that this is not possible, such as is the case for BNCI2014-004, the BTTDA model is truncated to the first HODA block and little performance is lost.

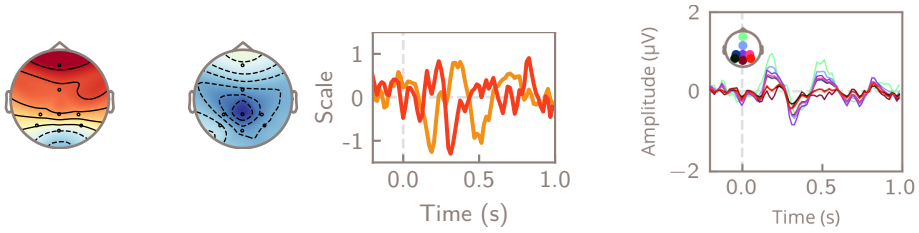
We assume the main benefit of BTTDA is that it can more easily discover relevant features while being more parsimonious due to its block-term structure compared to HODA's full Tucker structure, as illustrated by fig. 4.5. The same discriminative information captured by a relatively large Tucker-structured core tensor could be expressed more sparsely with a small number of block-terms, while avoiding redundant features. The PARAFAC structure employed in PARAFACDA is even more sparse, which could be a benefit or a drawback depending on the amount of regularization required, or on the true underlying structure of the data.

Alternatively, the enhanced performance could also stem from the modeled data covariance. Since HODA estimates one within-class scatter matrix $\mathbf{S}_{-k,\mathrm{w}} \in \mathbb{R}^{D_k \times D_k}$ per mode, its overall model of the data scatter is determined by these per-mode scatter matrices as a Kronecker product $\mathbf{S}_{-1,\mathrm{w}} \otimes \mathbf{S}_{-2,\mathrm{w}} \otimes \cdots \otimes \mathbf{S}_{-K,\mathrm{w}}$, which corresponds to the assumption that the EEG data is drawn from a multilinear normal distribution [189]. However, it is known that the EEG covariance cannot fully be expressed as a single Kronecker product, but rather is more accurately modeled by a sum of multiple Kronecker products [27, 244]. Since BTTDA iteratively fits HODA models to the residual error, it will be able to express the full covariance structure given sufficient blocks.

Additionally, the forward modeling step inherent to BTTDA results in an interpretable model since we can use the activation patterns or the forward projection to inspect the neural patterns corresponding to the relevant discriminatory information at each block [104]. Figure 4.6 shows the activation patterns of two blocks obtained from the BNCI2014-008 dataset as well as the forward projection of the difference between the averages of the mean latent tensor per



(a) The first block.



(b) The second block.

Figure 4.6: Spatial (left two columns) and temporal (middle column) activation patterns and condition contrasts (right column) obtained after forward projection of the latent features for 2 blocks of rank $(2, 2)$ of BTTDA fit on the full dataset BNCI2014-008. The separate blocks approximately model different ERP components.

class (*contrasts*) after forward projection. While the weights of the backward projection are uninterpretable [104], the activation patterns and contrasts after forward projection clearly show that ERP components can be recognized and separated into different BTTDA blocks. Given informed or correctly tuned hyperparameters, this method could be used to e.g., separate ERP components or neural processes based on the task-related information in the class labels.

4.4.2 Model selection and dimensionality

BTTDA trades in the rigid HODA model for increased model complexity with more hyperparameters to tune, which could open up a setting where performance can be improved. This turns tensor discriminant analysis into a model selection procedure instead of a projection optimization procedure. Despite favorable results in BCI decoding, the applications of the proposed BTTDA model are limited mainly by this model selection approach used to determine the individual block ranks. While our proposed greedy model selection algorithm is a step in the right direction, the high computational cost of setting hyperparameters through cross validation can still hinder the portability of decoders relying on BTTDA.

Due to its heuristic nature, the greedy algorithm does not always result in the set of ranks with the highest achievable performance. In combination with the

fact that the feature selection cutoff parameter was fixed somewhat arbitrarily, it is clear that thorough hyperparameter optimization could improve performance. Furthermore, it is clear that our proposed model selection procedure does not necessarily result in an optimal set of blocks that group coherent projections within the same block, according to some desirable metric. Examples of this are sparsity, pattern interpretability, minimal or maximal between- or within-block feature correlation, decreasing discriminability, etc. Finally, the completeness of the presented results is limited by the artificial restriction in hyperparameter choice. We imposed ($r_1 = r_2 = \ldots = r_k$) to reduce the computational demand of the performed experiments. In a sense, this goes against the proposition of increased model flexibility. It might well be that BTTDA in some cases offers little added value over the Tucker-structured HODA, when both are given free choice of rank. Future efforts should focus on automatic parameter setting, e.g., using sparsity or information criteria such as the ones used in Block-Term Tensor Regression [68] or other statistical measures based on the model's application.

Another limitation is that BTTDA might yield a disproportionate improvement for datasets with a low number of features relative to sample size, while being less effective for datasets with more features. This is reflected in our ERP results (low dimensionality vs. high number of trials) compared to the MI results (higher dimensionality due to third-order tensorization vs. lower number of trials). We expect a dimensionality limit beyond which the forward modeling step cannot accurately regress from the low-dimensional latent tensors to the high- dimensional original tensors, introducing error in the input data for the next block, which can stack up over blocks. Since the forward multilinear least squares problem is underdetermined, it is prone to numerical instability, which calls for regularization of the forward modeling procedure, but this would introduce another hyperparameter. It should also be thoroughly investigated what the impact is of going beyond second- and third-order cases to higher-order tensors, since this could have a large impact on the model. Other tensorization methods of the EEG data, like time-lagged Hankel tensors [200], or tensors across subjects or sliding windows, etc., could also be of interest if they are appropriately chosen based on prior knowledge of the dataset.

4.5 Conclusion

We have introduced Block-Term Tensor Discriminant Analysis (BTTDA), a novel, tensor-based, supervised dimensionality reduction technique optimized for class discriminability, which adheres to the block-term tensor structure. BTTDA is a generalization of Higher Order Discriminant Analysis (HODA) and can also be applied as a special sum-of-rank-one tensors PARAFACDA model. The model is obtained by iteratively fitting HODA in a deflation scheme, leveraging a novel forward modeling step.

Via an accompanying heuristic model selection procedure, BCI decoders using BTTDA feature extraction can significantly outperform decoders based on HODA and reach state-of-the-art decoding performance on event-related

potential problems (second-order tensors) and scores on par with or higher than HODA in motor imagery problems (third-order tensors). Moving from the rigid Tucker tensor structure of HODA to the more flexible block-term structure shifts the problem from finding optimally constrained multilinear projections to model and feature selection.

Introducing a flexible block-term tensor model as the underlying structure reformulates tensor discriminant analysis as a model selection problem. This allows performance to be traded off for model complexity and the number of features, to find a setting that is more effective for decoding. Because of its general implementation and minimal assumptions on data structure, BTTDA can equally be applied to classification for other neuroimaging modalities (MEG, ECoG, fNIRS, fMRI, EMG, ...) or to tensor classification problems in other fields.

Chapter 5

ERP latency estimation and alignment

Section 5.2.3, sections 5.3.1 to 5.3.3 (including figs. 5.2 and 5.3), and the third-to-last paragraph of section 5.5 were published as part of Van Den Kerchove et al. [272].

5.1 Introduction

Electroencephalography (EEG)-based visual oddball brain-computer interfaces (BCIs) establish communication with paralyzed individuals by decoding ERPs obtained from time-modulated flashing stimuli (see fig. 1.1). These ERPs consist of multiple ERP-components, peaks and troughs in ERP waveform (see fig. 1.4d). These components each have their own distinct amplitude and latency. ERP component latency can reflect a wide range of cognitive, neurophysiological, and clinical properties, that can be confounding factors in the development of BCIs [157]. Latency is often assessed as the timing of an ERP component peak, obtained after averaging over trials. Latencies of non-averaged, single-trial ERPs should capture even more relevant information, yet they are very hard to measure due to the low signal-to-noise ratio (SNR) of single-trial ERPs. Within-session variability of ERP latency, called ERP latency *jitter*, has been shown to influence BCI accuracy [256]. This effect is especially prominent when bringing BCI development from the lab to the clinical setting. For instance, Zisk et al. [318] showed that ERP jitter was significantly higher in 6 Amyotrophic Lateral Sclerosis (ALS) patients (1350 ms^2) compared to neurotypical controls (553 ms^2). Aricò et al. [12] showed that P3 latency jitter negatively correlated ($\rho = 0.17$) with BCI accuracy in a covert GeoSpell interface. Here the user does not gaze directly at the targets, such as is the case for individuals with severe speech, physical and gaze impairment (SSPGI). Therefore, properly accounting for ERP latency jitter can contribute to the development of gaze-independent visual BCIs.

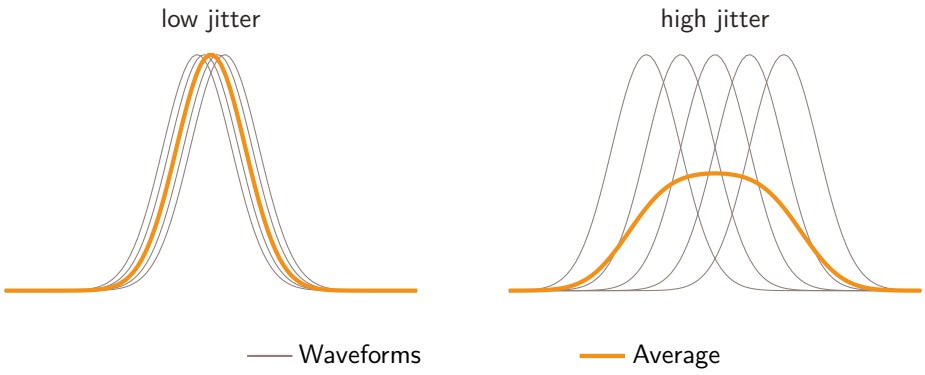


Figure 5.1: The smearing effect. When obtaining an ERP template through averaging, high jitter causes a decrease in peak amplitude and a deformation of the shape of the waveform.

5.1.1 Neural origins of ERP latency jitter

While the classic ERP model assumes a constant ERP response over trials, research has shown that cortical evoked potentials show significant trial-to-trial jitter [267]. Experiments show that latency jitter in ERPs is the result of the interaction of the evoked activity with ongoing dynamics in the brain [102, 131, 53, 13]. According to the Firefly model of event-related activity [39], the latency of an evoked potential results from a combination of the modulation of oscillating activity that is time-locked, but not necessarily phase-locked to stimulus onset, and a delay caused by ongoing neurophysiological processes that are not necessarily related to the task [249, 176]. If these processes are related to the BCI task at hand, such as attention and visual processing, jitter could have a significant impact on ERP decoding.

5.1.2 Implications for ERP analysis

The presence of latency jitter introduces multiple issues to ERP analysis, the most prominent of which is the smearing effect, as illustrated in fig. 5.1. In ERP analysis, single-trial ERPs are usually averaged per condition to cope with the unfavorable SNR. When averaging over multiple identical responses with jittered latencies, the shape and amplitude of the average do not reflect the properties of the original signal. Because the amplitude of the average is lower, the SNR will also be negatively impacted. More generally, the smearing effect leads to difficulties interpreting ERP results and making inferences about amplitude or latency effects because these are entangled when using averaging [157].

The smearing effect equally impacts the SNR of information captured by a classifier's parameters when training with a procedure that does not account for this jitter, and thus also affects its performance [256]. As an illustrative

example, think of an linear discriminant analysis (LDA) ERP classifier that uses class averages over ERP epochs to construct its between-class scatter matrix. These per-class averages will be affected by the smearing effect and will be less effective templates for ERP retrieval. Similarly, the noise model in LDA assumes stationarity and subtracts the class-wise averages from the signal to calculate the within-class scatter matrix, which will then also be affected. Hence, effective latency estimation and jitter compensation can contribute to higher decoding performance.

5.1.3 Benefits of latency estimation

As we will see later, a common approach to counteract the smearing effect is to correct single-trial ERPs for jitter before averaging. This can be done by estimating the latency of each single-trial ERP and aligning them before averaging. Latency estimation has multiple advantages: (1) Latency-based ERP alignment can help reveal the true shape and properties like peak amplitude and latency of ERPs across experimental conditions. These are variables of interest for testing electrophysiological and behavioral hypotheses. Additionally, latency estimation allows for the extraction of other descriptors from a set of single-trial ERPs, like amplitude or latency variance [116], which can also be of interest as studied variables. For instance, in the work of Saville et al. [231], the inter-trial latency variability itself is determined as a correlate of working memory performance. (2) Counteracting the smearing effect will boost the signal-to-noise ratio of averaged ERPs by increasing the amplitude of the ERP signal. This in turn can reveal smaller or more jittered ERP components which would otherwise be obscured or blurred due to desynchronization. These *deblurred* templates can then be used for analysis of the ERP waveform, or as templates of the signal of interest in BCIs [12]. (3) Finally, the introduction of latency estimation and alignment generally improves BCI classification performance. Single-trial latencies can be used as features to improve classification [100]. Furthermore, compensating for inter-trial differences, especially when extended to inter-session or inter-subject differences, can aid in transfer learning or generalization across subjects and protocols [119].

5.2 Literature

5.2.1 Single-component approaches

The simplest method to estimate single-trial latency is *peak picking*. Here, the latency of an ERP is determined as the time-point of its maximum (or minimum) amplitude relative to stimulus onset. While straightforward, this method does not perform well in low SNR conditions, unless combined with filtering. Filtering can suppress the noise contaminating EEG trials, lowering the risk of picking a noise peak instead of a true peak of the ERP. Several filtering approaches have been developed to be used in conjunction with peak-picking, in

the time domain [43, 163, 128, 188, 246, 83, 54] and in the time-frequency domain [219, 284, 111]. The aforementioned analysis and performance prediction method proposed by Aricò et al. [12] estimates the single-trial P3 latencies of attended epochs by P3 peak-picking after filtering in the time-frequency domain and corrects for jitter by temporally aligning all target epochs to the average latency, i.e., by shifting the epochs in time such that the P3 peaks all fall at the same time instance. Filtering can also be done in the spatial domain. The spatial filter can be constructed using principal component analysis (PCA) [196], independent component analysis (ICA) [262, 124, 168], or as a spatial LCMV beamformer [265].

Template matching is generally preferred to peak picking. The algorithm proposed by Woody [301] in 1976 is a simple and elegant latency estimation method that is still being used and is even considered among the more performant techniques for ERPs [258, 196]. It starts with an average ERP as a rough estimation of the aligned template ERP, and iteratively refines this template by determining at each step the time point of maximum cross-correlation of this template with all single-trial ERPs and aligning them on these time points to form the next template. Similarly, cross-correlation template matching can be combined with filtering. Souloumiac and Rivet [245] use the XDAWN technique to construct a spatial filter before template matching the spatially filtered signal. Iturrate et al. [119] uses a spatiotemporal filter. Instead of aligning the latencies to obtain the template for the next iteration, weighted averaging can also be used [82].

Despite good performance, the iterative scheme of Woody's algorithm has a risk of converging to a local optimum. To avoid this, a fitness function can be defined to jointly optimize the set of single-trial ERP latencies, which can be optimized with a genetic algorithm [207]. Alternative methods are based on graph optimization [62], hidden process models [130], or by fitting an ERP model with maximum likelihood estimation [90, 268, 170, 218].

5.2.2 Multi-component approaches

In the previous section, we have abstracted away the fact that ERPs are multi-component signals, which is not taken into account by the aforementioned methods. Multiple components can be time-locked to different events or neurophysiological processes and therefore are not necessarily time-locked to each other. They each can have different latency variabilities. The presence of more than one component can hamper the performance of some latency estimation algorithms, and it is often only possible to extract the latencies of the largest component present [196]. When aligning trials to one component and averaging, this will introduce blurring in the other components [193].

Most of the algorithms above can be adapted to work on multiple ERP components by carefully selecting peaks or pre-determined regions of interest. Hardiansyah, Pergher, and Van Hulle [100] use an SNR boosting method, applied to a specific time-window for each ERP component to extract a latency per component. Methods based on spatial decomposition filters like PCA, ICA,

or XDAWN can be leveraged to utilize multiple of their components as filters, resulting in multiple time series where peak picking or Woody's algorithm can be applied. In general, however, these algorithms lack an integrated approach to deal with multiple ERP components. A clear evaluation of the performance of these methods in multi-component settings is lacking.

Some algorithms separate components into (sets of) stimulus- and response-locked component clusters based on their latency distributions [124, 251, 315, 309]. These algorithms do, however, require a response button trigger signal as input to determine the response time, which is not always present or applicable in an ERP experiment, especially in the case of BCIs. The RIDE algorithm [197, 195, 281, 194, 193] is able to do this without entirely relying on the response time and can separate stimulus-locked, response-locked, and multiple central component clusters. Others, like spatial filtering methods that use decomposition methods to determine the spatial filter, can be modified to separate components [196].

Finally, the smearing problem can also be overcome by algorithms that are not based on ERP component latency estimation but use time-warping-based methods to enhance the shape of the averaged template ERP, like Dynamic Time Warping [94, 282, 319], Correlation-Optimized Warping [241], or Fast Variational Alignment [76]. These methods work fundamentally differently from the latency estimation methods, and cannot directly be used to extract latency features.

5.2.3 Contribution

Most of these methods suffer from a common drawback: they cannot be used in a decoding scheme to improve performance for incoming epochs with unknown class labels that contain either a target or non-target ERP response. They can be used offline on a set of labeled epochs for testing hypotheses concerning latency and jitter, for aligning templates, or for BCI performance prediction, but not for ERP classification. While some of the aforementioned methods could be adapted to perform classification tasks, few studies investigate how to exploit this latency estimation for jitter-resistant decoding. Hardiansyah, Pergher, and Van Hulle [100] incorporated single-trial latencies in classification by peak-picking within a given ERP time window, unaware of the class of the epoch under investigation. The Classifier-based Latency Estimation (CBLE) algorithm by Thompson, Warschausky, and Huggins [256] also explicitly applies latency estimation in a decoding setting. Thompson, Warschausky, and Huggins [256] initially formulated CBLE as an offline performance prediction method. Later, its output was successfully adapted to compensate for jitter to improve decoder performance [177, 317].

Time-series classification algorithms [1] that are robust to jitter can be used in a decoding setting, but, in general, have scarcely been applied to ERP decoding. Data augmentation involving jittering the training data [133, 317] and Riemannian Geometry methods using spatial covariances as features [16] have both been shown to perform well in the presence of ERP jitter. In this work, we

opted to apply CBLE because it has successfully been applied to classify jittered ERP. We adapt the CBLE to an iterative method akin to Woody's scheme that can better compensate for jitter, to improve covert VSA decoding performance.

5.3 Materials & methods

5.3.1 Classifier-based Latency Estimation

Consider a training set of N EEG epochs with C channels of S samples $\{\mathbf{X}_n^{\text{train}} \in \mathbb{R}^{C\times S}\}_{n=1}^{N}$ This dataset has a corresponding set of training labels $\mathbf{l}^{\text{train}} \in \{\text{target}, \text{non-target}\}_{n=1}^{N}$, and a similar testing set of M epochs $\{\mathbf{X}_m^{\text{test}} \in \mathbb{R}^{C\times S}\}_{m=1}^{M}$. We assume that the sampling period is T, i.e., that the sample with index s is sampled at time sT. In the following, we use the matrix slicing notation to denote row or column intervals extracted from a matrix. For instance, $\mathbf{X}[:, s_1 : s_2]$ denotes all columns of $\mathbf{X}$ with indices between s_1 (included) and s_2 (excluded).

CBLE, summarized in section 5.A, works by training a *first-stage* classifier $\mathcal{C}(\theta, f)$ defined within a time period $[s_1 : s_2]$ with a set of parameters θ and a decision function $f(\mathbf{X}[:, s_1 : s_2], \theta) \to y$ outputting a classification score $y \in \mathbb{R}$ for a given epoch $\mathbf{X}$, such that

$$\theta = \text{train}_{\mathcal{C}}(\{\mathbf{X}_n^{\text{train}}[:, s_1 : s_2]\}_{n=1}^{N}, \mathbf{l}^{\text{train}}) \tag{5.1}$$

Then, f can be applied to all (possibly overlapping) slices of length $s_2 - s_1$ of an epoch $\mathbf{X}$, resulting in a vector of score values $\mathbf{y} = [y_1 \ \dots \ y_R]^T \in \mathbb{R}^R$ such that

$$y_s = f(\mathbf{X}[:, s : s + (s_2 - s_1)], \theta) \quad \forall s \in 1, \dots, R \tag{5.2}$$

with $R = S - (s_2 - s_1)$. To leverage CBLE for ERP classification, the score vectors $\mathbf{y}$ can be arranged in matrices $\mathbf{Y}^{\text{train}} \in \mathbb{R}^{N\times R}$ and $\mathbf{Y}^{\text{test}} \in \mathbb{R}^{M\times R}$. These can be further classified by training a *second-stage* classifier on $\mathbf{Y}^{\text{train}}$ and class labels $\mathbf{l}^{\text{train}}$. However, the resulting score-over-time vectors per epoch still suffer from jitter. For classification, we follow the approach of Mowla, Huggins, and Thompson [177], using a maximum-level hierarchical Daubechies-4 wavelet transform to reduce dimensionality before classification with the second-stage classifier. In the CBLE-decoder, it is this wavelet transform that decreases the sensitivity to latency differences, actively compensating for ERP latency jitter.

When using a simple spatiotemporal linear classifier as first-stage classifier, CBLE is equivalent to the first iteration of Woody's algorithm with the spatiotemporal classifier weights as template. Thompson, Warschausky, and Huggins [256], Mowla, Huggins, and Thompson [177], and Mowla et al. [178] show that CBLE is relatively independent of the first-stage classifier for BCI accuracy prediction and for ERP classification. Therefore, we opt to use the variant of Linear Discriminant Analysis with block-Toeplitz regularized covariance matrix (tLDA) proposed by Sosulski and Tangermann [244], as the first-stage classifier and logistic regression as second stage.

5.3.2 Robust latency features

The ERP decoding method based on CBLE introduced by Mowla, Huggins, and Thompson [177] does not make use of the extracted latencies, only passing score matrix $\mathbf{Y}$ on to the second-stage classifier. Furthermore, previous CBLE works [256, 178] only used these latencies to correlate them with neurophysiological processes or to predict decoder performance. We noticed that, while CBLE performance was unaffected, the classification performance of our proposed method can be improved if the estimated latencies are also made available as features to the second-stage classifier, after a square transform for linear separability [256].

However, including these latency features gives rise to the following issue when classifying unseen data. The CBLE latency estimate is defined only for target epochs as $s_{\text{target}} = \arg\max_s y_s$. This is the point in time where the target class reaches the largest separation from the background noise and non-target class, indicating the target ERP is most likely to occur here. However, in a classifier test phase, it is not known a priori whether an unseen epoch is a target or a non-target epoch. This problem is solved by defining an estimated latency per class s_{target} and $s_{\text{non-target}}$ for every epoch, regardless of its actual class. The estimated class latencies can then be used as features for training and testing the second-stage classifier. This way, the latencies of the testing data can be presented to the second-stage classifier without knowledge of the testing data class labels, making them useful in decoding.

In a similar manner to the target latency, the non-target latency could be defined as $s_{\text{non-target}} = \arg\min_s y_s$. However, this is problematic since it is not evident how to estimate the latency of, e.g., a P3 ERP component for a non-target epoch, since the non-target class is characterized by the absence of this component. [1]. In fact, y can have multiple local minima or entirely lack distinct peaks for non-targets, rendering the minimum estimate meaningless.

Instead, we opt for a more robust, probabilistic definition of class latencies. This robust estimation method yields latencies that (1) are more meaningful as input for the second-stage classifier, and (2) lead to smoother convergence in our proposed iterative alignment scheme for Classifier-based Latency Estimation with Woody iterations (WCBLE), which heavily relies on exact latency estimation. Assume classifier $\mathcal{C}(\theta, f, \Pr)$ now can also output a probability per class $\Pr(l|\mathbf{X}[:, s_1 : s_2)], \theta)$ for a given epoch $\mathbf{X}$, a feature of many common classifiers. Analogous to equation 5.2, we can now write

$$\Pr(\mathbf{X}, \theta, l, s) = \frac{1}{R} \Pr(\mathbf{X}[:, s : s + (s_2 - s_1)], \theta, l) \quad \forall s \in 1, \ldots, R \tag{5.3}$$

The latency features assuming the epoch belongs to a given class given by $l \in \{\text{target}, \text{non-target}\}$ are then defined as the median of the corresponding

[1] Depending on the stimulation and experimental design, ERP components can be present in one experimental condition or class and missing in another, or appear in multiple classes at different amplitudes. In the latter case, it could be possible to estimate the latency of an ERP in the conditions where it appears, but if its amplitude is lower or negligible in some conditions, latency estimates will be less accurate, hence this case is still problematic.

distributions

$$s_l = \text{median}\left[\Pr(s|\mathbf{X}, \theta, l)\right] \tag{5.4}$$

Note that $\Pr(s|\mathbf{X}, \theta, \text{non-target}) = 1 - \Pr(s|\mathbf{X}, \theta, \text{target})$. The median of the probability distribution over time is more robust to outliers and noise than the maximum or minimum score. For the non-target case, the median approach tends towards the center of a near-uniform distribution, resulting in a more consistent latency estimate over trials as compared to the minimum approach.

5.3.3 Classifier-based Latency Estimation with Woody iterations

To improve performance over CBLE, we propose WCBLE a new algorithm inspired both by CBLE and the aforementioned Woody iteration scheme. Instead of using CBLE to estimate the features of a second-stage classifier directly, CBLE latency estimation is used as a step in a Woody iteration scheme. While the Woody algorithm iteratively enhances the SNR of an ERP template to cross-correlate with the data, WCBLE iteratively re-estimates the parameters of the first-stage classifier. To improve convergence and perform well in a classification setting, WCBLE aligns both targets and non-targets to their corresponding estimated latencies.

The WCBLE algorithm is presented in section 5.B. Its training phase is visualized in fig. 5.2. The initial training epochs $\{\mathbf{X}_n^{(1)}\}_{n=1}^N$ are set to $\{\mathbf{X}_n^{\text{train}}\}_{n=1}^N$. At every iteration, classifier $\mathcal{C}$ is trained like in CBLE:

$$\theta^{(i)} = \text{train}_{\mathcal{C}}(\{\mathbf{X}_n^{(i-1)}[:, s_1 : s_2]\}_{n=1}^N, \mathbf{l}^{\text{train}}) \tag{5.5}$$

Next, latency $s_{l_n}^{(i)}$ is determined for every epoch $\mathbf{X}^{(i)}$ corresponding to its class label l_n using eq. (5.4). Finally, the training epochs $\mathbf{X}^{(i+1)}$ for the next iteration are determined by aligning each original training epoch to the latency $s_{l_n}^{(i)}$ corresponding to its respective class label.

$$\mathbf{X}_n^{(i+1)} = \text{align}(\mathbf{X}_n^{\text{train}}, s_{l_n}^{(i)}) \quad \forall n = 1, \ldots, N \tag{5.6}$$

Aligning is performed by shifting and zero-padding the signal to the right if the latency is negative relative to the time window onset, and to the left if positive, by the difference between the latency and the window onset. The process halts after a fixed number of iterations or when the estimated set of latencies has been encountered before, indicating it ended up in a loop. In the end, the procedure should result in enhanced classifier parameters θ^*, closer to those when there would be no jitter between epochs. Note that using the median approach for robust latency estimation results in a smoother yet longer convergence process compared to the maximum/minimum approach. We can then apply the classifier with enhanced parameters θ^* in a CBLE manner to unseen epochs as illustrated in fig. 5.3 to obtain a vector of scores over time as in section 5.3.1 and the estimated latencies as in section 5.3.2.

