Linköping Studies in Science and Technology
Licentiate Thesis No. 2006

Direction of Arrival Estimation for Wildlife Protection

Gustav Zetterqvist

LINKÖPING UNIVERSITY

Linköping Studies in Science and Technology.
Licentiate Thesis No. 2006

Direction of Arrival Estimation for Wildlife Protection

Gustav Zetterqvist

LINKÖPING UNIVERSITY

This is a Swedish Licentiate's Thesis.

Swedish postgraduate education leads to a Doctor's degree and/or a Licentiate's degree.
A Doctor's Degree comprises 240 ECTS credits (4 years of full-time studies).
A Licentiate's degree comprises 120 ECTS credits,
of which at least 60 ECTS credits constitute a Licentiate's thesis.

Linköping Studies in Science and Technology.
Licentiate Thesis No. 2006

Direction of Arrival Estimation for Wildlife Protection

Gustav Zetterqvist

gustav.zetterqvist@liu.se
https://www.control.isy.liu.se/
Department of Electrical Engineering
Linköping University
SE-581 83 Linköping
Sweden

ISBN 978-91-8075-829-1 (print)
ISBN 978-91-8075-830-7 (PDF)
https://doi.org/10.3384/9789180758307
ISSN 0280-7971

Printed by LiU-Tryck, Linköping, Sweden 2024

To my pillars of strength!

Abstract

Direction of arrival (DOA) estimation is a well-established problem in signal processing. It involves determining the direction from which a signal reaches a sensor array, and is fundamental in applications like radar, sonar, and acoustics. Traditionally, DOA estimation relies on comparing the time of arrival of the signal across different sensors in the array. However, this approach is sensitive to the *time difference of arrival* (TDOA) between sensors, which can be challenging to estimate accurately. Additionally, precise synchronization among the sensors is essential, but this can be difficult to achieve in certain environments or applications.

In this thesis, we explore a novel approach to DOA estimation based on the received signal power at the sensors. The method exploits the directional sensitivity of the microphones in the array, which defines how effectively each microphone captures sound from different directions. To model the directional sensitivity, we use a *Fourier series* (FS) model. The model is then used to estimate the DOA of a sound source across various environments, and for different types of signals. The parametric model enables *Cramér-Rao lower bound* (CRLB) analysis of the DOA estimation problem.

Our findings demonstrate that the directional sensitivity exhibits a significant variation in accordance with the frequency content of the signal, and we exploit this to estimate the DOA for different types of sounds. The proposed method has been validated with a range of signals, including gunshots, elephant trumpets, sirens, and female screams.

The results show that the developed method achieves high accuracy in estimating the DOA for the above-mentioned signals. Furthermore, the method performs similarly well in outdoor scenarios with realistic background noise levels. When compared to state-of-the-art DOA estimation techniques, our approach performs better or equally well for the investigated sounds.

A key advantage of this method is that it does not require any TDOA measurement between the microphones, enabling the design of smaller, more compact devices. This opens up new possibilities for estimating DOA in environments where traditional methods are impractical. A limitation, however, is that the method requires knowledge of the microphone's directional sensitivity, which necessitates calibration in an anechoic chamber. Nevertheless, this calibration has proven to be robust, and only needs to be performed once to create a model applicable across different environments.

Additionally, this thesis explores a different application of DOA estimation, where geophones are used to estimate the DOA to elephants. As elephants move, they generate ground vibrations, and these signals can be captured by geophones. We show that a traditional delay-and-sum beamformer can accurately estimate the DOA of elephants at distances up to 40 meters. By determining when elephants are approaching and from which direction, park rangers can take early measures to avoid conflicts between humans and elephants, which is a major problem in some parts of the world.

Populärvetenskaplig sammanfattning

Förmågan att höra var ett ljud kommer ifrån, något vi ofta tar för givet, kallas för riktningsuppfattning. Den gör det möjligt för oss att snabbt avgöra om någon ropar på oss och från vilket håll ljudet kommer. Denna förmåga är viktig för att kunna orientera sig i omgivningen och uppfatta hot eller andra viktiga ljud. Våra öron samarbetar genom att jämföra hur ljud når varje öra, både när det gäller ljudets intensitet och hur lång tid det tar för ljudet att nå dem. Det här kallas för interaural tids- och nivåskillnad. Vissa ljud kan dock vara svåra att uppfatta, till exempel om ljudet är kort och impulsivt, eller om det är i en stadsmiljö med mycket bakgrundsljud och reflektioner.

I den här avhandlingen undersöker vi nya metoder för att uppskatta ljudets riktning. Vi använder mikrofoner för att mäta ljudet och beräknar därefter riktningen som ljudet kommer ifrån. Traditionella metoder fokuserar på tidsskillnaden mellan ljud som registreras i olika mikrofoner. Vi tar istället en annan väg och undersöker hur ljudets styrka kan användas för att avgöra riktningen, oavsett tidsskillnader mellan mikrofonerna.

Vår metod bygger på att vi skapar en modell av mikrofonernas riktningskänslighet, det vill säga hur väl de uppfattar ljud från olika håll. Modellen skapas genom att mäta mikrofonens riktningskänslighet i ett ekofritt rum. Genom att först mäta detta i en kontrollerad miljö, utan ekon, kan vi sedan använda modellen för att beräkna ljudriktningen i mer varierande miljöer och för olika typer av ljud. Till exempel har vi använt ljud såsom pistolskott, elefanttrumpeter, sirener och skrik för att testa vår metod.

Resultaten visar att vår metod kan beräkna riktningar med hög noggrannhet för de ovan nämnda ljuden, även i en utomhusmiljö med mer realistiska nivåer av bakgrundsljud. När vi jämfört vår metod med traditionella metoder, presterar vår lösning lika bra eller bättre för de testade ljuden.

En stor fördel med vår metod är att den inte kräver att mikrofonerna är placerade på ett visst avstånd från varandra, vilket innebär att vi kan bygga mindre och mer kompakta enheter. Detta kan leda till nya typer av produkter för att identifiera ljudriktningar i olika situationer. En nackdel är dock att mikrofonernas riktningskänslighet måste kalibreras i ett ljudlabb, men denna kalibrering har visat sig vara robust och det räcker att utföra en kalibrering som kan användas i flera olika miljöer.

I avhandlingen inkluderas även en annan tillämpning av riktningsskattning, nämligen att uppskatta riktningen till elefanter med hjälp av geofoner som mäter vibrationer i marken. Elefanter är stora djur som skapar tydliga vibrationer i marken när de går. Genom att mäta dessa vibrationer med geofoner kan vi uppskatta riktningen till elefanten. Vi visar att traditionella metoder kan uppskatta riktningen med hög noggrannhet på ett avstånd upp till 40 meter. Genom att avgöra när elefanter närmar sig människor och varifrån de kommer kan parkvakter vidta åtgärder för att undvika konflikter mellan människor och elefanter, vilket är ett stort problem i vissa delar av världen.

Acknowledgments

First and foremost, I would like to thank my supervisor, Prof. Fredrik Gustafsson, for his guidance and support throughout this project. I am grateful for the opportunity to work with you and learn from your vast knowledge and experience. Thank you for believing in me!

I would also like to thank my co-supervisor, Prof. Gustaf Hendeby, for his valuable input and feedback. Your insights have been invaluable to me. Thank you for always being around to assist with any questions I have had, from LaTeX support to technical questions.

Thanks to Prof. Martin Enqvist, head of division at Automatic Control, for providing me with the opportunity to work on this project. I am grateful for the resources and support you have offered me. Also, I would like to thank Ninna Stensgård for her administrative support. You are always there to assist with a big smile.

This thesis has been greatly improved by the feedback and suggestions from my colleagues at the Automatic Control division, a special thanks to Dr. Anton Kullberg, Johanna Wilroth and Daniel Bossér for proofreading parts of this thesis, thank you for your valuable input.

This work was partially funded by the Wallenberg AI, Autonomous Systems and Software Program (WASP) funded by the Knut and Alice Wallenberg Foundation. Thank you for providing me with the opportunities to expand my network, both nationally and internationally. The people I have met and the experiences I have gained have been invaluable to me.

I'm also grateful for the financial support from ELLIIT, thank you for funding my research and providing me with the resources I needed to get this far.

Since my first day at the division, I have had the pleasure of sharing an office with Johanna Wilroth. Thank you for all the fun experiences, and for always being there to support me when I needed it. I'm not sure if I had made it without you!

To my family and friends, thank your for the support and encouragement throughout this project. Your love and understanding have been my rock during tough times. I appreciate your constant support and presence. A special thanks to my parents, Anna and Lars, who always believe in me and support me in everything I do. You mean the world to me! To my brother, Erik, thank you for introducing me to Linköping, and for all the crazy adventures we have had together. I'm grateful for all the fun times we have shared!

Lastly, I would like to thank my partner, Trulsa, for always being there for me, for your love and support. You have given me the strength and courage to keep going when things got tough, and I am grateful for everything you do for me. I love you!

Linköping, September 2024
Gustav Zetterqvist

Contents

Notation

Abbreviations

Abbreviation	Description
ADC	analog-to-digital converter
BIC	Bayesian information criterion
CI	confidence interval
CRLB	Cramér-Rao lower bound
DFT	discrete Fourier transform
DOA	direction of arrival
FFT	fast Fourier transform
FIM	Fisher information matrix
FS	Fourier series
KF	Kalman filter
LS	least squares
MCCC	multi channel cross-correlation
MSE	mean square error
MUSIC	multiple signal classification
MVDR	minimum variance distortionless response
PDF	probability density function
RMSE	root mean square error
RSS	recived signal strength
SLS	separable least squares
SNR	signal to noise ratio
TDOA	time difference of arrival
UCA	uniform circular array
ULA	uniform linear array

1 Introduction

Since the first signs of life, uncountable species have gone extinct. Some vanish due to natural disasters, others due to evolutionary coincidences or, more recently, from the impact of human activities on the planet. In the history of our Earth, there have been five mass extinctions. The most recent one occurred approximately 66 million years ago, which wiped out 75 % of all species, including the dinosaurs [35].

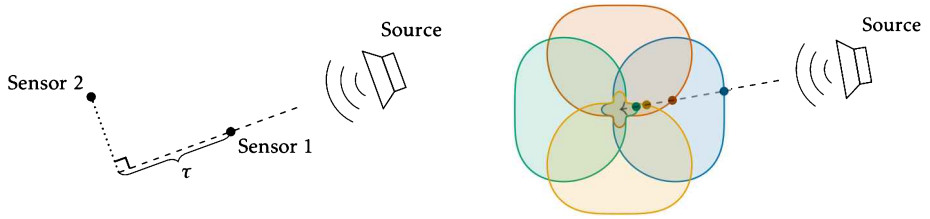
The impact of human activities through illegal hunting, overfishing and deforestation has led to a significant increase in the extinction rate of species. Researchers are warning that we are entering the sixth mass extinction, caused by human population growth and increased consumption [11].

One example is the African elephant, where the population has decreased by 98 % since the 1500s [38]. Even though elephants were recovering from the great poaching era of the 1970s and 1980s, the population has been decreasing since 2007 [38]. The decreasing rate between 2010–2014 was 8% [38], and if this rate continues, the African elephant population will decrease by 50 % every 9 years.

In this thesis, we develop technologies that are able to mitigate poaching, as well as the human-elephant conflict. This is achieved through two distinct applications. The first application involves estimating the *direction of arrival* (DOA) of elephant footsteps using seismic sensors, with the purpose to avoid elephants entering villages and to monitor the elephant population. The second application focuses on estimating the DOA estimation of poachers or other sound sources using a microphone array.

1.1 Background

The problem of estimating the DOA is an essential parameter in various applications, including surveillance, tracking, and localization. The estimation of the



(a) Illustration of classical DOA estimation, where the time-delays between the sensors τ are utilized to estimate the DOA.

(b) Visualization of the proposed method with four microphones. The colorful lines represent the directional sensitivity of each microphone.

Figure 1.1: *Visualization of data used for classical DOA estimation method and the proposed method.*

DOA has been extensively studied for the past decades [24, 26, 44]. For a more recent study, refer to [34]. Traditional DOA estimators rely on sensor arrays and are based on time-delays of the received signal at each sensor, known as *time difference of arrival* (TDOA) measurements, as shown in Figure 1.1a.

Array processing techniques, such as *minimum variance distortionless response* (MVDR) beamforming [10], *multiple signal classification* (MUSIC) [39], and *multi channel cross-correlation* (MCCC) [7, 8], are commonly used for DOA estimation. These principles can be applied to essentially any array geometry and signal waveform. The rule of thumb for classical methods is to separate the sensors with half the wavelength. With a fixed amount of sensors, placing them farther away will give aliasing (ambiguity in DOA), and placing them too close will decrease the frequency resolution and accuracy. As an example, for a broadband signal with frequency content in the interval 340–3400 Hz (corresponding to human speech), the sensor separation should be 5–50 cm. That is, it is inherently difficult to design one sensor array for broadband signals when using time-delay based methods [41, 45, 54].

To circumvent this, we are proposing a method to estimate the DOA using only the received power at each sensor [54, 55]. This method enables simple hardware, low sampling frequency, and smaller arrays. A visualization of the proposed method is shown in Figure 1.1b.

A brief overview of the proposed method is illustrated in Figure 1.2, and the main steps are as follows. First, data is collected using a microphone array, where the received power at each microphone is measured. The directional sensitivity of each microphone is thereafter estimated for different frequencies. For the DOA estimation, the frequency content of the received signal is calculated, and the directional sensitivity is estimated by summing the estimated directional sensitivity for each frequency, with the frequency content as weights. Thereafter, the DOA is estimated by comparing the measured power profile to the estimated directional sensitivity using the *least squares* (LS) method.

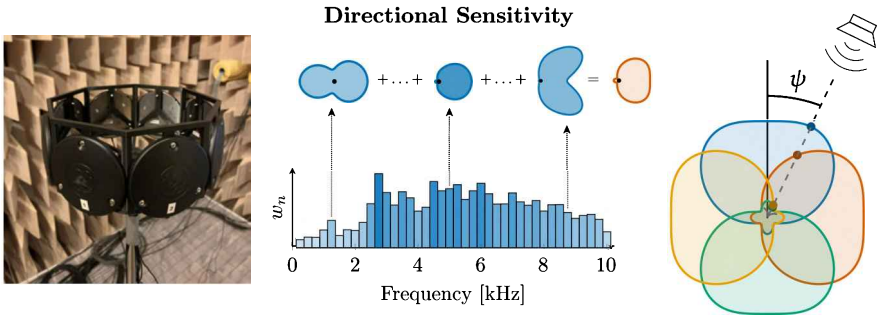


Figure 1.2: *Illustration of the proposed method. Left: The microphone array used for data collection. Middle: The estimated directional sensitivity of each frequency, summed with the frequency content as weights to estimate the directional sensitivity of the signal. Right: The received power profile is compared to the estimated directional sensitivity to estimate the* DOA ψ.

Another important aspect of signal processing is the detection and extraction of the signal of interest from the background noise. In the case of elephant footsteps, the signal of interest is the seismic waves generated by the footsteps, which are typically low-frequency signals. The detection of the signal of interest is crucial for the DOA estimation, as the DOA estimate is based on the detected signal. In this thesis, we elaborate on the detection of elephant footsteps using seismic sensors. For further enhancements of the detection algorithm, we refer to [16]. However, this thesis mainly focuses on the DOA estimation problem, and for the acoustic signals we assume the signal of interest to be detected and extracted from the background noise.

1.2 Problem Formulation

Classical DOA estimation methods rely on the time-delays between the sensors in the array. The main challenge with small arrays is the sensor separation, which leads to a loss of DOA resolution and accuracy for classical methods. To tackle this problem, this thesis aims to answer the following research questions:

- How can a conventional DOA estimator be used to estimate the DOA of elephant footsteps? And how can we differentiate between an elephant footstep and other seismic signals?
- How can the directional sensitivity of the sensors be utilized to estimate the DOA, and how should the directional sensitivity be modeled?
- What are the theoretical limits of the directional sensitivity based DOA estimator, and how does the frequency content of the signal affect the directional sensitivity?
- How does the directional sensitivity based DOA estimate compare to state-of-the-art methods using narrowband and realistic signals?

1.3 Contributions and Publications

The main contributions of this thesis are:

- A novel DOA estimation method based on the received power at each microphone in the array.
- Data collection from an anechoic chamber as well as an outdoor environment to evaluate the proposed method.
- A comparison of the proposed method to state-of-the-art methods using the collected data.
- An application of seismic sensors for DOA estimation of elephant footsteps.

Included Publications

The content of this Licentiate thesis is based on the following publications. All papers have been either peer-reviewed and published or submitted to a conference or journal. The contribution to each paper is described in the text.

The following abbreviations are used to refer to the author and co-authors: Gustav Zetterqvist (GZ), Fredrik Gustafsson (FG), Gustaf Hendeby (GH), Erik Wahledow (EW) and Philip Sjövik (PS).

Elephant DOA Estimation using a Geophone Network

> Gustav Zetterqvist, Erik Wahledow, Philip Sjövik, Fredrik Gustafsson, and Gustaf Hendeby. Elephant DOA Estimation using a Geophone Network. In *2023 IEEE 26th International Conference on Information Fusion (FUSION)*, Charleston, SC, USA, June 2023. IEEE. doi: 10.23919/FUSION52260.2023.10224115.

Abstract: Human-wildlife conflicts are a global problem which is central to the Global Goal 15 (life on land). One particular case is elephants, that can cause harm to both people, property and crops. An early warning system that can detect and warn people in time would allow effective mitigation measures. The proposed method is based on a small local network of geophones that sense the seismic waves of elephant footsteps. It is known that elephant footsteps induce low frequency ground waves that can be picked up by geophones in the ground. First, a method is described that detect the particular signature of such footsteps, and then the detections are used to estimate the *direction of arrival* (DOA). Finally, a *Kalman filter* is applied to the measurements in order to track the elephant. Field tests performed at a local zoo shows promising results with accurate DOA estimates at 15 meters distance and acceptable accuracy at 40 meters.

Contribution: FG had the idea. EW and PS developed the detection method and DOA estimation with input from GZ and FG. EW and PS implemented the method, collected the data and performed the experiments with support from GZ and FG. GZ improved the methods and developed the *Kalman filter* and tracking

with input from FG and GH. GZ wrote the paper. FG and GH provided editorial feedback.

Using Received Power in Microphone Arrays to Estimate Direction of Arrival

> Gustav Zetterqvist, Fredrik Gustafsson, and Gustaf Hendeby. Using Received Power in Microphone Arrays to Estimate Direction of Arrival. In *ICASSP 2023 - 2023 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, Rhodes Island, Greece, June 2023. IEEE. doi: 10.1109/ICASSP49357.2023.10097197.

Abstract: Conventional *direction of arrival* (DOA) estimators are based on array processing using either time differences or beamforming. The proposed approach is based on the received power at each microphone, which enables simple hardware, low sampling frequency and small arrays. The problem is recast into a linear regression framework where the *least squares* method applies, and the main drawback is that different sound sources are not readily separable. Our proposed approach is based on a training phase where the directional sensitivity of each microphone element is estimated. This model is then used as a fingerprint of the observed power vector in a real-time estimator. The learned power vector is here modeled by a *Fourier series* expansion, which enables *Cramér-Rao lower bound* computations. We demonstrate the performance using a circular array with eight microphones with promising results.

Contribution: FG had the idea for the paper. GZ developed the method with input from FG and GH. GZ implemented the method and performed the experiments. GZ wrote the paper. FG and GH provided editorial feedback.

Directional Sensitivity-Based DOA Estimation Using a Fourier Series Model

> Gustav Zetterqvist, Fredrik Gustafsson, and Gustaf Hendeby. Directional Sensitivity-Based DOA Estimation Using a Fourier Series Model. *IEEE Sensors Journal*, August 2024. Submitted, under review.

Abstract: *Direction of arrival* (DOA) estimation is a fundamental problem in signal processing and has applications in various fields such as radar, sonar, and acoustic. In this paper, we propose a method for DOA estimation using the received power at each microphone. The method is based on the directional sensitivity of the sensor elements at various frequencies. We model the directional sensitivity using a *Fourier series* (FS) model, where the parametric model enables *Cramér-Rao lower bound* (CRLB) computations. The FS model is estimated from measurements of a wideband noise signal. To estimate the DOA, the measured power profile is compared to the FS model using the *least squares* method.

We compare the proposed method to state-of-the-art methods in both an anechoic chamber and an outdoor scene. The results show that the proposed method performs better or equally well for all signals with a frequency content above 1000 Hz. Signals with low frequency content are not well suited for the proposed method, as the directional sensitivity becomes uniform for low frequencies. The

proposed method does not degrade with the size of the array since it utilizes the received power of the signal instead of the time-delays, enabling small arrays with great DOA resolution. Also, the outdoor data shows that the performance is great even a year after calibration, showcasing the robustness of the trained model.

Contribution: The idea originated from discussions between GZ, FG and GH. GZ developed the methodology with input from FG and GH. GZ implemented the method and performed the experiments. GZ wrote the paper. FG and GH provided editorial feedback.

Not Included Publication

The following publication is not included in the thesis to narrow the scope of the thesis. The contribution is described below.

The following abbreviations are used to refer to the author and co-authors: Gustav Zetterqvist (GZ), Fredrik Gustafsson (FG), Gustaf Hendeby (GH), Daniel Goderik (DG) and Albin Westlund (AW).

Seismic Detection of Elephant Footsteps

> Daniel Goderik, Albin Westlund, Gustav Zetterqvist, Fredrik Gustafsson, and Gustaf Hendeby. Seismic Detection of Elephant Footsteps. In *2024 IEEE 27th International Conference on Information Fusion (FUSION)*, Venice, Italy, July 2024. IEEE.

Abstract: As human settlement expands into the natural habitats of wild animals, the conflicts between humans and wildlife increases. The human-elephant conflict causes a tremendous amount of damage, often to poor villages close to the savannah. In this paper, we continue our earlier reported research on a geophone network aimed for elephant localisation by focusing on the detection challenge. We have now collected larger sets of seismic data with footsteps from both elephants and other big animals including humans. To detect the footsteps, a method is developed that analyses features of the geophone signal, which are then compared to those of an elephant footstep. The method detects 54% of the footsteps and has a classification accuracy of 89%. Subsequently, the detected elephant footstep is used to calculate the *direction of arrival* (DOA) angle using a delay-and-sum beamformer. The direction to an elephant is estimated with good precision on distances ranging from 8 to 30 meters. This research, not only, showcases a practical solution for mitigating human-elephant conflicts, but also underscores the potential of seismic technology in wildlife management and conservation efforts.

Contribution: FG had the idea. DG and AW developed the methods with input from GZ and FG. DG and AW implemented the method, collected the data and performed the experiments with support from GZ and FG. GZ wrote the majority of the paper with assistance from DG and AW. FG and GH provided editorial feedback.

1.4 Thesis Outline

Chapter 2 of this thesis presents the hardware setup as well as the different datasets used for the evaluation, covering both the seismic sensors and the microphone array. This is followed by Chapter 3, where the theory of DOA estimation is presented with different array structures, and the classical methods are derived, including the MUSIC and MVDR beamformer. In Chapter 4, we explore the potential use of seismic sensors for DOA estimation of elephant footsteps. The chapter presents the signal model, followed by the detection and DOA estimation method and concludes with the results.

The main contribution of this thesis is presented in Chapter 5, the DOA estimation method based on the directional sensitivity of the microphone array. Here, the signal model is presented, followed by the proposed method and concluded with a *Cramér-Rao lower bound* (CRLB) analysis. Chapter 6 presents the application of the proposed method, where it is evaluated both in an anechoic chamber and an outdoor environment, and compared to state-of-the-art methods. Also, the CRLB is evaluated for the proposed method. Finally, Chapter 7 concludes the thesis and outlines future work.

2

Experimental Setup

In this thesis we have carried out experiments in two different domains. The first one is the seismic domain, where we have used a sensor array with three geophones to estimate the DOA of elephant footsteps. The second is the acoustic domain, where a microphone array consisting of eight microphones has been used to estimate the DOA of different sound sources.

2.1 Seismic Sensor Array

In this section, the hardware design for the seismic sensor array is presented together with the experimental setup. Throughout the thesis, the following example will be considered.

Example 2.1: Elephant tracking

The human-elephant conflict poses significant risks to people, property, and crops worldwide. An effective solution involves an early warning system capable of detecting and alerting people of approaching elephants in time to implement mitigation measures.

We propose using a small local network of geophones to sense the seismic waves generated by elephant footsteps. These seismic waves are then used to estimate the direction of the elephant, as illustrated in Figure 2.1.

The primary goal is to detect and localize the elephant in real-time and to track its movement. This information can be used to warn nearby residents and help guide the elephant back to its natural habitat.

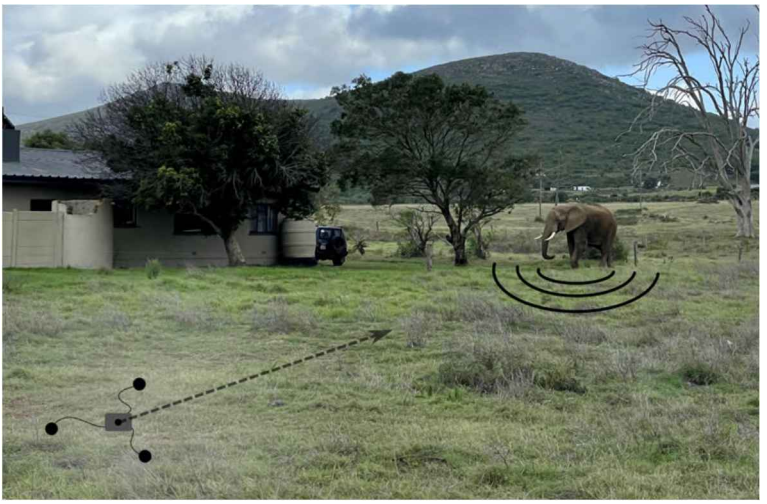


Figure 2.1: *Illustration of the first example – detection and localization of an elephant using a geophone network.*

2.1.1 Hardware

A prototype has been developed consisting of three geophones, SM-24 Geophone Element, with a bandwidth of 10 to 240 Hz [40]. A geophone is a sensor that converts ground vibrations into a voltage, and is typically used in seismology to detect earthquakes. The voltage from the analog geophones pass an *analog-to-digital converter* (ADC), ADS1256 [48], before being processed by a microcontroller board, Adafruit HUZZAH32 (ESP32-based, dual-core Tensilica LX6 microcontroller) [1]. The microcontroller is powered by a powerbank, and a handheld computer (Raspberry Pi 3 Model B) with a touchscreen is used to control the device. All components are illustrated in Figure 2.2a.

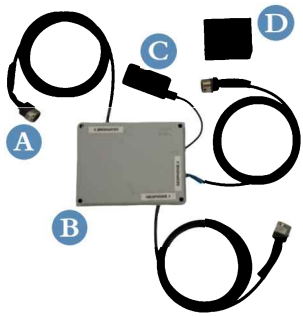
The geophones were positioned such that they formed the vertices in an equilateral triangle with a side length of 4 meters, as depicted in Figure 2.2b. This is typically known as a *uniform circular array* (UCA) and results in the geometric time-delays

$$\tau_m(\psi) = -\frac{r}{c}\cos\left(\frac{2\pi m}{3} - \frac{\pi}{2} + \psi\right), \quad \text{for } m = 1, 2, 3, \tag{2.1}$$

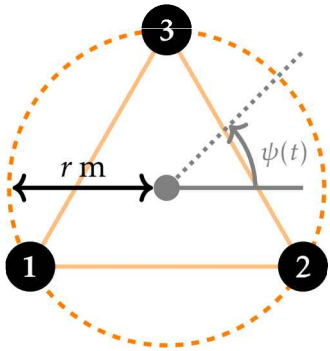
with respect to the center of the array. Where r is the radius of the array, m is the index of the geophone, ψ is the DOA angle, and c is the propagation speed of the seismic waves.

2.1.2 Experiments

The developed hardware prototype was then used to collect data at Kolmården Wildlife Park, with a sampling frequency of 474 Hz. This is high enough to avoid



*(a) Hardware overview of the complete system with the following items. **A** — Three geophones. **B** — Rain and dust resistant box containing a microcontroller, ADC and the circuitry for the geophones. **C** — Powerbank that supplies electricity to the microcontroller. **D** — Portable Raspberry Pi with a touchscreen and battery.*



(b) The setup of the geophone network, where the sensors are placed around a circle with a radius of r meters. This creates a UCA in the form of an equilateral triangle, the DOA angle is denoted $\psi(t)$.

***Figure 2.2:** The hardware components used and the setup of the geophone network.*

aliasing and almost entirely covers the bandwidth of the geophones.

First, the wave propagation speed was measured by dropping a large boulder at $\psi = 0°$ for one pair of geophones. Using cross-correlation to retrieve the time-delays, the wave propagation speed c was computed to be $c = 165\,\text{m/s}$.

Thereafter, a male Asian elephant walked from one side of the enclosure to the other, corresponding to an angle ψ of 30 to 150 degrees. The elephant kept a distance of approximately 15 to 40 meters from the array, as seen in Figure 2.3. The data was thereafter divided into training and validation data, where the training data was used to tune the detection algorithm and the validation data was used to verify the performance of the algorithm.

2.2 Acoustic Sensor Array

A hardware prototype consisting of eight microphones has been designed and used to collect data in an anechoic chamber and in an outdoor scenario. These datasets have also been utilized in previous work presented in [54] and [55], respectively. While we will examine various sound sources, our primary focus is on the natural sounds in the following example.

Figure 2.3: *The setup of the geophone network at Kolmården. The DOA angle is denoted $\psi(t)$. The elephant walks in front of the geophones network at a distance of 15–40 m.*

Example 2.2: Acoustic Surveillance for Savanna Protection

Effective wildlife protection on the savanna involves accurately localizing poachers and responding to various sound events. In this scenario, we use microphone arrays strategically placed across the savanna to monitor and analyze sounds.

The primary focus is on a rifle shot fired by a poacher. By analyzing the sound captured by the microphone arrays, we estimate the direction of the shot and thus the poacher's location.

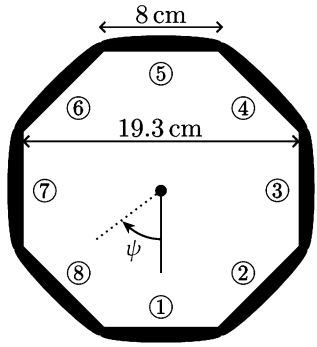
Additionally, we monitor other significant sounds: an elephant trumpet, a woman's scream, and a police siren. Each of these events is crucial for comprehensive wildlife protection. The elephant's location can inform conservation efforts, the woman's scream might indicate a need for emergency assistance, and the police siren could assist in coordinating the response.

The goal is to not only localize the poacher, but also to map out the locations of the elephant, the distressed individual, and the police to provide a complete situational overview and enhance overall response effectiveness.

2.2.1 Hardware

The microphone array used is a *uniform circular array* (UCA) in the form of an octagon with eight microphones. The microphones used are the hemispherical CBL99 from AKG [9] connected to a Behringer UMC1820 preamplifier [6]. The preamplifier is connected to a computer which allows for a sampling frequency of up to 96 kHz. The array and its dimensions are shown in Figure 2.4.

(a) The microphone array in the anechoic chamber.



(b) Top view of the dimension and microphone placement of the microphone array.

Figure 2.4: *An overview of the microphone array used in the acoustic parts of the thesis.*

2.2.2 Anechoic Measurements

Data were collected in an anechoic chamber, at the Swedish Defence Research Agency (FOI) in Linköping in May 2022, ISO rated from 100 Hz to 10 kHz. The microphone array was put on a turntable, and a Genelec 1029A speaker [13] was placed at a distance of 5 meters from the center of the array. In the setup, microphone 1 was facing the speaker at $\psi = 0°$, microphone 2 at $\psi = -45°$, and so on. The setup is illustrated in Figure 2.5. Thereafter, data were recorded with a sampling frequency of 48 kHz using 19 different types of signals. The signals and their frequency content are shown in Figure 2.6.

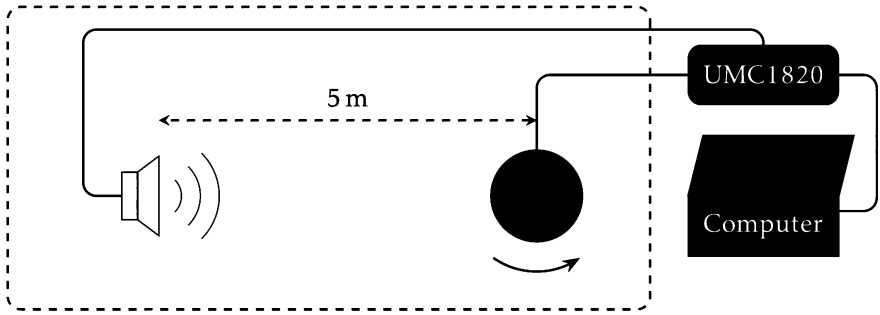
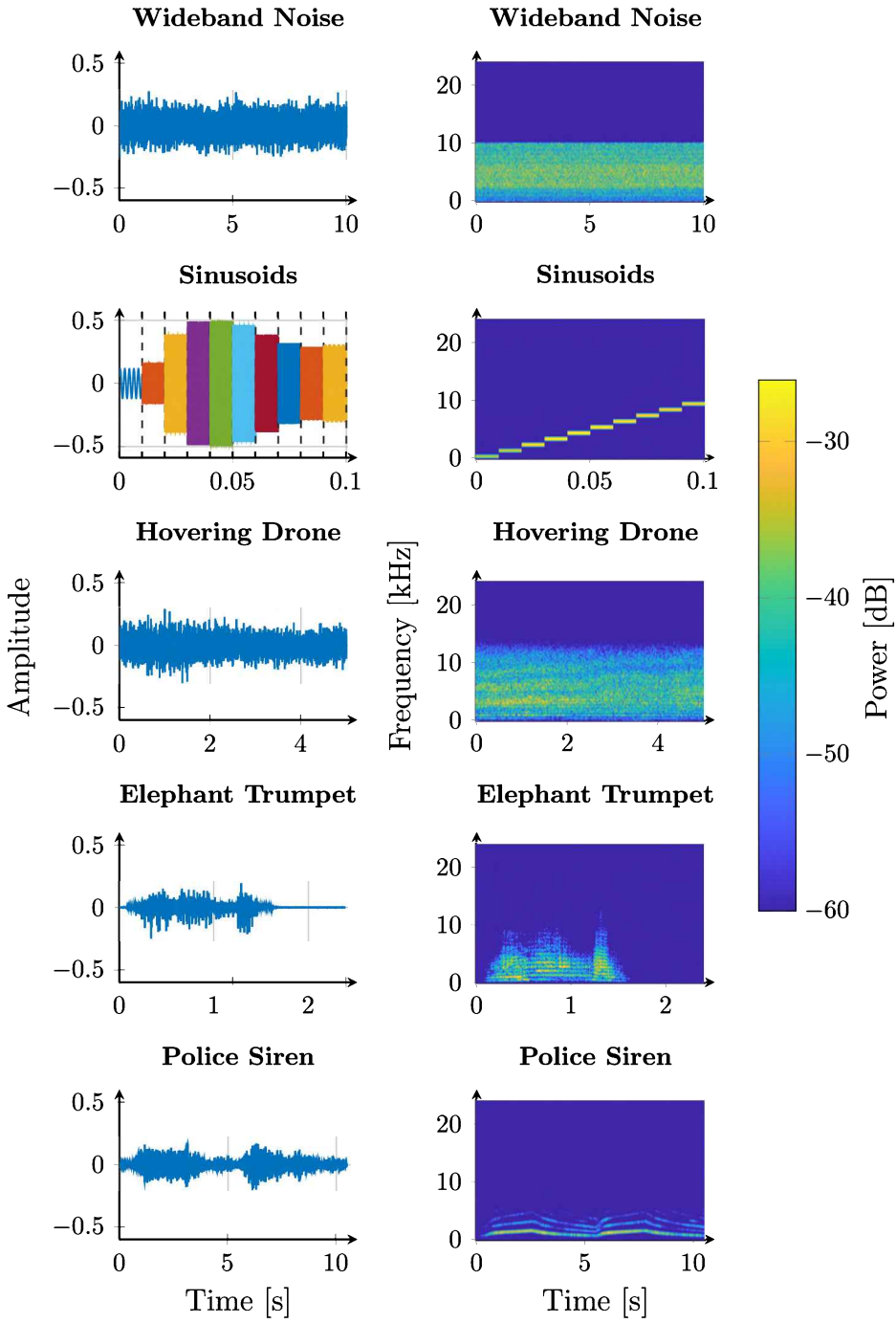


Figure 2.5: *The experimental setup. The microphone array is placed on a turntable 5 meters away from the speaker. The speakers and microphone array is connected to the preamplifier, which is connected to a computer.*


Wideband Noise
Wideband Noise
Sinusoids
Sinusoids
Hovering Drone
Hovering Drone
Elephant Trumpet
Elephant Trumpet
Police Siren
Police Siren
Amplitude
Frequency [kHz]
Power [dB]
Time [s]
Time [s]

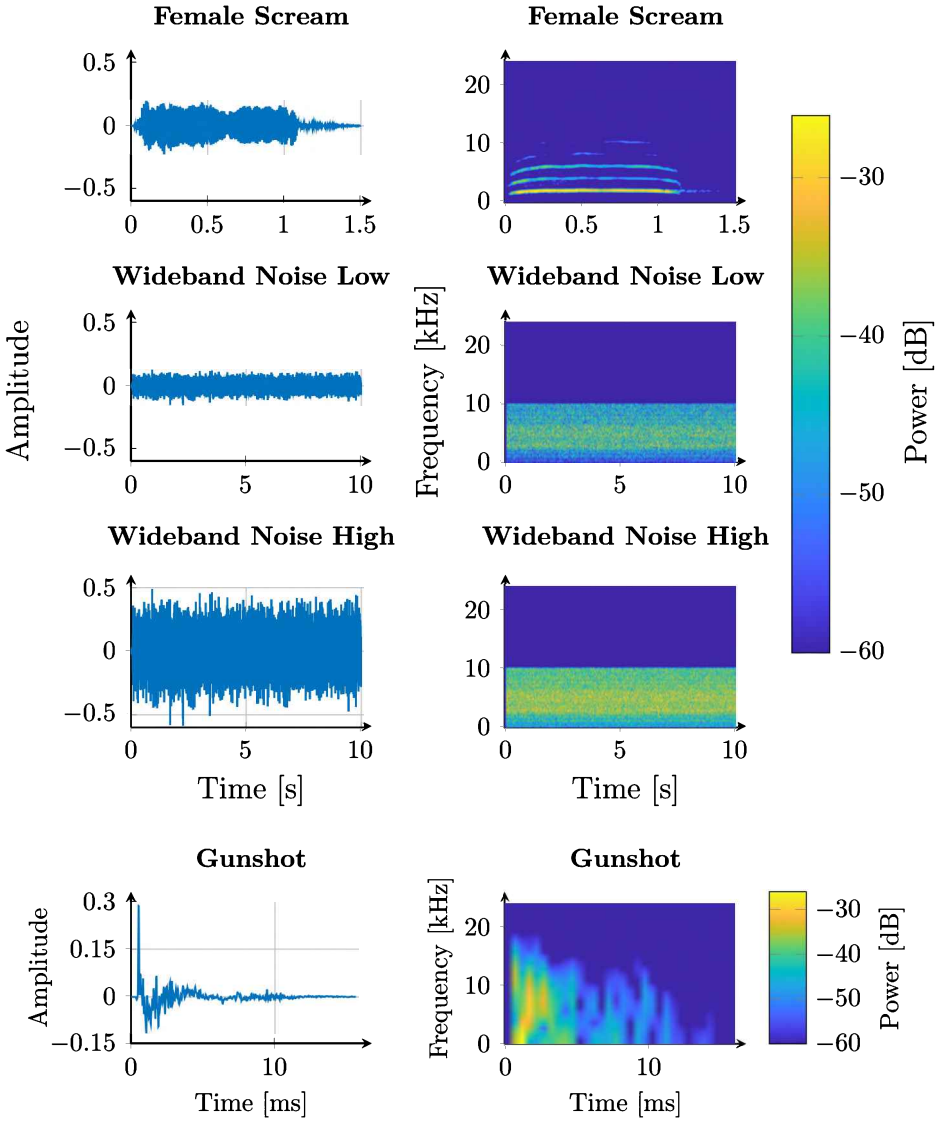


Figure 2.6: *Overview of signals 1–18 collected from microphone 1 in the anechoic chamber, positioned facing the speaker. The left column shows the signals in the time domain, while the right column shows their frequency content. The signals are arranged in the same order as described in the text. Note that only 10 ms of the sinusoids are shown, and the shot signal has a different amplitude and timescale.*

1) *Wideband noise* with a bandwidth of 100 Hz–10 kHz that lasted for 10 seconds.

2–11) *Sinusoids* with frequencies of 500–9500 Hz separated with 1 kHz for 1 second each.

12) *Hovering drone* sound during 5 seconds.

13) *Elephant trumpet* of an Indian elephant lasting 2.39 seconds [33].

14) *Police siren* lasting 10.48 seconds [50].

15) *Female scream* lasting 2.2 seconds [43].

16) *Attenuated wideband noise, i.e.,* the wideband noise signal with half the amplitude.

17) *Amplified wideband noise, i.e.,* the wideband noise signal with twice the amplitude.

18) *Gunshot* of a Smith & Wesson Chief's Special 9 mm revolver shot with blanks inside the anechoic chamber.

19) *Background noise* collected during 26.41 seconds.

First, the wideband signal (signal 1) was played from the speaker and data from 24 different angles were collected. This data was used as training data to estimate the directional sensitivity. Then, signal 1–17 were played from 24 different angles, not coincident with the training angles, to form the validation data. Signal 18 was recorded by shooting a blank from a revolver inside the anechoic chamber from 8 different angles, and was then included in the validation data. Finally, the background noise (signal 19) was recorded in the chamber. The data is available to listen to using the QR code on the bottom left of this page [53].

2.2.3 Outdoor Measurements

Data were also collected in an outdoor environment at Linköping University during December 2023. The weather conditions were calm with a light breeze and the temperature was around -4°C, with approximately 10 centimeters of snow on the ground. The microphone array was put in the middle, and then a circle was drawn around it with a radius of 5 m. Thereafter, points separated by 3.09 meters were then measured and marked, leading to 10 different points separated by 36°. The speaker, a Genelec 8030C [14], was then moved to these points. In total, data from 10 different angles was collected using signals 1–17 from the anechoic chamber. Background noise was also recorded for 26.41 seconds.

https://doi.org/10.5281/zenodo.13833589

3

Classical DOA Estimation

In this chapter, the problem of estimating the *direction of arrival* (DOA) of a signal is introduced. The DOA estimation problem is a well-studied topic, and there exists a wide range of methods for estimating the DOA of a signal with different characteristics. The most common methods are described in this chapter, including the delay-and-sum beamformer, the MUSIC algorithm, the Bartlett beamformer, the MVDR beamformer, and the MCCC algorithm. Finally, the algorithms are compared to each other using a simulation study, where the characteristics of the algorithms are demonstrated.

3.1 Signal Model

Assume a single source signal $s(t)$ measured by M sensors, the measured signal of the mth sensor at time t can then be expressed as

$$y_m(t) = \alpha_m s(t - \tau_m(\psi)) + e_m(t), \quad m = 1, 2, \ldots, M, \tag{3.1}$$

where α_m is the sensor dependent gain, $\tau_m(\psi)$ is the geometric time-delay to sensor m with respect to the local origin of the array (where $\tau(\psi) = 0$ for all ψ). The term $e_m(t)$ is referred to as the ambient noise, assumed to be uncorrelated with the signal source and the noise signals of the other sensors. Similarly, the discrete-time measurements can be expressed as

$$y_m[k] = \alpha_m s[k - \delta_m(\psi)] + e_m[k], \quad m = 1, 2, \ldots, M, \tag{3.2}$$

where $\delta_m(\psi) = f_s \tau_m(\psi)$ is the sample-delay. The setup of the estimation problem is seen in Figure 3.1.

The DOA problem is to find the incident angle of the waveform, ψ, given the measured signals. For most of the algorithms, the position of each sensor also

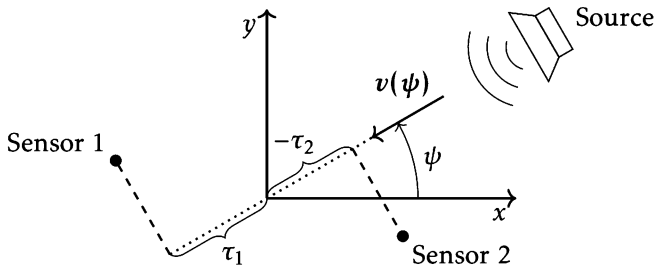


Figure 3.1: *Array geometry, where ψ is the DOA angle of the source and τ_m is the time-delay to sensor m. The direction vector of the signal is denoted $\boldsymbol{v}(\psi)$.*

needs to be known. Many algorithms are based on the complex-valued analytic signal. To transform a signal from real-valued to the analytic signal, the Hilbert transform can be applied [28].

Assume that the sensor gain coefficients are all equal to 1, and that the signal is narrowband, *i.e.*, it occupies a small bandwidth relative to its central frequency f. The Fourier transform of (3.1) can then be expressed as

$$Y_m(f) = e^{-j2\pi\tau_m(\psi)f} S(f) + E_m(f), \tag{3.3}$$

where $Y_m(f)$, $S(f)$ and $E_m(f)$ are the Fourier transform of $y_m(t)$, $s(t)$ and $e_m(t)$, respectively. Then, if the frequency-domain vector is defined as

$$\boldsymbol{Y} = \begin{bmatrix} Y_1(f) & Y_2(f) & \dots & Y_M(f) \end{bmatrix}^T, \tag{3.4}$$

this yields

$$\boldsymbol{Y} = \boldsymbol{a}(\psi)S(f) + \boldsymbol{E}, \tag{3.5}$$

where

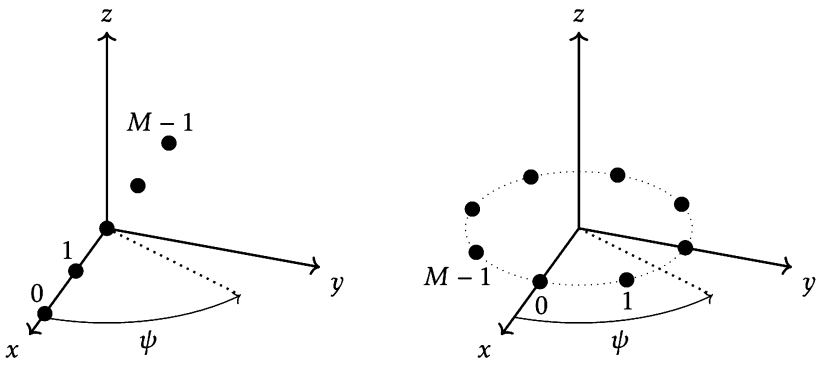
$$\boldsymbol{a}(\psi) = \begin{bmatrix} e^{-j2\pi\tau_1(\psi)f} & e^{-j2\pi\tau_2(\psi)f} & \dots & e^{-j2\pi\tau_M(\psi)f} \end{bmatrix}^T, \tag{3.6}$$

and $\boldsymbol{E}$ is a frequency-domain vector of the noise. The vector $\boldsymbol{a}(\psi)$ is known as the array manifold vector, or steering vector, and is used in the majority of the DOA estimators in the literature [45, 49].

3.2 Array Structures

DOA estimation can be performed using different array structures. In general, if the positions of the sensor elements are known in the xy-plane

$$\boldsymbol{p}_m = \begin{bmatrix} p_x & p_y & 0 \end{bmatrix}^T, \quad m = 0, 1, \dots, M-1, \tag{3.7}$$



(a) A ULA with 5 sensor elements along the x-axis.

(b) A UCA with 8 sensor elements in the xy-plane.

Figure 3.2: *Different array geometries for DOA estimation, with ψ defined as the DOA of the source signal.*

where $\boldsymbol{p}_m$ represents the position of the mth sensor element. The time-delay is then expressed as

$$\tau_m(\psi) = \frac{\boldsymbol{p}_m^T \boldsymbol{v}(\psi)}{c} \tag{3.8}$$

$$\boldsymbol{v}(\psi) = -\begin{bmatrix}\cos(\psi) & \sin(\psi) & 0\end{bmatrix}^T, \tag{3.9}$$

where $\boldsymbol{v}(\psi)$ is the direction vector of the signal arriving from ψ, as depicted in Figure 3.1, and c is the speed of the signal.

3.2.1 Uniform Linear Array

One common array is the *uniform linear array* (ULA). A ULA is a sensor array with M elements uniformly spaced, as depicted in Figure 3.2a.

The center of this array is conveniently defined to be in the origin, which will have computational advantages. Denoting the distance between the sensor elements as d, the position of each sensor in a ULA can be described as

$$\boldsymbol{p}_m = \begin{bmatrix}\left(\frac{M-1}{2} - m\right)d & 0 & 0\end{bmatrix}^T, \quad m = 0, 1, \ldots, M-1. \tag{3.10}$$

This gives the time-delay

$$\tau_m(\psi) = -\left(\frac{M-1}{2} - m\right)\frac{d\cos(\psi)}{c}, \tag{3.11}$$

and thus the array manifold vector

$$\boldsymbol{a}(\psi) = \begin{bmatrix}e^{-j2\pi f\left(\frac{M-1}{2}\right)\frac{d\cos(\psi)}{c}} & e^{-j2\pi f\left(\frac{M-1}{2}-1\right)\frac{d\cos(\psi)}{c}} & \cdots & e^{j2\pi f\left(\frac{M-1}{2}\right)\frac{d\cos(\psi)}{c}}\end{bmatrix}^T. \tag{3.12}$$

A limitation of ULAs is that the angle ψ is constrained to the interval $[0°, 180°]$. This is due to the fact that sources symmetrical with respect to the x-axis will produce identical sets of delays, resulting in ambiguity in direction estimation. Denote the signal wavelength λ as

$$\lambda = \frac{c}{f}, \tag{3.13}$$

and define

$$f_s = \frac{d\cos\psi}{\lambda} = \frac{f d\cos\psi}{c}, \tag{3.14}$$

and

$$\omega_s = 2\pi f_s. \tag{3.15}$$

The array manifold vector $\boldsymbol{a}(\psi)$ can then be rewritten as

$$\boldsymbol{a}(\psi) = \left[e^{-j\omega_s\left(\frac{M-1}{2}\right)} \quad e^{-j\omega_s\left(\frac{M-1}{2}-1\right)} \quad \dots \quad e^{j\omega_s\left(\frac{M-1}{2}\right)}\right]^T, \tag{3.16}$$

also known as a Vandermonde vector [45]. By this definition of the array manifold vector, ω_s is called the spatial frequency because of the similarity with the *discrete Fourier transform* (DFT).

When a continuous-time sinusoidal with frequency f is sampled, the sampling frequency f_0 must satisfy

$$f_0 \geq 2f, \tag{3.17}$$

to avoid aliasing effects according to the Shannon sampling theorem [41]. In the ULA case it is seen that the array manifold vector $\boldsymbol{a}(\psi)$ is uniquely defined (*i.e.*, there is no spatial aliasing) if and only if

$$|\omega_s| \leq \pi, \tag{3.18}$$

which is equivalent to

$$|f_s| \leq \frac{1}{2} \iff d\left|\cos\psi\right| \leq \frac{\lambda}{2}. \tag{3.19}$$

This condition depends on the DOA angle ψ, and in particular if the source is at $\psi = \pm 90°$, no constraints apply on the element spacing d. However, as the DOA is generally unknown, ensuring that the condition is satisfied for any ψ leads to the constraint

$$d \leq \frac{\lambda}{2}. \tag{3.20}$$

This result may be interpreted as the spatial Shannon sampling theorem [45].

3.2.2 Uniform Circular Array

Another common array design is the *uniform circular array* (UCA). A UCA is a sensor array with M elements evenly spaced around a circle, as depicted in Figure 3.2b. The position of each sensor in a UCA, assuming the radius of the circle is r, can be described as

$$\boldsymbol{p}_m = r\left[\cos(2\pi\tfrac{m}{M}) \quad \sin(2\pi\tfrac{m}{M}) \quad 0\right]^T, \quad m = 0, 1, \dots, M-1. \tag{3.21}$$

The time-delay is then given by

$$\tau_m(\psi) = -\frac{r}{c}\cos(2\pi\frac{m}{M} - \psi), \tag{3.22}$$

and thus the array manifold vector is

$$\boldsymbol{a}(\psi) = \begin{bmatrix} e^{-j2\pi f\tau_1(\psi)} & e^{-j2\pi f\tau_2(\psi)} & \dots & e^{-j2\pi f\tau_M(\psi)} \end{bmatrix}^T. \tag{3.23}$$

3.3 Estimation Methods

The DOA estimation problem is a well-studied topic, and during the years several methods have been developed with different characteristics. In this section, the most common ones are described.

3.3.1 Delay-and-Sum

The delay-and-sum beamformer, also known as the conventional beamformer, is a well-known and simple approach. This method is based on the geometry of the array and calculates the DOA by measuring the power of the summed signals at each possible arrival angle. The maximum of the power corresponds to the estimated angle of arrival

$$y_{\text{DS}}[k, \psi] = \frac{1}{M}\sum_{m=1}^{M} y_m[k + \delta_m(\psi)] \tag{3.24a}$$

$$V_{\text{DS}}(\psi) = \sum_k \left|y_{\text{DS}}[k, \psi]\right|^2 \tag{3.24b}$$

$$\hat{\psi}^{\text{DS}} = \arg\max_{\psi} V_{\text{DS}}(\psi). \tag{3.24c}$$

The general idea of the delay-and-sum beamformer is to improve the *signal to noise ratio* (SNR) of the signal. An overview of the algorithm is seen in Figure 3.3.

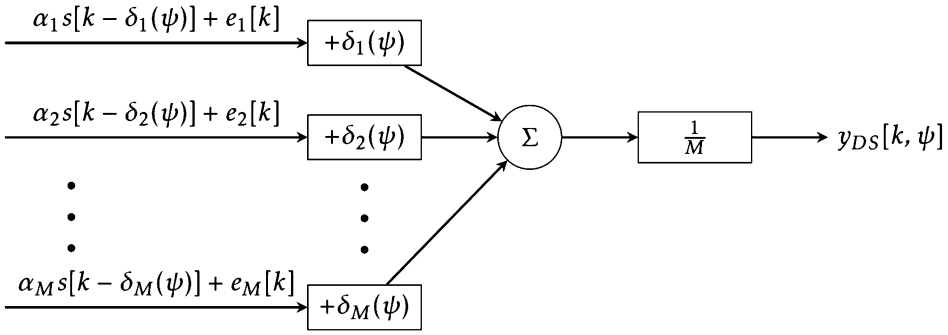


Figure 3.3: *An overview of the delay-and-sum beamformer in the time domain. The signals from the different sensors are delayed according to the DOA angle ψ and then summed together to improve the SNR of the signal.*

For computational efficiency, the same method can also be performed in the frequency domain for narrowband signals. From (3.5), we have

$$\boldsymbol{Y} = \boldsymbol{a}(\psi)S(f) + \boldsymbol{E}.$$

The reconstructed signal can therefore be calculated as

$$\boldsymbol{Y}_{\text{DS}}(\psi) = \boldsymbol{a}(\psi)^H \boldsymbol{Y} \tag{3.25}$$

$$y_{\text{DS}}(\psi) = \frac{1}{M}\mathcal{F}^{-1}\left\{\boldsymbol{Y}_{\text{DS}}(\psi)\right\}, \tag{3.26}$$

where $(\cdot)^H$ denotes the Hermitian transpose. The DOA is then estimated as in (3.24) [49]. The loss function can also be calculated in the frequency domain using Parseval's theorem as

$$V_{\text{DS}}(\psi) = \frac{1}{N}\sum_f |\boldsymbol{Y}(f)|^2, \tag{3.27}$$

where $\boldsymbol{Y}(f)$ is the DFT of the signal and N is the number of time-domain samples.

3.3.2 MUSIC

One popular algorithm to estimate the DOA is *multiple signal classification* (MUSIC), and was invented by Schmidth in 1986 [39]. MUSIC is a subspace algorithm, especially designed for narrowband signals. Assume that $e_m(t)$ are mutually independent Gaussian random processes with the same variance, and that the sensor gain coefficients are all equal to 1. From the formulation of the signal in the frequency-domain in (3.5)

$$\boldsymbol{Y} = \boldsymbol{a}(\psi)S(f) + \boldsymbol{E},$$

the covariance matrix of $\boldsymbol{Y}$ is

$$R_{\boldsymbol{Y}} = E\left\{\boldsymbol{Y}\boldsymbol{Y}^H\right\} \tag{3.28}$$

$$= \boldsymbol{a}(\psi)E\left\{S(f)S^H(f)\right\}\boldsymbol{a}^H(\psi) + E\left\{\boldsymbol{E}\boldsymbol{E}^{\boldsymbol{H}}\right\} \tag{3.29}$$

$$= \sigma_S^2\boldsymbol{a}(\psi)\boldsymbol{a}^H(\psi) + \sigma_E^2\boldsymbol{I}, \tag{3.30}$$

where σ_S^2 and σ_E^2 are the signal and noise variances, respectively. It is easily seen that the rank of $\sigma_S^2\boldsymbol{a}(\psi)\boldsymbol{a}^H(\psi)$ is 1. By performing eigenvalue decomposition of the covariance matrix, we get

$$R_{\boldsymbol{Y}} = \boldsymbol{Q}\boldsymbol{\Lambda}\boldsymbol{Q}^H, \tag{3.31}$$

where $\boldsymbol{\Lambda}$ is a diagonal matrix with the eigenvalues and $\boldsymbol{Q}$ contains the eigenvectors of $R_{\boldsymbol{Y}}$, respectively

$$\boldsymbol{\Lambda} = \text{diag}\begin{bmatrix}\lambda_{Y,1} & \lambda_{Y,2} & \dots & \lambda_{Y,M}\end{bmatrix} \tag{3.32}$$

$$= \text{diag}\begin{bmatrix}\sigma_S^2\lambda_{\boldsymbol{a}} + \sigma_E^2 & \sigma_E^2 & \dots & \sigma_E^2\end{bmatrix} \tag{3.33}$$

$$\boldsymbol{Q} = \begin{bmatrix}\boldsymbol{q}_1 & \boldsymbol{q}_2 & \dots & \boldsymbol{q}_M\end{bmatrix}. \tag{3.34}$$

It is seen that for $m \geq 2$, the following holds

$$R_{\boldsymbol{Y}} \boldsymbol{q}_m = \sigma_E^2 \boldsymbol{q}_m. \tag{3.35}$$

Combined with (3.30) this yields

$$\left[\sigma_S^2 \boldsymbol{a}(\psi)\boldsymbol{a}^H(\psi) + \sigma_E^2 \boldsymbol{I}\right] \boldsymbol{q}_m = \sigma_E^2 \boldsymbol{q}_m, \tag{3.36}$$

which implies that

$$\sigma_S^2 \boldsymbol{a}(\psi)\boldsymbol{a}^H(\psi)\boldsymbol{q}_m = 0, \quad m \geq 2. \tag{3.37}$$

Or equivalently,

$$\boldsymbol{a}^H(\psi)\boldsymbol{q}_m = 0, \quad m \geq 2. \tag{3.38}$$

This means that the eigenvectors corresponding to the $M-1$ lowest eigenvalues of $R_{\boldsymbol{Y}}$ are orthogonal to the vector corresponding to the actual DOA. By formulating the Euclidean distance between these vectors as $\boldsymbol{a}^H(\psi)\boldsymbol{q}_m\boldsymbol{q}_m^H\boldsymbol{a}(\psi)$, the cost function is defined as

$$V_{\text{MUSIC}}(\psi) = \frac{1}{\sum_{m=2}^{M} \boldsymbol{a}^H(\psi)\boldsymbol{q}_m\boldsymbol{q}_m^H\boldsymbol{a}(\psi)}, \tag{3.39}$$

where the DOA that gives the maximum $V_{\text{MUSIC}}(\psi)$ corresponds to the estimated DOA

$$\hat{\psi}^{\text{MUSIC}} = \arg\max_{\psi} V_{\text{MUSIC}}(\psi). \tag{3.40}$$

The MUSIC algorithm can easily be extended to support multiple signal sources, as long as the number of sensors is greater than the number of sources, hence its name [49].

In the application, the covariance matrix, $R_{\boldsymbol{Y}}$, cannot be calculated directly. Instead, the sample covariance matrix $\tilde{R}_{\boldsymbol{Y}}$ is used

$$\tilde{R}_{\boldsymbol{Y}} = \frac{1}{N}\sum_{k=1}^{N} \boldsymbol{y}[k]\boldsymbol{y}^H[k], \tag{3.41}$$

where $\boldsymbol{y}[k]$ is a vector of the measured signals at time k. The sample covariance matrix is the ML estimator of $R_{\boldsymbol{Y}}$, and $\tilde{R}_{\boldsymbol{Y}} \to R_{\boldsymbol{Y}}$ as $N \to \infty$ [19].

3.3.3 Bartlett

The Bartlett beamformer is based on the amplitude of the signal multiplied with the array manifold vector $\boldsymbol{a}(\psi)$,

$$V_{\text{B}}(\psi) = \frac{1}{M}\sum_{m=1}^{M} \left|\boldsymbol{a}^H(\psi)Y_m\right|^2, \tag{3.42}$$

which can be rewritten as

$$V_{\text{B}}(\psi) = \boldsymbol{a}^H(\psi)R_{\boldsymbol{Y}}\boldsymbol{a}(\psi), \tag{3.43}$$

where the same covariance matrix $R_{\boldsymbol{Y}}$ as well as the array manifold vector $\boldsymbol{a}(\psi)$ is used as in the MUSIC algorithm [49]. The DOA is thereafter calculated as

$$\hat{\psi}^{\mathrm{B}} = \arg\max_{\psi} V_{\mathrm{B}}(\psi). \tag{3.44}$$

3.3.4 MVDR (Capon)

Another well-known algorithm is the *minimum variance distortionless response* (MVDR) beamformer, also known as the Capon beamformer after its creator [10]. This beamformer uses the same covariance matrix $R_{\boldsymbol{Y}}$ as in the MUSIC algorithm, as well as the array manifold vector $\boldsymbol{a}(\psi)$, to calculate the DOA [45].

$$V_{\mathrm{MVDR}}(\psi) = \frac{1}{M}\sum_{m=1}^{M}\left|\frac{\boldsymbol{a}^H(\psi)R_{\boldsymbol{Y}}^{-1}}{\boldsymbol{a}^H(\psi)R_{\boldsymbol{Y}}^{-1}\boldsymbol{a}(\psi)}Y_m\right|^2, \tag{3.45}$$

which can be reduced to

$$V_{\mathrm{MVDR}}(\psi) = \frac{1}{\boldsymbol{a}^H(\psi)R_{\boldsymbol{Y}}^{-1}\boldsymbol{a}(\psi)}. \tag{3.46}$$

The DOA is thereafter calculated as

$$\hat{\psi}^{\mathrm{MVDR}} = \arg\max_{\psi} V_{\mathrm{MVDR}}(\psi). \tag{3.47}$$

3.3.5 MCCC

Another popular method is the MCCC algorithm, which is based on the correlation between the signals from the different sensors [8, 23]. From the signal model in (3.1), the cross-correlation matrix of the signals can be expressed as

$$R_{\boldsymbol{Y}}(\tau(\psi)) = E\left\{\boldsymbol{Y}\boldsymbol{Y}^H\right\} = \begin{bmatrix} \sigma_{y_1}^2 & r_{y_1y_2}(\tau(\psi)) & \dots & r_{y_1y_M}(\tau(\psi)) \\ r_{y_2y_1}(\tau(\psi)) & \sigma_{y_2}^2 & \dots & r_{y_2y_M}(\tau(\psi)) \\ \vdots & \vdots & \ddots & \vdots \\ r_{y_My_1}(\tau(\psi)) & r_{y_My_2}(\tau(\psi)) & \dots & \sigma_{y_M}^2 \end{bmatrix}. \tag{3.48}$$

The correlation matrix $R_{\boldsymbol{Y}}$ can be factorized as

$$R_{\boldsymbol{Y}}(\tau(\psi)) = \Sigma\tilde{\boldsymbol{R}}(\tau(\psi))\Sigma, \tag{3.49}$$

where

$$\Sigma = \begin{bmatrix} \sigma_{y_1} & 0 & \dots & 0 \\ 0 & \sigma_{y_2} & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & \sigma_{y_M} \end{bmatrix}, \tag{3.50}$$

and

$$\tilde{\boldsymbol{R}}(\tau(\psi)) = \begin{bmatrix} 1 & \rho_{y_1y_2}(\tau(\psi)) & \dots & \rho_{y_1y_M}(\tau(\psi)) \\ \rho_{y_2y_1}(\tau(\psi)) & 1 & \dots & \rho_{y_2y_M}(\tau(\psi)) \\ \vdots & \vdots & \ddots & \vdots \\ \rho_{y_My_1}(\tau(\psi)) & \rho_{y_My_2}(\tau(\psi)) & \dots & 1 \end{bmatrix}, \tag{3.51}$$

where $\rho_{y_my_n}(\tau(\psi))$ is the normalized cross-correlation between the signals from sensor m and n

$$\rho_{y_my_n}(\tau(\psi)) = \frac{r_{y_my_n}(\tau(\psi))}{\sqrt{\sigma_{y_m}\sigma_{y_n}}}. \tag{3.52}$$

Since the matrix $\tilde{\boldsymbol{R}}(\tau(\psi))$ is symmetric and positive definite, and its diagonal elements are equal to 1, it can be shown that the eigenvalues of $\tilde{\boldsymbol{R}}(\tau(\psi))$ lie in the interval $[0, 1]$ [8], and hence

$$0 \leq \det \tilde{\boldsymbol{R}}(\tau(\psi)) \leq 1. \tag{3.53}$$

With only two sensors, the squared correlation coefficients are given by

$$\rho^2_{y_1y_2}(\tau(\psi)) = 1 - \det \tilde{\boldsymbol{R}}(\tau(\psi)). \tag{3.54}$$

Subsequently, the squared multichannel cross-correlation coefficient is defined as

$$\begin{aligned} \rho^2_{y_1:y_M}(\tau(\psi)) &= 1 - \det \tilde{\boldsymbol{R}}(\tau(\psi)) \\ &= 1 - \frac{\det \boldsymbol{R}(\tau(\psi))}{\prod_{m=1}^{M} \sigma^2_{y_m}}, \end{aligned} \tag{3.55}$$

The MCCC algorithm is based on the maximum of the multichannel cross-correlation coefficient, and the DOA is calculated as

$$\hat{\psi}^{\text{MCCC}} = \arg\max_{\psi} \rho^2_{y_1:y_M}(\tau(\psi)). \tag{3.56}$$

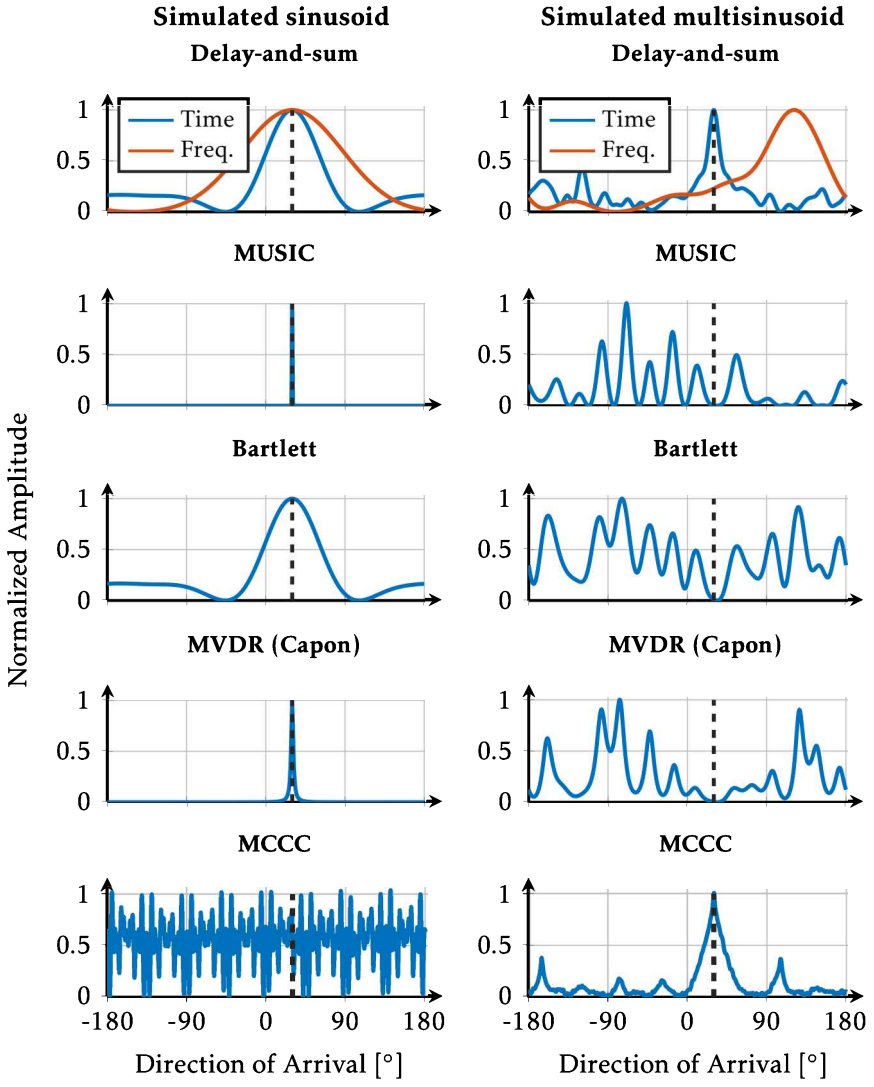
With (3.55) this can be rewritten as

$$\begin{aligned} \hat{\psi}^{\text{MCCC}} &= \arg\max_{\psi} \left\{1 - \det \tilde{\boldsymbol{R}}(\tau(\psi))\right\} \\ &= \arg\max_{\psi} \det \left\{1 - \frac{\det \boldsymbol{R}(\tau(\psi))}{\prod_{m=1}^{M} \sigma^2_{y_m}}\right\} \\ &= \arg\min_{\psi} \det \boldsymbol{R}(\tau(\psi)). \end{aligned} \tag{3.57}$$

The MCCC algorithm can easily be extended to support multiple signal sources, as long as the number of sensors is greater than the number of sources [8].

3.4 Comparision

The different algorithms have different characteristics, and the choice of algorithm depends on the application. In Figure 3.4, the cost functions of the algorithms are compared for two different signals. The first signal is a 1000 Hz pure sine signal, and the second signal is a multisine signal with frequencies 1, 2.5, 5 and 7 kHz. It is seen that the delay-and-sum algorithm in the time domain is able to estimate the correct DOA for both of the signals, as it does not rely on the signal being narrowband. However, the delay-and-sum algorithm in the frequency domain is not able to correctly estimate the DOA for the multisine signal, since it assumes a narrowband signal. The same applies for the MUSIC, Bartlett and MVDR algorithms. The MCCC algorithm fails to estimate the DOA for the sine signal, probably due to phase ambiguities, *i.e.*, the periodic nature of the sine signal causes multiple local maxima in the cost function.



(a) 1000 Hz sine signal

(b) Multisine signal: 1, 2.5, 5 and 7 kHz

Figure 3.4: *Cost functions of the DOA estimation algorithms for two simulated signals. The signals are generated using the microphone array described in Section 2.2, with a true DOA of 30° and a SNR of 20 dB. The signals are sampled at a frequency of 48 kHz and simulated over a duration of 0.1 seconds. The black dotted line indicates the actual DOA of the signals.*

4

Application — Elephant DOA

In this chapter we investigate an interesting application of DOA estimation, namely the localization of elephants. We use a geophone array to detect and estimate the DOA of an elephant. Due to the omnidirectional sensor elements, the DOA is estimated using a conventional delay-and-sum beamformer. This chapter is based on the work presented in [56].

4.1 Background

One of the main problems connected to national parks is human-wildlife conflicts. In 2020, over 90 people were killed by wildlife in Kenya, and six people died by elephants in Amboseli National Park alone [29]. Hundreds of casualties caused by elephants are reported in India each year and local communities suffer from destroyed fields and property [12]. It is costly to protect properties such as schools using fence lines, and even then it often fails due to lack of maintenance [18]. Thus, an automated warning system capable of detecting and localizing elephants long before they arrive could prove to be very useful. Here, the seismic waves from elephants' footsteps are used to detect and locate the elephants. Mitigation measures may include lights, sounds and smells disliked by the elephants, or even release of bee swarms.

Detection of seismic events using geophones is a well-studied topic. Studied seismic events include human footsteps [5, 25], vehicles [2, 15], landslides [46], etc. Detecting elephants using geophones has also been carried out in the past. The general idea is to utilize that the pressure waves generated by elephants can travel very far [30, 31, 36]. However, in most cases it is not the footsteps that are detected, but rather vocalization, also known as rumbles, that generates seismic components [31, 36]. Rumbles are produced by the elephants for various reasons, *e.g.*, to warn other elephants about a threat, as a greeting or to advertise mating

to the other sex [32]. The detection of elephant footsteps has previously been studied, *e.g.*, [52] uses a frequency analysis approach and [47] employs a machine learning approach. However, the level of detail in the description make these methods difficult to replicate.

The seismic waves produced by elephant vocalization have additionally been used for localization of elephants [36]. When it comes to safeguarding a village, it is essential to localize the elephants as they are moving. Therefore, the approach here is to use their footsteps for localization. Tracking a seismic source by using a network of geophones has been done before in a seismic security system [3]. However, the application to elephants is novel.

This chapter presents a well-documented approach for the tracking of elephant footsteps using seismic signatures. Our methodology incorporates a frequency-based approach to detect the signature of elephant footsteps. Thereafter, the DOA of the detected footstep is accurately estimated using a sensor fusion approach. Finally, the movement of elephants is effectively tracked using a *Kalman filter* (KF) and a gating technique that rejects unlikely DOA estimates. Our results demonstrate the effectiveness of this approach, validated at distances ranging from 15 to 40 meters, which is promising for practical application in the field.

4.2 Signal Processing Chain

In this section the signal model and characteristics are presented as well as the methods for detection, DOA estimation and tracking.

4.2.1 Signal Model

It is assumed that geophone m measures the signal

$$y_m(t) = \sum_n s_m(t - \kappa_n) + e_m(t) \tag{4.1}$$

$$s_m(t) = s(t - \tau_m(\psi)), \tag{4.2}$$

where

- $s(t)$ is the seismic signature of an elephant footstep at time $t = 0$.
- κ_n denotes the time for footstep number n.
- $\psi(t)$ is the DOA at time t. The array is assumed to satisfy a far-field assumption, *i.e.*, the DOA is the same for all geophones.
- $\tau_m(\psi)$ is the geometric time-delay to geophone m with respect to the local origin of the array (where $\tau(\psi) = 0$ for all ψ).
- $e_m(t)$ is all the non-interesting signals measured by geophone m, *e.g.*, electrical noise, sensor bias or other ambient signals.

4.2.2 Signal Characteristics

Seismologic sources produce both body and surface waves. The surface waves tend to have a low frequency, usually not higher than 200 Hz, and spread in a two-dimensional space which results in a decay of $1/\sqrt{r}$. On the other hand, body waves spread in a three-dimensional space, resulting in a decay of $1/r$, where r is the traveled distance of the wave. There are two types of surface waves, Rayleigh and Love waves. Rayleigh waves cause an elliptical motion in the vertical plane while Love waves cause horizontal ground movement. Rayleigh waves tend to be stronger from near-surface seismic sources as opposed to deeper underground sources. Furthermore, Rayleigh waves tend to be stronger than other wave forms on the Earth's surface, hence most of the shaking felt from earthquakes originates from Rayleigh waves [4].

Measurements of elephant rumbles and "foot stomps" were conducted by O'Connell-Rodwell et. al. [31] using both microphones and geophones. The findings indicate that a "foot stomp" has a duration of 103–250 ms, a mean frequency of 20.04 Hz, and a wave propagation speed of 248–264 m/s. The slow propagation speed suggests that it is a Rayleigh wave [31].

An important aspect of the seismic signals are the noise sources. Every location has different noise characteristics, but the most common noise sources around 0.5–50 Hz are wind, thunder and cultural noise. Wind usually creates seismic waves close to the surface at around 0.5–5 Hz [51], while the spectra of thunder is broadband around 5–100 Hz with a peak around 6–13 Hz [21]. Cultural noise is a term associated with man or man-made machines, *e.g.*, power plants, factories, trains, highways, etc. Power plants have a narrowband seismic signal around 50 or 60 Hz with harmonics and subharmonics dependent on the environment [51]. The seismic footprint of trains and cars is broadband with a peak around 15 Hz [37].

4.2.3 Pre-Filtering and Detection

To analyze if there are elephants present, and if so where, the signal is split into segments with overlapping samples. Since the purpose is to run the system in real time, it is not possible to run the algorithm for each new sample as the computational load would be too heavy. If an elephant footstep has a duration of $N/2$ samples, using a segment that is N samples long and has $N/2$ overlapping samples with adjacent segments, the full footstep is guaranteed to be present in at least one segment.

The method for detection of an elephant footstep in the signal segment is based on the frequency content of the signal. First, the signal segment is filtered with a band-pass filter to remove sensor bias, high-frequency noise and other ambient signals while preserving the elephant footstep. This filter is run both forwards and backwards in time to avoid phase shifting. As the Fourier transform assumes that the length of the signal is infinite, the signal is then windowed using a Hanning window to prevent leakage.

Subsequently, the *fast Fourier transform* (FFT) is computed whereupon the magnitude of the FFT (the periodogram) is computed. To make the algorithm

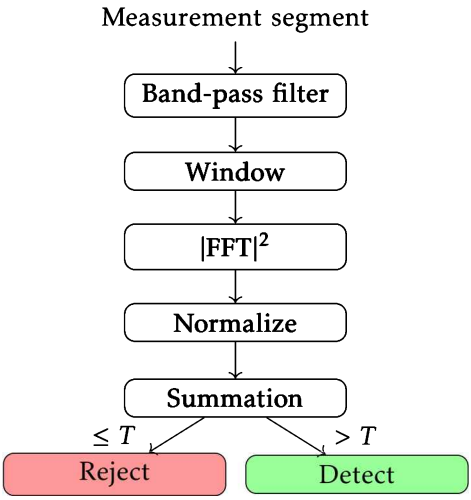


Figure 4.1: *An overview of the detection algorithm.*

independent of the strength of the signal, the periodogram is normalized with the total energy of the periodogram. Summing the energy of the normalized periodogram around the main frequencies of an elephant footstep thus provides an estimate of how strongly the signal resembles an elephant's footstep. This algorithm is applied for each geophone individually, and then for detection it is required that X out of M geophones receive a detection. An overview of the algorithm is seen in Figure 4.1.

There are a few parameters in this method that have to be tuned depending on what standing waves are present in the ground, and these may vary depending on location and weather conditions. The tuning parameters are the length of the signal segments N, the cut-off frequencies of the band-pass filter, the frequency range of the periodogram to sum, the threshold T for which the energy in the summed frequencies has to surpass for detection, and how many geophones need to get a detection.

4.2.4 DOA Estimation and Tracking

The seismic signal measured by geophone m is given by the signal model in (4.1). Using the frequency analysis explored in Section 4.2.3, a band-pass filter is applied to remove most of the noise $e_m(t)$ in the signal. Some noise may still be present after filtering, but since the noise is assumed to be small in amplitude compared to the signal of interest it is neglected in the following calculations.

DOA Estimation

To estimate the DOA of the elephant footstep, delay-and-sum is used on the measurements of $y_m(t)$ [49],

$$y_{DS}(t, \psi) = \frac{1}{M} \sum_{m=1}^{M} y_m(t + \tau_m(\psi)) \tag{4.3a}$$

$$V_N(\psi) = \int_{t \in W} \left| y_{DS}(t, \psi) \right|^2 dt \tag{4.3b}$$

$$\hat{\psi} = \arg\max_{\psi} V_N(\psi), \tag{4.3c}$$

where W denotes a sliding time window covering the detected footstep from all geophones. The computations are straightforward, and the optimization is performed in discrete time on a one-dimensional grid.

Variance Estimation and DOA refinement

To compute the variance and refine the DOA estimate, a second degree polynomial is fit locally around the maximum

$$V_N(\psi) \approx a + b\left(\psi - \hat{\psi}\right) + c\left(\psi - \hat{\psi}\right)^2, \tag{4.4}$$

where the parameters a, b and c are estimated using the LS method. The interval for which ψ to include around the maximum $\hat{\psi}$ is crucial to get a good estimate of the variance. Once this polynomial is estimated, it can be shown that a refined estimate of the angle $\hat{\hat{\psi}}$ and its variance can be computed as

$$\hat{\hat{\psi}} = \hat{\psi} - \frac{b}{2c} \tag{4.5a}$$

$$\mathrm{Var}(\hat{\hat{\psi}}) = \frac{\frac{b^2}{4c} - a}{Nc}, \tag{4.5b}$$

where N is the number of samples in the sliding window W.

Tracking

To track the movement of the elephant a constant position KF is used with the DOA angle as the state, $x_k = \psi_k$,

$$x_{k+1} = x_k + w_k, \qquad w_k \sim \mathcal{N}(0, Q_k) \tag{4.6a}$$

$$y_k = x_k + v_k, \qquad v_k \sim \mathcal{N}(0, R_k), \tag{4.6b}$$

where w_k is the process noise, y_k is the DOA estimate from (4.5a) and v_k is the noise of the DOA estimate [20]. A constant position model is used since the elephant moves slowly, and when no step is detected the elephant is most likely

standing still. The variance of v_k is R_k, which is estimated using (4.5b), and the variance of the process noise Q_k is a tuning parameter. Additionally, gating is used to remove false or unlikely DOA estimates

$$(y_k - \hat{y}_{k|k-1})^T S_k^{-1} (y_k - \hat{y}_{k|k-1}) < \gamma_G, \tag{4.7}$$

where γ_G is the gate threshold and S_k is the innovation covariance from the KF. The gate threshold is related to the probability to accept correct measurements $P_G = \int_0^{\gamma_G} \chi^2(\gamma; n_y)\, d\gamma$, where n_y is the degrees of freedom (here $n_y = 1$) [42].

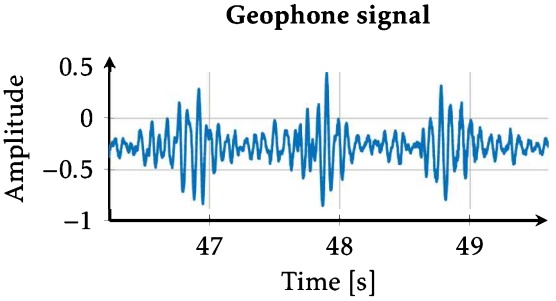
4.3 Results

In this section, the results of the signal analysis of the elephant footstep and ambient noise are presented to gain relevant information for the detection algorithm. This leads to the presentation of the detection, DOA estimate, and tracking results.

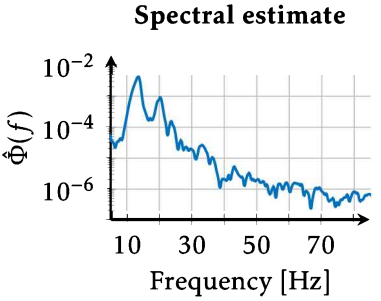
4.3.1 Elephant Footstep

The impulse of an elephant's footstep typically look like Figure 4.2a. It lasts for about 350 ms, which is a bit longer than stated in [31], as they claimed an elephant footstep to last for around 103–250 ms. By studying how the elephant walked, it was noticed that the elephant put his forefoot down around 100 ms after his hind foot. Thus, the signals originating from both the fore and hind foot have been captured within the 350 ms signal duration.

During the experiments, it was found that the frequency content of elephant footsteps depends on the composition of the ground. Early tests in the spring,



(a) Geophone signals in the time domain, where three elephant footsteps are present.



(b) A spectral estimate of the signal in Figure 4.2a using Blackman-Tukey's method with $\gamma = 140$.

Figure 4.2: *Properties of elephant footsteps in the time and frequency domain when the elephant is walking on dry sand.*

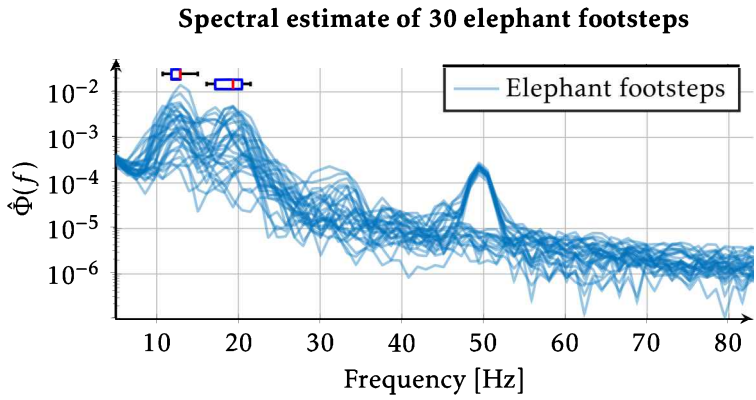


Figure 4.3: *Spectral estimate of 30 elephant footsteps, using Blackman-Tukey's method with $\gamma = 70$, plotted in blue lines with opacity. There are two main peaks of the elephant footstep where a box plot has been added to indicate the location of the peaks.*

with the ground still frozen, showed a higher frequency content of the signal compared to the vast majority of the experiments carried out later in the spring on dry sand. The experiments on the dry sand resulted in a main energy content around 8–23 Hz for the elephant footsteps, as seen in Figure 4.2b.

By looking at 30 different footsteps, it is observed in Figure 4.3 that the amplitude of the footstep varies, but the overall frequency response look similar with one peak at around 12 Hz and a second peak around 19 Hz. The most prominent peak varies between the footsteps, but the majority of the time the peak around 12 Hz has the highest amplitude. By looking at the box plot, it is seen that the first peak varies between 11–15 Hz with a median of 13 Hz, while the second peak varies between 16–21 Hz with a median of 19 Hz. Also, a lower peak around 50 Hz is present in half of the footsteps, this is due to a disturbance that appeared halfway into the experiments.

4.3.2 Background Noise

An important aspect of the detection method is to analyze the background noise of the geophones. In Figure 4.4, the frequency content of the background noise is shown. At first, during the data collection, the noise levels were quite low. But then, around 90 seconds in, a source of disturbance was turned on somewhere in the neighborhood, resulting in a clear peak around 50 Hz. Compared to the frequency content of the elephant footstep in Figure 4.3, there are no significant peaks around 11–21 Hz. This suggests that the main frequency content of the elephant footstep is around 8–23 Hz.

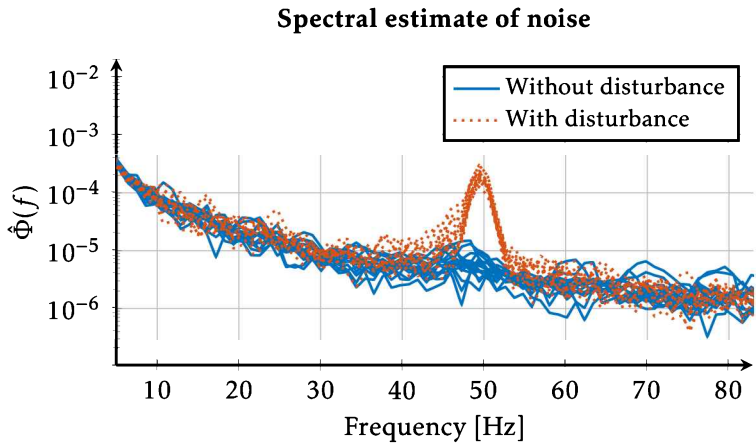


Figure 4.4: *Spectral estimate of 30 segments of the background noise using Blackman-Tukey's method with $\gamma = 70$. Each segment is the same length as the elephant footsteps in Figure 4.3. The blue solid lines are the frequency content before the disturbance appeared, and the orange dotted lines are the frequency content with the disturbance present.*

4.3.3 Detection, DOA Estimation, and Tracking

As the elephant footstep was found to last for around 350 ms, the segment length should be 700 ms. With a sample frequency of 474 Hz this corresponds to a segment length of roughly $N = 350$ samples. Since the main frequency of the elephant footstep was around 8–23 Hz, the cut-off frequencies of the band-pass filter in the detection algorithm were set to 4 and 30 Hz. Then, frequencies between 8–23 Hz were summed and the threshold T in the detection algorithm was tuned. Also, the number of geophones that were needed for detection was tuned, and 2 out of 3 geophones were found to give a satisfactory result. Thereafter, the detected signal was filtered with a band-pass filter with cut-off frequencies at 8 and 23 Hz before the delay-and-sum algorithm was applied. Lastly, the interval for ψ in the variance estimation was chosen to be $\hat{\psi} \pm 20°$. In Figure 4.5, the result using different thresholds is shown for the training data.

From the training data a good trade-off between false alarms and missed detections appears to be approximately $T = 97\,\%$. To verify that the parameters were not over-tuned to the specific data set, the tuning parameters were tested on a new validation data set. Additionally, tracking of the elephant was performed using the *Kalman filter*. In order to get a satisfying performance of the tracking, Q_k was set to 100 and the gating threshold to $\gamma_G = 10.83$, corresponding to a probability to accept correct measurements of 99.9 %. The results can be seen in Figure 4.6.

From the figure, it is seen that the elephant is tracked with acceptable accuracy. However, since the dotted line is estimated from a flying drone, it is hard to know the exact direction to the elephant, and it should not be considered as the ground truth.

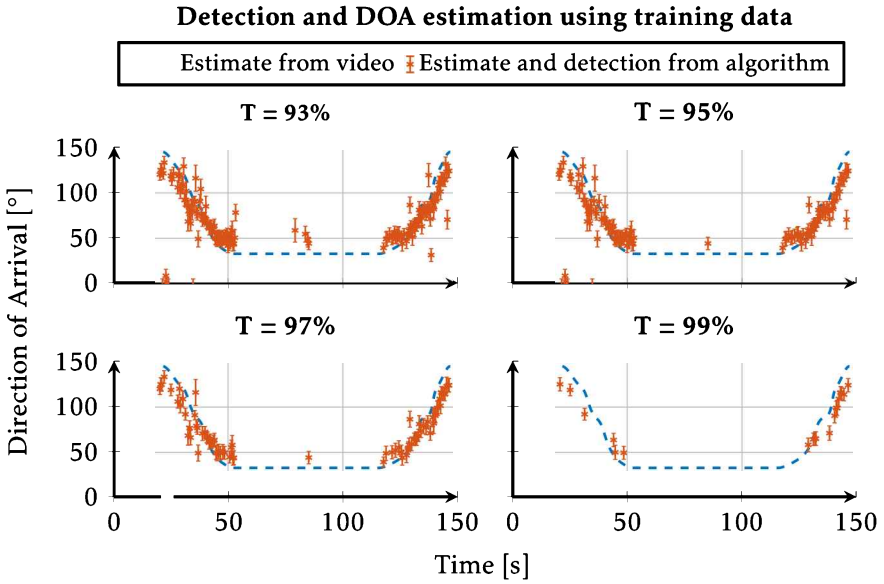


Figure 4.5: *The resulting detection and DOA estimation with different thresholds in the detection algorithm. The blue dotted line is the direction to the elephant estimated from a drone flying above the elephant. The measurements are presented as an orange star with a 95 %* confidence interval *(CI).*

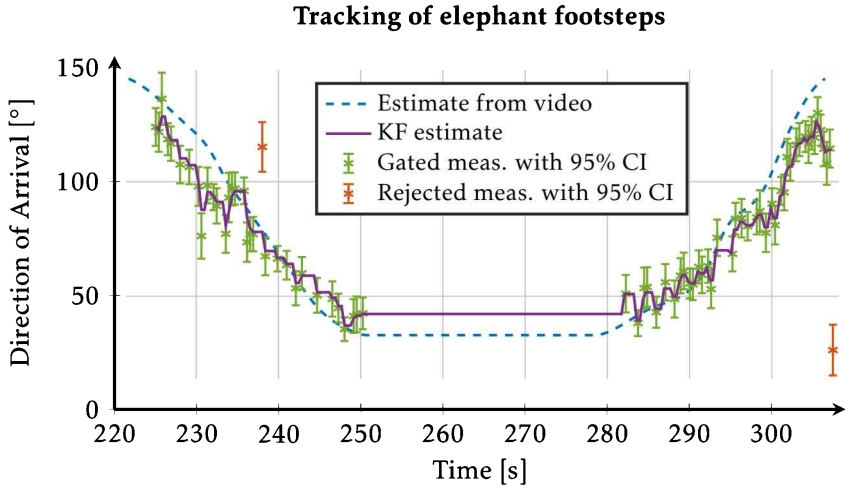


Figure 4.6: *The algorithm performance on the validation data. The blue dotted line is the direction to the elephant estimated from a drone flying above the elephant. The purple solid line is the KF estimate. The measurements are presented as a star with a 95 % CI, green (✳) indicates gated measurements and orange (✳) indicates rejected measurements from the gating.*

4.4 Summary

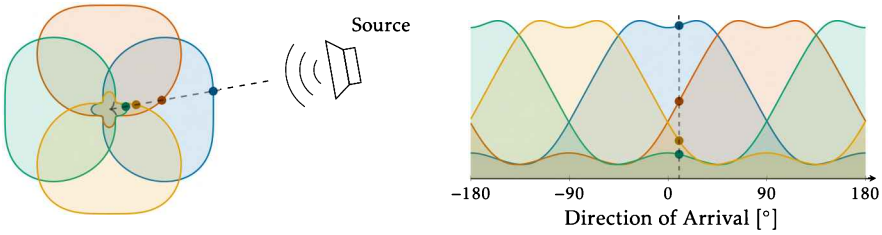
In this chapter, an algorithm for detection, localization, and tracking of elephants has been developed with the purpose to mitigate human-wildlife conflicts. The algorithm has been tested and validated at a nearby zoo using a prototype with three geophones. The results show that the suggested methods work to detect and track an elephant with promising accuracy within a range of 15 to 40 meters. However, in order to further validate the tracking performance a better estimate of the ground truth is necessary.

This chapter shows a proof of concept of a device capable of protecting the communities living close to national parks. The range of 40 meters would be sufficient if multiple arrays were deployed around the protected area, where the distance between the arrays is dependent on the range. In order to deploy this device in a real scenario, the algorithms must be adapted to cope with different weather conditions on the savanna, as well as the ambient noise.

5

Directional Sensitivity DOA

Recived signal strength (RSS) of a microphone array can also be used for DOA estimation. The method is based on the directional sensitivity of the microphone elements and requires a calibration step to estimate the directivity pattern of the microphones. The estimated directional sensitivity as a function of the DOA is then used to estimate the DOA of a sound source, as illustrated in Figure 5.1. The chapter is based on work that has been submitted in [55], as well as some earlier work published in [54].



(a) *Overview of the proposed method.*

(b) *The resulting directional sensitivity as a function of the DOA.*

Figure 5.1: *Visualization of the concept with 4 microphones facing away from each other. The colorful lines represent the directional sensitivity of each microphone. In the figure, the sound is arriving with a DOA of 10° with respect to the blue microphone, where the dashed line represents the direction of the sound. Note that each microphone receives a different sound level, marked with dark-colored dots in the figure. The data for the directional sensitivity is from a wideband noise signal.*

5.1 Signal Model

The signal model, based on [54], denotes the measured signal from microphone m at discrete time l with the scalar

$$y_m[l] = s_m[l] + e_m[l] \tag{5.1}$$
$$s_m[l] = s[l - \delta_m], \tag{5.2}$$

where $s_m[l]$ is the received signal, $s[l]$ is the emitted signal, δ_m is the sample-delay to microphone m, and $e_m[l]$ is the measurement noise, which is assumed to be uncorrelated with $s_m[l]$, and normally distributed, *i.e.*, $e_m[l] \sim \mathcal{N}(0, \sigma_m^2)$. The power of the measured signal from microphone m can then be calculated as

$$P_m = \frac{1}{L}\sum_{l=1}^{L} y_m^2[l], \tag{5.3}$$

where L is the number of samples in the signal of interest. By inserting (5.1) into (5.3), the signal power can thus be expressed in three terms

$$P_m = \underbrace{\frac{1}{L}\sum_{l=1}^{L} s_m^2[l]}_{P_m^s} + \underbrace{\frac{1}{L}\sum_{l=1}^{L} 2s_m[l]e_m[l]}_{P_m^{sw}} + \underbrace{\frac{1}{L}\sum_{l=1}^{L} e_m^2[l]}_{v_m}, \tag{5.4}$$

where P_m^s is the power of the received signal, P_m^{sw} is a cross-term from the signal and the noise, and v_m is the power of the measurement noise. Here, the number of samples L is assumed to be large, hence $P_m^{sw} \to 0$. Further, since $e_m(t)$ is normally distributed, v_m will be chi-squared distributed with L degrees of freedom, *i.e.*, $\frac{L}{\sigma_m^2} v_m \sim \chi_L^2$. However, since the degrees of freedom is assumed to be large, the chi-squared distribution can be approximated with a normal distribution

$$\frac{L}{\sigma_m^2} v_m \sim \chi_L^2 \overset{\text{Approx}}{\to} v_m \sim \mathcal{N}\left(\sigma_m^2, \frac{2\sigma_m^4}{L}\right) = \mathcal{N}\left(\sigma_m^2, \eta_m\right). \tag{5.5}$$

We assume that each microphone has a directional sensitivity in the power attenuation by design or from the construction of the array. Further, the absolute level of the received power, assumed to be the same at all sensors, is denoted α and is considered to be a nuisance parameter from a DOA estimation perspective. Thus, the power P_m measured by microphone m can be expressed as

$$P_m(\psi) = \alpha g_m h(\psi, \boldsymbol{\theta}_m) + v_m, \tag{5.6}$$

where g_m is the microphone gain, $h(\psi, \boldsymbol{\theta}_m)$ is the directional sensitivity of the microphone dependent on the DOA angle ψ and parameterized by the parameters $\boldsymbol{\theta}_m$, and v_m is the observation error as described in (5.5).

During data collection, it was evident that the mean and variance of the noise depend on the sound level of the recorded signal, *i.e.*, a higher α yields a larger

noise mean and variance. To account for this dependency, we introduce the signal model in decibel scale

$$\bar{P}_m(\psi) = \bar{\alpha} + \bar{g}_m + h\left(\psi, \bar{\boldsymbol{\theta}}_m\right) + \bar{v}_m, \tag{5.7}$$

where $\bar{P}_m(\psi) = 10\log_{10}(P_m(\psi))$, and the notation $\bar{x}$ is introduced to refer to the parameter x in decibel scale. We assume that the noise is distributed according to

$$\bar{v}_m \sim \mathcal{N}\left(\bar{\sigma}_m^2, \bar{\eta}_m\right). \tag{5.8}$$

5.2 Method

The approach is based on two steps, a training phase to estimate the directional sensitivity of the microphones, and an estimation phase to estimate the DOA of the signal. First, the method is described in the absolute scale using $P_m(\psi)$, followed by the decibel scale using $\bar{P}_m(\psi)$.

5.2.1 Training

In the training phase, the array is exposed to wideband noise from different directions ψ, and the measured signal y_m is observed in a controlled environment. Then, the power of the measured signal, $P_m(\psi)$, is computed according to (5.3).

Absolute scale

The parameters of the directional sensitivity $\boldsymbol{\theta}_m$ as well as the microphone gain g_m are thereafter estimated using YALMIP with the FMINCON solver [27]. The optimization problem is formulated as

$$\begin{aligned} \underset{\boldsymbol{x}}{\text{minimize}} \quad & V(\boldsymbol{x}) \\ \text{subject to} \quad & \alpha > 0 \\ & g_1 = 1 \\ & h(\psi_m, \boldsymbol{\theta}_m) = 1 \quad \forall\, m = 1, 2, \ldots, M, \end{aligned} \tag{5.9}$$

where ψ_m is the angle when microphone m is facing the sound source, M is the number of microphones and $V(\boldsymbol{x})$ is the loss function

$$V(\boldsymbol{x}) = \sum_{m=1}^{M} \frac{1}{\eta_m} \sum_{k=1}^{K} \left(P_m(\psi_k) - \left(\alpha g_m h(\psi_k, \boldsymbol{\theta}_m) + \sigma_m^2\right)\right)^2, \tag{5.10}$$

where $\boldsymbol{x}$ contains the optimization variables α, $\{g_1, \ldots, g_M\}$ and $\{\boldsymbol{\theta}_1, \ldots, \boldsymbol{\theta}_M\}$, and K is the number of observed directions. The term σ_m^2 appears since the noise does not have zero mean, and is estimated from measurements of the background noise.

Decibel scale

Similarly, the decibel scale model can be estimated using the power of the measured signal, $\bar{P}_m(\psi) = 10\log_{10}(P_m(\psi))$, and the corresponding optimization problem is formulated as

$$\begin{aligned} \underset{\bar{\boldsymbol{x}}}{\text{minimize}} \quad & V(\bar{\boldsymbol{x}}) \\ \text{subject to} \quad & \bar{g}_1 = 0 \\ & h\left(\psi_m, \bar{\boldsymbol{\theta}}_m\right) = 0 \quad \forall\, m = 1, 2, \ldots, M, \end{aligned} \tag{5.11}$$

with the loss function

$$V(\bar{\boldsymbol{x}}) = \sum_{m=1}^{M} \frac{1}{\bar{\eta}_m} \sum_{k=1}^{K} \left(\bar{P}_m(\psi_k) - \left(\bar{\alpha} + \bar{g}_m + h\left(\psi_k, \bar{\boldsymbol{\theta}}_m\right) + \bar{\sigma}_m^2 \right) \right)^2, \tag{5.12}$$

where $\bar{\boldsymbol{x}}$ contains the optimization variables $\bar{\alpha}$, $\{\bar{g}_1, \ldots, \bar{g}_M\}$ and $\{\bar{\boldsymbol{\theta}}_1, \ldots, \bar{\boldsymbol{\theta}}_M\}$.

5.2.2 Fourier Series Model

To model the directional sensitivity $h(\psi, \boldsymbol{\theta}_m)$ for microphone m, we adopt a *Fourier series* (FS) model

$$\begin{aligned} h(\psi, \boldsymbol{\theta}_m) &= \theta_0^m + \sum_{d=1}^{D} \theta_{d,c}^m \cos(d\psi) + \theta_{d,s}^m \sin(d\psi) \\ &= \boldsymbol{\Phi}(\psi)\boldsymbol{\theta}_m, \end{aligned} \tag{5.13}$$

where D is the order of the FS. The choice of D is a trade-off between model complexity and model fit. A higher order FS model can capture more complex directional sensitivity patterns, but also increases the risk of overfitting. However, a low order FS model can be too simple to capture the directional sensitivity and thus yield a poor model fit.

5.2.3 Estimation

In the estimation phase, the DOA is estimated using the directional sensitivity model and the measured power of the signal. The estimation method is applicable both in the absolute and decibel scale, and is based on the *separable least squares* (SLS) method [17, 22].

Absolute scale

From the training, the model for microphone m can be expressed as

$$\hat{P}_m(\psi) = \alpha \underbrace{\hat{g}_m \boldsymbol{\Phi}(\psi)\hat{\boldsymbol{\theta}}_m}_{\hat{h}_m(\psi)} + \sigma_m^2, \tag{5.14}$$

where $\hat{g}_m$ and $\hat{\boldsymbol{\theta}}_m$ are the estimated parameters from the training phase. This model can be seen as linear in α and nonlinear in ψ, and thus the SLS approach can be applied to estimate these parameters. The DOA estimator is then computed using SLS as

$$\hat{\psi} = \arg\max_{\psi} \frac{\left(\sum_{m=1}^{M} \hat{h}_m(\psi)(P_m - \sigma_m^2)\right)^2}{\sum_{m=1}^{M} \hat{h}_m^2(\psi)} \tag{5.15a}$$

$$\hat{\alpha} = \frac{\sum_{m=1}^{M} \hat{h}_m(\hat{\psi})(P_m - \sigma_m^2)}{\sum_{m=1}^{M} \hat{h}_m^2(\hat{\psi})}, \tag{5.15b}$$

where $\hat{h}_m(\psi) = \hat{g}_m \boldsymbol{\Phi}(\psi)\hat{\boldsymbol{\theta}}_m$ denotes the nonlinear part of the model [22].

Decibel scale

Similarly, the decibel model for microphone m can be expressed as

$$\hat{\bar{P}}_m(\psi) = \bar{\alpha} + \underbrace{\hat{\bar{g}}_m + \boldsymbol{\Phi}(\psi)\hat{\bar{\boldsymbol{\theta}}}_m}_{\hat{\bar{h}}_m(\psi)} + \bar{\sigma}_m^2, \tag{5.16}$$

where $\hat{\bar{g}}_m$ is the estimated microphone gain and $\hat{\bar{\boldsymbol{\theta}}}_m$ is the estimated parameters from the training phase. This model can be seen as linear in $\bar{\alpha}$ and nonlinear in ψ.

The DOA estimator is then computed using SLS as

$$\hat{\psi} = \arg\min_{\psi} \left(\bar{P} - \left(\hat{\bar{\alpha}}(\psi) + \hat{\bar{h}}_m(\psi) + \bar{\sigma}_m^2\right)\right)^2 \tag{5.17a}$$

$$\hat{\bar{\alpha}}(\psi) = \frac{1}{M} \sum_{m=1}^{M} \left(\bar{P} - \hat{\bar{h}}_m(\psi) - \bar{\sigma}_m^2\right), \tag{5.17b}$$

where $\hat{\bar{h}}_m(\psi) = \hat{\bar{g}}_m + \boldsymbol{\Phi}(\psi)\hat{\bar{\boldsymbol{\theta}}}_m$ contains the nonlinearity of the model [17].

5.2.4 Frequency Dependency

During data collection it was observed that the directional sensitivity highly varies with the frequency content of the signal. Therefore, a dependency of the frequency is introduced in the parameters of the directional sensitivity, *i.e.*, $\boldsymbol{\theta}_m(f)$, where f is the frequency content of the signal.

Training

In the training phase, when learning the directional sensitivity, we thus need to include the frequency dependency. One way to do this is by band-pass filtering a

wideband noise signal through F adjacent band-pass filters in order to span the whole spectrum

$$y_m[l, f_n] = h_{bp}(f_n) \star y_m[l], \quad \forall f_n = \{f_1, \ldots, f_F\} \tag{5.18a}$$

$$P_m(f_n) = \frac{1}{L} \sum_{l=1}^{L} y_m^2[l, f_n], \tag{5.18b}$$

where $h_{bp}(f_n)$ denotes a band-pass filter with frequency content f_n and $\star$ denotes the convolution operation. Thereafter, the optimization problems (5.9) and (5.11) can be solved for the signal power in each band-pass filtered signal, separately. Thus, for each band-pass frequency and microphone, a model for the directional sensitivity is learned.

Estimation

For the frequency dependency the estimation can be addressed in various ways, and one way is to create a model of the directional sensitivity by computing the frequency content of the received sound. Another way is to compare a band-pass filtered version of the received sound with the directional sensitivity for the corresponding band-pass filter, for all band-pass filters. Both these methods are visualized in Figure 5.2.

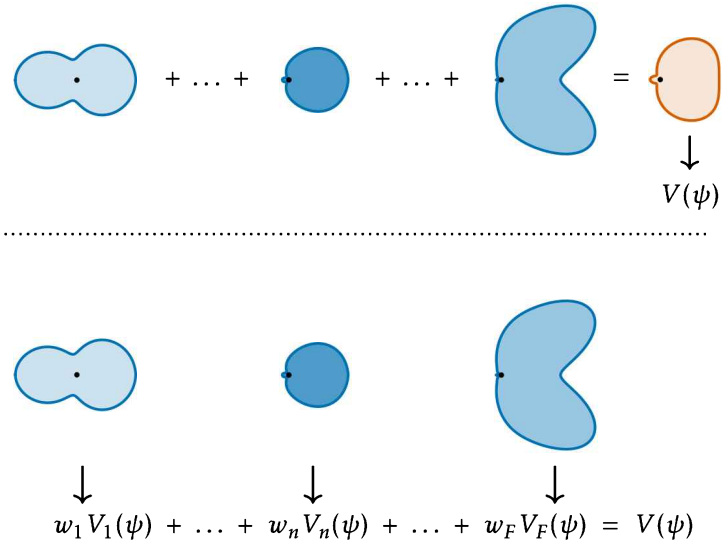


Figure 5.2: *Visualization of the different methods. In the upper part of the figure, we sum all the band-pass filtered models with weights w_n to create a model of the incoming sound based on its frequency content. In the lower part, we calculate the cost of each band-pass separately and sum these terms with their corresponding weight w_n. The opacity represents the weights w_n.*

Combine into one directional sensitivity First, we compute the frequency content of the incoming sound in the bins of the filter bank, denoted w_n

$$w_n = \sum_{m=1}^{M} \frac{1}{M} \left(\sum_{f \in f_n} \left(\sum_{l=1}^{L} \left| y_m[l] e^{-j2\pi f \frac{l}{L}} \right|^2 \right) \right). \tag{5.19}$$

To model the directional sensitivity of the incoming sound, we sum the models for the different frequencies with the weight of the frequency content, that is

$$\hat{P}_m(\psi) = \alpha \underbrace{\sum_{n=1}^{F} w_n \hat{g}_{m,n} \mathbf{\Phi}(\psi) \hat{\boldsymbol{\theta}}_m(f_n)}_{\hat{h}_m(\psi)} + \sigma_m^2. \tag{5.20}$$

This can be seen as a SLS problem, which is linear in α and nonlinear in ψ. Therefore, (5.15a) can be applied.

Each band-pass separately The second approach is to treat each band-pass filter separately, and then sum them in the cost function. The signal model for microphone m and band-pass filter n is expressed as

$$\hat{P}_{m,n}(\psi) = \alpha_n \underbrace{\hat{g}_{m,n} \mathbf{\Phi}(\psi) \hat{\boldsymbol{\theta}}_m(f_n)}_{\hat{h}_{m,n}(\psi)} + \sigma_{m,n}^2. \tag{5.21}$$

The DOA estimate is then calculated according to

$$\hat{\psi} = \arg\max_{\psi} \sum_{n=1}^{F} w_n V_n(\psi) \tag{5.22a}$$

$$V_n(\psi) = \frac{\left(\sum_{m=1}^{M} \hat{h}_{m,n}(\psi) \left(P_{m,n} - \sigma_{m,n}^2 \right) \right)^2}{\sum_{m=1}^{M} h_{m,n}^2(\psi)} \tag{5.22b}$$

$$\hat{\alpha}_n = \frac{\sum_{m=1}^{M} \hat{h}_{m,n}(\hat{\psi}) (P_{m,n} - \sigma_{m,n}^2)}{\sum_{m=1}^{M} \hat{h}_{m,n}^2(\hat{\psi})}. \tag{5.22c}$$

This method can also be performed in the decibel scale using the signal model

$$\hat{\bar{P}}_{m,n}(\psi) = \bar{\alpha}_n + \underbrace{\hat{\bar{g}}_{m,n} + \mathbf{\Phi}(\psi) \hat{\bar{\boldsymbol{\theta}}}_m(f_n)}_{\hat{\bar{h}}_{m,n}(\psi)} + \bar{\sigma}_{m,n}^2. \tag{5.23}$$

The DOA estimate is then calculated as

$$\hat{\psi} = \arg\min_{\psi} \sum_{n=1}^{F} w_n V_n(\psi) \tag{5.24a}$$

$$V_n(\psi) = \sum_{m=1}^{M} \left(\bar{P}_{m,n} - \left(\hat{\bar{\alpha}}_n(\psi) + \hat{\bar{h}}_{m,n}(\psi) + \bar{\sigma}^2_{m,n} \right) \right)^2 \tag{5.24b}$$

$$\hat{\bar{\alpha}}_n(\psi) = \frac{1}{M} \sum_{m=1}^{M} \left(\bar{P}_{m,n} - \hat{\bar{h}}_{m,n}(\psi) - \bar{\sigma}^2_{m,n} \right). \tag{5.24c}$$

5.3 Cramér-Rao Lower Bound

The CRLB for the DOA estimation can be derived for the two different scales, the absolute scale and the decibel scale.

5.3.1 Absolute scale

To calculate the CRLB, a *probability density function* (PDF) of the measurement is required. Since the noise v_m is assumed to be normally distributed as $v_m \sim \mathcal{N}(\sigma_m^2, \eta)$, the PDF of the measurement is denoted as

$$p(P|\psi, \boldsymbol{\theta}_m, \alpha, \eta) = \mathcal{N}(\alpha g_m h(\psi, \boldsymbol{\theta}_m) + \sigma_m^2, \eta), \tag{5.25}$$

where $\eta = {}^{2\sigma_m^4}/_{L}$ is the noise variance and assumed to be the same for all microphones.

This yields the *Fisher information matrix* (FIM)

$$\begin{aligned} \mathcal{I}_m(\psi, \alpha) &= -\mathrm{E}\left[\begin{pmatrix} \frac{\partial^2}{\partial \psi^2} & \frac{\partial^2}{\partial \psi \partial \alpha} \\ \frac{\partial^2}{\partial \alpha \partial \psi} & \frac{\partial^2}{\partial \alpha^2} \end{pmatrix} \log p(P|\psi, \boldsymbol{\theta}_m, \alpha, \eta) \right] \\ &= \frac{1}{\eta} \begin{bmatrix} \left(\alpha g_m h'_{\psi,m}\right)^2 & \alpha g_m^2 h_m h'_{\psi,m} \\ \alpha g_m^2 h_m h'_{\psi,m} & g_m^2 h_m^2 \end{bmatrix}, \end{aligned} \tag{5.26}$$

where h_m denotes $h(\psi, \boldsymbol{\theta}_m)$ and $h'_{\psi,m} = \frac{\partial h(\psi, \boldsymbol{\theta}_m)}{\partial \psi}$.

Assuming that the estimate in (5.15a) is unbiased and that the PDF $p(\boldsymbol{P}|\psi)$ satisfies the regularity conditions, the CRLB can be used to evaluate the performance of the estimation method [22]. Thus, the CRLB of ψ is computed as

$$\begin{aligned} \mathrm{var}(\hat{\psi}) \geq \mathrm{CRLB}(\psi) &= \left[\left(\sum_{m=1}^{M} \mathcal{I}_m(\psi, \alpha) \right)^{-1} \right]_{1,1} \\ &= \frac{\eta}{\alpha^2} \frac{\|\boldsymbol{h}\|^2}{\left(\|\boldsymbol{h}\|^2 \, \|\boldsymbol{h}'_{\psi}\|^2 - \langle \boldsymbol{h}, \boldsymbol{h}'_{\psi} \rangle^2 \right)}, \end{aligned} \tag{5.27}$$

where $\|\cdot\|$ denotes the Euclidean norm and $\langle \cdot, \cdot \rangle$ denotes the inner product, and

$$\boldsymbol{h} = \begin{bmatrix} g_1 h_1 & g_2 h_2 & \dots & g_M h_M \end{bmatrix}^T \tag{5.28}$$

$$\boldsymbol{h}'_\psi = \begin{bmatrix} g_1 h'_{\psi,1} & g_2 h'_{\psi,2} & \dots & g_M h'_{\psi,M} \end{bmatrix}^T. \tag{5.29}$$

Note that the denominator is strictly positive as long as $\boldsymbol{h}$ and $\boldsymbol{h}'_\psi$ are not parallel, according to Cauchy–Schwarz inequality. The term α^2/η can be interpreted as the SNR of the signal.

5.3.2 Decibel scale

Using the decibel scale, the PDF of the measurement is described by

$$p(\bar{P}|\psi, \bar{\boldsymbol{\theta}}_m, \bar{\alpha}, \bar{\eta}) = \mathcal{N}(\bar{\alpha} + \bar{g}_m + h(\psi, \bar{\boldsymbol{\theta}}_m) + \bar{\sigma}_m^2, \bar{\eta}), \tag{5.30}$$

where $\bar{\eta} = 2\bar{\sigma}_m^4/L$ is the noise variance assumed to be the same for all microphones.

This yields the FIM

$$\begin{aligned} \mathcal{I}_m(\psi, \bar{\alpha}) &= -\mathrm{E}\left[\begin{pmatrix} \frac{\partial^2}{\partial \psi^2} & \frac{\partial^2}{\partial \psi \partial \bar{\alpha}} \\ \frac{\partial^2}{\partial \bar{\alpha} \partial \psi} & \frac{\partial^2}{\partial \bar{\alpha}^2} \end{pmatrix} \log p(\bar{P}|\psi, \bar{\boldsymbol{\theta}}_m, \bar{\alpha}, \bar{\eta}) \right] \\ &= \frac{1}{\eta} \begin{bmatrix} \left(\bar{h}'_{\psi,m}\right)^2 & \bar{h}'_{\psi,m} \\ \bar{h}'_{\psi,m} & 1 \end{bmatrix}, \end{aligned} \tag{5.31}$$

where $\bar{h}'_{\psi,m} = \frac{\partial h(\psi, \bar{\boldsymbol{\theta}}_m)}{\partial \psi}$.

Thus, the CRLB of ψ is computed as

$$\begin{aligned} \mathrm{var}(\hat{\psi}) \geq \overline{\mathrm{CRLB}}(\psi) &= \left[\left(\sum_{m=1}^{M} \mathcal{I}_m(\psi, \bar{\alpha}) \right)^{-1} \right]_{1,1} \\ &= \bar{\eta} \frac{1}{\left(\sum_{m=1}^{M} \left(\bar{h}'_{\psi,m}\right)^2 - \frac{1}{M} \left(\sum_{m=1}^{M} \bar{h}'_{\psi,m} \right)^2 \right)}. \end{aligned} \tag{5.32}$$

Note that the CRLB does not depend on the SNR of the signal in this case, only the noise variance and the derivatives of $h(\psi, \bar{\boldsymbol{\theta}}_m)$. This is due to the fact that we have different assumptions on the noise distribution in the two cases.

6

Application — Directional Sensitivity DOA

In this chapter, the acoustic data presented in Chapter 2 is used to evaluate the proposed method for DOA estimation. First, we show that it is possible to estimate the DOA of a wideband signal using the proposed method without frequency dependency. Thereafter, we show that basically any signal can be used to estimate the DOA of a source when we introduce the frequency dependent model. For the frequency dependent model, the method is evaluated using both anechoic and outdoor data, and the results are compared to state-of-the-art methods.

6.1 Wideband Noise Model

The results with a wideband noise model, *i.e.*, no frequency dependency, are based on [54]. In this section, the method is evaluated solely using anechoic data and is restricted to broadband signals, *i.e.*, wideband noise at various levels and the drone sound from Section 2.2.2. Additionally, only the absolute model is assessed, with no comparisons to state-of-the-art methods. This serves as a proof of concept for the method before introducing frequency dependency in the subsequent section.

6.1.1 Training

For the training, the wideband noise signal is measured from 24 angles uniformly distributed around the array. Then, the signal power of the measured signals is calculated according to (5.3). Thereafter, the noise variance σ_m^2 is calculated from measurements of the background noise. Finally, the optimization problem (5.9) is solved to estimate the model parameters.

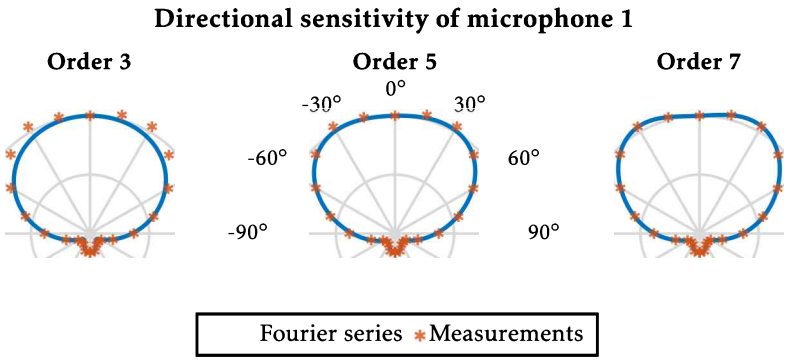


Figure 6.1: *The measured power of the wideband noise signal normalized with the power at $\psi = 0°$, and the corresponding FS as a function of the DOA for microphone 1, using the absolute scale.*

6.1.2 Fourier Series Fit

In the optimization problem, different orders of the FS are examined to find the best one. The best model order is determined using the *Bayesian information criterion* (BIC), which indicates that a model order of 7 is a good trade-off between model fit and model complexity. The measured power of the wideband noise signal normalized with the power at $\psi = 0°$ as well as the corresponding FS of orders 3, 5 and 7 are illustrated in Figure 6.1. It is seen that for orders 3 and 5, the FS does not fully capture the true directional sensitivity, especially around $\pm 30°$. However, for order 7 the FS fits the measurements well for all angles.

6.1.3 DOA Estimation

For the DOA estimation, only the wideband noise signal and the drone sound are used. The validation data is collected at 24 angles uniformly distributed around the microphone array, but at different angles than for the training data. Then, the power of the signals is calculated and compared to the estimated FS model using SLS. A histogram of the estimation error with a FS of order 7 is presented in Figure 6.2.

Here, it is seen that all signals are estimated with good precision. It is also noted that the SNR does not seem to affect the result in any significant way. This suggests that the main source of estimation error is not the measurement noise, but it could be due to a model error. Also, the drone performs worse than the wideband noise signals. The reason for this is due to the difference in the frequency content of the signal.

6.1.4 Cramér-Rao Lower Bound

By using the calculated noise variance and the resulting α from the training step, the term α^2/η becomes 130 dB, and the resulting CRLB is shown in Figure 6.3.

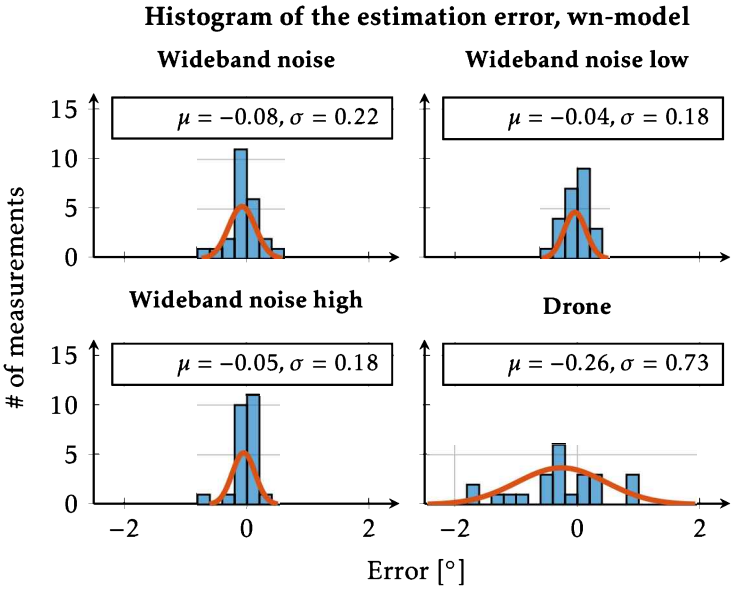


Figure 6.2: *Histogram of the estimation error for the validation data with a wideband noise model and FS expansion of order 7. The estimation is performed in the absolute scale using SLS. The red line represents a normal distribution fit to the estimation error.*

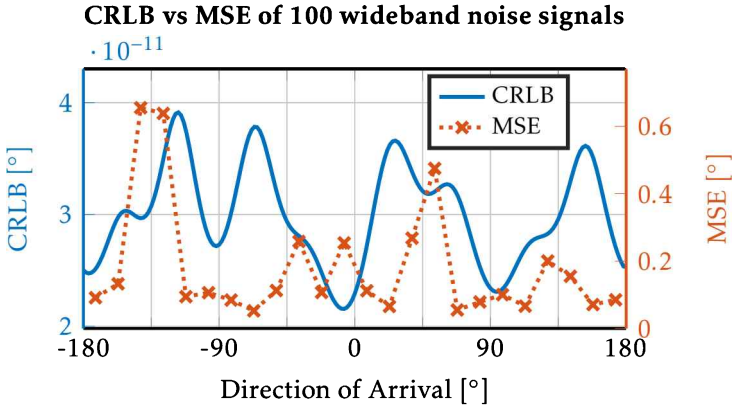


Figure 6.3: *The calculated CRLB for the wideband noise signal as well as the* mean square error *(MSE) of the wideband noise signal divided into 100 subsignals using a model order of 7.*

By looking at the MSE of the wideband noise signal in Figure 6.3, it can be seen that the squared estimation error does not follow the CRLB as expected. This suggests that the source of the estimation error is mainly caused by model error.

6.2 Frequency Dependent Model

For the results with frequency dependency, the method is evaluated using both anechoic and outdoor data, and the results are compared to state-of-the-art methods. This section is based on the work presented in [55].

6.2.1 Training

For the training, the same signals as in Section 6.1.1 are used. In this case, the signals are band-pass filtered into 40 different frequency bands with a bandwidth of 250 Hz.

Thereafter, the noise variances, $\sigma^2_{m,n}$ and $\bar{\sigma}^2_{m,n}$, are computed from measurements of the band-pass filtered background noise. Finally, the optimization problems are solved for each frequency band to estimate the model parameters.

6.2.2 Fourier Series Fit

As mentioned, the choice of model order is a trade-off between model fit and model complexity. Higher model orders also leads to overfitting, which is undesirable. By experimental evaluation, a model order of 7 is a good trade-off for both the absolute and the decibel model. The measured power of the band-pass filtered wideband noise signal as well as the corresponding FS of order 7 in decibel scale is illustrated in Figure 6.4. It is worth noting that below 1000 Hz the directional sensitivity is uniform for all angles. Hence, in this interval there is hardly any information to estimate the DOA from.

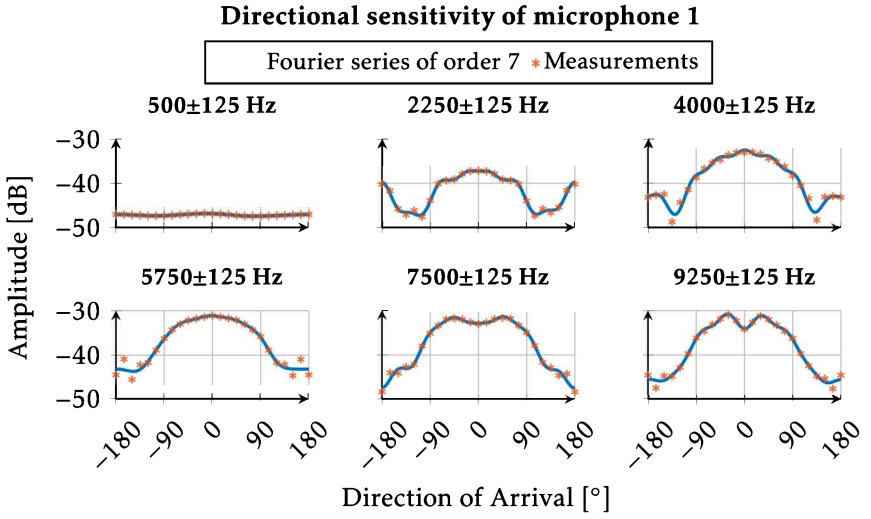


Figure 6.4: The measured power of the band-pass filtered wideband noise signal, and the corresponding FS of order 7 as a function of the DOA for microphone 1, both in the decibel scale.

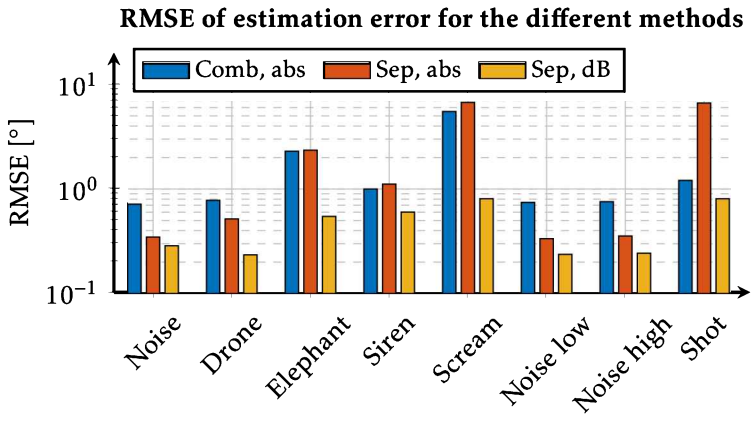


Figure 6.5: *Bar graph of the performance for the different methods for all signals except the sinusoids, all models have an order of 7. Comb refers to the combined method in Section 5.2.4, and Sep refers to the separate method in Section 5.2.4. The y-axis is in logarithmic scale.*

6.2.3 DOA Estimation

For the frequency dependent model, all signals are used for DOA estimation, except the background noise. We also include the data from the outdoor experiments.

In the training, 40 band-pass filters with a bandwidth of 250 Hz were used in order to get the results. The performance for the different methods using the natural signals are presented in Figure 6.5.

Here, it is seen that all signals are estimated with good precision. It is clearly noted that the separate method in decibel scale performed best for all signals, where we sum the cost function of the band-passed signal. With this method, all signals have a *root mean square error* (RMSE) value below 1°. This method will thus be referred to as the proposed method in the following sections.

We also compare the performance from the proposed method to state-of-the-art methods, where the results for the anechoic data are presented in Figure 6.6, and the data from the outdoor experiments is presented in Figure 6.7.

For the narrowband methods, *i.e.*, MUSIC and MVDR, the same band-pass filtering steps as in our proposed method has been performed

$$y_m^n[l] = h_{bp}(f_n) \star y_m[l] \tag{6.1}$$

$$P_{\text{MUSIC}}^n(\psi) = \text{MUSIC}(y_{1:M}^n) \tag{6.2}$$

$$P_{\text{MUSIC}}(\psi) = \sum_n w_n P_{\text{MUSIC}}^n(\psi), \tag{6.3}$$

where $h_{bp}(f_n)$ is a narrowband band-pass filter around frequency f_n, and w_n is the same frequency weights as in our method.

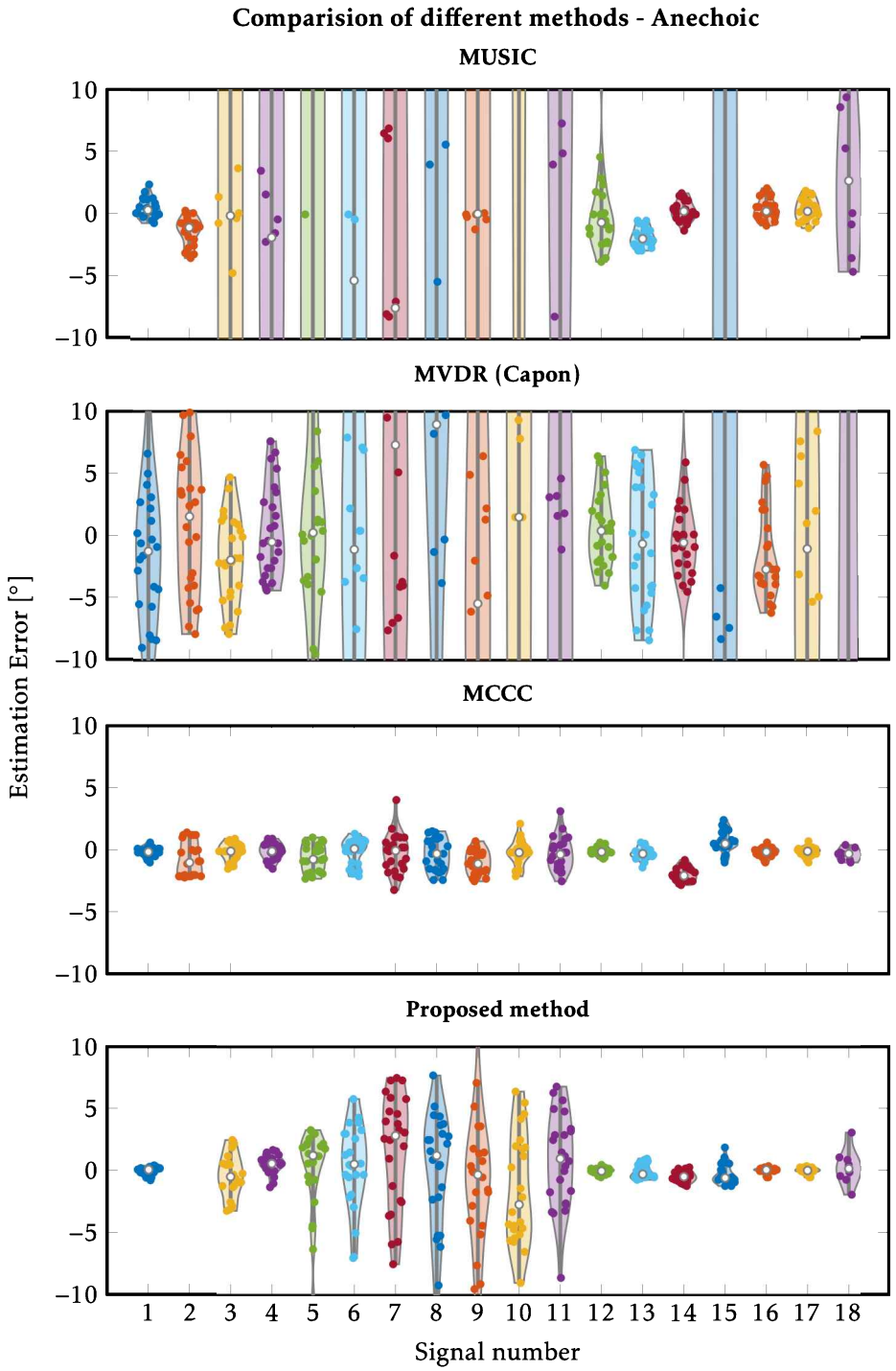


Figure 6.6: *Violin plot of the estimation error for the different signals using various methods for the data collected in the anechoic chamber. The signal number corresponds to the signals described in Section 2.2.2, where 1, 12–18 are natural signals.*

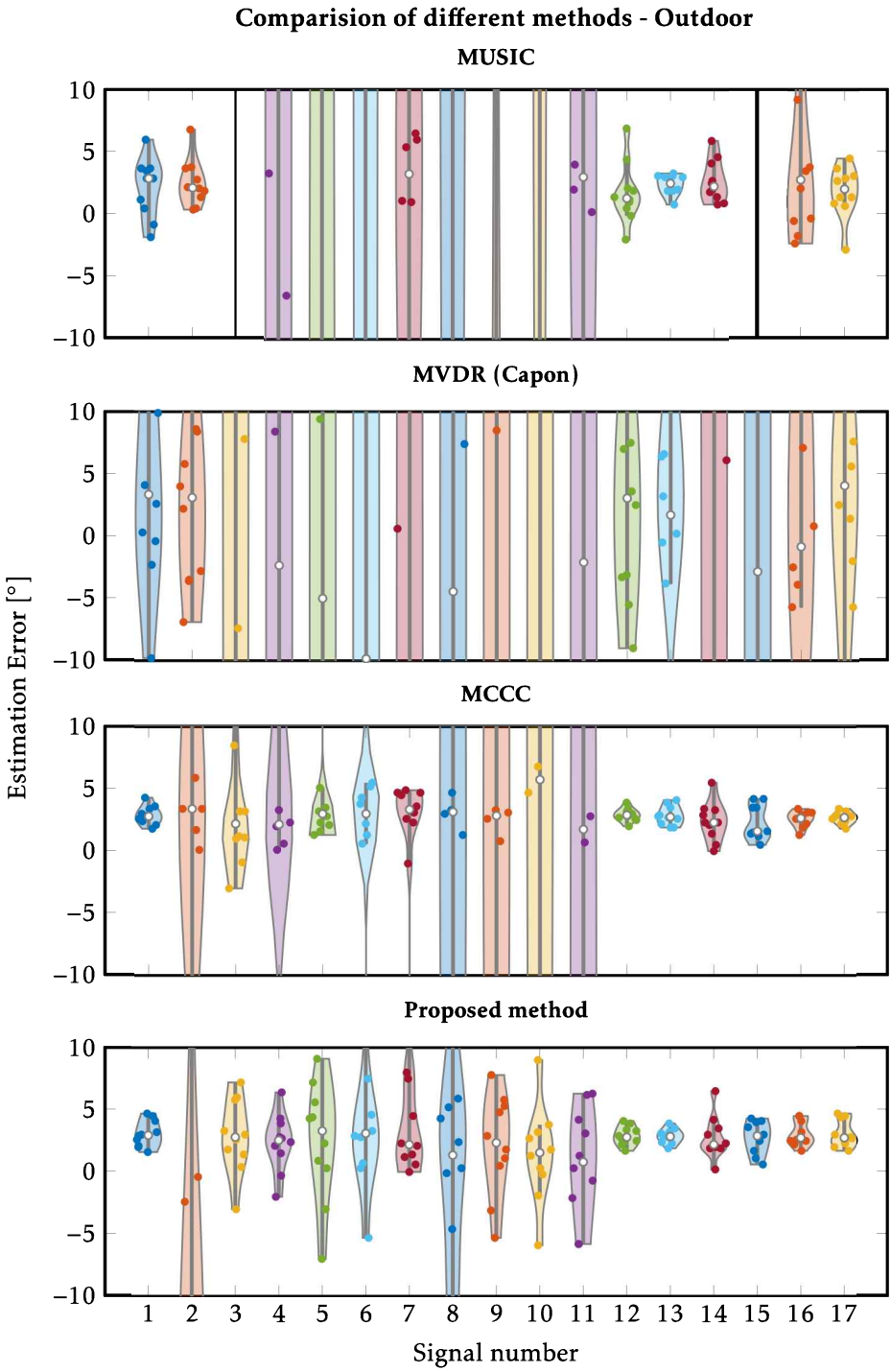


Figure 6.7: *Violin plot of the estimation error for the different signals using various methods for the data collected in the outdoor settings. The signal number corresponds to the signals described in Section 2.2.2, where 1, 12–17 are natural signals.*

For the anechoic data, the proposed method performs similar to the MCCC for the natural signals, *i.e.*, signals 1 and 12–18. While for the pure sinusoids, MCCC performs better. Compared to MUSIC and MVDR beamformer, the proposed method outperforms these methods for all signals, except for the low frequency sinusoid. It is clearly noted that the narrowband methods perform worse for higher frequencies, especially with frequencies that do not fulfil the spatial Shannon sampling theorem, *i.e.*, $d \leq \frac{\lambda}{2} \rightarrow f \lesssim 2100\,\text{Hz}$ [45].

In the results for the outdoor data, it is seen that the proposed method performs similar to the MCCC for most of the signals. Our method outperforms MCCC for some sinusoids, especially the high frequency sinusoids. For the outdoor data, both MUSIC and MVDR beamformer perform worse than the proposed method for all signals, except for the low frequency sinusoid. The resulting RMSE values for the different methods are presented in Table 6.1.

The reason why the estimation error is higher for the pure sinusoids is probably caused by the choice of band-pass filters. With a bandwidth of 250 Hz, the band-pass filters are too wide for the pure sinusoids, and thus the directional sensitivity is not well estimated. However, for the natural signals, the band-pass filters are well suited, and the directional sensitivity is well estimated. It is worth noting that all natural signals have a majority of their frequency content above 1000 Hz, which includes more information in the directional sensitivity.

It is also noted that for the outdoor data the estimation error of all methods has a bias of around 2.5°. This is most likely due to the fact that the setup is not perfectly calibrated, which is hard to achieve in an outdoor environment. This bias could be removed by retraining the model with additional outdoor data, but this is not the focus of this thesis.

The SNR for the anechoic data ranged between 22.2 dB and 68.0 dB, where the gunshot had the highest SNR, and otherwise the highest SNR was 59.7 dB. For the outdoor data, the SNR ranged between -14.1 dB and 32.8 dB, and thus the outdoor data had approximately 30 dB lower SNR than the anechoic data.

Table 6.1: *Mean RMSE values in degrees for the different methods across the signals, both for the Anechoic and Outdoor data. The signal number corresponds to the signals described in Section 2.2.2, where 1, 12–18 are natural signals.*

	Signal No.	Proposed	MUSIC	MVDR	MCCC
Anechoic	1-17	11.62	60.04	27.07	**0.99**
	1, 12-17	**0.45**	15.33	22.51	0.79
	18	1.42	9.68	92.43	**0.58**
Outdoor	1-17	**5.14**	74.25	76.08	42.03
	1, 12-17	3.10	29.15	59.23	**2.78**

6.2.4 Cramér-Rao Lower Bound

The CRLB is evaluated at each filter-bank separately and is then compared to the MSE of the collected sinusoids from the outdoor data. The results are presented in Figure 6.9. To calculate the MSE of the sinusoids, the collected signals are split into 50 subsignals, and the MSE is calculated based on these estimation errors. We can also calculate the CRLB for the wideband noise signal as

$$\text{CRLB}_{wn}(\psi) = \left[\left(\sum_{n=1}^{N} \bar{w}_n \mathcal{I}_n(\psi, \alpha_n)\right)^{-1}\right]_{1,1}, \tag{6.4}$$

where $\bar{w}_n$ is a normalized version of the weights w_n, such that $\sum_n \bar{w}_n = 1$. The resulting CRLB along with the MSE from the outdoor data are presented in Figure 6.8.

From the comparison between the CRLB and the MSE it can be noted that the MSE does not seem to follow the CRLB. Also, the magnitude of the MSE is roughly 10^3 times higher than the CRLB. The reason for this is most likely model errors at the validation angles, which are not included in the training data. Also, the bias of the estimation error for the outdoor data yields a higher MSE, as the CRLB is calculated for unbiased estimators. Further, the magnitude of the CRLB is mostly below the measurement accuracy of the angles, which is roughly 0.1°.

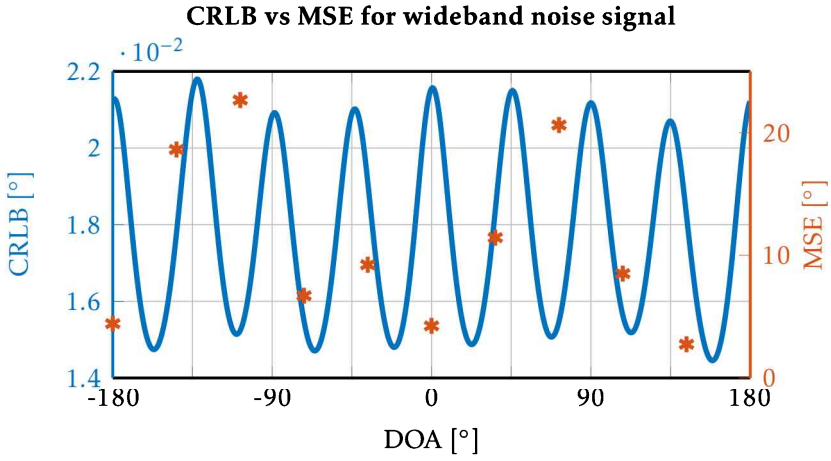


Figure 6.8: *CRLB compared to the MSE of the wideband noise signal for the outdoor data. The solid blue line represents the CRLB and the red stars the MSE.*

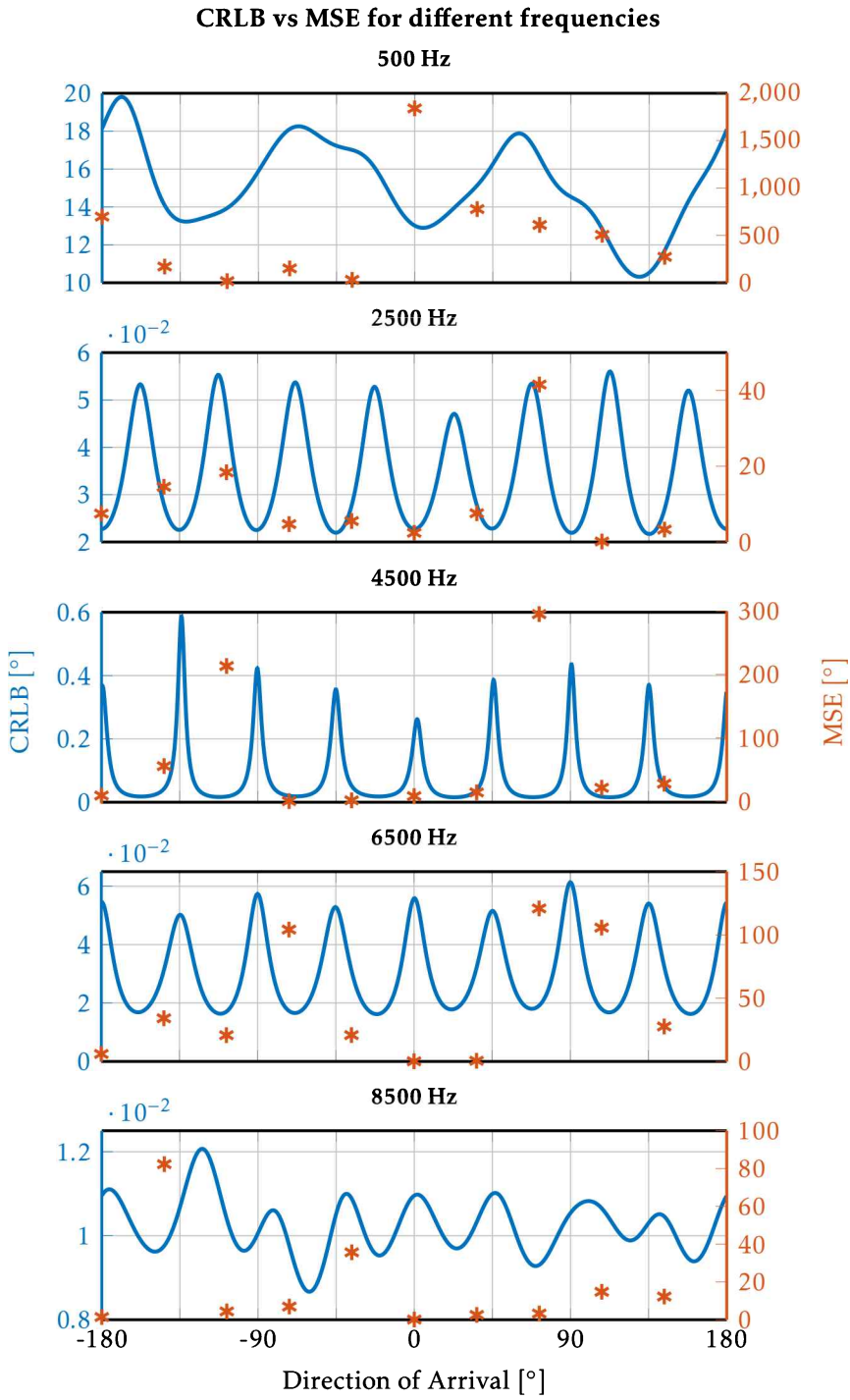


Figure 6.9: *CRLB compared to the MSE of the sinusoidal signals for the outdoor data. The solid blue line represents the CRLB and the red stars the MSE.*

6.3 Summary

In this chapter, we have presented the results of the proposed method for DOA estimation based on the directional sensitivity of microphones. The method is based on the directional sensitivity of the microphone elements and requires a calibration step to estimate the directivity pattern of the microphones. To show its potential, a sensor array with eight microphones has been used to estimate the DOA of 18 different types of signals.

First, the directional sensitivity of the microphones was estimated using a wideband noise signal. It was shown that the directional sensitivity highly varies depending on the frequency content of the signal, and we exploited this to estimate the DOA for different types of sounds.

In an anechoic setting, the method performs well for the natural signals evaluated, all with the majority of their frequency content above 1000 Hz. Less favorable results are obtained for signals with lower frequency content, as the directivity pattern becomes uniform for low frequencies. Also, the outdoor data shows that the model of the directional sensitivity is fairly robust over time and temperature changes, as the performance is great even after a year.

In the experimental setup used, the size of the array is still large enough to get reasonable estimates using state-of-the-art methods. However, these methods will degrade in performance as the array size decreases. The proposed method avoids this degradation, enabling virtually infinitesimal arrays with great DOA resolution.

7

Conclusions & Future Work

In this thesis, we have investigated the problem of DOA estimation using microphone arrays and geophones, focusing on applications such as detection and localization of elephant footsteps. The elephant footsteps are detected by exploiting the unique frequency content of the footsteps, which varies depending on the ground composition. The classical techniques for DOA estimation were reviewed and successfully applied to geophone data, demonstrating that a traditional delay-and-sum beamformer can accurately estimate the DOA of elephant footsteps at distances up to 40 meters. By detecting when elephants are approaching and from which direction, park rangers can proactively take measures to avoid conflicts, contributing to both human safety and elephant conservation.

The primary contribution of this work is the development of a DOA estimation algorithm based on the RSS of a microphone array. Unlike traditional methods, this algorithm leverages the directional sensitivity of the microphone elements, which is modeled using a FS model. This approach enables the use of the CRLB to evaluate the theoretical performance limits of the estimator.

Our derivation of the CRLB for DOA estimation indicates that the lower bound of the estimation error is significantly below the measurement accuracy observed in our experiments. We found that the estimator's performance is frequency-dependent; lower frequencies result in more uniform directional sensitivity, leading to a higher CRLB and estimation error. The method is particularly sensitive to narrowband signals, and narrowing the filterbank could potentially reduce estimation errors, although at the cost of increased computational complexity.

The robustness of the proposed method was validated through experiments with both anechoic and outdoor data. The method performs better or similar to state-of-the-art techniques for a variety of sound sources, particularly in natural environments. Notably, the method's consistent performance across different environments underscores its practical applicability.

A significant advantage of the developed algorithm is its independence from TDOA measurements between microphones, allowing for the design of smaller and more portable devices. These smaller devices could be used in a variety of new applications and environments where traditional methods might be impractical. However, the reliance on the calibration of the microphone's directional sensitivity presents a limitation, as it necessitates an initial anechoic chamber calibration. Nonetheless, this calibration has been shown to be robust, with only a single calibration required to produce a model applicable across various environments.

Future Work

The proposed method for DOA estimation based on directional sensitivity has shown promising results, but there are several avenues for further investigation to enhance its performance and applicability.

One important area for future research is the effect of reverberations on the algorithm. Reverberations, which are typically seen as noise in DOA estimation, can also provide valuable information about the environment, potentially aiding in mapping or improving source localization. Investigating how the proposed method can be adapted or extended to exploit reverberations could lead to significant improvements in its robustness and accuracy in complex acoustic environments.

Another critical area is the model error associated with the directional sensitivity of the microphone elements. Understanding and quantifying this model error is crucial for determining the limitations of the method and identifying opportunities for refinement. By improving the accuracy of the directional sensitivity model, the overall performance of the DOA estimation can be enhanced, particularly in diverse environmental conditions.

Moreover, it is important to evaluate the method using more complex signals, such as speech, to assess its effectiveness in real-world scenarios. Speech signals pose unique challenges due to their wide frequency range and variability, making them an excellent test case for the robustness of the DOA estimation algorithm in practical applications.

In addition, the application using geophones to detect elephant footsteps has demonstrated potential. Future research should explore its applicability to different individuals of elephants, considering variations in size, weight, and walking patterns. Investigating the impact of a herd of elephants on the detection and DOA estimates is also essential, as this scenario presents a more complex and dynamic environment. Applying a multi-target tracking framework to simultaneously monitor multiple elephants could significantly enhance the method's utility for wildlife management.

Another interesting area for future study is the effect of ground vehicles, as they generate noise at frequencies similar to elephant footsteps. Understanding how nearby vehicles influence the detection and DOA estimates is crucial for developing a robust system that can operate effectively in environments with hu-

man activity. Additionally, exploring the disturbance caused by other large mammals, such as humans, giraffes, and rhinos, would provide insights into the accuracy of the method and its potential for animal classification.

Lastly, since the current method performs well within a 40-meter range, it would be intriguing to explore its range limitations. Understanding the maximum effective distance of the DOA estimation method could lead to enhancements in its design, potentially extending its applicability in larger or more open environments.

Bibliography

[1] *Adafruit HUZZAH32 - ESP32 Feather.* Adafruit Industries, June 2024. URL https://cdn-learn.adafruit.com/downloads/pdf/adafruit-huzzah32-esp32-feather.pdf. Accessed: 2024-09-06.

[2] Jürgen Altmann, Sergey Linev, and Axel Weiß. Acoustic–seismic detection and classification of military vehicles—developing tools for disarmament and peace-keeping. *Applied Acoustics*, 63(10):1085–1107, October 2002. ISSN 0003-682X. doi: 10.1016/S0003-682X(02)00021-X.

[3] Sergey Alyamkin and ES Nezhevenko. Comparative analysis of the efficiency of the Kalman filter and particle filter in solving the problem of object tracking in a seismic security system. *Optoelectronics, Instrumentation and Data Processing*, 50(1):54–60, January 2014. ISSN 1934-7944. doi: 10.3103/S8756699014010075.

[4] Charles J. Ammon, Aaron A. Velasco, Thorne Lay, and Terry C. Wallace. Chapter 1 - An overview of global seismology. In Charles J. Ammon, Aaron A. Velasco, Thorne Lay, and Terry C. Wallace, editors, *Foundations of Modern Global Seismology (Second Edition)*, pages 3–37. Academic Press, 2021. ISBN 978-0-12-815679-7. doi: 10.1016/B978-0-12-815679-7.00008-2.

[5] Petre Anghelescu, Gabriel Vasile Iana, and Ionut Tramandan. Human footstep detection using seismic sensors. In *2015 7th International Conference on Electronics, Computers and Artificial Intelligence (ECAI)*, pages AE–1, Bucharest, Romania, June 2015. IEEE. doi: 10.1109/ECAI.2015.7301179.

[6] *U-PHORIA UMC1820.* Behringer. URL https://www.behringer.com/product.html?modelCode=P0B2J. Accessed: 2024-09-06.

[7] Jacob Benesty, Jingdong Chen, and Yiteng Huang. Time-delay estimation via linear interpolation and cross correlation. *IEEE Transactions on Speech and Audio Processing*, 12(5):509–519, September 2004. ISSN 1063-6676. doi: 10.1109/TSA.2004.833008.

[8] Jacob Benesty, Jingdong Chen, and Yiteng Huang. *Microphone Array Signal Processing*. Springer-Verlag, Berlin, Heidelberg, 2008. ISBN 978-3-540-78611-5.

[9] *AKG CBL99 Hemispherical boundary layer microphone*. Canford. URL https://www.canford.co.uk/Products/92-253_AKG-CBL99-MICROPHONE-Boundary-layer-hemispherical-9-52V-phantom-power-black. Accessed: 2024-09-06.

[10] J. Capon. High-resolution frequency-wavenumber spectrum analysis. *Proceedings of the IEEE*, 57(8):1408–1418, 1969. ISSN 0018-9219. doi: 10.1109/PROC.1969.7278.

[11] Gerardo Ceballos, Paul R. Ehrlich, Anthony D. Barnosky, Andrés García, Robert M. Pringle, and Todd M. Palmer. Accelerated modern human–induced species losses: Entering the sixth mass extinction. *Science Advances*, 1(5), June 2015. ISSN 2375-2548. doi: 10.1126/sciadv.1400253.

[12] Sanjana Ganesh. Human-elephant conflict kills 1,713 people, 373 pachyderms in 3 years, 2019. URL https://www.thehindu.com/news/national/human-elephant-conflict-kills-1713-people-373-pachyderms-in-3-years/article26225515.ece. Accessed: 2024-09-06.

[13] *Genelec 1029A Data Sheet*. Genelec, 2003. URL https://assets.ctfassets.net/4zjnzn055a4v/7547Y1zhZuQeaO68yMCYGe/98dd2e2995803eae08cddc2709797e38/1029A_datasheet.pdf. Accessed: 2024-09-06.

[14] *Genelec 8030C Manual*. Genelec, 2017. URL https://assets.ctfassets.net/4zjnzn055a4v/1my4dzvqUoe4oKYwS6KsUu/a8f964c3ce07eb802b88b8d861799a07/8030c_opman_gb_fi.pdf. Accessed: 2024-09-06.

[15] Ripul Ghosh, Aparna Akula, Satish Kumar, and HK Sardana. Time–frequency analysis based robust vehicle detection using seismic sensor. *Journal of Sound and Vibration*, 346:424–434, June 2015. ISSN 0022-460X. doi: 10.1016/j.jsv.2015.02.011.

[16] Daniel Goderik, Albin Westlund, Gustav Zetterqvist, Fredrik Gustafsson, and Gustaf Hendeby. Seismic Detection of Elephant Footsteps. In *2024 IEEE 27th International Conference on Information Fusion (FUSION)*, Venice, Italy, July 2024. IEEE.

[17] Fredrik Gustafsson. *Statistical Sensor Fusion*. Studentlitteratur AB, 2018. ISBN 9789144127248.

[18] Richard Hoare. Lessons from 15 years of human-elephant conflict mitigation: Management considerations involving biological, physical and governance issues in Africa. *Pachyderm*, 51:60–74, June 2012.

[19] Richard A. Johnson and Dean W. Wichern. *Applied Multivariate Statistical Analysis*. Applied Multivariate Statistical Analysis. Pearson Prentice Hall, Upper Saddle River, NJ, 6. ed. edition, 2007. ISBN 9780131877153. Literaturverz. S. 755 - 756.

[20] Rudolf E. Kalman. A New Approach to Linear Filtering and Prediction Problems. *Journal of Basic Engineering*, 82(1):35–45, March 1960. ISSN 0021-9223. doi: 10.1115/1.3662552.

[21] Mary E. Kappus and Frank L. Vernon. The acoustic signature of thunder from seismic records. *The Journal of the Acoustical Society of America*, 88 (S1):S191–S191, November 1990. ISSN 1520-8524. doi: 10.1121/1.2028860.

[22] Steven M. Kay. *Fundamentals of Statistical Signal Processing: Estimation Theory*, volume 1. Prentice-Hall, Inc., Upper Saddle River, NJ, 20. pr. edition, 1993. ISBN 0133457117.

[23] Udo Klein and Trinh Quoc Vo. Direction-of-arrival estimation using a microphone array with the multichannel cross-correlation method. In *2012 IEEE International Symposium on Signal Processing and Information Technology (ISSPIT)*, pages 251–256, Ho Chi Minh City, Vietnam, December 2012. IEEE Computer Society. doi: 10.1109/ISSPIT.2012.6621296.

[24] Charles Knapp and G. Carter. The generalized correlation method for estimation of time delay. *IEEE Transactions on Acoustics, Speech, and Signal Processing*, 24(4):320–327, August 1976. ISSN 0096-3518. doi: 10.1109/TASSP.1976.1162830.

[25] Gökhan Koç and Korkut Yegin. Footstep and vehicle detection using seismic sensors in wireless sensor network: Field tests. *International Journal of Distributed Sensor Networks*, 9(10):120386, October 2013. ISSN 1550-1477. doi: 10.1155/2013/120386.

[26] Hamid Krim and Mats Viberg. Two decades of array signal processing research: the parametric approach. *IEEE Signal Processing Magazine*, 13(4): 67–94, July 1996. ISSN 1053-5888. doi: 10.1109/79.526899.

[27] Johan Löfberg. YALMIP : A Toolbox for Modeling and Optimization in MATLAB. In *2004 IEEE International Conference on Robotics and Automation*, CACSD-04, Taipei, Taiwan, 2004. IEEE. doi: 10.1109/CACSD.2004.1393890.

[28] L. Marple. Computing the discrete-time "analytic" signal via FFT. *IEEE Transactions on Signal Processing*, 47(9):2600–2603, 1999. ISSN 1053-587X. doi: 10.1109/78.782222.

[29] Aljazeera Newsfeed. Kenya's human-elephant conflict, 2020. URL `https://www.aljazeera.com/program/newsfeed/2020/11/26/kenyas-human-elephant-conflict`. Accessed: 2024-08-16.

[30] Caitlin E. O'Connell-Rodwell. Keeping an "ear" to the ground: seismic communication in elephants. *Physiology*, 22(4):287–294, August 2007. ISSN 1548-9221. doi: 10.1152/physiol.00008.2007.

[31] Caitlin E. O'Connell-Rodwell, B. T. Arnason, and Lynette A. Hart. Seismic properties of Asian elephant (Elephas maximus) vocalizations and locomotion. *The Journal of the Acoustical Society of America*, 108(6):3066–3072, December 2000. ISSN 1520-8524. doi: 10.1121/1.1323460.

[32] Joyce H. Poole. *Behavioral Contexts of Elephant Acoustic Communication*, pages 125–159. University of Chicago Press, 2011. ISBN 9780226542263. doi: 10.7208/chicago/9780226542263.003.0009.

[33] Quicksounds. Indian Elephant 1. URL `https://quicksounds.com/sound/3446/indian-elephant-1`. Accessed: 2024-09-06.

[34] Caleb Rascon and Ivan Meza. Localization of sound sources in robotics: A review. *Robotics and Autonomous Systems*, 96:184–210, October 2017. ISSN 0921-8890. doi: 10.1016/j.robot.2017.07.011.

[35] David M. Raup and J. John Sepkoski. Mass extinctions in the marine fossil record. *Science*, 215(4539):1501–1503, March 1982. ISSN 1095-9203. doi: 10.1126/science.215.4539.1501.

[36] Michael Reinwald, Ben Moseley, Alexandre Szenicer, Tarje Nissen-Meyer, Sandy Oduor, Fritz Vollrath, Andrew Markham, and Beth Mortimer. Seismic localization of elephant rumbles as a monitoring approach. *Journal of The Royal Society Interface*, 18(180):20210264, July 2021. ISSN 1742-5662. doi: 10.1098/rsif.2021.0264.

[37] