Solving the Cocktail Party Problem

Spectral Estimation and Linear Modelling

Keding, Oskar

2024

Document Version:
Publisher's PDF, also known as Version of record

LUND UNIVERSITY

PO Box 117
221 00 Lund
+46 46-222 00 00

– CENTRUM SCIENTIARUM MATHEMATICARUM –

Solving the Cocktail Party Problem

Spectral Estimation and Linear Modelling

OSKAR KEDING

Lund University
Faculty of Engineering
Centre for Mathematical Sciences
Mathematical Statistics

Solving the Cocktail Party Problem
Spectral Estimation and Linear Modelling

Solving the Cocktail Party Problem
Spectral Estimation and Linear Modelling

by Oskar Keding

Licentiate Thesis
Thesis advisors: Maria Sandsten, Emina Alickovic, Martin A. Skoglund,
Faculty opponent: Senior Assoc. Prof. Fredrik Lindsten

To be presented, with the permission of the Faculty of Engineering of Lund University, for public criticism in the room MH:309A at the Centre for Mathematical Sciences

Mathematical Statistics on Tuesday, the 28th of May 2024 at 13:15.

DOKUMENTDATABLAD enl SIS 61 41 21

Organization **LUND UNIVERSITY** Centre for Mathematical Sciences Mathematical Statistics Box 118 SE–221 00 Lund, Sweden	Document name **Licentiate thesis**
	Date of presentation 2024-05-28
Author(s) Oskar Keding	Sponsoring organization The research in this thesis was supported by the ELLIIT strategic research environment.

Title and subtitle

Solving the Cocktail Party Problem Spectral Estimation and Linear Modelling

Abstract

By measuring brain activity, through techniques such as electroencephalography (EEG), it is possible to decode which sound source a person is listening to, called auditory attention decoding (AAD). This can either be done investigating the relation between speech sources and corresponding brain responses over time, or by discriminatively estimating directions to which auditory attention is focused. Spectral, temporal and spatial information are all useful and each essential for understanding how the brain processes sounds in a multi-talker scenario. Key challenges with EEG analysis are high levels of noise from various sources, as well as utilizing methods that infer onto the processing happening in the brain. Therefore, the work part of this thesis focuses on linear and fairly non-complex methods. This thesis explores spectral estimation based methods and linear modelling methods and their application to AAD. The linear correlation measure of coherence is investigated and improved for use in EEG and AAD, showing that it can differ between attended speech and ignored speech. The commonly applied method of common spatial patterns (CSP) within EEG-data is employed specifically for AAD. We are able to show how different CSP algorithms perform within the field of AAD, and that performance for CSP carries over from decoding auditory attention of individuals with normal hearing compared to individuals with hearing impairment. Independent Component Analysis-based (ICA) methods of removing noise components of EEG data are evaluated for AAD on a dataset with participants hearing impaired. Automatic noise cleaning methods are shown to perform equally as well as the traditional manual method on the given dataset. Finally, a phase estimation technique for transient components based on spectrogram reassignment is developed, which can estimate phase difference of signal components in multi-channel measurements such as EEG. Using the methods described, it is possible to draw interesting conclusions in the field of AAD. However, future work entails further improvement and exploration of useful methods for analysis of the system that is the hearing brain.

Key words

Classification system and/or index terms (if any)

Supplementary bibliographical information		Language English
ISSN and key title 1404-028X		ISBN 978-91-8104-068-5 (print) 978-91-8104-069-2 (pdf)
Recipient's notes	Number of pages 132	Price
	Security classification	

Signature ______________________ Date 2024-05-02

Solving the Cocktail Party Problem

Spectral Estimation and Linear Modelling

by Oskar Keding

LUND UNIVERSITY

Funding information: The research in this thesis was supported by the ELLIIT strategic research environment.

Mathematical Statistics
Centre for Mathematical Sciences
Lund University
Box 118
SE-221 00 Lund
Sweden

www.maths.lu.se

Licentiate Thesis in Mathematical Sciences 2024:1
ISSN: 1404-028X

ISBN: 978-91-8104-068-5 (print)
ISBN: 978-91-8104-069-2 (pdf)
LUTFMS-2020-2024

Printed in Sweden by Media-Tryck, Lund 2024

Jag har stannat vid vägen
För att vila mig ett tag
Och blev fångad i det gränsland
Som förenar natt och dag
Py Bäckman

Abstract

By measuring brain activity, through techniques such as electroencephalography (EEG), it is possible to decode which sound source a person is listening to, called auditory attention decoding (AAD). This can either be done investigating the relation between speech sources and corresponding brain responses over time, or by discriminatively estimating directions to which auditory attention is focused. Spectral, temporal and spatial information are all useful and each essential for understanding how the brain processes sounds in a multi-talker scenario. Key challenges with EEG analysis are high levels of noise from various sources, as well as utilizing methods that infer onto the processing happening in the brain. Therefore, the work part of this thesis focuses on linear and fairly non-complex methods. This thesis explores spectral estimation based methods and linear modelling methods and their application to AAD. The linear correlation measure of coherence is investigated and improved for use in EEG and AAD, showing that it can differ between attended speech and ignored speech. The commonly applied method of common spatial patterns (CSP) within EEG-data is employed specifically for AAD. We are able to show how different CSP algorithms perform within the field of AAD, and that performance for CSP carries over from decoding auditory attention of individuals with normal hearing compared to individuals with hearing impairment. Independent Component Analysis-based (ICA) methods of removing noise components of EEG data are evaluated for AAD on a dataset with participants hearing impaired. Automatic noise cleaning methods are shown to perform equally as well as the traditional manual method on the given dataset. Finally, a phase estimation technique for transient components based on spectrogram reassignment is developed, which can estimate phase difference of signal components in multi-channel measurements such as EEG. Using the methods described, it is possible to draw interesting conclusions in the field of AAD. However, future work entails further improvement and exploration of useful methods for analysis of the system that is the hearing brain.

Publications

Publications concerning the work of this thesis have been made as follows:

A **Improved coherence measure for EEG-based speech tracking in hearing impaired listeners**
Oskar Keding, Emina Alickovic, Martin A. Skoglund, Maria Sandsten
Submitted April 2024 to Frontiers in Neuroscience (in review)

B **Coherence Expectation Minimisation and Combining Weighted Multitaper Estimates**
Oskar Keding, Maria Åkesson, Maria Sandsten
31st European Signal Processing Conference (EUSIPCO), Helsinki, Finland, 2023, pp. 1993-1997

C **A Comparison of Common Spatial Patterns Algorithms for Auditory Attention Decoding**
Oskar Keding
Manuscript

D **Effect of Independent Component Artifact Rejection on EEG-Based Auditory Attention Decoding**
Oskar Keding, Johanna Wilroth, Martin A. Skoglund, Emina Alickovic
Submitted to the 32nd European Signal Processing Conference (EUSIPCO), Lyon, France, 2024

E **Robust Phase Difference Estimation of Transients in High Noise Levels**
Oskar Keding, Maria Sandsten
30th European Signal Processing Conference (EUSIPCO), Belgrade, Serbia, 2022, pp. 2271-2275

Publication not included in this thesis:

A **Coherence Estimation Tracks Auditory Attention in Listeners with Hearing Impairment**
Oskar Keding, Emina Alickovic, Martin A. Skoglund, Maria Sandsten
Proc. INTERSPEECH 2023, 2023, pp. 5162-5166

B **The Closed-Form Multitaper Spectrogram of a Gaussian Enveloped Linear Chirp Signal**
Markus Ydreskog, Oskar Keding, Maria Sandsten
Submitted October 2023 to IEEE Transactions on Signal Processing (in review)

C **Highly Accurate and Noise-Robust Phase Delay Estimation using Multitaper Reassignment**
Maria Åkesson, Oskar Keding, Maria Sandsten
31st European Signal Processing Conference (EUSIPCO), Helsinki, Finland, 2023, pp. 1763-1767

Acknowledgements

I would like to express gratitude to my supervisor Maria Sandsten for unwavering support and guidance during these three years. Furthermore, I would like to thank my two co-supervisors Emina Alickovic and Martin A. Skoglund for always contributing with valuable knowledge and experience, as well as providing guidance within the research field. Also, thanks to the administrative staff for all help in practical matters.

Additionally, thanks also to Johanna Wilroth, who is always ready to discuss matters PhD students alike. Thanks to David Sundström, for inspiration and spontaneous discussions during work days. Finally, the largest of thanks to my girlfriend Jonna, my father Mats and my sister Astrid who are always there for me.

Funding

Work part of this licentiate has been funded by the ELLIIT strategic research environment, as part of project C04 with project name Brain-Based Monitoring of Sound.

Popular summary

When we are kids we are taught that the ear is the organ that does the hearing in our body. In reality, this is not the case. Our perception of what we hear is a result of cognitive processing in the brain. In everyday life the auditory experiences are highly varied. One particularly straining scene is when multiple speakers are present in the same room as the person you are trying to listen to. This can happen in meetings at work, during morning commutes or a cocktail-party. The last example has inspired the name of the natural ability of our auditory system to subliminally focus on a particular speaker: the *cocktail party effect*.

Although picking out single speakers in noisy environments can be straining, normal hearing individuals are able to solve the cocktail party problem. However, for people with deteriorating hearing and hearing aid users this ability can be heavily encumbered. Tiredness and wanting to disconnect is common for hearing aid users, where the user turns the device off after being in a multiple speaker situation for too long. This is because the cocktail party effect does not come as naturally. Perhaps you even have experienced this with an older relative.

This thesis aims to develop methods for looking at measurements of brain activity when people are attending specific people in multi-speaker scenarios. With novel analysis methods it is possible to track the time and location of brain activity, in relation to the speech heard. In addition to this, it is also possible to work out which speaker a certain person is listening to.

The first paper of this thesis, Paper A, looks at methods for detecting connections between the speech heard and concurrent measured brain signals. These connections are different for attended and ignored sounds, which makes it possible to work out where a listeners attention is targeted. Paper B tackles some of the common problems with this set-up, and makes method improvements. Paper C utilizes spatial information of brain measurements to work out whether listeners are attending speech coming from the left or the right of the listener. Paper D investigates how natural sources of disturbances in brain signals, such as eye and muscle movements, affect the tracking of auditory attention. Paper E presents ways of extracting time lags between two different signals.

Looking forward, there are many challenges before these models can be applied to help hearing aid users. For example, there is a further need for methods that can track how auditory attention dynamically changes over time, as well as taking into account all the hard-to-predict peculiarities of everyday life.

Contents

Chapter 1

Introduction to Hearing and Auditory Attention

Most people are taught that the ear is the organ that does the hearing in our body. In reality, which has been known for a very long time, this is not the case. Our perception of what we hear is a result of cognitive processing in the brain as well as combination of multiple sensory networks. The biological process from detection of the physical sound to hearing and understanding meaning is complicated to say the least.

In everyday life the auditory environments experiences are highly varied. One particularly straining scene is when multiple speaker are present in the same room as the person one is trying to listen to. This can happen in meetings at work, during morning commutes or a cocktail-party. The last example has inspired the name of the natural ability of our auditory system to subliminally focus on a particular speaker: the *cocktail party effect* [6].

Although picking out single speakers in noisy environments can be straining, normal hearing individuals are able to solve the cocktail party problem. However, for people with deteriorating hearing and hearing aid (HA) users this ability can be heavily encumbered. To intuitively understand the problem, one can imagine a situation that most people have been in nowadays. In an online video call meeting, it is often the case that multiple people on different ends of the call start talking at the same time. If you try distinguishing a certain speaker from the others, this will require a high amount of effort compared to a normal meeting. This is due to all speakers are jumbled in the speakers of the computer, in an unnatural way. At the end of the meeting, you can feel very tired and wanting to disconnect. This tiredness is also quite common for HA users, were the user turns the hearing device off after being in a multiple speaker situation for too long.

The goal of auditory attention decoding is to help individuals with this problem. By working out which sound a HA user is focusing on, one can attenuate other disruptive sounds in the HA, helping the wearer. This requires continuously measuring brain activity, relating this to the sounds of which the individual hears and work out which source to enhance. To decipher this cognitive processing, understanding the processing of sounds in the auditory system is important.

1 The cocktail party effect

The cocktail party problem, first named by Cherry in the 1950s [6], was formalized as the task to solve by humans to recognize what one person is saying when others are speaking at the same time. Multiple aspects can alleviate or exacerbate the difficulty of this task. If the attended speech is relatively soft compared to other speakers, then focusing becomes harder. The separation is harder if speakers are similar, for example in fundamental frequency, accent or talking speed. Although spatial clues and information of the sound is useful for separating speakers, separation is not purely a result of these types of inferences. Interestingly, although it is harder to separate speakers that share a direction to the listener, it is still very much possible [16].

The auditory cortex is involved in the processing of sounds and speech [31]. In particular this involvement can be measured during listening in multi-talker scenarios [24]. In the primary auditory cortex both attended and ignored sounds can be tracked. Moving from the primary auditory cortex, outwards to other parts of the auditory cortex, the observed connection to sounds is reduced to attended speech alone [37]. In auditory processing of speech in a multi-talker scenario, tracking is stronger in the left hemisphere compared to the right [30]. This is true although there certainly is a bilateral aspect to hearing and listening processes in the brain.

Auditory attention decoding (AAD) is the process and research field of determining which, out of multiple concurrent speakers, a certain listener is attending. This is done by measuring brain signatures in some way. Multiple analogue approaches are available for AAD. Attention can be spatially decoded, working out the locus of attention, i.e. which direction the person is devoting mental resources to interpret [14, 34, 7]. Attention decoding can also focus more about the connection between measured brain responses and the content of what is spoken in each speech source. Significant research has been made relating both acoustic and linguistic speech features to corresponding brain responses [1, 18, 2, 9].

2 Harder cocktail party problem for hearing impaired

For individuals with hearing impairment, the cocktail party problem is harder. Due to partial hearing loss, the auditory pathway deteriorates in its ability of processing sounds in noisy environments [23]. Exactly what mechanisms triggers this decline is not fully understood. Both the ability of individuals with hearing impairment to discern sounds, as well as the fusing of sound information binaurally can play part. Hearing loss is also in and of itself an early indicator for both depression and dementia [19, 22], another reason to alleviate the perceptual barriers for people with hearing impairment.

HAs help individuals in selective attention in noisy environments. However, the mechanisms in play can be affected. Since HAs themselves augment the listening experience for the user, this adds a degree of freedom in estimation. The same algorithms that can decode attention of normal hearing individuals, may fail when applied to HA users. Measuring how HAs help users in cocktail party scenarios is important, as this can be used clinically. HAs have to be fitted which means setting appropriate enchancement levels of sounds at different frequencies. Additionally, modern HAs are employed with noise reduction schemes and different settings for different audio environments. Objective evaluation of performance of settings and algorithms in HAs are constantly sought after to help users in systematic ways [13].

Ideally, future HAs fused with some brain monitoring system can decode the attention of a listener and attenuate disturbing sounds. A schematic of this is shown in Figure 1.1. This is simplification of a potential system. For example, in any real application one has to be able to let enough of other speakers through so users can change attention naturally in a conversion. Also, the HA needs to let through salient or potentially threatening sounds.

3 The path from sound to neuronal activity

Sound is a series of changes in sound pressure that propagate through all mediums [29]. These mechanical pressure waves travel through multiple biological structures before they are converted into signals in the brain. These structures can be seen in the schematic of the human ear shown in Fig. 1.2. Vibrations in air reach the outer ear and travel into the ear canal. At the end of the ear canal, the sound waves push on the eardrum. The eardrum separates the outer ear and the middle ear. In the middle ear lies three bones, called the *malleus*, *incus* and *stapes*. The eardrum pushes on the malleus and the pressure from sound is propagated through the bones. At the end of the stapes the concentrated pressure (due to a very small surface area) is transferred into the *cochlea*.

When sound enters the fluid-filled cochlea, it enters one side of it [29]. The cochlea is

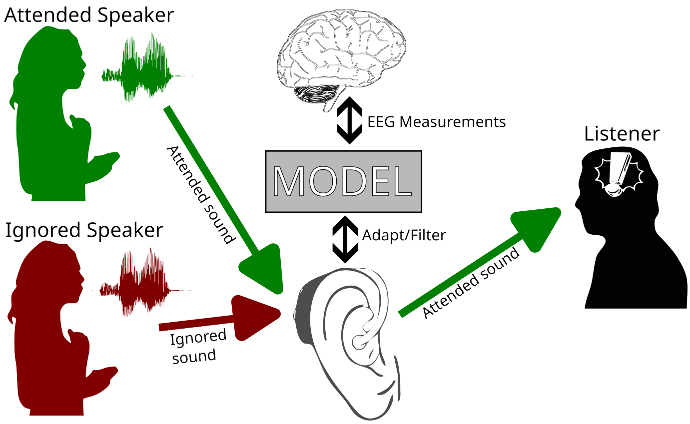

Figure 1.1: A conceptual future hearing aid that concurrently measures brain activity as a wearer (in black) listens to an attended speaker (in green) when a disturbing speaker (in red) is present. By analyzing the brain, the hearing aid adapts and only lets through sound from the speaker the listener is paying attention to.

split in two by the *basilar membrane*, which is a tapered wall that separates two tubes that make up the spiral-shaped cochlea. As pressure must be dispersed, sound waves move from one side of the basilar membrane to the other. Because of the tapered structure, different frequencies of sounds will take different paths, separating themselves along the basilar membrane. Within the basilar membrane, these mechanical oscillations of different frequencies are extracted and converted into electrical signals. These electrical signals are then carried by neurons through the auditory nerve to the central auditory pathway. These signals are frequency dependant as well as amplitude dependant of the oscillations caused by sounds.

As the fibres of the auditory nerve leaves each left and right cochlea they enter the corresponding cochlear nucleus, where they bifurcate. Roughly speaking, the nerves travel to a range of different nuclei that make up the cochlear nuclei, before they travel to various parts of the ascending auditory pathway. Signals reach the superior olivary complex, the nuclei of the lateral lemniscus, the inferior colliculus and the medial geniculatae body before arriving to the auditory cortex. The auditory cortex is the most recent evolutionary part of the auditory system and thus is important for processing of sound that is specific to humans [29]. High-level processing of the sounds heard is done here in connection to other parts of the brain. It is located bilaterally, at the upper sides of the temporal lobes in areas such as the Heschl's gyri and superior temporal gyri. The location of the auditory cortex is shown in Fig. 1.3.

Although this is the main path leading from sound coming into our ears to us actually hearing and listening, other paths and areas are included in the hearing process. These areas can be connected in undefined ways or be connected to specific biological processes, such as a connection from the inferior colliculus to the gaze control centre for reflexive eye

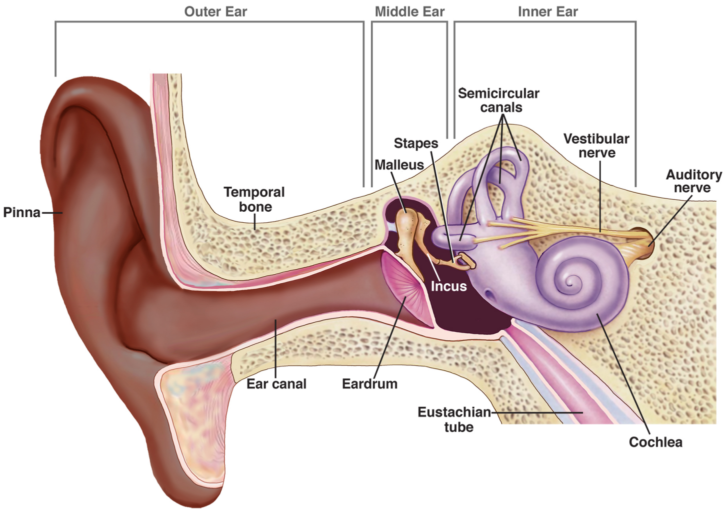

Figure 1.2: Parts that make up the human ear. Taken from the website of the National Institute on Deafness and Other Communication Disorders, National Institutes of Health (public domain).

movements towards sounds. Additionally, the auditory pathway is not a one way path for sound processing. In reality, feedback loops are incorporated into the biological structures, with descending projections back as far as the cochlea.

4 Probing brain activity with EEG

Measuring what goes on in the brain is not trivial, to say the least. Inferring neural activity has been done in multiple different ways. Techniques such as functional magnetic resonance imaging (fMRI) measure the blood flow in the brain. Blood flow is coupled to neurological activity, i.e. blood flows to regions that are activated in neuro-processing. The technique of fMRI has a high spatial resolution, but lacks temporal resolution to see fast cortical changes. Other techniques, such as the electroencephalography (EEG) and magnetoencephalography (MEG), measure electrical activity in the brain. Both techniques have a significantly higher temporal accuracy. Although both techniques are measured outside the head (or are so-called non-invasive) and MEG has better source-estimation qualities, EEG is significantly cheaper and easier to set-up. EEG is currently seen as the only technique that can have possible use in future HAs.

The first human EEG was captured 100 years ago, in 1924, by Hans Berger [15]. By placing as many as 32, 64 or 128 electrodes on the scalp, and measuring the electrical potential of

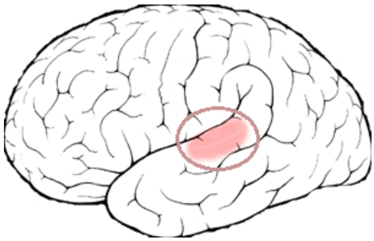

Figure 1.3: Location of the auditory cortex, highlighted in pink, on the surface of the brain. The anterior of the brain is to the left, and the posterior to the right.

each electrode, one can observe a spatial map of electrical potentials at each sample during EEG. An example of an EEG cap is shown in Figure 1.4. The sample rate is in the range of thousands of Hz, although this usually is reduced during analogue to digital conversion. Measuring electrical potential requires setting a reference potential, which is a disputed question. Commonly one designates a reference electrode where no cortical activity should be observed.

After measurement, data can either be analyzed or visualized. Commonly, EEG activity is interpreted through its spectral signature. Activity is split into the following frequency bands; $\sim [1, 4]$ Hz as delta, $\sim [4, 8]$ Hz as theta, $\sim [8, 12]$ Hz as alpha, $\sim [12, 30]$ Hz as beta, $\sim [30, 64]$ Hz as low gamma, and $\sim [64, 128]$ Hz as high gamma. The EEG caps used for measurements mentioned in this thesis have 64 channels spread out over the head of the wearer. They can be grouped into frontal, central, parietal, and occipital, as well as into the left (L) and right (R) hemispheres.

Although multiple limitations plague EEG analysis, perhaps the main limitation is the poor spatial resolution. EEG cannot sense activation of individual neurons. Instead, EEG captures the synchronized activation of thousands of individual neurons. EEG is also influenced heavily by so-called pink noise $1/f$-shaped in the frequency domain. Noise sources originate from ocular activity, heart beats, line noise, channel noise in electrodes, muscles artifacts, as well as brain activity that is not of interest to the particular study or clinical trial [33]. These noise sources disturb measurements both spatially and temporally.

5 Research questions of AAD

Historically, either steady state analysis, waveform analysis or event-related potential analysis have been been the most common application of EEG. To fully understand the neural response to speech in a realistic setting, one has to analyze the response of the brain to continuous stimulus. This together with the usual aspects of EEG brain function imagining poses a plethora of challenges in the field.

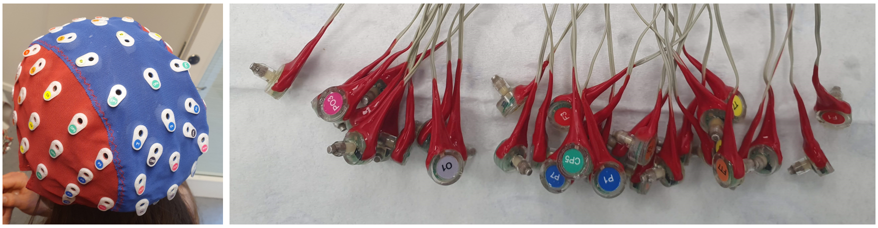

Figure 1.4: An EEG cap (to the left), and EEG electrodes (to the right).

Multiple goals are relevant in AAD. Inferring spatial and temporal location of sound/speech processing in the brain, identifying differences between individuals with different preconditions, and determining neurological processing are just some examples of research questions of interest. These are besides the obvious goal of creating algorithms that can decode auditory attention with the most efficacy.

In large, it is not known what happens when we listen to natural speech in multi-talker environments. This is key to keep in mind when designing algorithms. They should enable us to learn more about neural processing of speech. Sound processing in the brain in itself is highly complex and include non-linear relations as well as stochastic processes. If non-linear, complex models are applied there is a problem that due to relatively low levels of data and high levels of noise models can be hard to fit. Additionally, if we create models that are too hard to understand we are in some sense recreating the same problem again, trying understand complex processing of sounds but this time it is machine processing. Therefore, the work part of this thesis focuses on linear and fairly non-complex methods.

Since EEG results in highly noisy measurements, this is a great challenge, and extracting as much information as possible is pertinent. There is no set procedure that is regarded as the generally accepted path. This opens up multiple possibilities of preprocessing and handling noise before analysis is made. Specifying the effects of different steps during preprocessing of EEG data in specifically AAD is of high interest.

Chapter 2

The Role of Signal Processing in Auditory Attention Decoding

Consider recorded EEG with N_c channels, measured during an auditory attention experiment, as $y_c(t)$, during N_{tr} trials lasting T seconds. During experiments, subjects are presented with speech from multiple speakers. Time-varying features can be extracted from these, for instance the envelope of speech in our case. The speech feature for attended speech is denoted $x(t)$, but this can also be concurrent ignored speech. The measured EEG can be decomposed as $y_c(t) = r_c(t) + e_c(t)$, where $e_c(t)$ is noise.

The EEG response is assumed to be the output of some processing of the attended speech feature

$$r_c(t) = f(x(t), t) \tag{2.1}$$

This processing is assumed to be stationary in time, and only $x(t)$ and $y_c(t)$ is observable.

$$y_c(t) = r_c(t) + e_c(t) = f(x(t)) + e_c(t) \tag{2.2}$$

Our system of signals can be summarized as

$$\begin{array}{lcr} \text{Speech feature} & : & x(t) \\ \text{EEG measured, channel } c & : & y_c(t) \\ \text{EEG response, channel } c & : & r_c(t) \\ \text{EEG noise, channel } c & : & e_c(t) \end{array}$$

There is no reason to believe f is a linear operator. As described in the previous chapter, there is a large amount of different processes that connect sound heard to measurable brain processes. However, with this being said, there are multiple examples where linear modelling have been successful in describing the system in 2.2 for making meaningful inference.

Additionally, although it may be hard to define the non linear f, one can analyze the output alone to circumvent the problem. This introduces other problems instead. One has to be careful when defining models, to not overfit towards other patterns in the EEG data, such as artifacts from eye movement. Additionally, if one rejects to use the speech signals at all, how does one gauge where in time EEG data is interesting and carries information? Linear modelling can adequately provide insights of the system of interest. Presented here are four different signal analysis methods, that explore linear aspects of signals. Two use speech envelope information, and two only looks at EEG data.

1 Coherence

A real valued stationary random processes, $x(t)$, and its Fourier transform $X(f)$ is related through

$$X(f) = \mathcal{F}\left[x(t)\right] = \int x(t)e^{-i2\pi ft}dt \tag{2.3}$$

The magnitude squared coherence, or simply coherence, between two signals $x(t)$ and $y(t)$ is then defined

$$C(f) = \frac{|S_{xy}(f)|^2}{S_{xx}(f)S_{yy}(f)} \tag{2.4}$$

where a cross-spectrum $S_{xy}(f)$ is normalized by the two respective auto-spectra of the signals.

The common interpretation of the coherence measure is that it measures the relative linearity between the two processes. If $y(t)$ can be explained by a linear filtering of $x(t)$,

$$y(t) = f(t) * x(t), \tag{2.5}$$

then $C(f) = 1$ at the frequencies where x and y have power. If there is no linear connection between the processes, coherence is zero instead. Another way of putting this is that coherence can detect phase-locked spectral components between channel $x(t)$ and $y(t)$.

Estimating coherence, simply reduces to estimating the corresponding spectra of signals included,

$$\hat{C}(f) = \frac{|\widehat{S}_{xy}(f)|^2}{\widehat{S}_{xx}(f)\widehat{S}_{yy}(f)} \tag{2.6}$$

This however, is non-trivial. Perhaps the most common approach, using the Welch method, utilizes averaging of multiple data segments to estimate each cross- and auto-spectrum [36].

$$\widehat{S}_{xy}(f) = \frac{1}{L}\sum_{l=1}^{L} X_l(f)Y_l(f)^* = \frac{1}{L}\sum_{l=1}^{L} \mathcal{F}\left[x_l(t)h(t)\right]\mathcal{F}\left[y_l(t)h(t)\right]^* \tag{2.7}$$

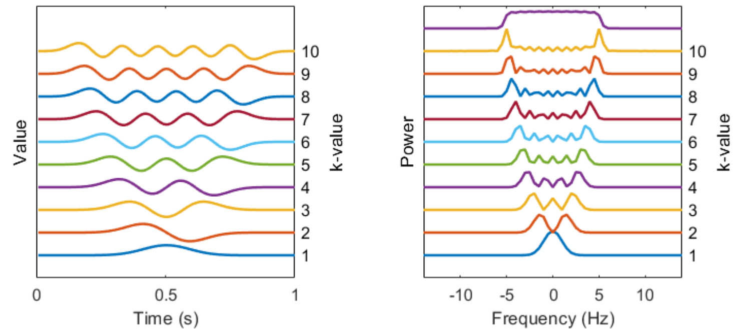

Figure 2.1: The first tenSlepian windows, for time lengths $T = 1$, and bandwidth $W = 5.5$, at a sample rate of 128.

where $l = 1...L$ indexes the data segment used in each Fourier transform, and $h(t)$ is a tapering window. The auto-spectra are calculated in the corresponding similar way. The statistics of coherence estimation using the Welch method has been thoroughly studied [36, 5, 28, 25]. As an alternative to this approach, Thomson introduced multi-taper spectrum estimation of processes [32]. Here K different data windows $h_k(t)$ are applied to the same data segment as

$$\widehat{S}_{xy}(f) = \sum_{k=1}^{K} \alpha_k X_k(f) Y_k(f)^* = \sum_{k=1}^{K} \alpha_k \mathcal{F}\left[x(t)h_k(t)\right] \mathcal{F}\left[y(t)h_k(t)\right]^* \tag{2.8}$$

Again, the auto-spectra are calculated in the corresponding similar way. The subestimates are weighted by α_k, which are commonly chosen as $1/K$. Commonly, the windows used are the set of Slepian functions, which maximally concentrates spectral power within a chosen bandwidth $[-W, W]$. The first ten Slepian windows for $W = 5.5$ is shown in Figure 2.1. The statistics of spectral estimation using the Thomson method have also been studied, although not to the same extent as the Welch method [35].

Both the Welch method and the Thomson method increases bias induced by widening of the bandwidth of the spectral kernel of the estimator. In this sacrifice, they gain a reduced variance in each frequency bin, approximately by factor of the number of terms in the estimates (L and K).

Naturally, one can also combine the two approaches to both average over multiple data segments, where multiple windows are applied to each. This can be beneficial when time series are long, but noise levels are high or narrowband information is not of interest. Coherence is applied in Paper A and Paper B.

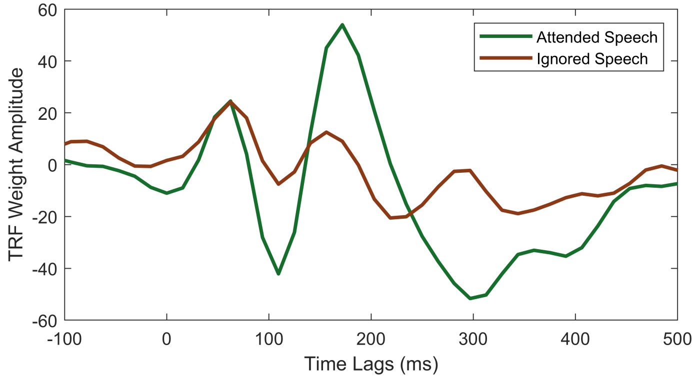

Figure 2.2: An example of TRF weights for channel Cz, extracted from multiple trials from one subject listening. Both envelopes from concurrent attended speech and ignored speech are used to fit the TRF. One can see that the two TRFs are different, and that attended speech results in a TRF with a higher amplitude.

2 Temporal response functions

As described in the previous chapter, linear methods have seen success within the field of AAD. The most commonly used method of analysis is using temporal response functions (TRFs) [1, 26, 18]. These are impulse response filters fitted between EEG and speech features. In this work, we are limiting speech features to be speech envelopes. Estimates of channel-wise EEG responses from speech envelopes presented to a person can be modelled using the forward (hence the superscript f) TRF as

$$\widehat{r}_c(t) = \sum_{\tau} w_c^f(\tau) x_a(t-\tau) \tag{2.9}$$

where $a = att, ign$, are labels corresponding to attended or ignored speech. TRFs, defined by the weights $w_c^f(\tau)$, are found by solving the regression problem

$$\underset{w_c^f}{\arg\min} \sum_t \|y_c(t) - \widehat{r}_c(t)\|_2^2 + \lambda^f \|w_c^f\|_2^2 \tag{2.10}$$

By solving 2.10 and comparing competing speech features, one can receive both temporal and spatial information about the system. As an example, Figure 2.2 shows the channel Cz TRF weights for attended speech and ignored speech envelopes. A clear stronger response is found for attended speech compared to ignored speech.

One can look at the system in the opposite way, and predict the envelope of the speech as a filtering of the EEG response using the backward TRF,

$$\widehat{x}_a(t) = \sum_c \sum_{\tau} y_c(t+\tau) w_c^b(\tau) \tag{2.11}$$

Then, this entails solving the following regression problem,

$$\arg\min_{w_c^b} \sum_t \|x_a(t) - \widehat{x_a}(t)\|_2^2 + \lambda^b \|w_c^b\|_2^2 \tag{2.12}$$

Using these models to predict the speech envelopes, one can look at the similarity of these to the true observed attended and ignored speech. The most commonly used metric is the Pearsons correlation metrics which gives the correlation sample-wise, but mean squared error or reconstruction are also used [8, 3].

Previous success with TRF analysis is relevant and a basis for the use of coherence in Paper A and Paper B. TRFs are also directly applied in Paper D to evaluate different steps of preprocessing EEG.

3 Common spatial patterns

Extracting the spatial patterns in EEG responses, only using the class information of which speaker a person is attending, can be solved using common spatial patterns (CSP) [14, 4]. An example of a two-class source discrimination is differentiating sources from left and sources from right of the listener. The class of CSP-based methods usually averages out the time aspect of data in some way.

For T samples of a C EEG channel measurement pertaining to two classes (class 1 and class 2) $\mathbf{y}_1, \mathbf{y}_2 \in \mathbb{R}^{C,T}$, the CSP algorithm aims to find weights $\mathbf{w}$ that maximise the ratio of the variance of the two classes,

$$\hat{\mathbf{w}} = \arg\max_{\mathbf{w}} \frac{\left\|\mathbf{w}^{\mathrm{T}}\mathbf{y}_1\right\|^2}{\left\|\mathbf{w}^{\mathrm{T}}\mathbf{y}_2\right\|^2} = \arg\max_{\mathbf{w}} \frac{\mathbf{w}^{\mathrm{T}}\mathbf{R}_1\mathbf{w}}{\mathbf{w}^{\mathrm{T}}\mathbf{R}_2\mathbf{w}} \tag{2.13}$$

which is solved by the generalised eigenvalue problem with matrices

$$\mathbf{R}_1 = \frac{\mathbf{y}_1\mathbf{y}_1^{\mathrm{T}}}{N} \quad \mathbf{R}_2 = \frac{\mathbf{y}_2\mathbf{y}_2^{\mathrm{T}}}{N} \tag{2.14}$$

which are the within-class covariance matrices normalised with N number of samples. The eigenvectors corresponding to the $M/2$ largest and $M/2$ smallest eigenvalues gives the estimates of optimal spatial filers $\hat{\mathbf{w}}_m$ (M filters in total). Each pair of projected vectors $\hat{\mathbf{w}}_m^{\mathrm{T}}\mathbf{y}_1$ and $\hat{\mathbf{w}}_m^{\mathrm{T}}\mathbf{y}_2$ can then be further analysed, for example in a classification task. Commonly the variance of the projected output is taken and then averaged over a whole trial. Thus the output of the transform of class trials are then $\mathbf{z}_1, \mathbf{z}_2 \in \mathbb{R}^{2M,1}$. Although multiple classification methods are available, a baseline option is linear discriminant analysis (LDA).

One can include spectral aspects to the CSP algorithm, by filtering time series before analysis, and performing CSP on each filtered time series [17]. Additionally, multiple regularized versions of the algorithm have been explored in different fields of EEG-based neuroscience [20, 21, 27], although there are only a few examples in the field of AAD [14, 4]. Long trials, compared to other fields of EEG analysis pose a problem, since useful information in the EEG data may be temporally hidden. CSP-based methods are analyzed in Paper C.

4 Phase Estimation of Transients

Estimating phase difference between EEG channels is generally of interest in neuroscience, since this can be used to infer causality of processing in the brain. EEG is non-stationary, which demonstrates the requirement for phase difference estimation of transient components in time series. Consider signals in two different channels c and d as $x_c(t)$ and $x_d(t)$. Spectral components of time-varying signals can be captured using the short time Fourier transform, with a window function $h(t)$,

$$F^h_{x_c}(t,f) = \int x_c(s)h^*(s-t)e^{-i2\pi fs}ds \tag{2.15}$$

The spectrogram and cross-spectrogram can then be defined as

$$S^h_{x_c x_c}(t,f) = F^h_{x_c}(t,f)F^h_{x_c}(t,f)^* \tag{2.16}$$

$$S^h_{x_c x_d}(t,f) = F^h_{x_c}(t,f)F^h_{x_d}(t,f)^* \tag{2.17}$$

respectively. The complex phase of the cross-spectrogram at different times and frequencies contains the time-shift of the component in the channels $x_c(t)$ and $x_d(t)$. Multiple options for extracting phase information from spectrograms are available, but noise-robust methods are critical since EEG data is so heavily disturbed by noise. Phase difference estimation between signals and spectrogram theory is used in Paper E.

5 Influence of preprocessing EEG

To mitigate the effects from the inherent limitations of EEG, the data is often preprocessed before analysis. Although there is no fixed standardised pipeline, the processes contain most of these steps. Bad channels are visually identified, removed and replaced by an interpolation from surrounding channels. Bad trials can also be removed, if there is a significant disturbance during the certain trial. Additionally, EEG data is highpassed filtered around $0.1-0.5$ Hz. Data is also lowpass filtered and downsampled to a desired sample frequency, usually somewhere between 64 Hz and 256 Hz. A notch filter at 50 Hz is also critical to apply in order to remove line noise in measurements. The role of these preprocessing steps is steadily discussed within the field, some researchers arguing for and against the necessity of each step. In particular, for event-related potential analysis of EEG, there has been extensive identification of the effect of preprocessing steps [12, 10, 11].

In addition to the previous steps, a common preprocessing step is independent component analysis (ICA) based artifact rejection [33]. ICA decomposes multivariate EEG into a series of independent and non-Gaussian components. Since different artifacts are sufficiently non-Gaussian and are generally independent from each other and from the brain activity of interest, these are possible to split apart. When EEG channels have been projected onto the independent components, artifacts can be identified and removed from data. Then EEG data in original sensor space is restored through reprojection.

Nuisance of identifying artifacts in EEG

One particularly time-consuming part of preprocessing is removing EEG artifacts. This is usually done by investigating the temporal, spectral and spatial structure of each component [33]. Using this one can identify the signatures of certain types of artifact. Identifiable artifacts can be grouped into muscle artifacts, eye gaze and blinks, heart beats, line noise, channel noise. Brain activity and other hard-to-identify artifact are contained in other components. Thus, it is very important to remove correct components, so data is not destroyed. For the domain of AAD, there has been little to none research towards understanding the effects of errors in ICA based artifact removal. So, experienced EEG researcher and scientists have to spend considerable amounts of time to perform these steps, in order for conclusions drawn from data to be reliable. Paper D aims to probe the effects of preprocessing in AAD, looking at manual and automatic alternatives of artifacts rejection in EEG.

6 Main results of research papers

Paper A: Improved coherence measure for EEG-based speech tracking in hearing impaired listeners

Due to success of linear input-to-output methods in AAD, the measure of coherence should be a suitable method to detect connections between speech features and EEG responses. High amounts of measurement noise in EEG and non-linearities cause very low values of coherence in this application, even though there is a linear connection. This paper presents a analytical bias expression to show an inherent bias of coherence peak towards higher frequencies, due to $1/f$– shaped noise. Additionally, we present a methodological improvement in coherence estimation compared to previous applied coherence methods in the field, which increases performance. The use of improved coherence is shown advantageous for decoding auditory attention of hearing impaired. Finally, an application of coherence in evaluation of hearing conditions is also presented.

Paper B: Coherence Expectation Minimisation and Combining Weighted Multitaper Estimates

This paper presents a novel expression for the expectation of a multitaper magnitude squared coherence estimate, for the signal model presented in this thesis. Minimisation of certain terms in this expression is used to find multitaper weights for coherence estimation. When weights are not equal, biases dependent spectral locations of signal power are induced. This is mitigated by combining two coherence estimates that are differently weighted. Performance, in terms of reduced mean squared error, is improved, giving better classification in simple EEG examples.

Paper C: A Comparison of Common Spatial Patterns Algorithms for Auditory Attention Decoding

The classification methods utilizing CSP are proven in EEG classification, in certain fields. However, applications of these methods within the field of AAD have been scarse. This paper employs and reports performance of a variety of CSP methods from other field of EEG classification to three different AAD datasets. These datasets are from EEG of listeners that are both normal hearing and hearing impaired, enabling us to compare performance of CSP between the two subsets. Classification of left or right locus of attention has higher accuracy for the dataset with normal hearing participants, while it is still possible to predict locus of attention for individuals with hearing impairment as well.

Paper D: Effect of Independent Component Artifact Rejection on EEG-Based Auditory Attention Decoding

Understanding the influence of EEG preprocessing is critical to make conclusions in analysis. There is a gap in the specific field of AAD in how reliable results are in relation to the methodology used when rejecting EEG artifacts with ICA. Here, multiple automatic labellers of artifacts are analyzed and compared to traditional manual inspection. If automatic artifact rejection schemes are satisfactory, this would save a lot of time. We show small deviations in the performance of TRF-based decoders on one investigated dataset, although the are some extrapolated patterns. Examples of these are that omitting artifact rejection can actually increase performance in backward modelling, and that manual artifact rejection was superior for frontal EEG channels compared to automatic artifact rejection schemes.

Paper E: Robust Phase Difference Estimation of Transients in High Noise Levels

Phase difference estimation of transient signals is challenging. This is particularly useful in EEG where differences in phase between channels contains information, that can be hard to extract from raw time signals. The phase difference estimation method presented in this paper is built upon previous methodology of reassignment and cross-spectrogram reassignment. Compared to previous methods, improvements to noise-robustness and mean squared error are shown. The method is not restricted to specific spectrogram window shapes. Additionally, the method does not make any assumptions regarding the envelope shapes of components, and is tunable to different signal characteristics.

7 Conclusions and outlook

In summary, this thesis explores multiple different ways linear models and methods can be applied in EEG analysis in general, and in AAD in specific. Although the methods are linear, they are in fact very capable of capturing interesting and important aspects of EEG when performing AAD. For example, coherence detects linear connections between speech envelopes and EEG response. CSP-based methods can differentiate left and right locus of attention from each other, using the spectral and spatial signatures of EEG. In both these cases, due to a very low signal-to-noise ratio in EEG data, methodological choices are important for reliable results. The thesis presents method improvements, building towards analysis tools of EEG that are applicable in clinical settings for hearing impaired persons. Additionally, an evaluation of current preprocessing methodology within the field of AAD shows comparable performance to new promising machine-based methodologies, which could save resources in research and clinics.

Looking ahead, multiple interesting paths of research are evident. Firstly, and perhaps most concretely, evaluation of artifact rejection schemes should be further expanded to encompass more evaluation methods as well as more datasets. This is necessary before any strong conclusions can be made. Additionally, one can investigate specific artifacts and their individual effect on TRF analysis. Additionally, involving temporal aspects more into models would be interesting. Detecting faltering attention or attention switches are important for applying AAD methods in hearing aid technologies. The problems of these temporal aspects can be approached in multiple ways, but this demands high standards on signal analysis methods as well as datasets used to make evaluations and conclusions. This sets the scene for multiple paths forward, research-wise.

References

[1] E. Alickovic, T. Lunner, F. Gustafsson, and L. Ljung. A Tutorial on Auditory Attention Identification Methods. *Frontiers in Neuroscience*, 13:153, Mar. 2019.

[2] C. Brodbeck, L. E. Hong, and J. Z. Simon. Rapid Transformation from Auditory to Linguistic Representations of Continuous Speech. *Current Biology*, 28(24):3976–3983.e5, Dec. 2018.

[3] C. Brodbeck, P. Das, M. Gillis, J. P. Kulasingham, S. Bhattasali, P. Gaston, P. Resnik, and J. Z. Simon. Eelbrain, a Python toolkit for time-continuous analysis with temporal response functions. *eLife*, 12:e85012, Nov. 2023.

[4] S. Cai, E. Su, Y. Song, L. Xie, and H. Li. Low Latency Auditory Attention Detection with Common Spatial Pattern Analysis of EEG Signals. In *Interspeech 2020*, pages 2772–2776. ISCA, Oct. 2020.

[5] G. Carter, C. Knapp, and A. Nuttall. Estimation of the magnitude-squared coherence function via overlapped fast Fourier transform processing. *IEEE Transactions on Audio and Electroacoustics*, 21(4):337–344, Aug. 1973.

[6] E. C. Cherry. Some Experiments on the Recognition of Speech, with One and with Two Ears. *The Journal of the Acoustical Society of America*, 25(5):975–979, Sept. 1953.

[7] G. Ciccarelli, M. Nolan, J. Perricone, P. T. Calamia, S. Haro, J. O'Sullivan, N. Mesgarani, T. F. Quatieri, and C. J. Smalt. Comparison of Two-Talker Attention Decoding from EEG with Nonlinear Neural Networks and Linear Methods. *Scientific Reports*, 9(1):11538, Aug. 2019.

[8] M. J. Crosse, G. M. Di Liberto, A. Bednar, and E. C. Lalor. The Multivariate Temporal Response Function (mTRF) Toolbox: A MATLAB Toolbox for Relating Neural Signals to Continuous Stimuli. *Frontiers in Human Neuroscience*, 10, Nov. 2016.

[9] C. Daube, R. A. Ince, and J. Gross. Simple Acoustic Features Can Explain Phoneme-Based Predictions of Cortical Responses to Speech. *Current Biology*, 29(12):1924–1937.e9, June 2019.

[10] A. De Cheveigné. *Is* EEG is better left alone? preprint, Neuroscience, June 2023.

[11] A. De Cheveigné and I. Nelken. Filters: When, Why, and How (Not) to Use Them. *Neuron*, 102(2):280–293, Apr. 2019.

[12] A. Delorme. EEG is better left alone. *Scientific Reports*, 13(1):2372, Feb. 2023.

[13] H. Dillon, J. A. Zakis, H. McDermott, G. Keidser, W. Dreschler, and E. Convery. The trainable hearing aid: What will it do for clients and clinicians? *The Hearing Journal*, 59(4):30, Apr. 2006.

[14] S. Geirnaert, T. Francart, and A. Bertrand. Fast EEG-Based Decoding Of The Directional Focus Of Auditory Attention Using Common Spatial Patterns. *IEEE Transactions on Biomedical Engineering*, 68(5):1557–1568, May 2021.

[15] L. F. Haas. Hans Berger (1873-1941), Richard Caton (1842-1926), and electroencephalography. *Journal of Neurology, Neurosurgery & Psychiatry*, 74(1):9–9, Jan. 2003.

[16] M. L. Hawley, R. Y. Litovsky, and H. S. Colburn. Speech intelligibility and localization in a multi-source environment. *The Journal of the Acoustical Society of America*, 105(6):3436–3448, June 1999.

[17] Kai Keng Ang, Zhang Yang Chin, Haihong Zhang, and Cuntai Guan. Filter Bank Common Spatial Pattern (FBCSP) in Brain-Computer Interface. In *2008 IEEE International Joint Conference on Neural Networks (IEEE World Congress on Computational Intelligence)*, pages 2390–2397, Hong Kong, China, June 2008. IEEE.

[18] E. C. Lalor and J. J. Foxe. Neural responses to uninterrupted natural speech can be extracted with precise temporal resolution. *European Journal of Neuroscience*, 31(1): 189–193, Jan. 2010.

[19] B. J. Lawrence, D. M. P. Jayakody, R. J. Bennett, R. H. Eikelboom, N. Gasson, and P. L. Friedland. Hearing Loss and Depression in Older Adults: A Systematic Review and Meta-analysis. *The Gerontologist*, 60(3):e137–e154, Apr. 2020.

[20] F. Lotte and Cuntai Guan. Regularizing Common Spatial Patterns to Improve BCI Designs: Unified Theory and New Algorithms. *IEEE Transactions on Biomedical Engineering*, 58(2):355–362, Feb. 2011.

[21] F. Lotte and C. Guan. Spatially Regularized Common Spatial Patterns for EEG Classification. In *2010 20th International Conference on Pattern Recognition*, pages 3712–3715, Istanbul, Turkey, Aug. 2010. IEEE.

[22] D. G. Loughrey, M. E. Kelly, G. A. Kelley, S. Brennan, and B. A. Lawlor. Association of Age-Related Hearing Loss With Cognitive Function, Cognitive Impairment, and

Dementia: A Systematic Review and Meta-analysis. *JAMA Otolaryngology–Head & Neck Surgery*, 144(2):115, Feb. 2018.

[23] N. Marrone, C. R. Mason, and G. Kidd. The effects of hearing loss and age on the benefit of spatial separation between multiple talkers in reverberant rooms. *The Journal of the Acoustical Society of America*, 124(5):3064–3075, Nov. 2008.

[24] N. Mesgarani and E. F. Chang. Selective cortical representation of attended speaker in multi-talker speech perception. *Nature*, 485(7397):233–236, May 2012.

[25] A. Nuttall and G. Carter. Bias of the estimate of magnitude-squared coherence. *IEEE Transactions on Acoustics, Speech, and Signal Processing*, 24(6):582–583, Dec. 1976.

[26] J. A. O'Sullivan, A. J. Power, N. Mesgarani, S. Rajaram, J. J. Foxe, B. G. Shinn-Cunningham, M. Slaney, S. A. Shamma, and E. C. Lalor. Attentional Selection in a Cocktail Party Environment Can Be Decoded from Single-Trial EEG. *Cerebral Cortex*, 25(7):1697–1706, July 2015.

[27] P. Rithwik, V. Benzy, and A. Vinod. High accuracy decoding of motor imagery directions from EEG-based brain computer interface using filter bank spatially regularised common spatial pattern method. *Biomedical Signal Processing and Control*, 72:103241, Feb. 2022.

[28] E. Scannell and G. Carter. Confidence bounds for magnitude-squared coherence estimates. *IEEE Transactions on Acoustics, Speech, and Signal Processing*, 26(5):475–477, Oct. 1978.

[29] J. Schnupp, I. Nelken, and A. J. King. *Auditory Neuroscience: Making Sense of Sound*. The MIT Press, 2010.

[30] S. K. Scott. Identification of a pathway for intelligible speech in the left temporal lobe. *Brain*, 123(12):2400–2406, Dec. 2000.

[31] S. K. Scott and I. S. Johnsrude. The neuroanatomical and functional organization of speech perception. *Trends in Neurosciences*, 26(2):100–107, Feb. 2003.

[32] D. Thomson. Spectrum estimation and harmonic analysis. *Proceedings of the IEEE*, 70(9):1055–1096, 1982.

[33] J. A. Urigüen and B. Garcia-Zapirain. EEG artifact removal—state-of-the-art and guidelines. *Journal of Neural Engineering*, 12(3):031001, June 2015.

[34] S. Vandecappelle, L. Deckers, N. Das, A. H. Ansari, A. Bertrand, and T. Francart. EEG-based detection of the locus of auditory attention with convolutional neural networks. *eLife*, 10:e56481, Apr. 2021.

[35] A. T. Walden. A unified view of multitaper multivariate spectral estimation. *Biometrika*, 87(4):767–788, Dec. 2000.

[36] P. Welch. The use of fast Fourier transform for the estimation of power spectra: A method based on time averaging over short, modified periodograms. *IEEE Transactions on Audio and Electroacoustics*, 15(2):70–73, June 1967.

[37] E. Zion Golumbic, N. Ding, S. Bickel, P. Lakatos, C. Schevon, G. McKhann, R. Goodman, R. Emerson, A. Mehta, J. Simon, D. Poeppel, and C. Schroeder. Mechanisms Underlying Selective Neuronal Tracking of Attended Speech at a "Cocktail Party". *Neuron*, 77(5):980–991, Mar. 2013.

Scientific Publications

Author contributions

Co-authors are abbreviated as follows: Maria Sandsten (MS), Maria Åkesson (MÅ), Emina Alickovic (EA), Martin A. Skoglund (MAS), Johanna Wilroth (JW)

Paper A: Improved coherence measure for EEG-based speech tracking in hearing impaired listeners

The idea for the paper was realized by all authors (EA, MAS, MS and I). I developed the method changes to multitaper coherence estimation presented in the paper. I also made the application of methods to data and was main contributor to the writing process.

Paper B: Coherence Expectation Minimisation and Combining Weighted Multitaper Estimates

MS and I came up with the idea for the paper. Together with aid from MS I developed the theory. Simulations were done by me. MÅ applied methods in a real data scenario. I wrote the main part of the paper, with an exception of the real data parts, which was done by MÅ.

Paper C: A Comparison of Common Spatial Patterns Algorithms for Auditory Attention Decoding

I, with support from my supervisors, developed the idea for the paper. I made the application of methods on the datasets. Results analysis and writing was done by me, in discussion with supervisors.

Paper D: Effect of Independent Component Artifact Rejection on EEG-Based Auditory Attention Decoding

I had the first idea for the paper, but the idea was then developed by all authors (JW, EA, MAS and I). JW and I wrote the code for the investigation and results. Result analysis was done by all authors (JW, EA, MAS and I). The majority of the paper was written by me, but with substantial parts written by JW.

Paper E: Robust Phase Difference Estimation of Transients in High Noise Levels

MS and I came up with the idea for the paper. Method development was done by me, although with invaluable guidance from MS. I made the simulations, real data example and the majority of the writing, although MS wrote selected parts.

Printed by Media-Tryck, Lund 2024 NORDIC SWAN ECOLABEL 3041 0903

LUND UNIVERSITY

Licentiate Theses in Mathematical Sciences 2024:1

ISBN 978-91-8104-068-5
ISSN 1404-028X