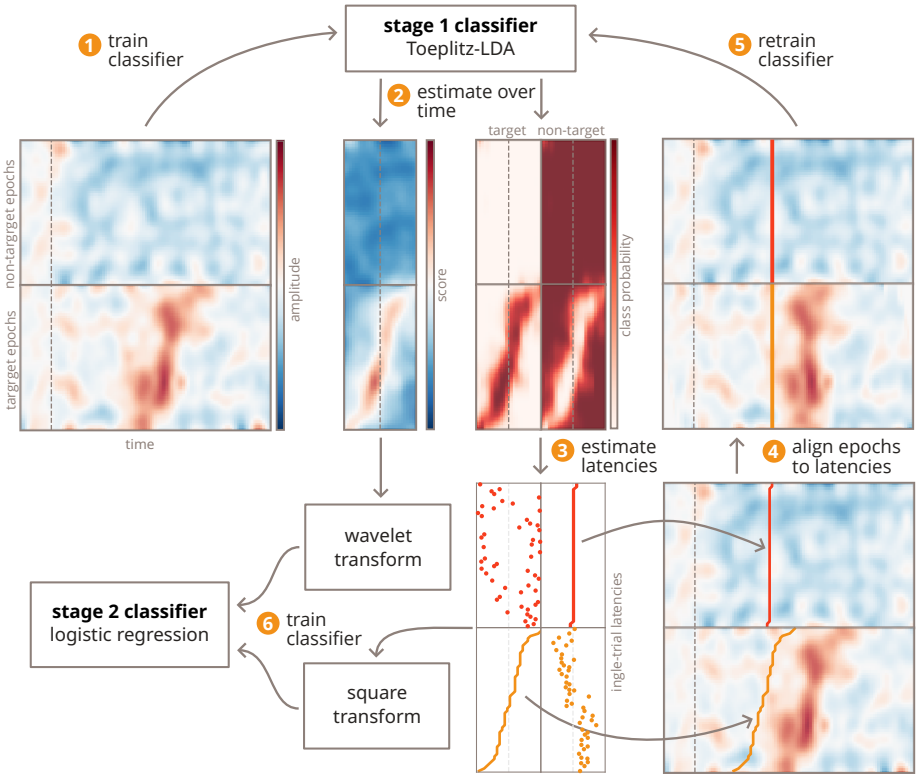


Figure 5.2: The Classifier-based Latency Estimation with Woody iterations (WCBLE) training phase. (1) The first-stage classifier is trained on a set of epochs. (2) It is then applied to time-shifted copies of the epochs, yielding scores and class probabilities over time. (3) The medians of the probability distributions are assumed as the class latencies. (4) The epochs are aligned by shifting in time based on their corresponding class latencies by shifting in time. (5) The spatiotemporal classifier is then retrained on the aligned epochs for the next iteration. (6) On convergence, the scores and latencies obtained from the last iteration are used to train the second-stage classifier.

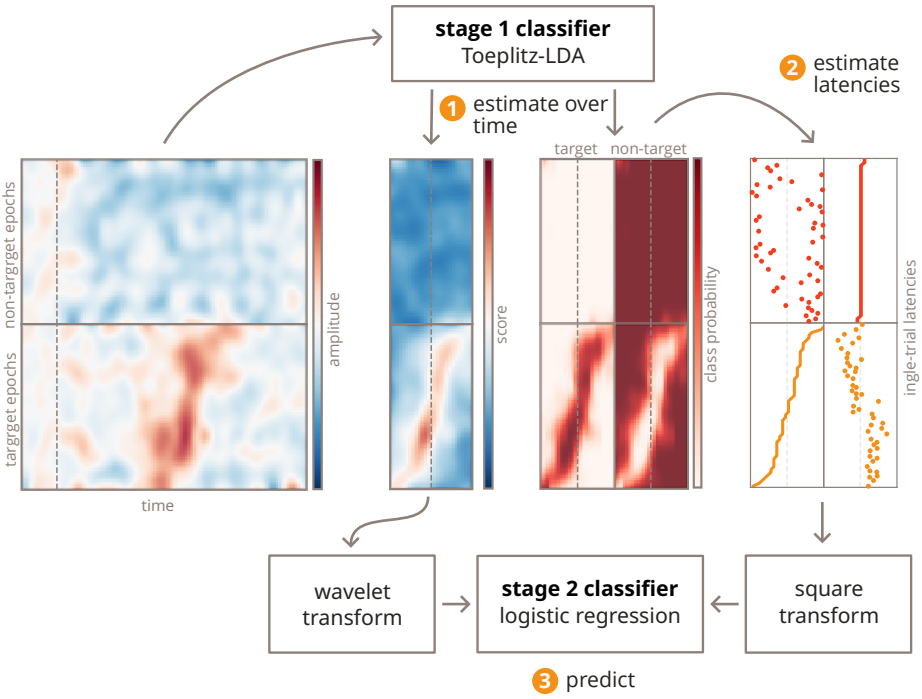


Figure 5.3: The Classifier-based Latency Estimation with Woody iterations (WCBLE) test phase. (1) The first-stage classifier obtained from the training phase is applied to time-shifted copies of the epochs, yielding scores and class probabilities over time. (2) The medians of these probability distributions are assumed as the new class latencies. (3) The scores and class latencies are input to the trained second-stage classifier, which predicts the label of the epochs.

5.3.4 Synthetic data generation

Since it is hard to obtain ground truth measures of ERP latencies, we first evaluate our approach with synthetic data. In chapter 6, the proposed algorithm will be applied to real EEG data.

To verify our approach, we simulated synthetic 16 channel EEG data with standard 10-20 electrode positions using the `simulate_evoked` method of the MNE (version 1.8) [88] software package. ERPs were generated by projecting a sine wave pulse source time course in a dipole in the left hemisphere to the scalp electrodes using a boundary element method forward model [174] and MNE's `fsaverage` source space and anatomy. Pink temporal noise is generated by passing Gaussian noise through an infinite impulse response filter with transfer function $\frac{B(z)}{A(z)} = \frac{1}{1 - 1z^{-1} + 0.15z^{-2}}$. Pink spatial noise is also added, using a noise covariance matrix constructed from the cosine distances between electrode positions and scaling it by the inverse of the resulting spatial covariance. We refer to Gramfort et al. [89] for implementation details.

The source time course is defined by the following function:

$$s(t, l) = \begin{cases} a \sin\left(2\pi f\left(t - l\right)\right) & \text{if } l - \frac{1}{2f} < t < l + \frac{1}{2f} \\ 0 & \text{otherwise} \end{cases} \tag{5.7}$$

over time t, with latency l, amplitude $a = 1 \times 10^{-7}$V, and frequency $f = 4$.

100 target epochs containing the evoked potential and noise and 100 non-target epochs containing only noise were simulated. The target epochs were jittered by setting the latency offset l of each target epoch to a random latency drawn from a normal distribution with 0 mean and standard deviations of respectively $\sigma = 0.1, 0.2, 0.3$s. The noise was scaled at 32 different noise levels, with the SNR of a target epoch ranging from 0 to -31 dB, with

$$\text{SNR} = 10 \log_{10} \frac{\text{Var}\left[s(t, 0)\right]}{\text{Var}\left(\text{noise}\right)} \tag{5.8}$$

A sample of the simulated evoked data is displayed in fig. 5.4.

5.4 Results

5.4.1 Latency estimation

We compared our proposed WCBLE algorithm to CBLE on a latency estimation task using the Pearson correlation coefficient ρ at different jitter and SNR levels. For both CBLE and WCBLE, target latencies were extracted in a 10-fold cross-validation scheme using the robust median latency method after fitting the model. WCBLE's maximum iteration number was set to 64. Results are presented in fig. 5.5. A 95% confidence interval on each measure was obtained through bootstrapping with 1000 permutations.

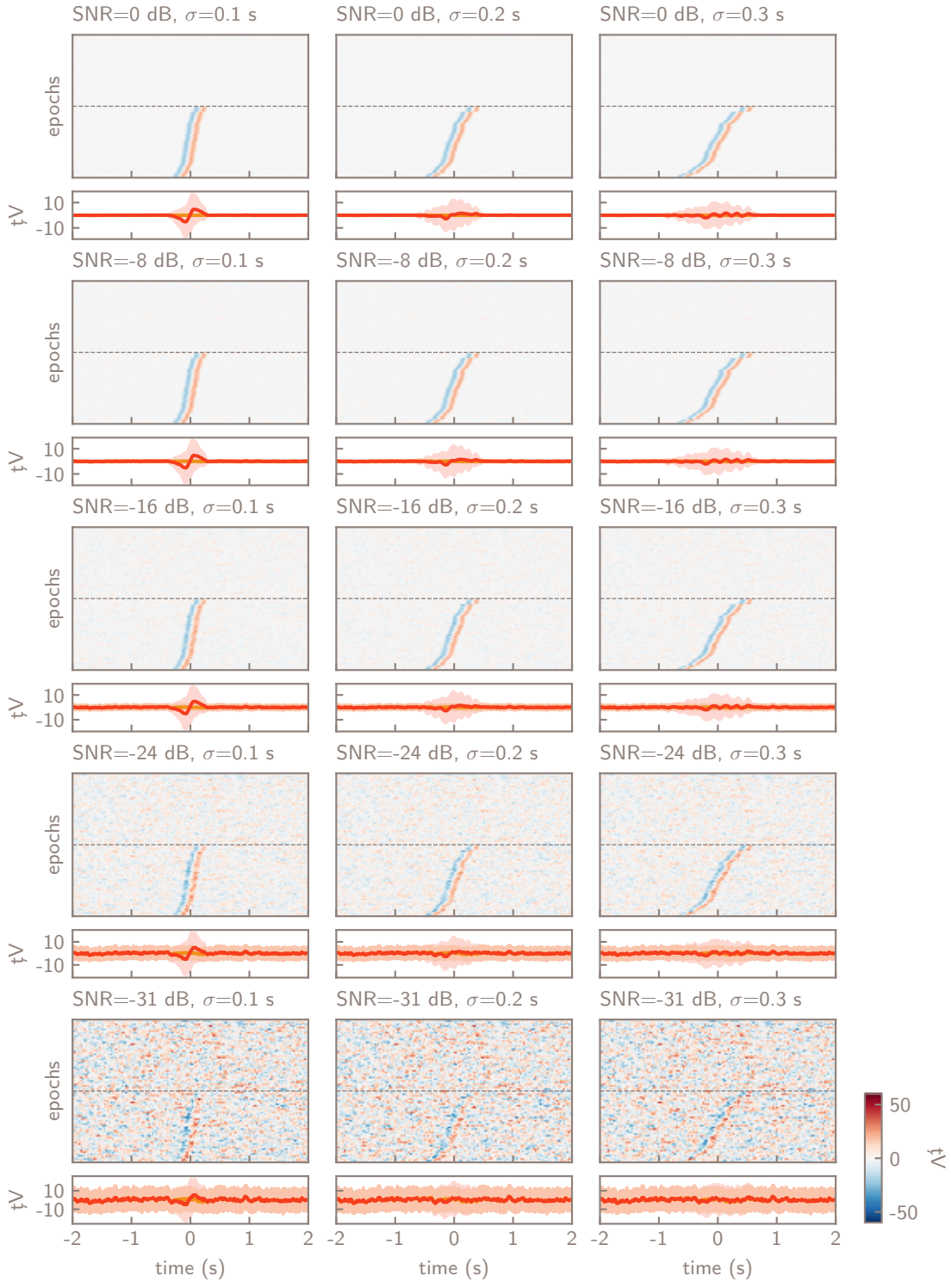


Figure 5.4: Simulated target (bottom half of epochs) and non-target (top half of epochs) evoked data and their average (red: target, yellow: non-target, shaded area: standard deviation). Targets contain a sine wave pulse, jittered by σ. 32 EEG channels with varying levels of noise were simulated using a forward model, channel C3 is plotted here.

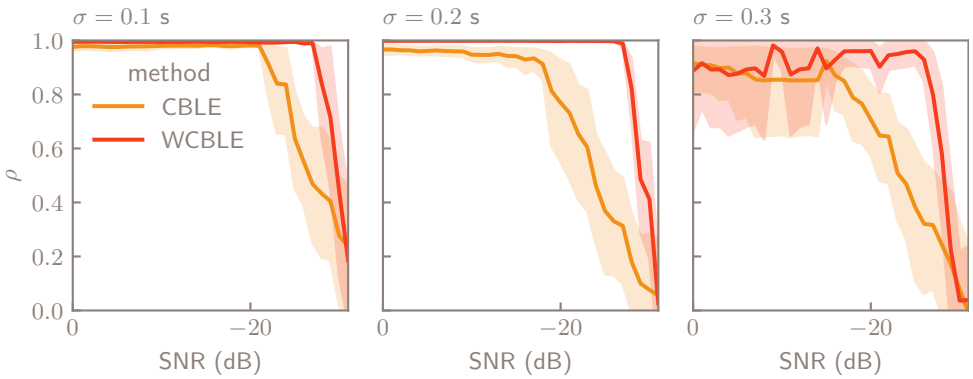


Figure 5.5: Correctness of estimated latency for varying jitter (σ) and SNR introduced in synthetic evoked target data, measured as Pearson correlation coefficient ρ between estimated and ground truth latencies. WCBLE estimates latencies more accurately in the presence of noise and high jitter.

As the SNR decreases, both CBLE and WCBLE can accurately estimate evoked potential latencies up to a given threshold. After this threshold, correlation between estimated and ground-truth latencies will start decreasing. However, we notice that this threshold occurs later for WCBLE than for CBLE for all evaluated jitter levels, indicating that WCBLE is more robust in the presence of pink spatiotemporal noise. At around -28 dB, WCBLE does not significantly outperform CBLE anymore, and both methods get overtaken by the strong presence of noise.

5.4.2 Classification

Analogous to the previous experiment, 10-fold cross-validated classification accuracy between target and non-target epochs was assessed. Here, WCBLE was compared to CBLE, but both also to their base classifier block-Toeplitz linear discriminant analysis (tLDA). CBLE and WCBLE classifiers were implemented as described in section 5.3, with both the wavelet-transformed score time series and square-transformed latencies as input to a second-stage logistic regression classifier with L_2-norm regularization ($C = 0.2$).

Figure 5.6 shows that the classification accuracy of all three methods decreases with σ and SNR, yet WCBLE outperforms both CBLE outperforms tLDA since its performance declines less steeply. This difference is particularly large when σ is higher ($\sigma = 0.2$ and $\sigma = 0.3$ s) and SNR is not yet high, with accuracies of respectively 77.50% for WCBLE, 17.00% for CBLE and 9.50% for tLDA when $\sigma = 0.3$ and SNR = -23 dB. WCBLE always scored better than CBLE and tLDA, except for $\sigma = 0.1$ and SNR $<$ 27dB, at which point both CBLE and WCBLE performed below the chance level of 50%.

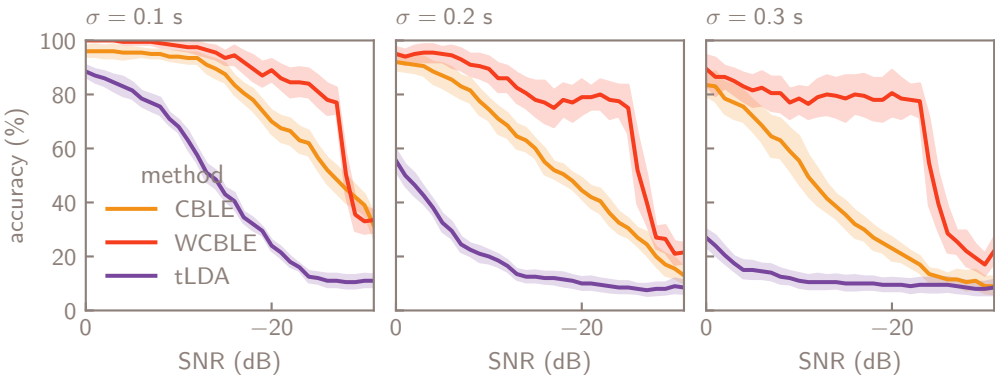


Figure 5.6: Decoding performance, measured as binary classification accuracy, for CBLE, WCBLE, and their first stage classifier tLDA for varying introduced jitter (σ) and SNR in synthetic evoked target and non-target data. WCBLE decoding performance is robust to more noise and jitter than the other methods.

5.5 Discussion & conclusion

In this work, we introduced a new ERP latency estimation and decoding method that applies a spatiotemporal classification iteratively along the time dimension of the data, refining the dataset at each iteration by aligning trials. The algorithm combines Classifier-based Latency Estimation (CBLE) with Woody's template matching iteration scheme and is therefore named Classifier-based Latency Estimation with Woody iterations (WCBLE). Our method attempts to model the spatial and temporal activation of each ERP component, as well as its latency distribution. Through the first-stage classifier, it takes into account the spatial and temporal structure of EEG background noise and the trials from the classes to discriminate.

The key finding of the simulation study presented in section 5.4 is that WCBLE is robust to higher evoked potential jitter and lower SNR than CBLE. It could therefore be a better candidate to apply in ERP analysis and decoding in those cases where jittered ERP data is of interest. At a certain noise threshold (-27 dB for $\sigma = 0.1$ s, -26 dB for $\sigma = 0.2$ s and -23 dB for $\sigma = 0.3$ s), both methods begin to fail, but WCBLE delivers higher accuracy ($> 75\%$) at intermediate SNR levels than CBLE, likely due to the iterative alignment process.

In general, noise beyond a certain level will overwhelm both CBLE and WCBLE performance, making it impossible to accurately estimate latencies or perform meaningful classification. When latencies cannot be properly estimated, aligning epochs using WCBLE does not help to improve classification.

As shown, the proposed method is suitable for BCI decoding settings as well as single-trial ERP latency analyses. In the case where there is only one data class present, the first-stage classifier can be replaced by a spatiotemporal filter. Instead of outputting a classification score, this filter should now output a metric representing the presence of a template response in the signal. A

suitable candidate is the spatiotemporal linearily constrained minimum-variance (LCMV)-beamformer [274], described in chapter 3. The activation pattern can be set to the average of the set of trials to analyze, and it can be refined at each WCBLE iteration by taking the average of aligned trials.

Finally, we argue that tLDA is a suitable first-stage classifier. Firstly, imposing a Toeplitz-covariance structure strongly regularizes the problem [244, 271], benefiting decoding performance. Secondly, this method has a synergy with CBLE since both make the same assumption about the short-time stationarity of the EEG background noise within an epoch. CBLE does not retrain the first-stage classifier for each time shift but rather trains it once within the given window. After training, the classifier parameters represent some information about the expected ERP waveform and background noise. By applying the trained classifier to different time shifts, it assumes this ERP waveform can be shifted in time, but since the classifier's information about the background noise was only obtained from the initial window, CBLE assumes its properties do not vary throughout the epoch. The block-Toeplitz covariance structure of tLDA also assumes that the background noise represented by this covariance after subtracting the class averages is stationary within the epoch [244].

The main limitation of the proposed method is its lack of capacity to handle multi-component ERP data. In real ERP analysis settings, multiple components are usually present, and each component has a distinct contribution to class discriminability and latency distribution. A single-component method will yield issues in interpretability and convergence in the presence of other components, as it can 'lock-on' to a given component cluster and flatten out the others. This can be avoided, however, by the choice of a proper region of interest isolating solely the component of interest. Current work focuses on extending WCBLE to a multi-component setting, on the one hand to improve latency estimation and decoding performance on real ERP data, and on the other hand to apply it as an ERP component separation method in the fashion of Ouyang et al. [196].

The proposed implementation theoretically supports processing multi-class data, but the convergence and correctness of the WCBLE solution is yet to be properly studied for this case.

5.A CBLE algorithm

Algorithm 5 CBLE

TRAIN

Input: $\{\mathbf{X}_n^{\text{train}}\}_{n=1}^N, \mathbf{l}^{\text{train}}, \mathcal{C}(\cdot, f, \Pr), s_1, s_2$

$\theta \leftarrow \text{train}_{\mathcal{C}}(\{\mathbf{X}_n^{\text{train}}[:, s_1 : s_2]\}_{n=1}^N, \mathbf{l}^{\text{train}})$ ▷ Train stage 1

for $n = 1 \dots N$ **do** ▷ Feature extraction for stage 2

  for $s = 1 \dots R$ **do**

    $y_{n,s}^{\text{train}} \leftarrow f(\mathbf{X}_n^{\text{train}}[:, s : s + (s_2 - s_1)], \theta)$

  end for

  $s_{n,\text{target}}^{\text{train}} \leftarrow \text{median}\left[\Pr(s|\mathbf{X}_n^{\text{train}}, \theta, \text{target})\right]$

  $s_{n,\text{non-target}}^{\text{train}} \leftarrow \text{median}\left[\Pr(s|\mathbf{X}_n^{\text{train}}, \theta, \text{non-target})\right]$

end for

Output: $\theta, \mathbf{Y}^{\text{train}}, \mathbf{s}_{\text{target}}^{\text{train}}, \mathbf{s}_{\text{non-target}}^{\text{train}}$

EVALUATE

Input: $\{\mathbf{X}_m^{\text{test}}\}_{m=1}^N, \mathcal{C}(\theta, f, \Pr), s_1, s_2$

for $m = 1 \dots M$ **do** ▷ Feature extraction for stage 2

  for $s = 1 \dots R$ **do**

    $y_{m,s}^{\text{test}} \leftarrow f(\mathbf{X}_n^{\text{test}}[:, s : s + (s_2 - s_1)], \theta)$

  end for

  $s_{m,\text{target}}^{\text{test}} \leftarrow \text{median}\left[\Pr(s|\mathbf{X}_m^{\text{test}}, \theta, \text{target})\right]$

  $s_{m,\text{non-target}}^{\text{test}} \leftarrow \text{median}\left[\Pr(s|\mathbf{X}_m^{\text{test}}, \theta, \text{non-target})\right]$

end for

Output: $\mathbf{Y}^{\text{test}}, \mathbf{s}_{\text{target}}^{\text{test}}, \mathbf{s}_{\text{non-target}}^{\text{test}}$

5.B WCBLE algorithm

Algorithm 6 WCBLE

TRAIN

Input: $\{\mathbf{X}_n^{\text{train}}\}_{n=1}^{N}, \mathbf{l}^{\text{train}}, \mathcal{C}(\cdot, f, \Pr), s_1, s_2$

$\mathbf{X}'_n \leftarrow \mathbf{X}_n^{\text{train}} \quad \forall \quad n = 1 \dots N$ ▷ Train stage 1
repeat
  $\theta^* \leftarrow \text{train}_{\mathcal{C}}(\{\mathbf{X}'_n[:, s_1 : s_2]\}_0^{N-1}, \mathbf{l}^{\text{train}})$
  for $n = 1 \dots N$ **do**
    $s_n \leftarrow \text{median}\left[\Pr(s|\mathbf{X}'_n, \theta^*, l_n)\right]$
    $\mathbf{X}'_n \leftarrow \text{align}(\mathbf{X}_n^{\text{train}}, s_n^*)$
  end for
until convergence or maximum iterations reached
for $n = 1 \dots N$ **do** ▷ Feature extraction for stage 2
  for $s = 1 \dots R$ **do**
    $y_{n,s}^{\text{train}} \leftarrow f(\mathbf{X}_n^{\text{train}}[:, s : s + (s_2 - s_1)], \theta^*)$
  end for
  $s_{n,\text{target}}^{\text{train}} \leftarrow \text{median}\left[\Pr(s|\mathbf{X}_n^{\text{train}}, \theta^*, \text{target})\right]$
  $s_{n,\text{non-target}}^{\text{train}} \leftarrow \text{median}\left[\Pr(s|\mathbf{X}_n^{\text{train}}, \theta^*, \text{non-target})\right]$
end for

Output: $\theta^*, \mathbf{Y}^{\text{train}}, \mathbf{s}_{\text{target}}^{\text{train}}, \mathbf{s}_{\text{non-target}}^{\text{train}}$

EVALUATE

Input: $\{\mathbf{X}_m^{\text{test}}\}_{m=1}^{N}, \mathcal{C}(\theta^*, f, \Pr), s_1, s_2$

for $m = 1 \dots M$ **do** ▷ Feature extraction for stage 2
  for $s = 1 \dots R$ **do**
    $y_{m,s}^{\text{test}} \leftarrow f(\mathbf{X}_n^{\text{test}}[:, s : s + (s_2 - s_1)], \theta^*)$
  end for
  $s_{m,\text{target}}^{\text{test}} \leftarrow \text{median}\left[\Pr(s|\mathbf{X}_m^{\text{test}}, \theta^*, \text{target})\right]$
  $s_{m,\text{non-target}}^{\text{test}} \leftarrow \text{median}\left[\Pr(s|\mathbf{X}_m^{\text{test}}, \theta^*, \text{non-target})\right]$
end for

Output: $\mathbf{Y}^{\text{test}}, \mathbf{s}_{\text{target}}^{\text{test}}, \mathbf{s}_{\text{non-target}}^{\text{test}}$

Chapter 6

Compensating ERP jitter for gaze-independence

This chapter was published as part of Van Den Kerchove et al. [272].

6.1 Introduction

The decoding methods described in chapter 3 and chapter 4 are general attempts at increasing ERP decoding accuracy. Given the problem statement of this work, we are also interested in targeted methods that would specifically advance the development of gaze-independent visual brain-computer interfaces (BCIs) (see section 2.2). Methods that directly enhance gaze-independent decoding through optimizing decoder design have not been widely studied (see section 2.2.3). We assume that advances in gaze-independent decoding can be made if properties of covert visuospatial attention (VSA) that affect performance can be identified and accounted for. The latency estimation and alignment approach proposed in chapter 5 was specifically designed for this purpose.

Aricò et al. [12] observed higher variability in single-trial P3 peak latencies relative to stimulus onset during covert VSA (529 ms^2) compared to overt VSA (256^2 ms). This latency variability contributes to reduced covert VSA decoding performance. While they proposed an analysis and performance prediction method, they did not provide a decoding solution. They suggested that compensating for latency jitter could enhance covert VSA decoding, but did not verify this hypothesis directly. Additionally, Hardiansyah, Pergher, and Van Hulle [100] developed a classifier for covert VSA ERPs, exploiting single-trial latency features in combination with amplitude features for classification with a support vector machine. They demonstrated the positive influence of single-trial ERP component latency features on covert VSA inference, yet did not attempt to correct the amplitude features for these latencies.

Furthermore, as mentioned in section 2.2.2 and section 2.4.2, our goal is also to explore more flexible gaze-independent settings than the usual covert

VSA with central gaze fixation reported in most literature. Frenzel, Neubert, and Bandt [78] introduced a similar protocol to the proposed split VSA setting (see fig. 2.2d). They showed that it is possible to perform split VSA and that, in this case, visuospatial attention and gaze direction can be decoded separately using classical ERP techniques. To the best of our knowledge, this is the sole study that investigated split attention in ERP-based BCIs. However, Frenzel, Neubert, and Bandt [78] considered their interface only for the case where the user actively intends to select both targets determined by the gaze and the VSA. In the split VSA setting considered in our work, we instruct the participant to ignore the distractor and only attend to the cued target, as we are interested in decoding the visuospatial attention only.

6.2 Materials & methods

6.2.1 Data collection

We evaluate our approach on a dataset specifically recorded for this study, designed to probe different modalities of covert VSA, and on the publicly available BNCI2014-009 dataset [9].

CVSA-ERP dataset

We recorded a dataset to validate our approach. The Covert Visuospatial Attention ERP (CVSA-ERP) dataset consists of 15 participants, mean age 26.34±3.04 years. This study was approved by the Ethics Commission of University Hospital Leuven (S62547). Each subject performed different VSA conditions (overt, covert, and split), illustrated in Figure 6.1 (top row). Using a hexagonal layout interface, similar to the visual Hex-o-Spell proposed by Treder and Blankertz [264], we presented six flashing targets (without letters or symbols) to the participant while the electroencephalography (EEG), electrooculogram (EOG), and the participant's eye gaze using eye tracking were recorded. The VSA conditions described in the first row of Figure 6.1 are considered.

In contrast to the protocol proposed by Frenzel, Neubert, and Bandt [78], split VSA was performed by instructing the participant to attend to the intensifications of the cued target and ignore the intensifications of the distractor target. Since we assume there will be an effect depending on the distance between the attended target and the distractor, we discern three split VSA sub-conditions: the distractor is either clockwise or counterclockwise directly next to the attended target ($d = 1$), there is one other target between the attended target and the distractor ($d = 2$), or the distractor is opposite the intended target ($d = 3$).

EEG for the CVSA-ERP dataset was recorded using a SynAmps RT amplifier (Compumedics Neuroscan, Australia) at 2048 Hz and 62 Ag/AgCl active electrodes arranged in the international 10-10 layout fitted to a standard electrode cap (EASYCAP GmbH, Germany), with electrodes located at AFz and FCz as ground and reference, respectively. Using electrolyte gel, electrode

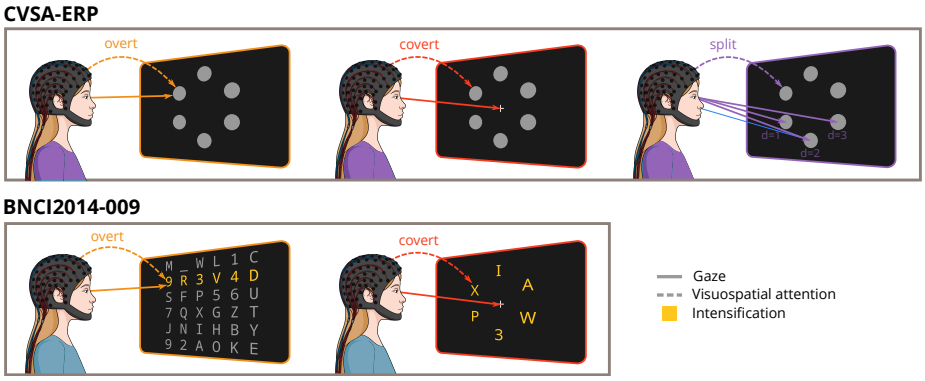


Figure 6.1: Interfaces and visuospatial attention (VSA) conditions in the CVSA-ERP and BNCI2014-009 datasets. In the CVSA-ERP oddball BCI interface, screen targets are intensified one after the other in pseudorandom order while the participant can either pay overt, covert, or split VSA to the cued target. In the BNCI2014-009 overt VSA interface, entire rows and columns are intensified at once. In its covert counterpart, groups of 6 letters are intensified one after the other, partly relying on feature attention.

impedances were brought below 5kΩ. Electrodes TP9 and TP10, used for off-line re-referencing, were directly attached to the skin using stickers for better contact. The power line frequency in Belgium is 50 Hz. The participant's eye gaze was registered using an EyeLink 1000 Plus eye tracker (SR Research, Canada) in non-fixation mode.

Participants signed an informed consent form and were seated at a distance of 60 cm from a CRT-emulating monitor (VPixx Technologies, Canada) operating at a refresh rate of 120 Hz, displaying 6 circular white targets with a diameter of 4.15ř visual angle and laid out on a hexagon with a radius of 12.28ř of visual angle centered on the midpoint of the screen, conforming to the interface proposed by Treder and Blankertz [264] (fig. 6.3a). A hexagonal layout interface with an empty center and a low number of targets counteracts target crowding, and as long as the subjects gaze is within the hexagon of targets, no other target can be between the subjects gaze and a covertly attended target. Targets are full-contrast white and were intensified by scaling them to a larger size (5.60ř of visual angle, see fig. 6.3b) instead of changing the contrast to avoid Troxler-fading[1] [264] in the peripheral visual field. Stimuli were presented using PsychoPy (version 2023.1.3) [206].

The participant was instructed to press the space bar when ready for a block of stimulations. Then, one target was indicated as the cue, and the participant was instructed to count the number of intensifications of the cued target during the following block of stimulations. After pressing the space bar again, a blue

[1]The optical illusion of disappearing unchanging stimuli experienced when visually fixating [266].

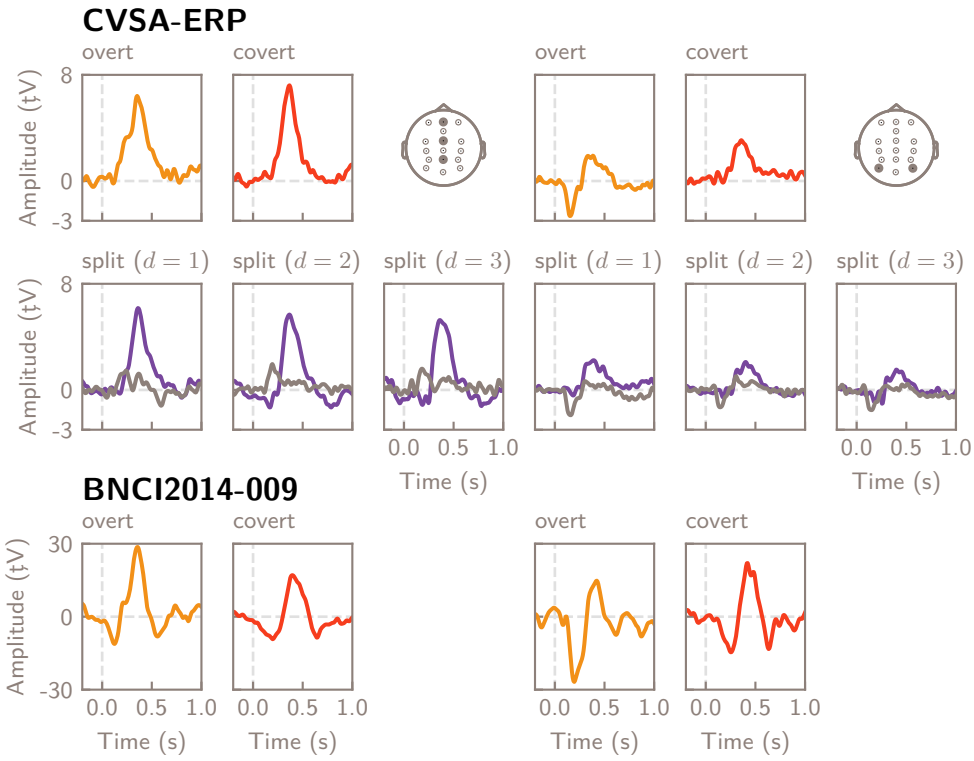
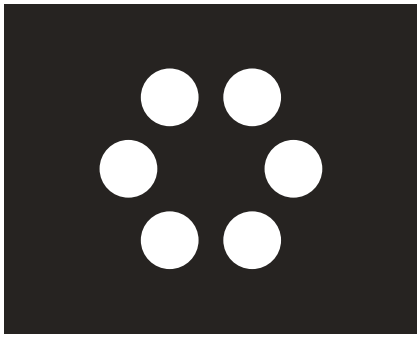


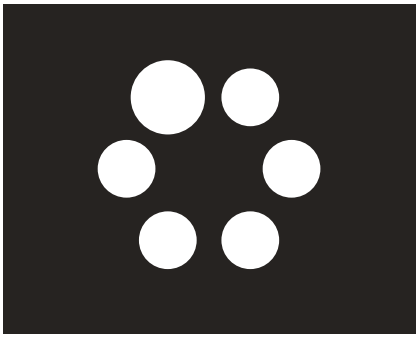
Figure 6.2: Contrast between target (color) and non-target, and distractor (gray) and non-target grand average event-related potentials per VSA condition and dataset. Overt VSA yields a strong modulation of the N1 component in both datasets; the P3 amplitude decreases with the degree of split VSA. In split VSA, N1 and P2 are more prominently evoked by the distractor, while the P3 is evoked by the target.

crosshair appeared, and the subject was instructed to fixate their gaze on the blue crosshair for the duration of the stimulation block (fig. 6.3c and fig. 6.3d). The position of this crosshair determined the VSA condition for this trial: overt VSA when the crosshair was at the same location as the cued target, covert VSA when the crosshair appeared in the center of the screen, and split VSA when the crosshair appeared on a different target than the cued one. After pressing the space bar again and a delay of 5 seconds, the stimulation block started.

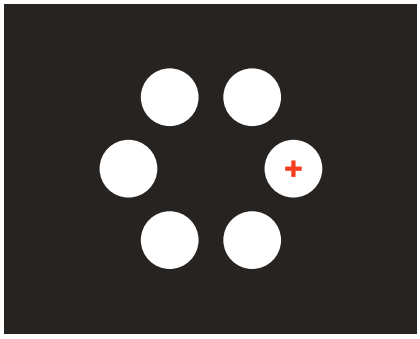
All targets were intensified for a duration of 100 ms in pseudorandom order. The inter-stimulus-interval (inter-stimulus interval (ISI)), the time between the onsets of subsequent intensifications, was variable and consisted of a fixed 300 ms interval (of which 100 ms with an intensified target onscreen) with 200 ms of uniform jitter added, resulting in an ISI between 200 and 400 ms. ISIs



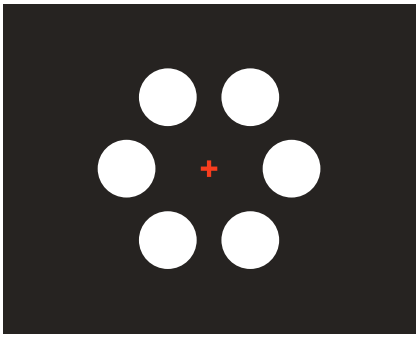
(a) The stimulation interface based on the visual Hex-o-Spell BCI [264].



(b) When a target is intensified, it is enlarged for a very short time.



(c) In the overt and split VSA settings, the participant is cued to fixate on one of the targets.



(d) In the covert VSA setting, the participant is cued to fixate at the center of the interface.

Figure 6.3: Layout and visual elements of the experimental stimulation interface used in recording the CVSA-ERP dataset.

were jittered to counteract steady-state effects and residue in averaging. A longer ISI will increase component amplitude and aid in counteracting temporal autocorrelation for higher statistical test precision. In a block of stimulations, each target was intensified a pseudorandom number of times between 10 and 15. This led to stimulation blocks with an average duration of 26.25 seconds. After a block of stimulations, an input prompt appeared to enter the mentally counted number of intensifications. After inputting this number, the subject was allowed to pause until pressing the space bar again. In total, six blocks were presented for overt VSA, six blocks for covert VSA, 12 blocks for split ($d = 1$) VSA, 12 blocks for split ($d = 2$) VSA, and 6 blocks for split ($d = 3$) VSA, covering all possible combinations of VSA conditions, cued targets, and crosshair locations.

The experiment started with a sequence of five non-recorded practice stimulation blocks, one for each of the five VSA conditions. During these practice blocks, the participant received feedback about their gaze position and counting

accuracy. Counting the instructions and the participant's response to the input prompts, a block lasted about 30 seconds. In sum, the experiment featured approximately 45 minutes of stimulation time. After blocks 14 and 28, the participant was allowed to take a longer break. Including these longer breaks, the experiment lasted approximately one hour.

BNCI2014-009 dataset

The BNCI2014-009[2] dataset [9] was used in the analysis performed by Aricò et al. [12]. It contains data from 10 subjects (median age 24.5 ± 1.9 years) who performed two spelling tasks illustrated in the second row of fig. 6.1: using the P3 Matrix speller interface to exploit overt VSA, and the GeoSpell covert VSA interface. To use the GeoSpell interface, the participant gazes at the fixation point at the center of the screen, while groups of characters flash simultaneously in a circular layout around the fixation point. The user directs their visuospatial attention to the location where the intended letter is expected to appear, and when it does, a P3 ERP component is expected to be evoked. This results in a specific setting where both visuospatial attention and feature attention (the attended letter) are exploited. For a detailed description of the paradigm and dataset, we refer to Aloise et al. [9].

6.2.2 Data processing and analysis

Preprocessing

Analysis was performed using Python and the MNE software package (version 1.3.1) [88]. All datasets were band-pass filtered between 0.1 Hz and 20 Hz with a 4th-order Butterworth filter. Bad channels in the data were automatically detected using the RANSAC method [74] and rejected. The recorded EEG was re-referenced offline to the average of the mastoid electrodes TP9 and TP10. Next, the EEG signals were corrected for eye movement artifacts using Independent Component Analysis (ICA). Since we have access to electrooculogram (EOG) data for the CVSA-ERP dataset, components correlating significantly with the EOG were rejected. For the BNCI2014-009 dataset, ICA components were manually rejected. Finally, the EEG signal was divided into epochs ranging from 100 ms before stimulus onset to 700 ms after stimulus onset and down-sampled to 128 Hz. In both datasets, only 16 channels were kept for analysis (Fz, FCz, Cz, CPz, Pz, Oz, F3, F4, C3, C4, CP3, CP4, P3, P4, PO7, and PO8).

Decoders

We compared the two latency-based methods Classifier-based Latency Estimation with Woody iterations (WCBLE) and Classifier-based Latency Estimation (CBLE) as described in section 5.3, with their base classifier block-Toeplitz linear discriminant analysis (tLDA) and with Riemannian Geometry approaches

[2] `https://bnci-horizon-2020.eu/database/data-sets`

that rely on spatial covariance as features. Together with tLDA, Riemannian Geometry generally achieves state-of-the-art decoding performance [155]. We implemented two Riemannian Geometry pipelines. The first estimates shrunk covariances from the ERPs filtered with 6 XDAWN filters, projects these covariances to a tangent space, and classifies the result using L_2-regularized logistic regression (XDAWNCov-TS-LR) [46]. Secondly, we adopted the pipeline from Aydarkhanov et al. [16], as their work shows favorable performance in the presence of single-trial ERP latency jitter. Shrunk spatial covariance matrices were estimated from epochs that were augmented by concatenating the average target and average non-target ERP as extra channels, projected to tangent space, and classified using L_2-regularized logistic regression (ERPCov-TS-LR).

To evaluate performance, 6-fold cross-validation without shuffling was performed for both datasets. At each fold, classifiers were trained on five target selection blocks (300 epochs) and tested on one block (60 epochs) without overlap for CVSA-ERP. For each subject and run in the BNCI2014-009 dataset, classifiers were trained on five symbol selections (480 epochs) and tested on one symbol selection (96 epochs) without overlap at each fold. A window ranging from 0 ms to 600 ms after stimulus onset was used for CBLE and WCBLE. With epochs ranging from -100 ms to 700 ms relative to stimulus onset, this allows for extracting latencies ranging from -100 ms to +100 ms.

6.3 Results

6.3.1 BCI decoding performance

We evaluated the BCI decoding performance in a single-trial classification experiment, as well as in a target selection experiment reflecting BCI operation.

Figure 6.4 shows a comparison of area under the receiver-operator characteristic curve (ROC-AUC) for all pairs of tLDA, CBLE, and WCBLE for single-trial classification to investigate the contributions of CBLE and WCBLE relative to their first-stage classifier tLDA. For this evaluation, epochs were rejected when the peak-to-peak amplitude exceeded 800 ţV, and for the CVSA-ERP dataset, if the user's gaze differed by more than 10 degrees of visual angle from the fixation crosshair.

Wilcoxon signed-rank tests controlled for multiple comparisons by Benjamini and Hochberg's False Discovery Rate procedure (FDR) revealed that for the BNCI2014-009 dataset, WCBLE significantly outperformed tLDA ($\Delta\text{ROC} - \text{AUC} = 0.019$, $p = 0.004$) and CBLE ($\Delta\text{ROC} - \text{AUC} = 0.016$, $p = 0.036$) for covert VSA, but was significantly outperformed by tLDA in overt VSA decoding ($\Delta\text{ROC} - \text{AUC} = -0.004$, $p = 0.040$). For the CVSA-ERP dataset, WCBLE also achieved significantly better covert VSA performance than tLDA ($\Delta\text{ROC} - \text{AUC} = 0.023$, $p = 0.041$) and CBLE ($\Delta\text{ROC} - \text{AUC} = 0.036$, $p = 0.024$). We found no significant difference in WCBLE performance over tLDA in the split VSA conditions in the CVSA-ERP dataset, but results show a clear trend of increased WCBLE performance over tLDA and CBLE as d increases. CBLE failed to

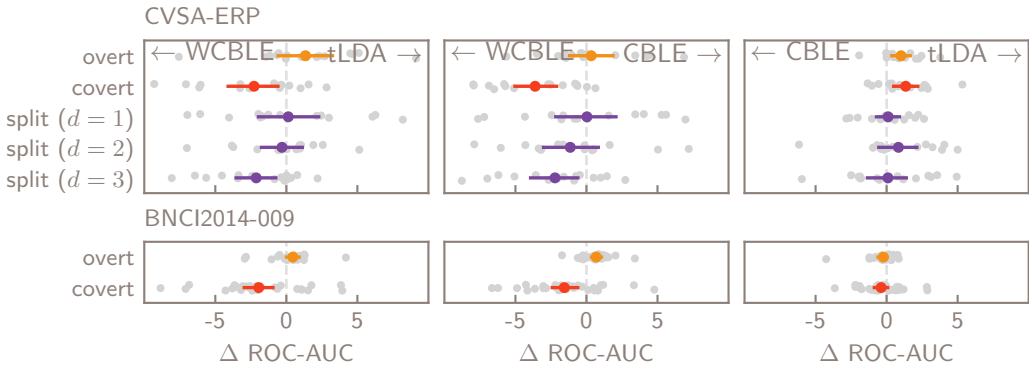


Figure 6.4: Difference in cross-validated single-trial classification area under the receiver-operator characteristic curve (%.) between CBLE, WCBLE, and their first-stage classifier tLDA. 95% confidence intervals were determined using $k = 1000$ bootstrapping. Our proposed WCBLE decoder outperforms tLDA and CBLE for covert and split ($d = 3$) VSA. CBLE scores on par with tLDA.

significantly outperform its first-stage classifier tLDA in all evaluated VSA conditions. Table 6.1 reports all single-trial classification scores for all considered models, datasets, and conditions.

Section 6.3.1 shows the cross-validated BCI selection accuracy on the BNCI2014-009 and the CVSA-ERP datasets for all investigated decoders. Accuracy was determined by, for each block, selecting the character with the highest (stage-two if applicable) classifier score and comparing it to the cued target. Significance was calculated using one-sided Wilcoxon signed-rank tests ($p = 0.05$) corrected for FDR over decoders. For this evaluation, no epochs were rejected to keep the trial-based structure of BCI operation intact. For all datasets and VSA conditions, CBLE scores approximately on par with tLDA. Yet, WCBLE yields an improved decoding accuracy for covert VSA in both datasets, which is greatest for smaller numbers of repetitions and decreases as the number of repetitions increases. This covert VSA accuracy increase over tLDA is significant in the BNCI2014-009 dataset for 1 and 3 repetitions, and in CVSA-ERP for 1-5 and 10 repetitions. Furthermore, while we reported a relative decrease in single-trial ROC-AUC for WCBLE in overt VSA, this does not seem to result in a consistent decrease in target selection accuracy. No significant increase of WCBLE over other methods was found in split VSA. While Riemannian methods are significantly outperformed by tLDA, CBLE, and WCBLE in the BNCI2014-009 dataset, they perform approximately on par with tLDA and WCBLE in CVSA-ERP.

Overall, we observed a 5.10%pt. accuracy increase with WCBLE over tLDA for covert VSA in the BNCI2014-009 dataset and 5.55%pt. in the CVSA-ERP dataset. These results compare to the performance gain reported by Zisk et al. [317]. They observed a 5.63%pt. accuracy increase with 1-10 selection repetitions over step-wise Linear Discriminant Analysis (SWLDA) for 6 participants

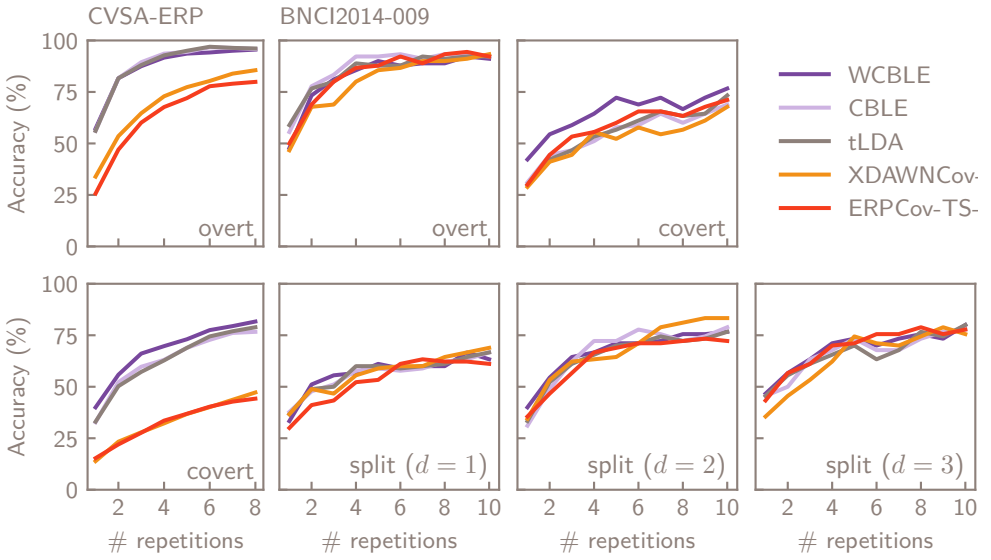
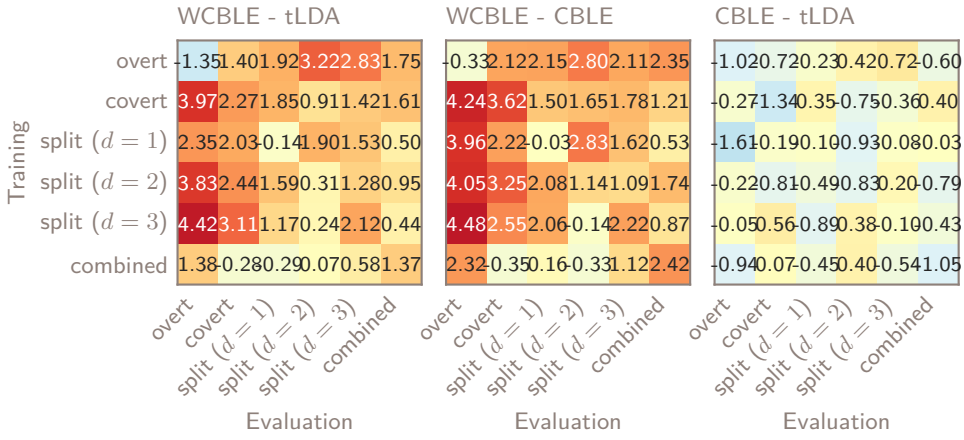


Figure 6.5: Cross-validated target selection accuracy for all decoders plotted as a function of the number of test repetitions in different VSA conditions. WCBLE generally achieves the highest covert VSA target selection accuracy.

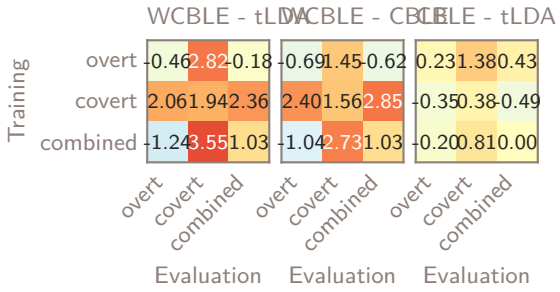
with Amyotrophic Lateral Sclerosis (ALS), whose SWLDA performance also suffered from jitter. Note that interpretation of this comparison may be challenging due to differences in interface design (number of targets, ISI), subject population, EEG recording procedure, and available training data.

6.3.2 Gaze-independence through cross-condition transfer

To further back our claim of gaze-independence in the case where eye motor control cannot be assumed, we evaluate our proposed decoder in a transfer learning setting between VSA conditions. While performing overt attention requires gaze redirection for each target selection, performing covert or split VSA continuously still requires sustained gaze fixation, which might be impaired in those who could benefit from such an application. Studying the transfer between conditions simulates what happens when the user performs different VSA conditions throughout the experimental session. Furthermore, if our decoder performs well in transfer-learning settings, it must capture some information about the ERP responses that is independent of the VSA condition, and hence does not depend on gaze redirection to perform these conditions. We introduce an additional setting of interest here, namely a combination of VSA conditions, which represents those cases where the user cannot redirect their gaze and hence can be in any one of the VSA conditions depending on the target they attend to. For BNCI2014-009, this is implemented as an equal mix of overt and covert VSA; for CVSA-ERP, the combined condition represents an equal mix of overt,



(a) Cross-condition performance for CVSA-ERP



(b) Cross-condition performance for BNCI2014-009

Figure 6.6: Difference in cross-validated ROC-AUC (%.) between Classifier-based Latency Estimation (CBLE) and Woody CBLE (WCBLE) across conditions for the CVSA-ERP and BNCI2014-009 datasets, respectively. A decoder is each time trained on a visuospatial attention (VSA) condition and tested on all VSA conditions. WCBLE yields an improvement in most non-overt VSA settings, indicating it is more invariant to eye gaze than CBLE and tLDA.

covert, and split VSA, disregarding parameter d.

Figure 6.6a and fig. 6.6b show the pairwise differences in area under the ROC curve (ΔROC-AUC) between the investigated decoders. In this evaluation, bad epochs were rejected as in section 6.3.1. When comparing CBLE and tLDA, we do not observe large differences in any of the evaluated settings, similar to the within-subject conditions, with the greatest ROC-AUC difference -2% for training in split ($d = 2$) VSA and overt VSA. On the contrary, when considering the comparisons between WCBLE and tLDA, we see that performance is on par or greater using WCBLE for most conditions, except for within-overt VSA decoding, with the greatest ROC-AUC difference +4% for training in split ($d = 2$) VSA and testing in overt VSA.

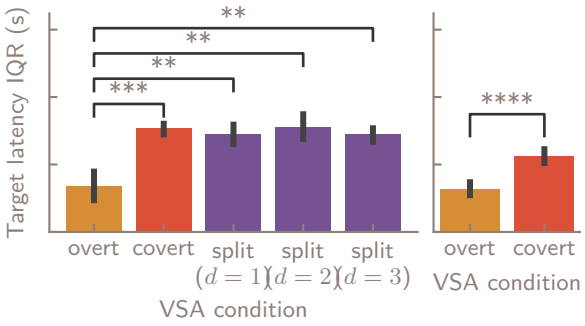


Figure 6.7: Jitter characterized as the interquartile range (IQR) of target epochs for different VSA conditions. Overt VSA exhibits lower jitter than other conditions. Significance of differences was determined with two-sided Wilcoxon signed-rank tests with False Discovery Rate correction on per-subject jitter ($* = p < 0.05$, $** = p < 0.01$, $*** = p < 0.001$, $**** = p < 0.0001$).

6.3.3 Jitter analysis

Finally, an analysis is performed to quantify the jitter presence in the different VSA conditions. To obtain comparable results across VSA conditions, WCBLE was evaluated per session and trained on all combined VSA conditions as in section 6.3.2. Since all conditions have the P3 in common, the estimated latencies can be interpreted as P3 latencies. Results are shown in fig. 6.7.

Two-sided Wilcoxon rank-sum tests on jitter, expressed as the interquartile range (IQR) of the estimated latencies of target trials, revealed that overt VSA exhibited significantly lower jitter than all other conditions for both datasets, with jitter equal to overt: 32 ms and covert: 56 ms ($p < 0.001$) for BNCI2014-009, and overt: 33 ms, covert: 77 ms ($p < 0.001$), split ($d = 1$): 72 ms ($p = 0.006$), split ($d = 2$): 77 ms ($p = 0.002$) and split ($d = 3$): 72 ms ($p = 0.002$) for CVSA-ERP. No other significant differences were found. p-values were corrected for the family-wise error rate using Bonferroni correction.

6.4 Discussion

Figure 6.4 shows that WCBLE significantly improves covert VSA decoding. This is advantageous for the development of a class of ERP-BCI interfaces for users who prefer to rest their gaze on a fixation cross on the screen, avoiding the effort of redirecting their eye gaze for every selection. Furthermore, the performance gain over the first-stage classifier in split VSA ($d = 3$) and between-VSA condition transfer settings are promising for users with even less eye motor control who may experience involuntary eye movements or fixation fatigue and hence cannot keep their gaze fixed throughout an entire BCI operation session. WCBLE would allow them to operate a BCI comfortably while directing their gaze to whichever portion of the screen they prefer, even when there is

another target present or this location varies during the course of operation. Although WCBLE did not significantly improve overt VSA single-trial decoding, section 6.3.1 shows that this does not negatively impact target selection accuracy. While target selection accuracy also did not improve for split VSA, the increase in single-trial performance in split ($d = 3$) shows that an iterative alignment procedure has the potential to improve over CBLE and its first-stage classifier in this case as well.

We believe the relative increase in performance of our proposed decoder in covert and split VSA, and the lack thereof in overt VSA, could stem from the following: (1) Covert and split VSA exhibit higher P3 jitter than overt VSA. In covert and split VSA, participants have to execute a dual task by dissociating their visuospatial attention and gaze fixation. Evidence shows that ERP latency variability is higher when attention is divided [215, 12]. Aricò et al. [12] also partly attribute higher latency jitter to the covert VSA task performed in the BNCI2014-009 dataset, since the GeoSpell interface requires both spatial and feature attention. (2) In overt VSA, the first-stage classifier can rely mostly on the modulation of early visually evoked potentials (VEPs) like N1 rather than on the P3 [264]. These VEPs are closely related to visual processing, hence exhibit lower jitter, contrary to P3, which is more prone to the effects of attention and workload [111], reducing the contribution of alignment. (3) This property can also result in the estimation of VEP latencies instead of the P3 latency, and WCBLE would in this case fail to increase the P3 signal-to-noise ratio (SNR), which still could be somewhat jittered in overt VSA. (4) Aligning to the P3 will lower the SNR of early VEPs, while aligning to VEPs will lower P3 SNR, since they are not time-locked to each other. (5) Covert and split VSA ERPs may exhibit lower SNR than overt VSA due to lower P3 amplitudes or even due to the presence of higher P3 jitter itself. Higher SNR in overt VSA results in higher decoding performance of state-of-the-art classifiers, leaving less room for relative improvement in this case.

Although it is not immediately clear if WCBLE actively corrects for higher P3 jitter present in covert and split VSA compared to overt VSA, we justify our approach in a similar manner to Hardiansyah, Pergher, and Van Hulle [100] by observing that the increased discrimination performance of a machine-learning model accounting for jitter forgoes the need for characterizing the underlying physiological processes, while still objectively quantifying the presence of jitter between data classes. Furthermore, some evidence points toward higher P3 jitter as the main contributing factor. While Figure 6.2 shows no visible smearing effect in the shape of the ERPs, the more quantitative analysis on latencies presented in section 6.3.3 indicates the opposite. Additionally, Aricò et al. [12] prove that P3 jitter does play a non-negligible role in covert VSA by comparing performance between overt and covert VSA, while including or excluding early VEPs. This analysis showed that the absence of the N1 and other early VEPs is not the only factor hampering covert VSA decoding performance. However, the large increase in accuracy for covert VSA reported by Aricò et al. [12] (60% to 95% with a window focusing on only the P3 component) could also be explained by overfitting on artifacts amplified by aligning, since their method is

not evaluated on unseen data, as opposed to ours.

We also argue that WCBLE does not only enhance the P3 signal, but effectively reduces the negative impact of noise that would otherwise mask the ERP response, thereby increasing the discriminative ability of the decoder. This process likely results in more reliable feature extraction, especially in conditions with high jitter.

We found that CBLE did not improve gaze-independent decoding performance significantly and also did not increase performance over its first-stage classifier in overt VSA, contrary to what was reported by Mowla, Huggins, and Thompson [177]. While they report that CBLE is relatively independent of the first-stage classifiers evaluated in their work, it is evident here that applying any given classifier in the CBLE scheme does not necessarily increase its performance. In our case, this could be due to characteristics of the tested dataset, e.g., the presence of jitter or the generally higher performance of tLDA compared to the first-stage classifiers tested by Mowla, Huggins, and Thompson [177], leaving less performance to be gained.

Thompson, Warschausky, and Huggins [256] already attempted applying CBLE in an iterative scheme but did not report any results due to convergence issues. We mitigated this by combining the robust latency estimation presented in section 5.3.2 with the alignment of both target and non-target epochs. Aligning only the target epochs containing the jittered P3 component is prone to overfitting by aligning non-discriminative noise that is present in both classes, such as environment noise, oscillatory background rhythms, or non-modulated VEPs. If SNR is low, residual noise varying slightly between classes could dominate the expected response of the first-stage classifier and subsequently dominate WCBLE from the start, preventing convergence to a meaningful solution. Our procedure circumvents this problem by aligning both classes to the time points where the expected separation between classes is greatest. This way, noise of which the latencies are estimated in a given iteration will be perfectly time-locked in all classes in the next iteration. The first-stage classifier can then more easily suppress this noise since it is now clear it is present in both classes and non-discriminative. This aids the method in converging to a more robust classifier by iteratively increasing SNR for both classes and class separation over time.

Zisk et al. [317] addressed P3 jitter inparticipants ALS by augmenting the training data once with time-shifted copies based on CBLE-estimated jitter. While we aim to train the first-stage classifier without the effects of jitter in its parameters, they do the opposite by intentionally jittering the training data. As their focus was on ALS and inter-session stability, they did not assess how their method interacts with visuospatial attention. We achieved a similar performance gain with our jitter compensation method, but argue that our method can cope with more granular latency differences, as Zisk et al. [317] augment the data with just one positive and negative time shift.

Hardiansyah, Pergher, and Van Hulle [100] decoded covert VSA more effectively by contributing single-trial latency and amplitude features for decoding. Contrary to our approach, they did not correct these amplitude features for

the jitter in their latencies by, e.g., aligning trials to achieve better separability. Hence, their approach would not, in principle, render the classifier more robust to jitter. Furthermore, we incorporated estimated latency features in both CBLE and WCBLE, yet only WCBLE improved covert VSA performance. This shows that the incorporation of latency features is not the only driver of covert VSA decoding performance increase.

Despite encouraging results, our study faces some limitations that we plan to tackle in the future. Firstly, multiple ERP components can be time-locked to different neural processes, each with its own jitter, hampering the performance of single-trial latency estimation and their interpretability. Adaptations could be made to incorporate prior time windows or probability distributions on the latency of specific ERP components or to simultaneously estimate a set of multiple component (clusters) latencies per ERP, such as in Residue Iteration Decomposition [196]. Future efforts should investigate how strong spatiotemporal filtering can be combined with methods that allow for a more flexible non-stationarity of the ERP, like Dynamic Time Warping (DTW) or other techniques borrowed from time series classification, or methods that explicitly model multiple time displacements present in one ERP.

Secondly, performance might be improved by venturing beyond the classical target/non-target binary classification problem. Due to, for example, the perifoveal stimulus cruciform model [275], covert and split VSA responses might differ based on the relative position in the field of view of their related stimulus, which could be exploited in a multi-class classification problem. Similarly, explicitly taking into account the characteristics of the distractor ERP response might have a beneficial effect. Thirdly, results were obtained in an offline and within-session evaluation, which does not reflect true BCI operation. Using multiple sessions with online feedback, the user could optimize their performance over time by controlling attention or gaze. Finally, since this work was conducted with applications for individuals with severe speech, physical and gaze impairment (SSPGI) in mind, we should highlight that the gaze of participants in the conducted experiments was cued and fixed, which is per definition impossible for the end user group we consider. In chapter 7, we present a study to further investigate whether the studied VSA conditions are appropriate and to what extent they occur when individuals with SSPGI operate a BCI.

6.5 Conclusion

Our aim was to improve gaze-independent BCI performance for spatially organized visual event-related potential (ERP) paradigms by using a suitable decoder. Earlier results on BCI performance in covert visuospatial attention (VSA) performance prediction have shown that accounting for single-trial latency jitter could improve gaze-independent decoding performance. We applied Classifier-based Latency Estimation (CBLE) as a decoder robust to latency jitter, but found no increase in gaze-independent decoding performance. To remedy this, we improved CBLE by adapting it into CBLE with Woody iterations (WCBLE),

an iterative scheme using probabilistic latency estimation. Results for WCBLE within and across VSA condition decoding show that gaze-independent BCI performance can be improved at the decoding stage. Overt decoding performance was not improved, but our proposed method can provide added value for users who are unable to operate a visual BCI in overt attention mode. Later studies should confirm whether our findings hold in individuals with severe physical impairment and a variety of eye-motor impairments, and develop a solution that is capable of properly handling multiple non-time-locked ERP components.

6.A Single-trial classification performances

dataset	BNCI2014-009		CVSA-ERP				
VSA condition	overt	covert	overt	covert	split ($d = 1$)	split ($d = 2$)	split ($d = 3$)
ERPCov-TS-LR	90.00	74.50	79.56	68.80	66.90	72.52	72.14
XDAWNCov-TS-LR	90.44	74.64	81.13	67.80	66.83	72.59	71.36
tLDA	94.74	78.90	**86.60**	71.11	**70.97**	75.50	74.74
CBLE	**94.97**	79.28	85.59	69.77	70.87	74.67	74.65
WCBLE	94.28	**80.84**	85.25	**73.38**	70.84	**75.81**	**76.87**

Table 6.1: Cross-validated single-trial classification ROC-AUC (%) for all evaluated models, visuospatial attention conditions, and datasets.

6.B Cross-condition transfer classification performances

classifier	test / train	overt	covert	split ($d = 1$)	split ($d = 2$)	split ($d = 3$)	combined
CBLE	overt	85.59	66.68	63.92	60.22	60.11	66.25
	covert	68.51	69.77	70.52	68.50	70.29	75.08
	split ($d = 1$)	64.11	67.81	70.87	69.60	69.44	75.04
	split ($d = 2$)	60.48	67.59	69.99	74.67	72.88	75.96
	split ($d = 3$)	59.32	69.66	68.51	72.95	74.65	75.18
	combined	68.77	73.71	74.36	76.41	75.14	72.73
tLDA	overt	86.60	67.40	64.15	59.80	59.39	66.85
	covert	68.78	71.11	70.17	69.25	70.64	74.69
	split ($d = 1$)	65.72	68.00	70.97	70.53	69.52	75.07
	split ($d = 2$)	60.69	68.40	70.48	75.50	72.69	76.75
	split ($d = 3$)	59.37	69.10	69.40	72.56	74.74	75.61
	combined	69.71	73.64	74.81	76.01	75.68	73.78
WCBLE	overt	85.25	68.80	66.07	63.02	62.22	68.60
	covert	72.75	73.38	72.02	70.16	72.06	76.30
	split ($d = 1$)	68.07	70.04	70.84	72.42	71.05	75.56
	split ($d = 2$)	64.53	70.84	72.07	75.81	73.97	77.70
	split ($d = 3$)	63.79	72.21	70.57	72.80	76.87	76.05
	combined	71.09	73.36	74.53	76.07	76.26	75.15

Table 6.2: Cross-condition transfer classification performance (cross-validated ROC-AUC) for all evaluated models in the CVSA-ERP dataset. Each decoder is trained on one VSA condition and tested across all VSA conditions to evaluate gaze-independence and transfer learning performance.

classifier	test train	overt	covert	combined
CBLE	overt	94.97	63.59	83.13
	covert	67.00	79.28	73.01
	combined	93.81	73.36	83.17
tLDA	overt	94.74	62.22	82.70
	covert	67.34	78.90	73.51
	combined	94.01	72.55	83.16
WCBLE	overt	94.28	65.04	82.51
	covert	69.40	80.84	75.87
	combined	92.77	76.10	84.20

Table 6.3: Cross-condition transfer classification performance (cross-validated ROC-AUC) for all evaluated models in the CVSA-ERP dataset. Each decoder is trained on one VSA condition and tested across all VSA conditions to evaluate gaze-independence and transfer learning performance.

Chapter 7

Case studies

7.1 Introduction

Brain-computer interface (BCI) assistive technologies for communication [167] target individuals with severe speech and physical impairment (SSPI) [209]. Visual BCIs, which rely on the interpretation of visual stimuli by the user, offer several advantages in this context. They can work with non-invasive recording technology and can use rapid stimulation. This makes them well-suited for real-time communication tasks.

Yet, there is a large comorbidity between SSPI and eye motor impairment [79]. Impairments such as nystagmus (uncontrolled eye movements), diplopia (double vision), and ophthalmoplegia (eye paralysis) can significantly hinder the ability to use visual BCIs. These impairments make it difficult for BCI users to focus on or track visual stimuli accurately, reducing their performance with BCIs that rely on visual cues [162, 79, 203]. Unfortunately, it is again for this group that eye tracking solutions also perform poorly, making them more reliant on potential developments in BCI that do not rely on eye gaze.

Eye motor impairments are presumed to reduce performance in operating visual oddball BCIs (see chapter 2 for an overview), since users cannot comfortably redirect their gaze at the desired target, i.e., perform overt visuospatial attention (VSA). This is usually circumvented by designing gaze-independent BCIs [223]. These interfaces either avoid visual stimulation or exploit some form of covert VSA, where the gaze and VSA do not coincide.

Several studies with visual oddball BCIs show that performance drops when not fixating the intended target [38, 264, 226], necessitating gaze-independent solutions. These studies build on the assumption that BCI users with severe speech, physical and gaze impairment (SSPGI) would feel comfortable operating an interface in pure covert VSA with central fixation. One could argue that a BCI that is only verified to work when central fixation is maintained could also be considered gaze-dependent. This does not account for the residual eye motor capabilities of most people with SSPI, the (dis)comfort they experience while

performing gaze fixation and other confounding factors resulting from their eye motility.

It is a striking constatation that studies reporting on gaze-independent visual BCI with people with SSPI and eye-motor impaired are very few. Results are usually different from those obtained with healthy control participants in the lab, due to difference in capabilities, brain response, equipment and environment.

Lesenfants et al. [142] tested a BCI using gaze-independent steady-state visually evoked potential (SSVEP) in six participants with Locked-in Syndrome (LiS) yet only exceeded chance level accuracy in two. More recently, Peters et al. [210] performed a trial with two participants with late-stage Amyotrophic Lateral Sclerosis (ALS) and visual impairment. Their SSVEP interface was not optimized for gaze-independence, but the system showed high accuracy, outperforming an eye tracking alternative. It would be of interest to verify if such results can be replicated with participants with other conditions, and with a visual oddball BCIs.

Orhan et al. [192] and Oken et al. [190] tested the rapid serial visual presentation (RSVP) speller with individuals with LiS.

Severens et al. [239] evaluated the visual Hex-o-Spell [264] on 5 participants with ALS and showed that this visual oddball interface optimized for gaze-independence can outperform a tactile BCI. While this speaks to the power of visual paradigms even in groups that are expected to have eye motor impairment, they did not verify the gaze direction of participants during the experiment. It was suggested that participants were performing overtly. Participants with ALS also had a substantially lower accuracy than healthy controls (58% vs. 88%).

Our previous study, presented in chapter 6 also used the visual Hex-o-Spell interface [272]. This work partially accounted for the idea that BCI users with SSPGI might not fully rely on central gaze fixation and evaluated settings that are not strictly dependent on this. We showed gaze-independent performance can be improved in healthy subjects using a suited decoding strategy that accounts for latency jitter in covert VSA responses. Yet, there is a strong need for verification of these results in an applied setting with people with SSPI.

Eventually, one of the end goals of this research line is to develop gaze-independent BCI for people that are fully locked-in and have no option left than to use a BCI. However, this group is very small and it is often a challenge to recruit them into a study and perform experiments with them [297]. Individuals with less severe paralysis or in less progressed disease stages that struggle with eye-tracking technology could also benefit from solutions tailored to their specific situation. Therefore, we aim to apply the concepts from earlier work and literature to people with SSPI and various degrees of motor impairment in a visual oddball BCI. The objectives of this case study are as follows: 1. Explore capabilities and experienced comfort of individuals with SSPGI, when operating a visual BCI, 2. evaluate the performance of a gaze-independent visual BCI for this group, 3. verify if this performance can be improved with a suitable decoding strategy.

7.2 Materials & methods

7.2.1 Recruitment

Participants were recruited across the Neuromuscular Reference center at University Hospital Leuven (Leuven, Belgium), TRAINM Neuro Rehab Clinics (Antwerp, Belgium), the Neurorehabilitation Unit at University Hospital Lille (Lille, France), and a specialized care home (France). Experiments were performed under the supervision of their treating physician. Participants were recruited based on the following criteria. To qualify for inclusion, participants must:

1. be at least 18 years old and no older than 60 years,
2. belong to class 2 or 3 according to the BCI user selection criteria presented by Wolpaw et al. [297],
3. have limitations to the extent or comfort of their eye motor control

Participants were excluded if they:

1. had a diagnosis of a major medical condition, including any major neurological or psychiatric disorder other than those of interest based on inclusion criteria 2, and 3
2. had a predisposition to or a history of any kind of epileptic seizures, including photosensitive epilepsy,
3. had a severe loss in vision or hearing that would significantly impair participation in the experiment,
4. are currently using specific psychoactive medications or substances that could affect the outcome.
5. were unable to understand the experiment instructions and cooperate,
6. had any other limitations preventing them from performing the given task.

In total, 11 individuals were contacted. Of these, one person with Multiple Sclerosis (MS) was excluded based on criterion 3. One person recovering from traumatic brain injury (TBI) was excluded based on both 2 and 4, and one person recovering from stroke based on 1. One further person recovering from a stroke was excluded due to technical difficulties during the experimental session. Vision was assessed using a LogMAR chart [17].

Ultimately, 7 participants were retained. Of these, one participant was diagnosed with bulbar-onset ALS, three with Friedreich's Ataxia (FRDA) an three were recovering from stroke. We refer to section 2.1.1 for a short description of these conditions. Table 7.1 lists the included participants and their diagnoses.

ID	Diagnosis	Age	Sex	Hand.	Speech	Trach.	Communication	W	KB
PA1	bulbar-onset ALS	58	M	L	anarthric	no	tablet	3	4
PB1	FRDA	41	M	L	dysarthric	no	verbal	3	3
PB2	FRDA	43	F	R	dysarthric	no	verbal	3	3
PB4	FRDA	48	M	R	dysarthric	no	verbal	3	3
PC2	brainstem stroke	43	M	R	anarthric	yes	prompting +eye movement	2	4
PC3	brainstem stroke	43	F	R	anarthric	yes	letterboard	2	3
PC4	left cerebellar stroke (trombosis of the basilar artery)	54	M	R	anarthric	yes	letterboard	2	3

Table 7.1: Included participants with their diagnosis and capabilities. Trach.: underwent a tracheostomy, W: classification according to Wolpaw et al. [297][1], KB: classification according to Kübler and Birbaumer [134][2].

7.2.2 Visual skills and eye tracking and eye motor examination

Self-reported eye motor and visual abnormalities were recorded according to the relevant visual BCI skills presented by Fried-Oken et al. [79]. These include visual acuity, visual fixation, eyelid function, ocular motility, binocular vision, and field of vision. Additionally, participants and their caregivers were asked about eye tremors (nystagmus or other) and other involuntary eye movements.

As an objective metric, we implemented and performed the automated NeuroEye eye movement test proposed by Hassan et al. [103] using calibration-free eye tracking to check if it revealed any further eye motor abnormalities. This was not the case.

Finally, we also recorded gaze position throughout the experimental session to register the participant's gaze relative to the stimulated BCI targets.

7.2.3 BCI stimulation

The BCI stimulation procedure was based on the Hex-o-Spell [264] implementation presented by Van Den Kerchove et al. [272]. Similar to this study, the task consists of counting the flashes of a cued target among 6 round, flashing targets laid out in a hexagonal pattern in the field of view of the user. We refer to section 6.2.1 for implementation details.

Three different VSA settings were explored. In the overt VSA setting, the participant was instructed to fixate on the cued target or try to the maximum extent of their visual skill, even if experiencing slight discomfort. In the covert

[1] See page 36 footnote 6.

[2] "With minor degree of impairment, we refer to patients who had only slightly impaired limb movement and normal speech. Under the category moderate impairment, we summarized those patients with restricted limb movement (wheelchair-bound) and unaffected speech or intact limb movement without speech. [...]Patients who were almost tetraplegic with restricted speech were considered majorly impaired. Categories four and five were the LIS and the CLIS, respectively." [134]

	PA1	PB1	PB2	PB4	PC2	PC3	PC4
Visual fixation	✕	✕	✕	✕	✕	✕	✕
Eyelid function						✕	✕
Ocular motility		✕		✕	◇	◇	✕
Binocular vision					✕	◇	◇
Field of vision						✕	✕
Involuntary movement		✕	◇	✕	✕	✕	
Visual acuity (logMAR)	0.0	0.0	0.6	0.2	0.0	0.7	0.6

Table 7.2: Visual skills of the included participants. Visual BCI skills [79] were assessed with a combination of self-reported issues by the subject and the NeuroEye [103] test. ✕ impaired, ◇ severely impaired. logMAR: lower is better.

VSA setting, the participant was instructed to fixate on the center of the screen, to the extent of their ability. An additional *free VSA* setting was introduced. Here, the participant was instructed to perform the task as they deemed most comfortable. This allowed us to investigate the user's natural way of operating the BCI given their individual set of visual skills. If the participant was not fully paralyzed, they were instructed not to move their head. The cued *split attention* setting proposed by Van Den Kerchove et al. [272] was not studied here, as we were interested in natural VSA operation settings for gaze-impaired individuals.

To make the interface suitable for use by individuals with SSPI [79], the number of blocks was decreased to 6 per VSA setting. Inter-stimulus interval (ISI) was increased to 200±50 ms to decrease task difficulty. The experiment also started with a training block in each condition, where the participant was instructed with feedback on their performance to ensure they understood and were able to perform the task.

7.2.4 Data collection & preprocessing

During the recording session, participants were positioned in their wheelchair in front of a table. Stimuli were presented on an Acer Predator Helios laptop with an 18" screen (Acer, Inc., Taiwan) placed at a 60 cm distance. A Cedrus Stim-Tracker (Cedrus Corp., CA, USA) ensured synchronization of stimuli with the recorded electroencephalography (EEG). Eye tracking was performed throughout using the Tobii X2-30 Compact (Tobii Technology AB, Sweden) portable

Figure 7.1: A participant with the stimulation and recording setup.

eye tracker placed at the bottom of the laptop screen.

EEG was recorded at 1000 Hz using the Neuroscan Neuvo portable amplifier (Compumedics Neuroscan, Australia) connected to a second laptop for registration. The EEG headset used 18 active AgCl electrodes (EASYCAP GmbH, Germany) placed on a cap according to the international 10-20 layout. Using electrolyte gel, electrode impedances were reduced below 10 kΩ. Additionally, the electrooculogram (EOG) was recorded.

The EEG was band-pass filtered between 0.5 and 16 Hz. Bad channels were rejected using the RANSAC algorithm [74] and visual inspection. Next, the EEG was re-referenced to the average of mastoid electrodes TP9 and TP10, and independent component analysis (ICA) was performed to reject artifactual components based on correlation with the EOG or by visual inspection. Epochs were cut from -0.1 to 0.9 s relative to stimulus onset, and no baseline correction was performed in order to meet the assumptions.

Eye tracking data was cleaned by fusing left and right gaze into one channel for the horizontal and vertical gaze position. If both were present for a given sample, the fused channel was the mean of both values. If at a given sample either the left or the right eye was not detected for a given channel, the value of the other one was adopted. If both were missing, the gaze position remained unset at that time point, and no interpolation was performed of the employed classifiers.

7.2.5 BCI decoding

We evaluated the recorded data using the Classifier-based Latency Estimation with Woody iterations (WCBLE) [272] and block-Toeplitz linear discriminant analysis (tLDA) [244] classifiers, as well as the Riemannian approach XDAWN-Cov+TS+LDA [46]. For WCBLE, a region of interest from 0 ms to 800 ms relative to stimulus onset was used while the epoch was cropped to -100 ms to

900 ms. For the other decoders, the epoch was cropped between 0 ms and 800 ms, which resulted in maximal performance. Decoding scores were obtained using 6-fold cross-validation where folds corresponded to stimulation blocks.

7.3 Results

7.3.1 Visual skill and eye tracking analysis

Table 7.2 details the eye motor impairments and vision of the included participants. All participants reported some degree of fatigue or discomfort when fixating. Participant PA1 had the mildest impairment, only reporting fatigue when fixating for prolonged times. The FRDA participants were mostly affected by eye tremors and impaired pursuit. PB2 suffered from especially severe horizontal oscillating involuntary eye movements. Eye motor function of participants PC2, PC3, and PC4 was most severely affected. Participant PC2 was only able to look up and down and had a deviation in the left eye causing diplopia, but this was corrected by a prism glass. Participant PC3 only retained partial motility of the right eye, while the left eye was permanently closed. Participant PC4 had one deviated eye with a corneal abscess affecting the motility and vision in the right eye, and reducing motility in the left.

Given these information, we aimed to shed more light on the actual capabilities of individuals with SSPGI regarding performing overt VSA and central gaze fixation, as well as to investigate how relevant these two settings are when the gaze is not cued. Figure 7.2 maps gaze position relative to the stimuli across conditions. These results should be interpreted with care, as the eye tracker to some degree relies on functioning eye motility. The participant's position relative to the eye tracker might have shifted throughout the experimental session despite our best efforts, e.g., because they needed aspiration of their tracheostomy.

PA1 had relatively intact gaze control and was able to correctly perform the cued overt and covert settings. When gaze was uncued, he fixated on the cued target. This was also mostly the case for PB1, although eye tracking revealed that he chose not to perform central gaze fixation when cued in at least one of the stimulation blocks. We were unable to record his gaze near the bottom-left stimulus position, either due to eye tracker failure or because the participant was not comfortable fixating on this position. Eye tracker calibration did not succeed for subject PB4, but given transformation of gaze positions to the stimulus space, they were assumed to be overtly performing the free task.

PB2 was able to perform overt VSA and central fixation to some extent, yet eye tracking shows a larger spread in gaze position compared to PA1 and PB1. In the free VSA condition, however, she preferred to attend stimuli covertly when the gaze was uncued. This was confirmed by the participant.

The overt and central gaze fixation settings were also not properly adapted to participant PC4. In the free VSA condition, eye tracker results show that his gaze was usually near the bottom two targets, indicating some degree of covert or split VSA.

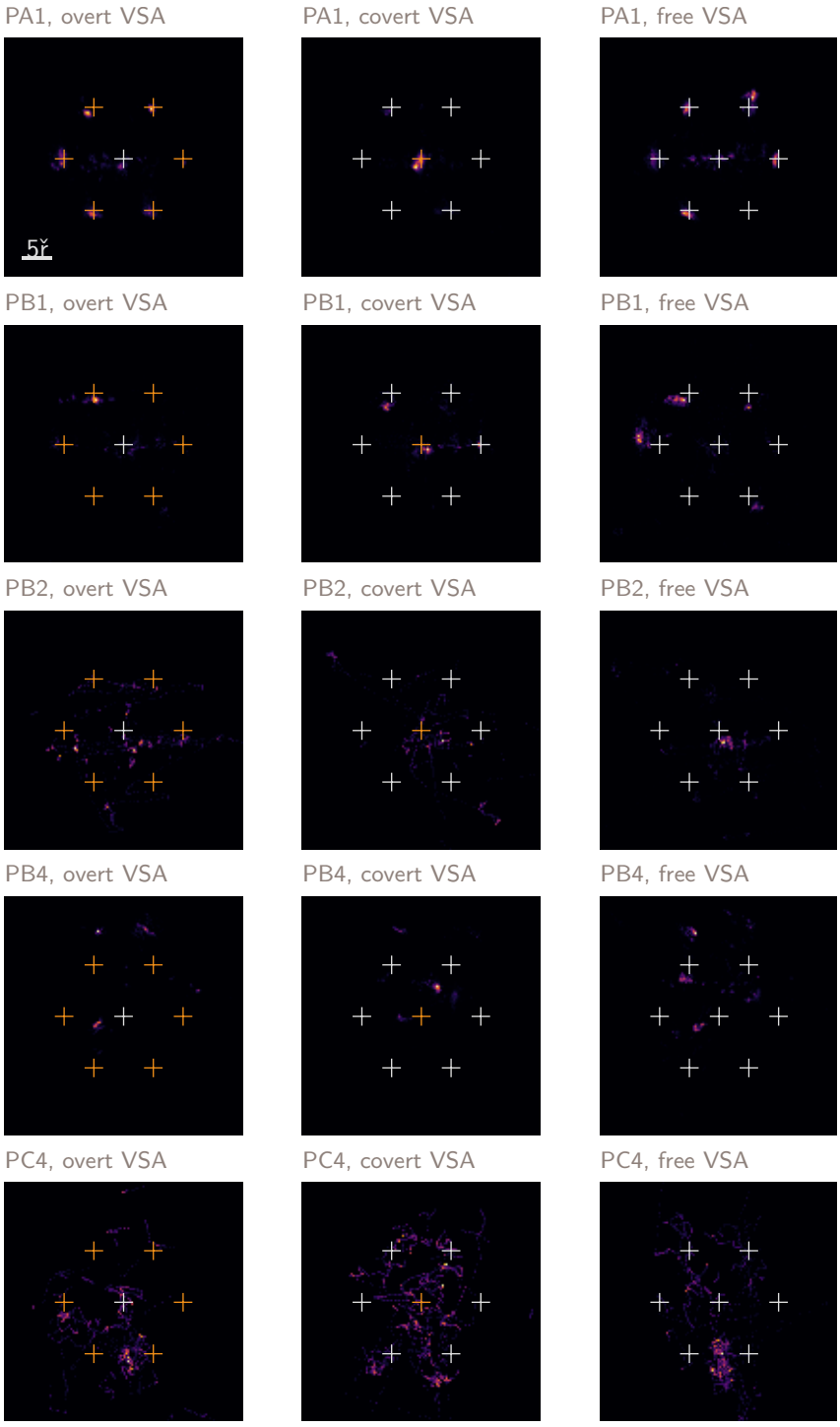


Figure 7.2: Distribution of the recorded gaze position during the experimental session in the three VSA conditions. Crosshairs represent stimulus positions, with the orange ones cued during the given condition. Subjects PB2 and PC4 preferred covert BCI operation, with PB2 resting gaze near the middle of the screen, and PC4 near the bottom.

It was technically impossible to register gaze position with the Tobii X2-30 Compact for participants PC2 and PC3 since they both had one eye that was occluded respectively by the prism glass and the eyelid. Both participants reported they could not fixate on some of the stimuli.

7.3.2 BCI decoding performance

Figure 7.3 shows single-trial area under the receiver-operator characteristic curve (ROC-AUC) in the evaluated VSA settings for the different decoders.

In the overt VSA setting, the evaluated decoders performed similarly on average (WCBLE 75.58%, XDAWNCov+TS+LDA 74.24%, tLDA 75.99%). In the covert VSA setting with cued central gaze fixation, performance deteriorated, but WCBLE significantly improved performance over the base classifier tLDA in this condition (WCBLE 62.49%, XDAWNCov+TS+LDA 59.42%, tLDA 59.05%). Decoding performance for this task was at chance level for participants PB4 and PC3.

However, WCBLE did not improve tLDA performance in the free VSA setting, but XDAWNCov+TS+LDA performance was slightly lower here (though not significantly). (WCBLE 74.15%, XDAWNCov+TS+LDA 71.88%, tLDA 74.27%). More interestingly, we noticed that performances of the decoders in free VSA were close to those in the overt VSA. A substantial decrease in performance from the overt setting to the free setting was observed for subjects PC3 (WCBLE: 70.31>62.14 %, XDAWNCov+TS+LDA: 65.78>62.18 %, tLDA: 70.49>63.76 %) and PC4 (WCBLE: 65.56>55.71 %, XDAWNCov+TS+LDA: 62.02>54.24 %, tLDA: 66.12>57.08 %). For PB2, who also relied on covert VSA during the uncued free VSA according to gaze tracking setting, the decrease in performance was also present, but not as substantial (WCBLE: 82.76>78.88 %, XDAWNCov+TS+LDA: 80.74>77.99 %, tLDA: 83.21>78.84 %).

7.3.3 Cross-condition calibration

As an alternative approach to selecting the most suitable decoder, we used tLDA as the base decoder and verified whether performance could be improved if BCI users with gaze impairment performed the calibration session relying maximally on their residual gaze control.

Figure 7.4 shows that, on average, covert VSA decoding improved when training with overt VSA. This was especially true for participants PA1, PB2, and PC3. Note that, according to eye tracking data, participants PB1, PB4, and PC4 did not always perform cued central gaze fixation in the covert VSA setting, which might have affected the results.

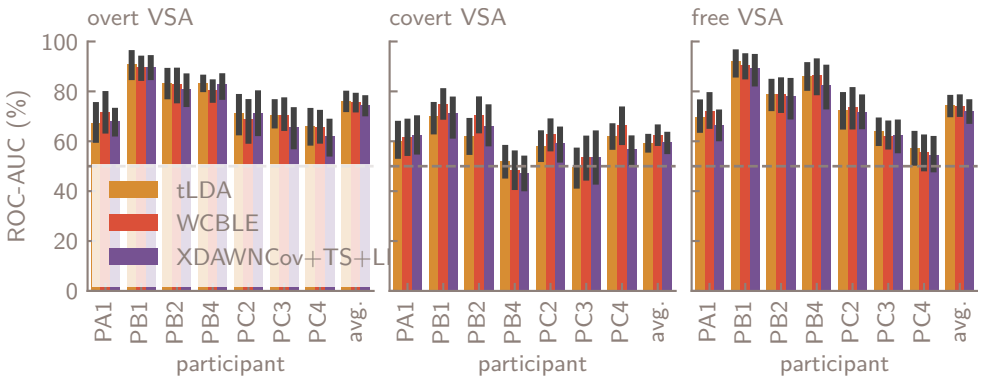


Figure 7.3: Decoding performance in different VSA settings reported as single-trial ROC-AUC. Free VSA is generally on par with performance in the overt VSA setting. Performance in the covert VSA setting with central gaze fixation is lower, but can be improved with the WCBLE decoder. 95% confidence intervals were calculated using 1000 bootstrapping repetitions.

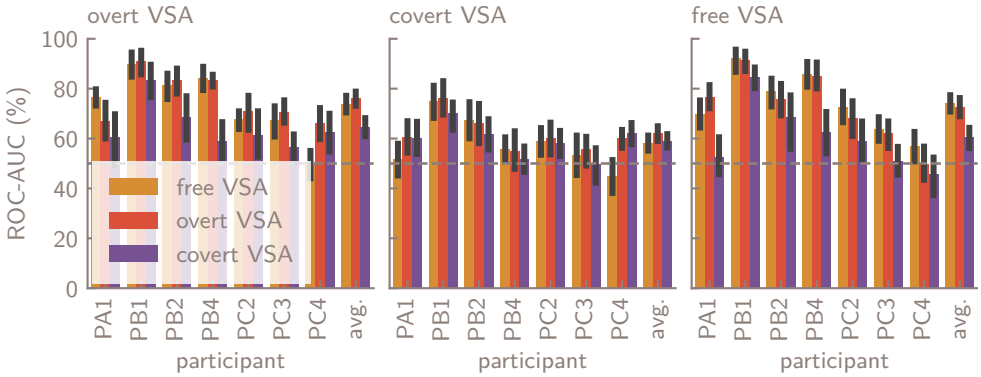


Figure 7.4: Decoding performance when calibrating the tLDA decoder in a given VSA setting, and evaluating it in another, reported as single-trial ROC-AUC. For participant PA1, free VSA performance improved when calibrated with overt gaze fixation. For participants PA1, PB2, and PC3, performance in the covert VSA setting with central gaze fixation improved when calibrating with overt gaze fixation.

7.4 Discussion

7.4.1 Gaze-independent operation & decoding

Due to the heterogeneous nature of the participants' conditions, it is difficult to draw general conclusions. This study should therefore be seen as a collection of case studies, highlighting different obstacles encountered in developing gaze-independent visual oddball BCIs for individuals with SSPGI. Nevertheless, we would like to highlight some aspects that might be of interest for the further development of this class of BCIs.

True gaze-independent visual BCIs should not rely on gaze fixation. Hence, our analysis centers around the free VSA condition. Eye tracking results presented in section 7.3.1 confirm our assumption that voluntary covert VSA can occur in individuals with SSPGI. We also confirmed part of the results from Van Den Kerchove et al. [272] presented in chapter 6, which state that decoding of covert VSA with central gaze fixation can be improved by accounting for latency jitter. We showed that this also holds for individuals with SSPGI.

Contrary to our assumptions, however, we have shown that this does not necessarily improve covert VSA when gaze fixation is not cued. One possible explanation is that actively performing central gaze fixation increases task load. This, in turn, can reduce overall performance, even though the participant might have otherwise performed covert VSA, but would not be occupied with maintaining strict central gaze fixation. This extra task demand is not present in the free VSA condition, so there is less performance to be gained. Furthermore, cued central gaze fixation combined with counting flashing stimuli in the visual periphery is an explicit example of a dual task. Dual tasks have been shown to increase P3 latency jitter [215, 12, 272], which is what WCBLE accounts for. Hence, increased P3 jitter might be more related to maintaining central gaze fixation than to the actual covert VSA aspect.

The seemingly stable performance across overt and free VSA could be misinterpreted as an indication that the Hex-o-Spell BCI already works well for individuals with SSPGI, and no optimization is needed. However, we assume that overt VSA performance was also decreased in some subjects or for some blocks if the participant was not able to comfortably perform the task. Nevertheless, the large difference between the free VSA setting and the covert VSA setting with central gaze fixation is food for thought about the applicability of solutions developed with central fixation in mind.

Individuals with all but the most severe gaze impairments will likely retain some degree of gaze direction in visual BCI operation, which can drastically boost performance. Subject PB2 exemplifies this: his free VSA performance is on par with his overt VSA performance, although eye tracking showed that he relied mostly on overt VSA when cued to do so, and mostly on covert VSA when gaze was uncued. This is also supported by our results on cross-condition calibration presented in section 7.3.3, which show that leveraging residual eye motor control to fixate targets during the calibration phase can improve performance in some settings. This is likely due to the increased P3 component

amplitude in overt VSA, which improves the discriminative power of a classifier trained on this data. Cueing this overt gaze fixation only during the calibration phase leaves the user free to operate in the manner that is most comfortable for them in the operation phase. Early VEPs in the training data could also contribute in those cases where the participant was not able to perform covert VSA with central gaze fixation.

7.4.2 Limitations

Despite results that prompt interesting reflections on gaze-independent BCI approaches, there are some limitations to the presented results that need to be addressed in current and future work.

First and foremost, this study works with a limited sample size, which does not represent the full spectrum of individuals with SSPI and SSPGI, and their specific symptoms and skills. Individuals with FRDA met the inclusion criteria, but they are usually not considered one of the typical interest groups for BCI communication assistive technology, partly due to the rarity of the disease and partly due to its progression. It would be most interesting to verify these results with individuals with LiS and no eye movement capability at all.

Another limiting factor is the difficulty experienced in correctly interpreting eye tracker results in studies with individuals with gaze impairments. If eye tracking is possible at all, it is not guaranteed that the user is able to successfully perform the calibration procedure. Further experiments should be carried out with a stationary eye tracker with more advanced capabilities, although systems using a head fixator or headrest should be avoided. This is not practical when working with wheelchair-bound individuals who might have undergone a tracheostomy and may suffer from spasticity.

In this study, user comfort in the different conditions was not objectively measured. Instead, it was assumed that participants operated most comfortably in the free VSA condition. To properly contextualize performance results, they should be coupled with metrics evaluating the user's requirements with a measure of usability, comfort and perceived effort, like the NASA Task Load Index [101] and other metrics proposed in the user-centered design framework for BCIs [135]. Performance might, after all, be traded off for user comfort. Eye motor disability could also have been assessed more objectively [79], using, e.g., the Revised Coma Recovery Scale [84] or the NSUCO oculomotor exam [160].

Finally, the stimulation procedure parameters from Van Den Kerchove et al. [272] were adapted to make the counting task accessible to the BCI users with SSPGI. However, the number of repetitions and ISI were not optimized to achieve maximal information transfer rate (ITR). An interface that aims to maximize ITR could necessitate more and faster gaze redirections, which might result in different conclusions regarding the comfort and the effect of visual skill.

Chapter 8

Conclusion & recommendations

Section 8.1.1 was published as part of Van Den Kerchove et al. [271].

8.1 Contributions

This work aimed to improve the performance of visual gaze-independent BCIs, in general and when applied to individuals with eye motor impairment, by developing novel decoding strategies and evaluating their effectiveness. We addressed the limitations of current gaze-dependent BCIs by proposing methods that exploit covert visuospatial attention (VSA) and reduce the reliance on eye gaze. The following key contributions were made:

8.1.1 Developed ERP decoders

We introduced a covariance estimator using adaptive shrinkage (STBF-shrunk) and an estimator exploiting prior knowledge about the spatiotemporal nature of the EEG signal (STBF-struct). We compared these estimators with the original formulation of the spatiotemporal (STBF-emp) beamformer and a state-of-the-art Riemannian Geometry method (XDAWN+RG) in an off-line P3 detection task on an existing dataset. Our results show that the structured estimator results in an accuracy increase of up to 4 %. compared to shrinkage regularization. y when training data are sparsely available. Results can be computed faster and with substantially less memory usage. Since these algorithms are not paradigm-specific, the conclusions can be generalized to other ERP-based BCI settings. These results have been published in Van Den Kerchove et al. [271].

Next, Block-Term Tensor Discriminant Analysis (BTTDA) was introduced as a tensor-based decoder that better captures the multidimensional nature of ERP data. By preserving the structure of neural data, BTTDA can cope with noise and other challenges arising from gaze-independent BCIs, yielding robust

results under certain conditions. Results show that BTTDA and its special sum-of-rank-one structured case improved over Higher Order Discriminant Analysis (HODA) and can reach state-of-the-art decoding performance for ERPs. This work is submitted as Van Den Kerchove et al. [270].

Finally, we developed the Classifier-based Latency Estimation with Woody iterations (WCBLE) decoder, specifically designed to address the challenges posed by P3 latency jitter in covert VSA settings. Latency variability is a well-known issue in ERP decoding, particularly when users engage in covert attention, where the timing of the P3 component fluctuates significantly across trials. Traditional decoders struggle in these settings due to the inconsistent timing of brain responses, which can reduce the robustness and accuracy of ERP decoding. The WCBLE decoder mitigates this issue by introducing an iterative process to estimate and align ERP latencies across trials. This is performed in such a way that discriminative power of the ERP signal is preserved and enhanced. The method has been published in Van Den Kerchove et al. [272]. We designed this method to improve decoding where unseen incoming test data is not yet known. WCBLE was first tested on synthetic data to evaluate its effectiveness under conditions of controlled latency variability and noise. These results showed that it is robust to higher noise and jitter compared to a non-iterative method in a latency estimation task. Decoding accuracy was higher in high jitter and noise settings compared to the non-iterative method and a state-of-the-art decoder.

8.1.2 Gathered datasets

The CVSA-ERP dataset consists of recordings of 15 healthy participants, mean age 26.38±3.15 years. The dataset was presented in Van Den Kerchove et al. [272]. The experiment in this dataset implemented a visual oddball BCI with six circular targets in a hexagonal layout. The gaze fixation of participants was carefully controlled to dissociate the visuospatial attention of the participant and their eye gaze. This allowed us to study the effects of gaze-independence on ERP decoding. The CVSA-ERP dataset gives us insight into the ERP dynamics in overt, covert, and the novel split VSA condition, and confirms our hypothesis that P3 jitter has a significant impact on performance in covert and split VSA. It also confirmed that the effects of covert VSA on ERP component amplitude hold for split VSA.

Additionally, we gathered data from 7 individuals with severe speech, physical and gaze impairment (SSPGI) with conditions such as Amyotrophic Lateral Sclerosis (ALS), Friedreich's Ataxia (FRDA) and stroke. These participants exhibited varying degrees of eye motor impairment, such as involuntary eye movements, ophthalmoplegia, and gaze fixation fatigue. The data provide invaluable real-world evidence on how gaze-independent BCIs perform in populations that experience eye motor difficulties. This dataset allowed us to evaluate the proposed decoding strategies in clinical settings and highlighted the practical challenges of implementing gaze-independent BCIs in these user groups.

8.1.3 Investigated gaze-independent visual BCIs

We evaluated our proposed WCBLE algorithm, its non-iterative counterpart, and state-of-the-art decoders on the CVSA-ERP dataset as well as on a publicly available dataset [8] that also contains the overt and covert VSA conditions. We evaluated the BCI decoding performance in a single-trial classification experiment, as well as in a target selection experiment reflecting BCI operation. Performance was significantly different between decoders, but this result was significantly dependent on the VSA condition and the dataset. While WCBLE was slightly outperformed by the state-of-the-art in overt VSA decoding (94.28<94.74 and 85.25<86.60 % ROC-AUC), it increased covert VSA decoding (80.84>78.90 and 73.38>71.11 % ROC-AUC).

For the split attention conditions in the CVSA-ERP dataset, WCBLE yielded a significant improvement over CBLE and the state-of-the-art method only in some cases. These results were corroborated by analyzing selection accuracy, which showed similar behavior for both datasets, except in overt VSA, where accuracy was not harmed by the lower single-trial selection performance of WCBLE.

To further study the gaze-independent performance of these algorithms, transfer learning between VSA conditions was studied to simulate conditions where an individual with gaze impairment can end up in different VSA settings within a BCI operation session due to their lack of proper motor control. When trained and evaluated on overt VSA data, our proposed WCBLE algorithm results in a small but significant decrease in performance compared to the state-of-the-art (85.25>86.60 and 94.74>94.28 % ROC-AUC), consistent with the within-condition results. For all other pairs of training and evaluation VSA conditions, however, WCBLE was equal to or significantly better with increases exceeding 4 % ROC-AUC.

Later case studies explored gaze-independent BCI performance in individuals with severe speech and physical impairment (SSPI) and varying eye motor impairments. Participants showed different levels of gaze control, with some using overt VSA, while others preferred covert VSA, resting their gaze and mentally attending to targets.

The WCBLE decoder improved decoding accuracy in covert VSA settings with cued central gaze fixation. This is of interest for gaze-independent decoding, but not as optimal as when the gaze fixation was uncued. The latter was tested in the free VSA setting. Here, participants were generally more comfortable when allowed to use their preferred gaze strategy. Decoding performance in free VSA was often comparable to overt VSA, showing the system's flexibility.

Revisiting the hypotheses put forward in section 2.3, we conclude that we have proposed several techniques that did, in effect, improve visual oddball ERP-based BCI decoding accuracy, and at least one technique that does so specifically under conditions relevant for gaze-indepent BCI. It was not conclusively proven, however, that these results have a meaningful impact on on-line BCI assistive technology use by individuals with SSPGI.

Together, the obtained results do show that there is an interest in developing

a new class of ERP-BCI interfaces for users that prefer to choose their own gaze strategy, to avoid the effort of redirecting their eye gaze to different spatial locations on the stimulation screen in manners that might be uncomfortable for them.

8.2 Current & future work

On the decoder development side, we have partially implemented some interesting extensions of the proposed methods. The Kronecker-structured beamformer can be generalized to the linear discriminant analysis (LDA) case. Furthermore, we obtained promising results when extending the single Kronecker product covariance model to a sum of Kronecker product terms, since the electroencephalography (EEG) covariance is probably better expressed by such a structure [27]. The Kronecker-LDA approach has also been extended to the combined space-time-frequency domain, where more information can be modeled at the cost of increased dimensionality. A properly structured covariance model can strongly regularize the problem.

To overcome the limitations introduced by the presence of multiple ERP components in latency estimation, we have also developed a multi-component version of the WCBLE algorithm. This algorithm should be able to separate mixed ERP component clusters based on their temporal coherence, in a similar manner to Ouyang et al. [196]. It could also yield favorable decoding results in a broader range of settings than overt VSA, since it should be able to account for the presence of both visual and attentional ERP components.

In gaze-independence decoding, our current efforts focus on – counterintuitively – integrating eye tracking into the decoding strategy. Since we are able to explicitly discern overt, covert, and split VSA from the ERP, and different decoders perform best in different settings, it could be helpful to derive the current setting from the gaze position and the ERP, and propose the most suited decoder for a given data point. This could allow us to select the best classifier using eye tracking.

Coincidentally, making advances in gaze-independent decoding also builds towards a solution for the Midas Touch Problem in visual BCI. Here, a BCI user sometimes accidentally selects a target while not intending to give any input. Decoding of true intention independent of eye gaze, with the option of gazing without paying visuospatial attention to a stimulus, would counteract this, and be a valuable addition to a BCI assistive technology device.

We also aim to apply the tensor approach to gaze-independent decoding. The problems introduced by jitter could be accounted for, as in WCBLE, by alignment, but also by other methods that model the possible time shifts. In ordering the data as a Hankel tensor[1], we should have a data representation that is more robust to jitter, at the cost of increased dimensionality. This increased dimensionality can then be countered by a tensor method like HODA or BTTDA.

[1] Adding an extra mode with time-shifted copies of the temporal response per channel.

Finally, we wish to expand our experimentation with individuals with SSPGI. As the work presented in this thesis started during the COVID-19 pandemic, experimentation was delayed and the envisioned application could not be validated within the time frame of the doctoral project. To complete this work, we aim to revisit the participants with an on-line experiment and satisfaction assessment to establish the usability of the proposed interface according to the principles of user-centered design (UCD) and compare its information transfer rate (ITR) to the state-of-the-art literature.

8.3 Limitations and recommendations

8.3.1 Decoding: keep it linear; structure it

Throughout this thesis, we noticed that linear or multilinear decoding methods consistently outperform non-linear methods, such as SVMs or those based on Riemannian Geometry. Within these linear models, regularization (shrinkage, covariance structure, tensor structure/rank) is of paramount importance, since neural data is inherently multivariate and BCI calibration time is minimized. Proper regularization should impose some specific structure on the linear model, ideally reflecting properties of the signal. When faced with a new problem and few data relative to dimensionality, it is a good first instinct to pick a simple, restrictive structure.

However, the brain is a complex organ, so neural data is not linear in origin. These regularizing assumptions based on prior knowledge are likely only superficially true. When optimizing for performance, it will soon become clear that, given a more flexible model, there are settings where these assumptions can be broken to better describe the ongoing interactions. For example, spatiotemporal beamforming or LDA with Kronecker-structured covariance improves performance in low sample sizes but is outperformed by more flexible methods, such as LDA using a Kronecker-sum covariance model or block-Toeplitz linear discriminant analysis (tLDA) when the training sample size increases. Similarly, BTTDA offers more flexibility than HODA.

This calls for the development of more flexible (multi-)linear models that still rely on some structure but are combined with efficient model selection, such as the methods proposed in conjunction with BTTDA. These models try to combine the best of both worlds.

8.3.2 Optimize user experience

While our work with BCI users with SSPGI resulted in interesting preliminary findings, it is currently impossible to make claims about the experience in long-term home use of the proposed systems as assistive technology. This work focused heavily on classification methods and solving problems arising from gaze-independence through decoding. Since these decoders were evaluated by collecting data and performing tests later using cross-validation, only off-line

performances have been reported.

To gain proper insight in the performance of a system that actually supports end-users, performance in an on-line experiment should be reported. The interface, now a collection of white circles, should be adapted to include meaningful selection elements, like groups of letters or icons and selection feedback should immediately be presented to the user. The performance metric then also captures effects of user engagement, learning and strategy adaptation. The measured brain activity and the proper execution of the task by the participant are, after all, partially dependent on their sense of completing a task that is tangible, entertaining and directly useful for them. The participant then becomes an active BCI assistive technology user, instead of just a near-passive observer in an abstract data collection experiment.

Furthermore, classification performance is only one aspect of what makes a BCI desirable as assistive technology. An interesting approach here is to consider the problem from a UCD [118] perspective. UCD is a framework for effective optimization and evaluation of usability in product design, which has successfully been applied to BCI development [235, 135, 99]. BCIs as assistive technologies for users with particular and individual physical and psychological requirements form a perfect use case for such a framework.

One of the main principles of UCD is to address the *entire* user experience. As mentioned before, this includes measuring *effectiveness*, in the form of e.g. selection accuracy. UCD goes beyond this by also assessing *efficiency* and user *satisfaction* Similarly to effectiveness, efficiency can objectively be measured as the ITR achieved in – importantly – on-line operation.

Satisfaction, on the other hand, is a more subjective quality. Nevertheless, optimizing user satisfaction is the penultimate goal in application design. Satisfaction depends on factors such as performance (in the form of efficiency and effectiveness), comfort, and user perception. It must be assessed through subjective questionnaires following realistic, on-line application use[135]. Only focusing on evaluating effectiveness through classification performance is a pragmatic approach, because a researcher can perform a single data collection and iterate on the results offline. Yet, this can be too limiting when seeking to design the most usable interface for a specific group of users.

The optimal approach should unify decoder development, choice of paradigm, and interface design, as advocated by Pan et al. [198] and Fouad et al. [77], by integrating both engineering and clinical perspectives. Performance and abilities likely depend to a great degree on the user, so off-line decoding performance in healthy controls is probably not a good predictor of on-line decoding performance and ITR in individuals with SSPI. Incremental gains in decoding performance for a specific paradigm might also be outweighed by how well that paradigm is adapted to the user, factoring in their skills and abilities. Longitudinal research should follow several user groups, specifically those with gaze impairment, and experiment with specific paradigms, interfaces, and decoders for a given BCI user.

8.3.3 Work *with* individuals with SSPI

A logical point arising from this is the question of how to set up an effective research project that aims to optimize BCI user experience, in the context of the topics presented in this work. Another valuable principle of UCD principle points the way: encouraging early and active involvement of end-users. This entails that, ideally, individuals with SSPI should be involved at every stage of the research project, from conceptualization to evaluation. A good rule of thumb is that research should solve problems *with* the users, not *for* them, implying that research projects should start from the needs of the BCI user population.

Instead of choosing a research target and then moving on to decoder development, data collection with healthy controls, patient studies, and finally, verification in on-line operation, it could be more effective to start the other way around. In a mature field like visual oddball BCI, the research project can, after a thorough literature review, immediately start with on-line experimentation involving individuals with SSPGI and an adequately chosen, existing BCI system. It should then become apparent where challenges for the envisioned user population lie. As an example, Fried-Oken et al. [79] share an overview of their experiences from this step, specifically concerning eye motor impairment in visual BCI. After that, hypotheses can be formed on how to mitigate these challenges through interface optimization or choice, taking into account the full picture of the capabilities and skills of the users. These hypotheses can then be verified in experiments with the users and/or a population of healthy controls. Finally, the decoder can be optimized, only if it is strongly suspected that effects will hold in on-line operation.

Unfortunately, this approach might also be one of the least practical ways to set up a 4-year research project, such as in the context of a doctoral thesis. It can only work if the proper facilities can be gathered immediately from the start, such as ethical approval, access to patients, and on-line BCIs experimentation systems. It is therefore important that research labs maintain a long-term working collaboration with patient centers, so that new projects can immediately be verified with the user population. This also allows interested users to compare multiple systems and guide development. The fact that a user has previous BCI experience is often seen as a confounding factor in off-line studies. When the research is in a phase that heavily depends on user experimentation, this is actually a strength, since the user can help the researcher gain insight into what works and what does not work for them, and where research efforts should be focused.

Similarly, there is value in maintaining a working, on-line, in-house BCI system that can be iterated upon and adapted if necessary. Given this system's design and data gathered from earlier use, experiments can be supported by advance calculations of hypothetical ITR, ensuring that on-line or off-line experiments reasonably reflect real-world operation, and no effort is wasted exploring settings that will not achieve satisfactory performance. For example, why study a specific number of repetitions or inter-stimulus interval (ISI), if they would not be practical to achieve a high ITR, but could still influence

decoding performance? Having access to (or implementing) a working, on-line BCI prototype early on allows for the implementation of the full UCD framework throughout the research project. It opens the door for iterative design with interaction between developers and end-users, yet another UCD principle.

This top-down or iterative approach is probably not suited for all types of BCI research If a novel BCI paradigm is still in an early phase, the focus should lie on creating limited, fundamental experiments to develop a method to decode the brain activity of interest and working bottom-up from there to a proof of concept of an end-user application. Yet, in visual oddball BCI, the challenge no longer lies in generating this proof of concept but in translating it into usable technology. Here, the top-down approach might help to avoid prematurely solving problems which may not necessarily arise further down the road. UCD could improve the quality of study outcomes by enabling the reporting of realistic, interpretable metrics, and by taking into account the human aspect of working *with* instead of *for* BCI users.

Personal contribution

Arne Van Den Kerchove conceptualized the works and performed the analyses presented in this thesis, under supervision and in consultation with prof. Hakim Si-Mohammed, prof. Marc Van Hulle and prof. François Cabestaing. Dr. Arno Libert and Dr. Benjamin Wittevrongel consulted on the analyses performed in chapter 3 Data collection was performed by Arne Van Den Kerchove for the dataset presented in chapter 6 and by Arne Van Den Kerchove and Juliette Meunier for the dataset presented in chapter 7. Participant recruitment at TRAINM was performed by Alixe Willemssens under supervision of dr. Edward Schietekatte, at Fondation Partage et Vie by Marie de Moura under supervision of prof. Etienne Allart and at the UZ Leuven Neuromuscular Reference Center under supervision of prof. Kristl Claeys and prof. Philip Van Damme. Prof. Etienne Allart provided clinical consultation.

Acknowledgments

We thank Arno Libert, Tjaa Mlinari, Yide Li, Qiang Sun and Mani Mirsaeedi for their help in collecting data for the works presented in chapter 6 and chapter 7. We would also like to thank the other colleagues at the KU Leuven Computational Neurosciences group at KU Leuven and the BCI team at the University of Lille, for their general support. Furthermore, we thank the teams at TRAINM Neuro Rehab Clinics Antwerp, the UZ Leuven Neuromuscular Reference Renter and the Fondation Partage et Vie specialized care home. We wish to acknowledge dr. Axel Faes for his inspiration in conceptualizing the work presented in chapter 4. Finally, we thank the Flemish Supercomputer Center (VSC) and the High-Performance Computing (HPC) center of the KU Leuven for allowing us to execute our computational experiments on their systems.

Conflict of interest & funding

The authors declare no conflict of interest.

Arne Van Den Kerchove is funded by a bilateral cooperation program between KU Leuven (special research fund of the KU Leuven GPUDL/20/031) and the University of Lille under the Global PhD Partnership program.

Prof. Marc Van Hulle is supported by research grants received from the European Unions Horizon Europe Marie Sklodowska-Curie Action program (grant agreement No. 101118964), the European Unions Horizon 2020 research and innovation program (grant agreement No. 857375), the special research fund of the KU Leuven (C24/18/098), the Belgian Fund for Scientific Research Flanders (G0A4118N, G0A4321N, G0C1522N), and the Hercules Foundation (AKUL 043).

The authors acknowledge the support by the RITMEA project co-financed by the European Union with the European Regional Development Fund, the French state, and the Hauts-de-France Region Council.

Curriculum vitæ

Arne Van Den Kerchove
arne@vandenkerchove.com

Education

Ph.D. in Biomedical Sciences, KU Leuven, 2024

Ph.D. in Control Science and Signal Processing, University of Lille, 2024

M.Sc. in Engineering Science: Computer Science, KU Leuven, 2020

B.Sc. in Informatics, KU Leuven, 2017

Publications

1. Arno Libert, Arne Van Den Kerchove, Benjamin Wittevrongel, and Marc M Van Hulle. "Analytic beamformer transformation for transfer learning in motion-onset visual evoked potential decoding". In: *J. Neural Eng.* 19.2 (Apr. 2022), p. 026040.
2. Arne Van Den Kerchove. "Linguistic transcription of EEG responses to sequences of visual stimuli". MA thesis. KU Leuven, 2020.
3. Arne Van Den Kerchove, Hakim Si-Mohammed, M M Van Hulle, and François Cabestaing. "Block-Term Tensor Discriminant Analysis for Brain-Computer Interfacing". In: *Journal of Neural Engineering* (submitted).
4. Arne Van Den Kerchove, Arno Libert, Benjamin Wittevrongel, and Marc M. Van Hulle. "Classification of Event-Related Potentials with Regularized Spatiotemporal LCMV Beamforming". en. In: *Applied Sciences* 12.6 (Jan. 2022), p. 2918.
5. Arne Van Den Kerchove, Hakim Si-Mohammed, MM Van Hulle, and François Cabestaing. "Correcting for ERP latency jitter improves gaze-independent BCI decoding". In: *J. Neural Eng.* 21.4 (July 2024), p. 046013.

Awards & funding

Clinical Trial Sponsorship Program, Zeto, 2024
Our research project titled "A gaze-independent Visual Brain-Computer

Interface for use by patients with limited or no eye control" won a spot in the Zeto Clinical Trial Sponsorship Program, granting us the resources to conduct research with the Zeto EEG headset.

Student BCI competition 4th place worldwide, NeuroTechX, 2022
Our NeuroTechLeuven BrainBrowsR project, a plug-and-play BCI system that lets you control social media applications trough SSVEP-BCI was awarded 90 out of 100 points by the jury of the NeuroTechX international student club competition.

Global PhD Partnership Grant, KU Leuven and University of Lille, 2021–2024

Conferences and presentations

Presented

CORTICO Days, CORTICO, Nancy, 2024
Poster presentation on covert visual attention BCIs for patients with oculomotor impairment.

10th International BCI Meeting, BCI Society, Brussels, 2023
Poster presentation on classifier-based latency estimation for gaze-independent BCIs.

Leuven.AI Scientific Workshop 2022, Leuven.AI, Leuven, 2022
Poster presentation on (multi-)Kronecker-structured linear discriminant analysis for low sample size event-related potential classification.

CORTICO Days, CORTICO, Grenoble, 2022
Presentation on Kronecker-structured LCMV-beamforming for event-related potential classification.

CORTICO Days, CORTICO, 2021
Presentation on a multi-component approach to spatiotemporal beamforming decoding of event-related potentials.

Attended

BCI & Neurotechnology Spring School 2024, g.tec, 2024

Closed Loop Neurotechnologies Autumn School, NeurotechEU, Lille, 2023

CORTICO Days, CORTICO, Paris, 2023

BCI & Neurotechnology Spring School 2023, g.tec, 2023

BCI & Neurotechnology Masterclass Belgium, g.tec, 2022

BCI & Neurotechnology Spring School 2022, g.tec, 2022

Teaching Experience

Classes

Teaching Assistant Brain-Computer Interfaces, KU Leuven, 2022–2024
: Teaching exercise sessions in BCI design and signal processing to students in the Master of Bioengineering and Advanced Master of Artificial Intelligence.

Teaching Assistant Fundamentals of Computer Science, KU Leuven, 2022
: Teaching exercise sessions in Python programming and algorithmic reasoning to students in the Bachelor of Engineering Sciences.

Master theses and internships

Eye-tracker and ERP data fusion for gaze-independent visual BCI,
: Juliette Meunier, University of Lille, 2024

LDA in the combined space-time-frequency domain for BCI decoding,
: Lunkyadi Sucipto, KU Leuven, 2023–2023

Tackling the Midas Touch problem in eye tracking with a BCI,
: Reniflal Ebenezer Sundaralal, KU Leuven, 2023–2023

An EEG connectivity analysis of Dementias, Zoe Barinaga, KU Leuven, 2023–2024

Single-trial ERP latencies as a predictor for the mode of visual attention,
: Yildiz Dilara Parry, KU Leuven, 2023

A hybrid P300-gaze BCI alternative for navigating virtual spaces,
: Gijs Claes, KU Leuven, 2021–2022

Extracurricular

Project supervisor, NeuroTech Leuven, 2022–ongoing
: Project supervisor and technical advisor for extracurricular neurotechnology student projects and a student team competing in the annual NeuroTechX BCI competition.

PAL tutor Principles of Computer Programming, KU Leuven, 2015–2016
: Organizing and teaching peer-assisted learning sessions in Python programming for first year students.

Professional experience

Ambulance EMT, Red Cross, FAST vzw, 2022-ongoing

Python developer, Mindspeller, 2019 Python Flask developer in a KU Leuven spin-off that provides neuromarketing services based on scientifically validated neuroscience and AI research.

Freelance full-stack web developer, self-employed, 2014–2020

Licenses and certifications

ICH Good Clinical Practice, E6 TransCelerate BioPharma Inc.

Ambulance EMT, FOD Volksgezondheid

References

1. Amaia Abanda, Usue Mori, and Jose A. Lozano. “A review on distance based time series classification”. en. In: *Data Min. Knowl. Disc.* 33.2 (Mar. 2019), pp. 378–412.
2. Laura Acqualagna and Benjamin Blankertz. “A gaze independent spelling based on rapid serial visual presentation”. In: *2011 Annual International Conference of the IEEE Engineering in Medicine and Biology Society.* ISSN: 1558-4615. Aug. 2011, pp. 4560–4563.
3. Laura Acqualagna and Benjamin Blankertz. “Gaze-independent BCI-spelling using rapid serial visual presentation (RSVP)”. en. In: *Clin. Neurophysiol.* 124.5 (May 2013), pp. 901–908.
4. Seyedeh Nadia Aghili et al. “A spatial-temporal linear feature learning algorithm for P300-based brain-computer interfaces”. In: *Heliyon* 9.4 (Apr. 2023), e15380.
5. Maryam Alimardani and Kazuo Hiraki. “Passive brain-computer interfaces for enhanced human-robot interaction”. In: *Frontiers in Robotics and AI* 7 (Oct. 2020), p. 125.
6. Brendan Allison et al. “BCI demographics: How many (and what kinds of) people can use an SSVEP BCI?” In: *IEEE transactions on neural systems and rehabilitation engineering* 18.2 (Apr. 2010), pp. 107–116.
7. Brendan Z Allison and Christa Neuper. “Could anyone use a BCI?” In: *Brain-computer interfaces: Applying our minds to human-computer interaction* (2010), pp. 35–54.
8. F. Aloise et al. “A comparison of classification techniques for a gaze-independent P300-based braincomputer interface”. en. In: *J. Neural Eng.* 9.4 (July 2012), p. 045012.
9. Fabio Aloise et al. “A covert attention P300-based braincomputer interface: Geospell”. In: *Ergonomics* 55.5 (May 2012), pp. 538–551.
10. Evangelos Anagnostou et al. “Preserved eye movements in adults with spinal muscular atrophy”. en. In: *Muscle Nerve* 63.5 (2021), pp. 765–769.
11. Stefan Appelhoff et al. “MNE-BIDS: Organizing electrophysiological data into the BIDS format and facilitating their analysis”. en. In: *Journal of Open Source Software* 4.44 (Dec. 2019), p. 1896.
12. P. Aricò et al. “Influence of P300 latency jitter on event related potential-based braincomputer interface performance”. en. In: *J. Neural Eng.* 11.3 (May 2014), p. 035008.

13. Amos Arieli et al. “Dynamics of Ongoing Activity: Explanation of the Large Variability in Evoked Cortical Responses”. In: *Sci* 273.5283 (Sept. 1996), pp. 1868–1871.
14. Bruno Aristimunha et al. *Mother of all BCI Benchmarks*. Version 1.0.0. 2023.
15. Mattia Arlotti et al. “Eight-hours adaptive deep brain stimulation in patients with Parkinson disease”. In: *Neurology* 90.11 (Mar. 2018), e971–e976.
16. Ruslan Aydarkhanov et al. “Spatial covariance improves BCI performance for late ERPs components with high temporal variability”. In: *J. Neural Eng.* 17.3 (June 2020), p. 036030.
17. Ian L Bailey and Jan E Lovie. “New design principles for visual acuity letter charts”. In: *Optom. Vis. Sci.* 53.11 (Nov. 1976), pp. 740–745.
18. Alexandre Barachant. “MEG decoding using Riemannian Geometry and Unsupervised classification”. In: *Grenoble University: Grenoble, France* (2014).
19. Alexandre Barachant and Marco Congedo. “A Plug&Play P300 BCI Using Information Geometry”. In: (Aug. 2014).
20. Alexandre Barachant et al. “Multiclass brain–computer interface classification by Riemannian geometry”. In: *IEEE Trans. Biomed. Eng.* 59.4 (Apr. 2011), pp. 920–928.
21. Suzanna Becker et al. “BCI illiteracy: its us, not them. Optimizing BCIs for individual brains”. In: *2022 10th International Winter Conference on Brain-Computer Interface (BCI)*. Vol. 20. IEEE. IEEE, Feb. 2022, pp. 1–3.
22. Leandro Beltrachini, Nicolás von Ellenrieder, and Carlos Horacio Muravchik. “Shrinkage Approach for Spatiotemporal EEG Covariance Matrix Estimation”. In: *IEEE Trans. Signal Process.* 61.7 (Apr. 2013), pp. 1797–1808.
23. Camille Benaroch et al. “Long-term BCI training of a tetraplegic user: adaptive riemannian classifiers and user training”. In: *Front. Hum. Neurosci.* 15 (Mar. 2021), p. 635653.
24. Abdeslem H Bentbib, Asmaa Khouia, and Hassane Sadok. “The LSQR method for solving tensor least-squares problems”. In: *Electron. Trans. Numer. Anal* 55 (2022), pp. 92–111.
25. Hans Berger. “Über das Elektrenkephalogramm des Menschen”. In: *Arch. Psychiatr. Nervenkr.* 87.1 (Dec. 1929), pp. 527–570.
26. M Bhuvaneshwari et al. “A comprehensive review on deep learning techniques for a BCI-based communication system”. In: *Demystifying big data, machine learning, and deep learning for healthcare analytics* (2021), pp. 131–157.
27. Fetsje Bijma, Jan C. de Munck, and Rob M. Heethaar. “The spatiotemporal MEG covariance matrix modeled as a sum of Kronecker products”. en. In: *NeuroImage* 27.2 (Aug. 2005), pp. 402–415.

28. Fetsje Bijma et al. “A mathematical approach to the temporal stationarity of background noise in MEG/EEG measurements”. en. In: *NeuroImage* 20.1 (Sept. 2003), pp. 233–243.
29. Benjamin Blankertz et al. “Single-trial analysis and classification of ERP components A tutorial”. In: *NeuroImage* 56.2 (2011). Multivariate Decoding and Brain Reading, pp. 814–825.
30. Julien Bogousslavsky and Otmar Meienberg. “Eye-Movement Disorders in Brain-Stem and Cerebellar Stroke”. In: *Arch. Neurol.* 44.2 (Feb. 1987), pp. 141–148.
31. Seyyed Bahram Borgheai et al. “Enhancing communication for people in late-stage ALS using an fNIRS-based BCI system”. In: *IEEE Trans. Neural Syst. Rehabil. Eng.* 28.5 (May 2020), pp. 1198–1207.
32. Davide Borra, Silvia Fantozzi, and Elisa Magosso. “Convolutional Neural Network for a P300 Brain-Computer Interface to Improve Social Attention in Autistic Spectrum Disorder”. en. In: *XV Mediterranean Conference on Medical and Biological Engineering and Computing MEDICON 2019*. Ed. by Jorge Henriques, Nuno Neves, and Paulo de Carvalho. IFMBE Proceedings. Cham: Springer International Publishing, Sept. 2020, pp. 1837–1843.
33. Mariana P Branco et al. “Brain-computer interfaces for communication: preferences of individuals with locked-in syndrome”. In: *Neurorehabilitation and neural repair* 35.3 (Feb. 2021), pp. 267–279.
34. Mariana P Branco et al. “Decoding hand gestures from primary somatosensory cortex using high-density ECoG”. In: *Neuroimage* 147 (Feb. 2017), pp. 130–142.
35. A. Breloy et al. “Robust rank constrained kronecker covariance matrix estimation”. In: *2016 50th Asilomar Conference on Signals, Systems and Computers*. Nov. 2016, pp. 810–814.
36. Sebastianus Petrus van den Broek et al. “Volume conduction effects in EEG and MEG”. In: *Electroencephalography and clinical neurophysiology* 106.6 (June 1998), pp. 522–534.
37. Anne-Marie Brouwer and Jan B Van Erp. “A tactile P300 brain-computer interface”. In: *Frontiers in neuroscience* 4 (2010), p. 1440.
38. P. Brunner et al. “Does the P300 speller depend on eye gaze?” en. In: *J. Neural Eng.* 7.5 (Sept. 2010), p. 056013.
39. Adrian P Burgess. “Towards a unified understanding of event-related changes in the EEG: the firefly model of synchronization through cross-frequency phase modulation”. In: (2012).
40. Qian Cai et al. “Multilinear Discriminative Spatial Patterns for Movement-Related Cortical Potential Based on EEG Classification with Tensor Representation”. In: *Computational Intelligence and Neuroscience* 2021 (May 2021). Ed. by Anastasios D. Doulamis, pp. 1–9.
41. Flavio Camarrone and Marc M Van Hulle. “Fast multiway partial least squares regression”. In: *IEEE Trans. Biomed. Eng.* 66.2 (2018), pp. 433–443.

42. Flavio Camarrone et al. “Accurate Offline Asynchronous Detection of Individual Finger Movement From Intracranial Brain Signals Using a Novel Multiway Approach”. In: *IEEE Trans. Biomed. Eng.* 68.7 (July 2021), pp. 2176–2187.
43. Eloise H. Carlton and Sidney Katz. “Is Wiener Filtering an Effective Method of Improving Evoked Potential Estimation?” In: *IEEE Trans. Biomed. Eng.* BME-27.4 (Apr. 1980), pp. 187–192.
44. Mario H Castaneda and Josef A Nossek. “Estimation of rank deficient covariance matrices with Kronecker structure”. In: *2014 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*. IEEE. IEEE, May 2014, pp. 394–398.
45. Giovanni Castelnovo et al. “Study of subclinical ocular motor disorders by video oculography in patients with clinically isolated syndromes”. en. In: *Multiple Sclerosis and Related Disorders* 9 (Sept. 2016), pp. 118–120.
46. Hubert Cecotti et al. “Single-trial detection of event-related fields in MEG from the presentation of happy faces: Results of the Biomag 2016 data challenge”. In: *2017 39th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*. IEEE. IEEE, July 2017, pp. 4467–4470.
47. Ujwal Chaudhary, Niels Birbaumer, and Ander Ramos-Murguialday. “Brain-computer interfaces for communication and rehabilitation”. In: *Nat. Rev. Neurol.* 12.9 (Aug. 2016), pp. 513–525.
48. Yilun Chen et al. “Shrinkage algorithms for MMSE covariance estimation”. In: *IEEE Trans. Signal Process.* 58.10 (Oct. 2010), pp. 5016–5029.
49. Yonghao Chen et al. “Implementing a calibration-free SSVEP-based BCI system with 160 targets”. In: *J. Neural Eng.* 18.4 (June 2021), p. 046094.
50. Srivas Chennu et al. “The cost of space independence in P300-BCI spellers”. en. In: *Journal of NeuroEngineering and Rehabilitation* 10.1 (July 2013), p. 82.
51. Sylvain Chevallier et al. “The largest EEG-based BCI reproducibility study for open science: the MOABB benchmark”. In: *arXiv preprint arXiv:2404.15319* (2024).
52. A Cook and P Giunti. “Friedreichs ataxia: clinical features, pathogenesis and management”. In: *British medical bulletin* 124.1 (Oct. 2017), pp. 19–30.
53. Carina Curto et al. “A simple model of cortical dynamics explains variability and state dependence of sensory responses in urethane-anesthetized auditory cortex”. In: *Journal of neuroscience* 29.34 (Aug. 2009), pp. 10600–10612.
54. Costanza DAvanzo et al. “A Bayesian method to estimate single-trial event-related potentials with application to the study of the P300 variability”. en. In: *J. Neurosci. Methods* 198.1 (May 2011), pp. 114–124.
55. Lieven De Lathauwer. “Decompositions of a higher-order tensor in block termsPart I: Lemmas for partitioned matrices”. In: *SIAM J. Matrix Anal. Appl.* 30.3 (Jan. 2008), pp. 1022–1032.

56. Lieven De Lathauwer. “Decompositions of a higher-order tensor in block termsPart II: Definitions and uniqueness”. In: *SIAM J. Matrix Anal. Appl.* 30.3 (Jan. 2008), pp. 1033–1066.
57. Lieven De Lathauwer, Bart De Moor, and Joos Vandewalle. “A Multilinear Singular Value Decomposition”. In: *SIAM J. Matrix Anal. Appl.* 21.4 (Jan. 2000), pp. 1253–1278.
58. Lieven De Lathauwer and Dimitri Nion. “Decompositions of a higher-order tensor in block termsPart III: Alternating least squares algorithms”. In: *SIAM journal on Matrix Analysis and Applications* 30.3 (Jan. 2008), pp. 1067–1083.
59. Jan C. De Munck and Bob W. Van Dijk. “The Spatial Distribution of Spontaneous EEG and MEG”. en. In: ed. by Christian Uhl. Springer Series in Synergetics. Berlin, Heidelberg: Springer, 1999, pp. 202–228.
60. Jan Casper De Munck et al. “Estimating stationary dipoles from MEG/EEG data contaminated with spatially and temporally correlated background noise”. In: *IEEE Trans. Signal Process.* 50.7 (July 2002), pp. 1565–1572.
61. Arnaud Delorme, Terrence Sejnowski, and Scott Makeig. “Enhanced detection of artifacts in EEG data using higher-order statistics and independent component analysis”. In: *Neuroimage* 34.4 (Feb. 2007), pp. 1443–1449.
62. Stavros I. Dimitriadis et al. “A Novel, Fast, Reliable, and Data-Driven Method for Simultaneous Single-Trial Mining and AmplitudeLatency Estimation Based on Proximity Graphs and Network Analysis”. In: *Front. Neuroinf.* 12 (2018).
63. Chuck Easttom et al. “A functional model for unifying brain computer interface terminology”. In: *IEEE Open Journal of Engineering in Medicine and Biology* 2 (2021), pp. 91–96.
64. Günter Edlinger, Brendan Z. Allison, and Christoph Guger. “How Many People Can Use a BCI System?” In: (Dec. 2014), pp. 33–66.
65. John M Egan et al. “A gaze independent hybrid-BCI based on visual spatial attention”. In: *J. Neural Eng.* 14.4 (May 2017), p. 046006.
66. M Eidel and A Kübler. “Identifying potential training factors in a vibrotactile P300-BCI”. In: *Sci. Rep.* 12.1 (2022), p. 14006.
67. Aybüke Erol and Borbála Hunyadi. “Tensors for neuroimaging: A review on applications of tensors to unravel the mysteries of the brain”. In: *Tensors for Data Processing* (2022), pp. 427–482.
68. A Faes and M M Van Hulle. “Finger movement and coactivation predicted from intracranial brain activity using extended block-term tensor regression”. In: *J. Neural Eng.* 19.6 (Nov. 2022), p. 066011.
69. Axel Faes, Flavio Camarrone, and Marc M. Van Hulle. “Single Finger Trajectory Prediction From Intracranial Brain Activity Using Block-Term Tensor Regression With Fast and Automatic Component Extraction”. In: *IEEE Trans. Neural Netw. Learn. Syst.* (2022), pp. 1–12.
70. Michael C Fahey et al. “Vestibular, saccadic and fixation abnormalities in genetically confirmed Friedreich ataxia”. In: *Brain* 131.4 (Feb. 2008), pp. 1035–1045.

71. Lawrence Ashley Farwell and Emanuel Donchin. “Talking off the top of your head: toward a mental prosthesis utilizing event-related brain potentials”. In: *Electroencephalography and clinical Neurophysiology* 70.6 (1988), pp. 510–523.
72. Elisa Fazzi et al. “Neuro-ophthalmological disorders in cerebral palsy: ophthalmological, oculomotor, and visual aspects”. In: *Developmental Medicine & Child Neurology* 54.8 (June 2012), pp. 730–736.
73. Thomas C Ferree, MT Clay, and Don M Tucker. “The spatial resolution of scalp EEG”. In: *Neurocomputing* 38 (June 2001), pp. 1209–1216.
74. Martin A. Fischler and Robert C. Bolles. “Random sample consensus: a paradigm for model fitting with applications to image analysis and automated cartography”. In: *Commun. ACM* 24.6 (June 1981), pp. 381–395.
75. Sharlene N Flesher et al. “A brain-computer interface that evokes tactile sensations improves robotic arm control”. In: *Sci* 372.6544 (May 2021), pp. 831–836.
76. Philipp Flotho et al. “Fast variational alignment of non-flat 1D displacements for applications in neuroimaging”. In: *J. Neurosci. Methods* 353 (Apr. 2021), p. 109076.
77. Islam A Fouad et al. “Improving the performance of P300 BCI system using different methods”. In: *Network Modeling Analysis in Health Informatics and Bioinformatics* 9.1 (Sept. 2020), p. 64.
78. S. Frenzel, E. Neubert, and C. Bandt. “Two communication lines in a 3 Œ 3 matrix speller”. en. In: *J. Neural Eng.* 8.3 (May 2011), p. 036021.
79. Melanie Fried-Oken et al. “Human visual skills for brain-computer interface use: a tutorial”. In: *Disability and Rehabilitation: Assistive Technology* 15.7 (June 2020), pp. 799–809.
80. Laura Frølich, Tobias Søren Andersen, and Morten Mørup. “Rigorous optimisation of multilinear discriminant analysis with Tucker and PARAFAC structures”. In: *BMC Bioinf.* 19.1 (May 2018), p. 197.
81. Joseph M Furman, Susan Perlman, and Robert W Baloh. “Eye movements in Friedreich’s ataxia”. In: *Archives of neurology* 40.6 (June 1983), pp. 343–346.
82. Theo Gasser, Joachim Möcks, and Rolf Verleger. “SELAVCO: A method to deal with trial-to-trial variability of evoked potentials”. In: *Electroencephalography and clinical Neurophysiology* 55.6 (June 1983), pp. 717–723.
83. S.D. Georgiadis et al. “Single-trial dynamical estimation of event-related potentials: a Kalman filter-based approach”. In: *IEEE Trans. Biomed. Eng.* 52.8 (Aug. 2005), pp. 1397–1406.
84. Joseph T Giacino, Kathleen Kalmar, and John Whyte. “The JFK Coma Recovery Scale-Revised: measurement characteristics and diagnostic utility”. In: *Archives of physical medicine and rehabilitation* 85.12 (Jan. 2004), pp. 2020–2029.
85. Paula Gonzalez-Navarro et al. “A kronecker product structured EEG covariance estimator for a language model assisted-BCI”. In: *Interna-*

tional Conference on Augmented Cognition. Springer. Springer International Publishing, 2016, pp. 35–45.
86. Paula Gonzalez-Navarro et al. "Spatio-temporal EEG models for brain interfaces". In: *Signal processing* 131 (2017), pp. 333–343.
87. M Graber et al. "Evaluation of the visual function of patients with locked-in syndrome: report of 13 cases". In: *Journal Français d'Ophtalmologie* 39.5 (May 2016), pp. 437–440.
88. Alexandre Gramfort et al. "MEG and EEG Data Analysis with MNE-Python". In: *Front. Neurosci.* 7.267 (2013), pp. 1–13.
89. Alexandre Gramfort et al. "MNE software for processing MEG and EEG data". In: *neuroimage* 86 (Feb. 2014), pp. 446–460.
90. Gabriele Gratton et al. "Simulation Studies of Latency Measures of Components of the Event-Related Brain Potential". en. In: *Psychophysiology* 26.2 (1989), pp. 233–248.
91. Kristjan Greenewald and Alfred O. Hero. "Regularized block Toeplitz covariance matrix estimation via Kronecker product expansions". In: *2014 IEEE Workshop on Statistical Signal Processing (SSP)*. ISSN: 2373-0803. June 2014, pp. 9–12.
92. Christoph Guger et al. "How many people are able to control a P300-based brain–computer interface (BCI)?" In: *Neuroscience letters* 462.1 (Sept. 2009), pp. 94–98.
93. Xintong Guo et al. "Eye Movement Abnormalities in Amyotrophic Lateral Sclerosis". In: *Brain Sciences* 12.4 (Apr. 2022), p. 489.
94. L. Gupta et al. "Nonlinear alignment and averaging for estimating the evoked potential". In: *IEEE Trans. Biomed. Eng.* 43.4 (1996), pp. 348–356.
95. Sebastian Halder et al. "An auditory oddball brain–computer interface for binary choices". In: *Clin. Neurophysiol.* 121.4 (Apr. 2010), pp. 516–523.
96. D Corydon Hammond. "What is neurofeedback: An update". In: *Journal of neurotherapy* 15.4 (Oct. 2011), pp. 305–336.
97. Jin Han et al. "A high-speed hybrid brain-computer interface with more than 200 targets". In: *J. Neural Eng.* 20.1 (Jan. 2023), p. 016025.
98. Xiangke Han, Jianye Niu, and Shijie Guo. "A tactile-based brain computer interface P300 paradigm using vibration frequency and spatial location". In: *J. Med. Biol. Eng.* 40.6 (June 2020), pp. 773–782.
99. Yiyuan Han et al. "Two sides of the same coin: adaptation of BCIs to internal states with user-centered design and electrophysiological features". In: *Brain-Computer Interfaces* 9.2 (Mar. 2022), pp. 102–114.
100. Irzam Hardiansyah, Valentina Pergher, and Marc M Van Hulle. "Single-trial EEG responses classified using latency features". In: *Int. J. Neural Syst.* 30.06 (June 2020), p. 2050033.
101. Sandra G Hart. "NASA-task load index (NASA-TLX); 20 years later". In: *Proceedings of the human factors and ergonomics society annual meeting*. Vol. 50. 9. Sage publications Sage CA: Los Angeles, CA. American Psychological Association (APA), 2006, pp. 904–908.

102. Andrea Hasenstaub, Robert NS Sachdev, and David A McCormick. “State changes rapidly modulate cortical neuronal responsiveness”. In: *J. Neurosci.* 27.36 (2007), pp. 9607–9622.
103. Mohamed Abul Hassan et al. “Approach to quantify eye movements to augment stroke diagnosis with a non-calibrated eye-tracker”. In: *IEEE Trans. Biomed. Eng.* 70.6 (June 2022), pp. 1750–1757.
104. Stefan Haufe et al. “On the interpretation of weight vectors of linear models in multivariate neuroimaging”. en. In: *Neuroimage* 87 (Feb. 2014), pp. 96–110.
105. Hideaki Hayashi, Shuuichi Kato, and Akihiro Kawada. “Amyotrophic lateral sclerosis patients living beyond respiratory failure”. en. In: *J. Neurol. Sci.* 105.1 (Sept. 1991), pp. 73–78.
106. Andreas Herweg and Andrea Kübler. “High performance with tactile P300 BCIs”. In: *2016 4th International Winter Conference on Brain-Computer Interface (BCI)*. IEEE. IEEE, Feb. 2016, pp. 1–2.
107. Andreas Herweg et al. “Wheelchair control by elderly participants in a virtual environment with a brain-computer interface (BCI) and tactile stimulation”. In: *Biological psychology* 121 (Dec. 2016), pp. 117–124.
108. Hiroshi Higashi et al. “Multilinear Discriminant Analysis With Subspace Constraints for Single-Trial Classification of Event-Related Potentials”. In: *IEEE J. Sel. Topics Signal Process.* 10.7 (Oct. 2016), pp. 1295–1305.
109. Frank L Hitchcock. “The expression of a tensor or a polyadic as a sum of products”. In: *J. Math. Phys. Camb.* 6.1-4 (Apr. 1927), pp. 164–189.
110. Darren R Hocking et al. “Ocular motor fixation deficits in Friedreich ataxia”. In: *The Cerebellum* 9.3 (May 2010), pp. 411–418.
111. L. Hu et al. “A novel approach for enhancing the signal-to-noise ratio and detecting automatically event-related potentials (ERPs) in single trials”. en. In: *NeuroImage* 50.1 (Mar. 2010), pp. 99–111.
112. Shoulin Huang et al. “Spectrum-weighted tensor discriminant analysis for motor imagery-based BCI”. In: *IEEE Access* 8 (2020), pp. 93749–93759.
113. Jane E Huggins, Patricia A Wren, and Kirsten L Gruis. “What would brain-computer interface users want? Opinions and priorities of potential users with amyotrophic lateral sclerosis”. In: *Amyotroph. Lateral Scler.* 12.5 (May 2011), pp. 318–324.
114. Jane E Huggins et al. “What would brain-computer interface users want: opinions and priorities of potential users with spinal cord injury”. In: *Archives of physical medicine and rehabilitation* 96.3 (Mar. 2015), S38–S45.
115. Hilde M Huizenga et al. “Spatiotemporal EEG/MEG source analysis based on a parametric noise covariance model”. In: *IEEE Trans. Biomed. Eng.* 49.6 (June 2002), pp. 533–539.
116. David F. Hultsch and Stuart W. S. MacDonald. “Intraindividual variability in performance as a theoretical window onto cognitive aging”. In: *New Frontiers in Cognitive Aging*. Ed. by Roger Dixon, Lars Backman, and Lars-Goran Nilsson. Oxford University Press, Apr. 2004, pp. 65–88.

117. Han-Jeong Hwang et al. “A Gaze Independent Brain-Computer Interface Based on Visual Stimulation through Closed Eyelids”. en. In: *Sci. Rep.* 5.1 (Oct. 2015), p. 15890.
118. International Organization for Standardization. “Ergonomics of Human System InteractionPart 210: Human-Centered Design for Interactive Systems (formerly known as 13407)”. In: *ISO FDIS* 9241.210 (2009), p. 2009.
119. I. Iturrate et al. “Latency correction of event-related potentials between different experimental protocols”. en. In: *J. Neural Eng.* 11.3 (Apr. 2014), p. 036005.
120. Mina Jamshidi Idaji, Mohammad B. Shamsollahi, and Sepideh Hajipour Sardouie. “Higher order spectral regression discriminant analysis (HOSRDA): A tensor feature reduction method for ERP detection”. In: *Pattern Recogn.* 70 (Oct. 2017), pp. 152–162.
121. Jing Jin et al. “Developing a novel tactile P300 brain-computer interface with a cheeks-stim paradigm”. In: *IEEE Trans. Biomed. Eng.* 67.9 (Sept. 2020), pp. 2585–2593.
122. Jing Jin et al. “The changing face of P300 BCIs: a comparison of stimulus changes in a P300 BCI involving faces, emotion, and movement”. In: *PloS one* 7.11 (Nov. 2012). Ed. by Laura Frishman, e49688.
123. Tania Jorajuría et al. “Oscillatory Source Tensor Discriminant Analysis (OSTDA): A regularized tensor pipeline for SSVEP-based BCI systems”. In: *Neurocomputing* 492 (July 2022), pp. 664–675.
124. Tzyy-Ping Jung et al. “Analysis and visualization of single-trial event-related potentials”. In: *Human brain mapping* 14.3 (2001), pp. 166–185.
125. Yu Kageyama et al. “Nationwide survey of 780 Japanese patients with amyotrophic lateral sclerosis: their status and expectations from brain–machine interfaces”. In: *J. Neurol.* 267 (2020), pp. 2932–2940.
126. Bong-Hui Kang et al. “Abnormal oculomotor functions in amyotrophic lateral sclerosis”. In: *J. Clin. Neurol.* 14.4 (2018), pp. 464–471.
127. Netiwit Kaongoen and Sungho Jo. “A novel hybrid auditory BCI paradigm combining ASSR and P300”. In: *Journal of neuroscience methods* 279 (Mar. 2017), pp. 44–51.
128. P.A. Karjalainen et al. “Subspace regularization method for the single-trial estimation of evoked potentials”. In: *IEEE Trans. Biomed. Eng.* 46.7 (July 1999), pp. 849–860.
129. C. Khatri and C.R. Rao. “Effects of estimated noise covariance matrix in optimal signal detection”. In: *IEEE Trans. Acoust., Speech, Signal Process.* 35.5 (May 1987), pp. 671–679.
130. Minju Kim, Jongsu Kim, and Sung-Phil Kim. “Enhancing ERP component detection by estimating ERP latency variability using hidden process model”. In: *2020 IEEE International Conference on Systems, Man, and Cybernetics (SMC)*. ISSN: 2577-1655. Oct. 2020, pp. 1262–1268.
131. Michael A Kisley and George L Gerstein. “Trial-to-trial variability and state-dependent modulation of auditory-evoked responses in cortex”. In: *J. Neurosci.* 19.23 (Dec. 1999), pp. 10451–10460.

132. Nicole Kraemer. “On the Peaking Phenomenon of the Lasso in Model Selection”. In: (Apr. 2009).
133. Mario Michael Krell, Anett Seeland, and Su Kyoung Kim. *Data Augmentation for Brain-Computer Interfaces: Analysis on Event-Related Potentials Data*. Tech. rep. arXiv:1801.02730 [cs, q-bio] type: article. arXiv, Jan. 2018.
134. Andrea Kübler and Niels Birbaumer. “Brain–computer interfaces and communication in paralysis: Extinction of goal directed thinking in completely paralysed patients?” In: *Clinical neurophysiology* 119.11 (Nov. 2008), pp. 2658–2666.
135. Andrea Kübler et al. “The user-centered design as novel perspective for evaluating the usability of BCI-controlled applications”. In: *PloS one* 9.12 (Dec. 2014). Ed. by Stefano Federici, e112392.
136. Simon Ladouce et al. “Improving user experience of SSVEP BCI through low amplitude depth and high frequency stimuli design”. In: *Sci. Rep.* 12.1 (May 2022), p. 8865.
137. Amy N. Langville and William J. Stewart. “The Kronecker product and stochastic automata networks”. en. In: *J. Comput. Appl. Math.* 167.2 (June 2004), pp. 429–447.
138. Olivier Ledoit and Michael Wolf. “A well-conditioned estimator for large-dimensional covariance matrices”. In: *J. Multivariate Anal.* 88.2 (Feb. 2004), pp. 365–411.
139. Olivier Ledoit and Michael NMI Wolf. “Honey, I Shrunk the Sample Covariance Matrix”. In: *SSRN Electronic Journal* (2003).
140. Robert Leeb et al. “Brain–computer communication: motivation, aim, and impact of exploring a virtual apartment”. In: *IEEE Trans. Neural Syst. Rehabil. Eng.* 15.4 (Dec. 2007), pp. 473–482.
141. Stephanie Lees et al. “A review of rapid serial visual presentation-based braincomputer interfaces”. en. In: *J. Neural Eng.* 15.2 (Jan. 2018), p. 021001.
142. Damien Lesenfants et al. “An independent SSVEP-based brain–computer interface in locked-in syndrome”. In: *Journal of neural engineering* 11.3 (May 2014), p. 035002.
143. Feng Li et al. “Transfer learning algorithm of P300-EEG signal based on XDAWN spatial filter and Riemannian geometry classifier”. In: *Applied Sciences* 10.5 (Mar. 2020), p. 1804.
144. Jianing Li et al. “An online P300 brain–computer interface based on tactile selective attention of somatosensory electrical stimulation”. In: *J. Med. Biol. Eng.* 39 (2019), pp. 732–738.
145. Yueqing Li et al. “A P300-based brain–computer interface: Effects of interface type and screen size”. In: *Intl. Journal of Human–Computer Interaction* 27.1 (Dec. 2010), pp. 52–68.
146. Arno Libert and Marc M Van Hulle. “Predicting premature video skipping and viewer interest from EEG recordings”. In: *Entropy-switz.* 21.10 (Oct. 2019), p. 1014.
147. Arno Libert, Benjamin Wittevrongel, and Marc M. Van Hulle. “Effect of stimulus direction on motion-onset visual evoked potentials decoded

using spatiotemporal beamforming Abstract". In: *2021 10th International IEEE/EMBS Conference on Neural Engineering (NER)*. ISSN: 1948-3554. May 2021, pp. 503–506.
148. Arno Libert et al. "Analytic beamformer transformation for transfer learning in motion-onset visual evoked potential decoding". In: *J. Neural Eng.* 19.2 (Apr. 2022), p. 026040.
149. Arno Libert et al. "Phase-spatial beamforming renders a visual brain computer interface capable of exploiting EEG electrode phase shifts in motion-onset target responses". In: *IEEE Trans. Biomed. Eng.* 69.5 (May 2021), pp. 1802–1812.
150. Zhimin Lin et al. "A novel P300 BCI speller based on the Triple RSVP paradigm". en. In: *Sci. Rep.* 8.1 (Feb. 2018), p. 3350.
151. Yadong Liu et al. "An tactile ERP-based brain–computer interface for communication". In: *International Journal Of Human–Computer Interaction* 35.7 (May 2019), pp. 559–567.
152. Ye Liu, Qibin Zhao, and Liqing Zhang. "Uncorrelated Multiway Discriminant Analysis for Motor Imagery EEG Classification". In: *Int. J. Neural Syst.* (2015).
153. Charles F. Van Loan. "The ubiquitous Kronecker product". en. In: *J. Comput. Appl. Math.* Numerical Analysis 2000. Vol. III: Linear Algebra 123.1 (Nov. 2000), pp. 85–100.
154. Henri Lorach et al. "Walking naturally after spinal cord injury using a brain–spine interface". In: *Nat* 618.7963 (May 2023), pp. 126–133.
155. Fabien Lotte et al. "A review of classification algorithms for EEG-based brain–computer interfaces: a 10 year update". In: *Journal of neural engineering* 15.3 (Apr. 2018), p. 031005.
156. Nelson Lu and Dale L. Zimmerman. "The likelihood ratio test for a separable covariance matrix". en. In: *Stat. Probabil. Lett.* 73.4 (July 2005), pp. 449–457.
157. Steven J Luck. *An introduction to the event-related potential technique*. MIT press, 2014.
158. Steven J Luck and Emily S Kappenman. *The Oxford handbook of event-related potential components*. Oxford university press, 2013.
159. Fausta Lui et al. "Saccadic eye movements are impaired in Duchenne Muscular Dystrophy". en. In: *Doc. Ophthalmol.* 103.3 (Nov. 2001), pp. 219–228.
160. WC Maples. "Northeastern State University College of Optometry'Oculomotor norms". In: *J Behav Optom* 3 (1992), pp. 143–150.
161. Edwin M. Maynard, Craig T. Nordhausen, and Richard A. Normann. "The Utah Intracortical Electrode Array: A recording structure for potential brain-computer interfaces". en. In: *Electroencephalogr. Clin. Neurophysiol.* 102.3 (Mar. 1997), pp. 228–239.
162. Lynn M McCane et al. "Brain-computer interface (BCI) evaluation in people with amyotrophic lateral sclerosis". In: *Amyotrophic lateral sclerosis and frontotemporal degeneration* 15.3-4 (Feb. 2014), pp. 207–215.

163. Gregory McCarthy and Emanuel Donchin. “A Metric for Thought: A Comparison of P300 Latency and Reaction Time”. In: *Sci* 211.4477 (Jan. 1981), pp. 77–80.
164. Jürgen Mellinger et al. “An MEG-based braincomputer interface (BCI)”. en. In: *NeuroImage* 36.3 (July 2007), pp. 581–593.
165. Sean L Metzger et al. “A high-performance neuroprosthesis for speech decoding and avatar control”. In: *Nat* 620.7976 (Aug. 2023), pp. 1037–1046.
166. Tomislav Milekovic et al. “Stable long-term BCI-enabled communication in ALS and locked-in syndrome using LFP signals”. In: *Journal of neurophysiology* 120.7 (July 2018), pp. 343–360.
167. J d R Millán et al. “Combining brain–computer interfaces and assistive technologies: state-of-the-art and challenges”. In: *Frontiers in neuroscience* 4 (2010), p. 161.
168. Elizabeth Milne. “Increased Intra-Participant Variability in Children with Autistic Spectrum Disorders: Evidence from Single-Trial Analysis of Evoked EEG”. In: *Front. Psychol.* 2 (2011).
169. Peter Mitchell et al. “Assessment of safety of a fully implanted endovascular brain-computer interface for severe paralysis in 4 patients: the stentrode with thought-controlled digital switch (SWITCH) study”. In: *JAMA neurology* 80.3 (Mar. 2023), pp. 270–278.
170. J. Möcks et al. “Novel approaches to the problem of latency jitter”. eng. In: *Psychophysiology* 25.2 (Mar. 1988), pp. 217–226.
171. Hakim Si-Mohammed et al. “Brain-computer interfaces and augmented reality: A state of the art”. In: *Graz Brain-Computer Interface Conference*. 2017.
172. Hakim Si-Mohammed et al. “Detecting system errors in virtual reality using EEG through error-related potentials”. In: *2020 IEEE Conference on Virtual Reality and 3D User Interfaces (VR)*. IEEE. IEEE, Mar. 2020, pp. 653–661.
173. Jorge Moncayo and Julien Bogousslavsky. “Oculomotor disorders in vertebrobasilar stroke”. English. In: *Expert Rev. Ophthalmol.* 4.3 (June 2009), pp. 259–281.
174. John C Mosher, Richard M Leahy, and Paul S Lewis. “EEG and MEG: forward solutions for inverse methods”. In: *IEEE Transactions on biomedical engineering* 46.3 (1999), pp. 245–259.
175. Heather E Moss et al. “Cross-sectional evaluation of clinical neuro-ophthalmic abnormalities in an amyotrophic lateral sclerosis population”. In: *Journal of the neurological sciences* 314.1-2 (2012), pp. 97–101.
176. A. Mouraux and G. D. Iannetti. “Across-trial averaging of event-related EEG responses and beyond”. en. In: *Magn. Reson. Imaging*. Proceedings of the International School on Magnetic Resonance and Brain Function 26.7 (Sept. 2008), pp. 1041–1054.
177. Md Rakibul Mowla, Jane E. Huggins, and David E. Thompson. “Enhancing P300-BCI performance using latency estimation”. In: *Brain-Computer Interfaces* 4.3 (July 2017), pp. 137–145.

178. Md Rakibul Mowla et al. “A Comparison of Classification Techniques to Predict Brain-Computer Interfaces Accuracy Using Classifier-Based Latency Estimation”. en. In: *Brain Sciences* 10.10 (Oct. 2020), p. 734.
179. Christian Mühl et al. “A survey of affective brain computer interfaces: principles, state-of-the-art, and challenges”. In: *Brain-Computer Interfaces* 1.2 (Apr. 2014), pp. 66–84.
180. J.C. de Munck, P.C.M. Vijn, and F.H. Lopes da Silva. “A random dipole model for spontaneous brain activity”. In: *IEEE Trans. Biomed. Eng.* 39.8 (Aug. 1992), pp. 791–804.
181. René M Müri and Otmar Meienberg. “The clinical spectrum of internuclear ophthalmoplegia in multiple sclerosis”. In: *Archives of neurology* 42.9 (1985), pp. 851–855.
182. Elon Musk et al. “An integrated brain-machine interface platform with thousands of channels”. In: *Journal of medical Internet research* 21.10 (July 2019), e16194.
183. Lorina Naci et al. “Braincomputer interfaces for communication with nonresponsive patients”. en. In: *Ann. Neurol.* 72.3 (2012), pp. 312–323.
184. Masaki Nakanishi et al. “Enhancing detection of SSVEPs for a high-speed brain speller using task-related component analysis”. In: *IEEE Trans. Biomed. Eng.* 65.1 (Jan. 2017), pp. 104–112.
185. Kianoush Nazarpour et al. “Parallel space-time-frequency decomposition of EEG signals for brain computer interfacing”. In: *2006 14th European Signal Processing Conference*. IEEE. 2006, pp. 1–4.
186. M. de Neeling and M. M. Van Hulle. “Single-paradigm and hybrid brain computing interfaces and their use by disabled patients”. en. In: *J. Neural Eng.* 16.6 (Oct. 2019), p. 061001.
187. Femke Nijboer et al. “An auditory brain–computer interface (BCI)”. In: *Journal of neuroscience methods* 167.1 (Jan. 2008), pp. 43–50.
188. Shigeto Nishida et al. “A morphological filter for extracting waveform characteristics of single-sweep evoked potentials”. en. In: *Automatica* 35.5 (May 1999), pp. 937–943.
189. Martin Ohlson, M Rauf Ahmad, and Dietrich Von Rosen. “The multilinear normal distribution: Introduction and some basic properties”. In: *J. Multivariate Anal.* 113 (Jan. 2013), pp. 37–47.
190. Barry S Oken et al. “Brain–computer interface with language model-electroencephalography fusion for locked-in syndrome”. In: *Neurorehabilitation and neural repair* 28.4 (2014), pp. 387–394.
191. Akinari Onishi et al. “Tensor classification for P300-based brain computer interface”. In: *2012 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*. ISSN: 2379-190X. Mar. 2012, pp. 581–584.
192. Umut Orhan et al. “RSVP keyboard: An EEG based typing interface”. In: *2012 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*. IEEE. IEEE, Mar. 2012, pp. 645–648.

193. Guang Ouyang. “ReSync: Correcting the trial-to-trial asynchrony of event-related brain potentials to improve neural response representation”. en. In: *J. Neurosci. Methods* 339 (June 2020), p. 108722.
194. Guang Ouyang, Werner Sommer, and Changsong Zhou. “Reconstructing ERP amplitude effects after compensating for trial-to-trial latency jitter: A solution based on a novel application of residue iteration decomposition”. en. In: *Int. J. Psychophysiol.* 109 (Nov. 2016), pp. 9–20.
195. Guang Ouyang, Werner Sommer, and Changsong Zhou. “Updating and validating a new framework for restoring and analyzing latency-variable ERP components from single trials with residue iteration decomposition (RIDE)”. en. In: *Psychophysiology* 52.6 (2015), pp. 839–856.
196. Guang Ouyang et al. “Exploiting the intra-subject latency variability from single-trial event-related potentials in the P3 time range: A review and comparative evaluation of methods”. en. In: *Neuroscience & Biobehavioral Reviews* 75 (Apr. 2017), pp. 1–21.
197. Guang Ouyang et al. “Residue iteration decomposition (RIDE): A new method to separate ERP components on the basis of latency variability in single trials”. en. In: *Psychophysiology* 48.12 (2011), pp. 1631–1647.
198. Jiahui Pan et al. “Advances in P300 brain–computer interface spellers: toward paradigm design and performance evaluation”. In: *Front. Hum. Neurosci.* 16 (Dec. 2022), p. 1077717.
199. Jerrin Thomas Panachakel and Angarai Ganesan Ramakrishnan. “Decoding covert speech from EEG-a comprehensive review”. In: *Front. Neurosci.* 15 (Apr. 2021), p. 642251.
200. Jean-Michel Papy, Lieven De Lathauwer, and Sabine Van Huffel. “Exponential data fitting using multilinear algebra: the single-channel and multi-channel case”. In: *Numerical linear algebra with applications* 12.8 (2005), pp. 809–826.
201. Sang-Hoon Park, David Lee, and Sang-Goog Lee. “Filter bank regularized common spatial pattern ensemble for small sample motor imagery classification”. In: *IEEE Trans. Neural Syst. Rehabil. Eng.* 26.2 (2017), pp. 498–505.
202. Lucas C. Parra et al. “Recipes for the linear analysis of EEG”. In: *NeuroImage* 28.2 (Nov. 2005), pp. 326–341.
203. Emanuele Pasqualotto et al. “Usability and workload of access technology for people with severe motor impairment: a comparison of brain-computer interfacing and eye tracking”. In: *Neurorehabilitation and neural repair* 29.10 (Mar. 2015), pp. 950–957.
204. James R Patterson and Martin Grabois. “Locked-in syndrome: a review of 139 cases.” In: *Stroke* 17.4 (1986), pp. 758–764.
205. Fabian Pedregosa et al. “Scikit-learn: Machine learning in Python”. In: *the Journal of machine Learning research* 12 (2011), pp. 2825–2830.
206. Jonathan Peirce et al. “PsychoPy2: Experiments in behavior made easy”. en. In: *Behav. Res. Methods* 51.1 (Feb. 2019), pp. 195–203.

207. P. Da Pelo et al. “Trial latencies estimation of event-related potentials in EEG by means of genetic algorithms”. en. In: *J. Neural Eng.* 15.2 (Feb. 2018), p. 026016.
208. Cyril R. Pernet et al. “EEG-BIDS, an extension to the brain imaging data structure for electroencephalography”. en. In: *Sci. Data* 6.1 (June 2019), p. 103.
209. Betts Peters et al. “A systematic review of research on augmentative and alternative communication brain-computer interface systems for individuals with disabilities”. In: *Frontiers in human neuroscience* 16 (July 2022), p. 952380.
210. Betts Peters et al. “SSVEP BCI and eye tracking use by individuals with late-stage ALS and visual impairments”. In: *Front. Hum. Neurosci.* 14 (Nov. 2020), p. 595890.
211. Jimmy Petit, José Rouillard, and François Cabestaing. “EEG-based brain–computer interfaces exploiting steady-state somatosensory-evoked potentials: a literature review”. In: *J. Neural Eng.* 18.5 (Oct. 2021), p. 051003.
212. Anh Huy Phan and Andrzej Cichocki. “Tensor decompositions for feature extraction and classification of high dimensional datasets”. In: *Nonlinear Theory and Its Applications, IEICE* 1.1 (2010), pp. 37–68.
213. Gabriel Pires, Urbano Nunes, and Miguel Castelo-Branco. “GIBS block speller: Toward a gaze-independent P300-based BCI”. In: *2011 Annual International Conference of the IEEE Engineering in Medicine and Biology Society*. ISSN: 1558-4615. Aug. 2011, pp. 6360–6364.
214. K. Polet et al. “Video-oculography in multiple sclerosis: Links between oculomotor disorders and brain magnetic resonance imaging (MRI)”. en. In: *Multiple Sclerosis and Related Disorders* 40 (May 2020), p. 101969.
215. John Polich. “Updating P300: An integrative theory of P3a and P3b”. en. In: *Clin. Neurophysiol.* 118.10 (Oct. 2007), pp. 2128–2148.
216. Alex Pollock et al. “Interventions for disorders of eye movement in patients with stroke”. In: *Cochrane Db. Syst. Rev.* 10 (2011).
217. Sashank Prasad and Steven L. Galetta. “Eye Movement Abnormalities in Multiple Sclerosis”. English. In: *Neurol. Clin.* 28.3 (Aug. 2010), pp. 641–655.
218. Aina Puce et al. “P3 latency jitter assessed using 2 techniques. I. Simulated data and surface recordings in normal subjects”. en. In: *Electroencephalography and Clinical Neurophysiology/Evoked Potentials Section* 92.4 (July 1994), pp. 352–364.
219. R. Quian Quiroga and H Garcia. “Single-trial event-related potentials with wavelet denoising”. en. In: *Clin. Neurophysiol.* 114.2 (Feb. 2003), pp. 376–390.
220. Sarunas Raudys and Robert P. W. Duin. “Expected classification error of the Fisher linear classifier with pseudo-inverse covariance matrix”. en. In: *Pattern Recogn. Lett.* 19.5 (Apr. 1998), pp. 385–392.
221. Christoph Reichert et al. “Decoding the covert shift of spatial attention from electroencephalographic signals permits reliable control of a brain-computer interface”. en. In: *J. Neural Eng.* 17.5 (Oct. 2020), p. 056012.

222. Aya Rezeika et al. “Brain–computer interface spellers: A review”. In: *Brain sciences* 8.4 (Mar. 2018), p. 57.
223. A. Riccio et al. “Eye-gaze independent EEG-based braincomputer interfaces for communication”. en. In: *J. Neural Eng.* 9.4 (July 2012), p. 045001.
224. Angela Riccio et al. “Attention and P300-based BCI performance in people with amyotrophic lateral sclerosis”. In: *Frontiers in human neuroscience* 7 (2013), p. 732.
225. Bertrand Rivet et al. “xDAWN algorithm to enhance evoked potentials: application to brain–computer interface”. In: *IEEE Trans. Biomed. Eng.* 56.8 (Aug. 2009), pp. 2035–2043.
226. Ricardo Ron-Angevin et al. “Impact of speller size on a visual P300 brain-computer interface (BCI) system under two conditions of constraint for eye movement”. In: *Computational Intelligence and Neuroscience* 2019 (2019).
227. Athanasios A Rontogiannis, Eleftherios Kofidis, and Paris V Giampouras. “Block-term tensor decomposition: Model selection and computation”. In: *IEEE J. Sel. Topics Signal Process.* 15.3 (2021), pp. 464–475.
228. Fiona J. Rowe et al. “High incidence and prevalence of visual problems after acute stroke: An epidemiology study with implications for service delivery”. en. In: *PLOS ONE* 14.3 (Mar. 2019), e0213035.
229. Tomasz M Rutkowski and Hiromu Mori. “Tactile and bone-conduction auditory brain computer interface for vision and hearing impaired users”. In: *J. Neurosci. Methods* 244 (Apr. 2015), pp. 45–51.
230. Simanto Saha and Mathias Baumert. “Intra-and inter-subject variability in EEG-based sensorimotor brain computer interface: a review”. In: *Frontiers in computational neuroscience* 13 (Jan. 2020), p. 87.
231. C. W. N. Saville et al. “COMT Val158Met genotype is associated with fluctuations in working memory performance: converging evidence from behavioural and single-trial P3b measures”. en. In: *NeuroImage* 100 (Oct. 2014), pp. 489–497.
232. Sulamith Schaeff et al. “Exploring motion VEPs for gaze-independent communication”. en. In: *J. Neural Eng.* 9.4 (July 2012), p. 045006.
233. Juliane Schäfer and Korbinian Strimmer. “An empirical Bayes approach to inferring large-scale gene association networks”. In: *Bioinformatics* 21.6 (Oct. 2004), pp. 754–764.
234. Gerwin Schalk and Eric C. Leuthardt. “Brain-computer interfaces using electrocorticographic signals”. eng. In: *IEEE reviews in biomedical engineering* 4 (2011), pp. 140–154.
235. Martijn Schreuder et al. “User-centered design in brain-computer interfaces – A case study”. In: *Artificial intelligence in medicine* 59.2 (Oct. 2013), pp. 71–80.
236. Perrine Séguin, Emmanuel Maby, and Jérémie Mattout. “Why BCIs work poorly with the patients who need them the most?” In: *arXiv preprint arXiv:2302.06312* (2023).

237. Eric W Sellers et al. “A P300 event-related potential brain–computer interface (BCI): the effects of matrix size and inter stimulus interval on performance”. In: *Biological psychology* 73.3 (Oct. 2006), pp. 242–252.
238. Alessandro Serra, Clara G. Chisari, and Manuela Matta. “Eye Movement Abnormalities in Multiple Sclerosis: Pathogenesis, Modeling, and Treatment”. In: *Front. Neurol.* 9 (2018).
239. M Severens et al. “Comparing tactile and visual gaze-independent brain–computer interfaces in patients with amyotrophic lateral sclerosis and healthy users”. In: *Clinical neurophysiology* 125.11 (Nov. 2014), pp. 2297–2304.
240. Mehdi Shokoueinejad et al. “Progress in the field of micro-electrocorticography”. In: *Micromachines* 10.1 (Jan. 2019), p. 62.
241. Thomas Skov et al. “Automated alignment of chromatographic data”. en. In: *J. Chemometr.* 20.11-12 (2006), pp. 484–497.
242. Eimear Smith and Mark Delargy. “Locked-in syndrome”. In: *BMJ* 330.7488 (2005), pp. 406–409.
243. Jordi Sole-Casals et al. “Brain-computer interface with corrupted EEG data: a tensor completion approach”. In: *Cogn. Comput.* 10.6 (July 2018), pp. 1062–1074.
244. Jan Sosulski and Michael Tangermann. “Introducing Block-Toeplitz Covariance Matrices to Remaster Linear Discriminant Analysis for Event-related Potential Brain-computer Interfaces”. In: *arXiv:2202.02001 [cs, q-bio]* (Feb. 2022). arXiv: 2202.02001.
245. Antoine Souloumiac and Bertrand Rivet. “Improved estimation of EEG evoked potentials by jitter compensation and enhancing spatial filters”. In: *2013 IEEE International Conference on Acoustics, Speech and Signal Processing*. IEEE. IEEE, May 2013, pp. 1222–1226.
246. Giovanni Sparacino et al. “A Bayesian approach to estimate evoked potentials”. en. In: *Comput. Methods Programs Biomed.* 68.3 (June 2002), pp. 233–248.
247. Charles Stein. “Inadmissability of the usual estimator for the mean of a multivariate normal distribution”. en. In: *Volume 1 Contribution to the Theory of Statistics*. University of California Press, Sept. 1956, pp. 197–206.
248. Nicholas A Steinmetz et al. “Neuropixels 2.0: A miniaturized high-density probe for stable, long-term brain recordings”. In: *Sci* 372.6539 (Apr. 2021), eabf4588.
249. Mark Stokes and Eelke Spaak. “The Importance of Single-Trial Analyses in Cognitive Neuroscience”. en. In: *Trends Cogn. Sci.* 20.7 (July 2016), pp. 483–486.
250. Qingyu Sun et al. “A 120-target brain-computer interface based on code-modulated visual evoked potentials”. In: *J. Neurosci. Methods* 375 (June 2022), p. 109597.
251. Yusuke Takeda, Kentaro Yamanaka, and Yoshiharu Yamamoto. “Temporal decomposition of EEG during a simple reaction time task into stimulus-

and response-locked components". en. In: *NeuroImage* 39.2 (Jan. 2008), pp. 742–754.
252. Michael Tangermann et al. "Review of the BCI competition IV". In: *Frontiers in neuroscience* 6 (2012), p. 55.
253. Salma Tayeb et al. "Efficient detection of P300 using Kernel PCA and support vector machine". In: *2014 Second World Conference on Complex Systems (WCCS)*. Nov. 2014, pp. 17–22.
254. The BCI Society. https://bcisociety.org/bci-definition. Accessed on 2024-09-09.
255. Thomas Thiery et al. "Decoding the Locus of Covert Visuospatial Attention from EEG Signals". en. In: *PLOS ONE* 11.8 (Aug. 2016), e0160304.
256. David E Thompson, Seth Warschausky, and Jane E Huggins. "Classifier-based latency estimation: a novel way to estimate and predict BCI accuracy". In: *Journal of neural engineering* 10.1 (Dec. 2012), p. 016006.
257. Margaret C Thompson. "Critiquing the concept of BCI illiteracy". In: *Science and engineering ethics* 25.4 (Aug. 2019), pp. 1217–1233.
258. A. Roger D. Thornton. "Evaluation of a technique to measure latency jitter in event-related potentials". en. In: *J. Neurosci. Methods* 168.1 (Feb. 2008), pp. 248–255.
259. Jun Tong et al. "Linear shrinkage estimation of covariance matrices using low-complexity cross-validation". en. In: *Signal Process.* 148 (July 2018), pp. 223–233.
260. L. Tonin et al. "An online EEG BCI based on covert visuospatial attention in absence of exogenous stimulation". en. In: *J. Neural Eng.* 10.5 (Aug. 2013), p. 056007.
261. Edgar P Torres et al. "EEG-based BCI emotion recognition: A survey". In: *Sensors-basel.* 20.18 (Sept. 2020), p. 5083.
262. Jeanne Townsend, Eric Courchesne, and Terrence J Sejnowski. "Analyzing and visualizing single-trial event-related potentials". In: *Advances in neural information processing systems* 11 (1999), pp. 118–24.
263. Mathias. S. Treder, N. M. Schmidt, and B. Blankertz. "Gaze-independent braincomputer interfaces based on covert attention and feature attention". en. In: *J. Neural Eng.* 8.6 (Oct. 2011), p. 066003.
264. Matthias S Treder and Benjamin Blankertz. "(C)overt attention and visual speller design in an ERP-based brain-computer interface". In: *Behavioral and brain functions* 6.1 (2010), pp. 1–13.
265. Matthias S. Treder et al. "The LDA beamformer: Optimal estimation of ERP source time series using linear discriminant analysis". In: *NeuroImage* 129 (2016), pp. 279–291.
266. D. Troxler. "Ueber das Verschwinden gegebener Gegenstande innerhalb unseres Gesichtskreises". In: *Ophthalmologische Bibliothek* 2 (1804), pp. 1–119.
267. Wilson A Truccolo et al. "Trial-to-trial variability of cortical evoked responses: implications for the analysis of functional connectivity". In: *Clinical neurophysiology* 113.2 (Feb. 2002), pp. 206–226.

268. Pham Dinh Tuan et al. “Variable latencies of noisy signals: estimation and testing in brain potential data”. In: *Biometrika* 74.3 (1987), pp. 525–533.
269. Jose Antonio Urigüen and Begoña Garcia-Zapirain. “EEG artifact removalstate-of-the-art and guidelines”. In: *Journal of neural engineering* 12.3 (Apr. 2015), p. 031001.
270. Arne Van Den Kerchove et al. “Block-Term Tensor Discriminant Analysis for Brain-Computer Interfacing”. In: *Journal of Neural Engineering* (submitted).
271. Arne Van Den Kerchove et al. “Classification of Event-Related Potentials with Regularized Spatiotemporal LCMV Beamforming”. en. In: *Applied Sciences* 12.6 (Jan. 2022), p. 2918.
272. Arne Van Den Kerchove et al. “Correcting for ERP latency jitter improves gaze-independent BCI decoding”. In: *J. Neural Eng.* 21.4 (July 2024), p. 046013.
273. Barry D Van Veen et al. “Localization of brain electrical activity via linearly constrained minimum variance spatial filtering”. In: *IEEE Transactions on biomedical engineering* 44.9 (1997), pp. 867–880.
274. Marijn Van Vliet et al. “Single-trial erp component analysis using a spatiotemporal lcmv beamformer”. In: *IEEE Trans. Biomed. Eng.* 63.1 (2015), pp. 55–66.
275. M Isabel Vanegas, Annabelle Blangero, and Simon P Kelly. “Exploiting individual primary visual cortex geometry to boost steady state visual evoked potentials”. In: *Journal of neural engineering* 10.3 (2013), p. 036003.
276. Luká Vaeka. “Evaluation of convolutional neural networks using a large multi-subject P300 dataset”. In: *Biomed. Signal Proces.* 58 (Apr. 2020), p. 101837.
277. Jacques J Vidal. “Toward direct brain-computer communication”. In: *Annual review of Biophysics and Bioengineering* 2.1 (June 1973), pp. 157–180.
278. Pauli Virtanen et al. “SciPy 1.0: Fundamental Algorithms for Scientific Computing in Python”. In: *Nat. Methods* 17 (2020), pp. 261–272.
279. Marijn van Vliet and Riitta Salmelin. “Post-hoc modification of linear models: Combining machine learning with domain information to make solid inferences from noisy data”. en. In: *NeuroImage* 204 (Jan. 2020), p. 116221.
280. Marjolein van der Waal et al. “Introducing the tactile speller: an ERP-based brain–computer interface for communication”. In: *J. Neural Eng.* 9.4 (July 2012), p. 045002.
281. Fang Wang et al. “Re-Examination of Chinese Semantic Processing and Syntactic Processing: Evidence from Conventional ERPs and Reconstructed ERPs by Residue Iteration Decomposition (RIDE)”. en. In: *PLOS ONE* 10.1 (Jan. 2015), e0117324.

282. Kongming Wang, Henri Begleiter, and Bernice Porjesz. “Warp-averaging event-related potentials”. In: *Clin. Neurophysiol.* 112.10 (2001), pp. 1917–1924.
283. Yijing Wang et al. “Spatial localization in target detection based on decoding N2pc component”. en. In: *J. Neurosci. Methods* 369 (Mar. 2022), p. 109440.
284. Zhisong Wang et al. “Single-trial evoked potential estimation using wavelets”. In: *Comput. Biol. Med.* 37.4 (Apr. 2007), pp. 463–473.
285. Nikolaus Weiskopf et al. “Principles of a brain-computer interface (BCI) based on real-time functional magnetic resonance imaging (fMRI)”. In: *IEEE transactions on biomedical engineering* 51.6 (June 2004), pp. 966–970.
286. Karl Werner, Magnus Jansson, and Petre Stoica. “On estimation of covariance matrices with Kronecker product structure”. In: *IEEE Trans. Signal Process.* 56.2 (Feb. 2008), pp. 478–491.
287. Ami Wiesel. “Geodesic Convexity and Covariance Estimation”. In: *IEEE Trans. Signal Process.* 60.12 (Dec. 2012), pp. 6182–6189.
288. Ami Wiesel. “On the convexity in Kronecker structured covariance estimation”. In: *2012 IEEE Statistical Signal Processing Workshop (SSP)*. ISSN: 2373-0803. IEEE. IEEE, Aug. 2012, pp. 880–883.
289. Francis R. Willett et al. “High-performance brain-to-text communication via handwriting”. en. In: *Nat* 593.7858 (May 2021), pp. 249–254.
290. Petter Wirfält and Magnus Jansson. “On Toeplitz and Kronecker structured covariance matrix estimation”. In: *2010 IEEE Sensor Array and Multichannel Signal Processing Workshop*. IEEE. IEEE, Oct. 2010, pp. 185–188.
291. Benjamin Wittevrongel and Marc M. Van Hulle. “Frequency- and Phase Encoded SSVEP Using Spatiotemporal Beamforming”. en. In: *PLOS ONE* 11.8 (Aug. 2016), e0159988.
292. Benjamin Wittevrongel and Marc M Van Hulle. “Faster p300 classifier training using spatiotemporal beamforming”. In: *International journal of neural systems* 26.03 (Apr. 2016), p. 1650014.
293. Benjamin Wittevrongel and Marc M. Van Hulle. “Spatiotemporal Beamforming: A Transparent and Unified Decoding Approach to Synchronous Visual Brain-Computer Interfacing”. In: *Front. Neurosci.* 11 (Nov. 2017), p. 630.
294. Benjamin Wittevrongel, Elia Van Wolputte, and Marc M. Van Hulle. “Code-modulated visual evoked potentials using fast stimulus presentation and spatiotemporal beamformer decoding”. en. In: *Sci. Rep.* 7.1 (Nov. 2017), p. 15037.
295. Benjamin Wittevrongel et al. “Optically pumped magnetometers for practical MEG-based brain-computer interfacing”. In: *Brain-Computer Interface Research: A State-of-the-Art Summary 10* (2021), pp. 35–46.
296. Jonathan R Wolpaw, Jose Del R Millan, and Nick F Ramsey. “Brain-computer interfaces: Definitions and principles”. In: *Handbook of clinical neurology* 168 (2020), pp. 15–23.

297. Jonathan R Wolpaw et al. “BCI meeting 2005-workshop on signals and recording methods”. In: *IEEE Transactions on neural systems and rehabilitation engineering* 14.2 (June 2006), pp. 138–141.
298. Jonathan R Wolpaw et al. “Independent home use of a brain-computer interface by people with amyotrophic lateral sclerosis”. In: *Neurology* 91.3 (2018), e258–e267.
299. Jonathan R. Wolpaw et al. “Brain–computer interfaces for communication and control”. In: *Clinical neurophysiology* 113.6 (2002), pp. 767–791.
300. Dong-Ok Won et al. “Motion-Based Rapid Serial Visual Presentation for Gaze-Independent Brain-Computer Interfaces”. In: *IEEE Trans. Neural Syst. Rehabil. Eng.* 26.2 (Feb. 2018), pp. 334–343.
301. Charles D Woody. “Characterization of an adaptive filter for the analysis of variable latency neuroelectric signals”. In: *Medical and biological engineering* 5.6 (1967), pp. 539–554.
302. Xiaolong Wu et al. “A Review of Motor Brain-Computer Interfaces using Intracranial Electroencephalography based on Surface Electrodes and Depth Electrodes”. In: *IEEE Trans. Neural Syst. Rehabil. Eng.* 32 (2024), pp. 2408–2431.
303. Lei Xie et al. *Regularized Estimation of Kronecker-Structured Covariance Matrix*. 2021.
304. Lichao Xu et al. “Review of brain encoding and decoding mechanisms for EEG-based brain–computer interface”. In: *Cognitive neurodynamics* 15.4 (Apr. 2021), pp. 569–584.
305. Minpeng Xu et al. “Fast detection of covert visuospatial attention using hybrid N2pc and SSVEP features”. In: *Journal of neural engineering* 13.6 (2016), p. 066003.
306. Wei Xu et al. “Improving the performance of a gaze independent P300-BCI by using the expectancy wave”. en. In: *J. Neural Eng.* 19.2 (Apr. 2022), p. 026036.
307. Shuicheng Yan et al. “Discriminant analysis with tensor representation”. In: *2005 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR’05)*. Vol. 1. ISSN: 1063-6919. June 2005, 526–532 vol. 1.
308. Erwei Yin et al. “An auditory-tactile visual saccade-independent P300 brain–computer interface”. In: *International journal of neural systems* 26.01 (Jan. 2016), p. 1650001.
309. Gang Yin et al. “A multi-component decomposition algorithm for event-related potentials”. en. In: *J. Neurosci. Methods* 178.1 (Mar. 2009), pp. 219–227.
310. Thorsten O Zander and Christian Kothe. “Towards passive brain–computer interfaces: applying brain–computer interface technology to human–machine systems in general”. In: *Journal of neural engineering* 8.2 (Mar. 2011), p. 025005.
311. Renato Zanetti et al. “Real-time EEG-based cognitive workload monitoring on wearable devices”. In: *IEEE transactions on biomedical engineering* 69.1 (Jan. 2021), pp. 265–277.

312. Paolo Zanini et al. “Transfer learning: A Riemannian geometry framework with applications to brain–computer interfaces”. In: *IEEE Trans. Biomed. Eng.* 65.5 (May 2017), pp. 1107–1116.
313. Boyang Zhang, Zongtan Zhou, and Jing Jiang. “A 36-class bimodal ERP brain-computer interface using location-congruent auditory-tactile stimuli”. In: *Brain sciences* 10.8 (Aug. 2020), p. 524.
314. Dan Zhang et al. “An independent braincomputer interface using covert non-spatial visual selective attention”. en. In: *J. Neural Eng.* 7.1 (Jan. 2010), p. 016010.
315. Jun Zhang. “Decomposing stimulus and response component waveforms in ERP”. en. In: *J. Neurosci. Methods* 80.1 (Mar. 1998), pp. 49–63.
316. Hairong Zheng et al. “The emergence of functional ultrasound for noninvasive brain–computer interface”. In: *Research* 6 (2023), p. 0200.
317. Alyssa Hillary Zisk et al. “Improving longitudinal P300-BCI performance for people with ALS using a data augmentation and jitter correction approach”. In: *Brain-Computer Interfaces* 9.1 (Jan. 2022), pp. 49–66.
318. Alyssa Hillary Zisk et al. “P300 latency jitter and its correlates in people with amyotrophic lateral sclerosis”. In: *Clin. Neurophysiol.* 132.2 (Feb. 2021), pp. 632–642.
319. A. Zoumpoulaki et al. “Latency as a region contrast: Measuring ERP latency differences with Dynamic Time Warping”. en. In: *Psychophysiology* 52.12 (2015), pp. 1559–1576.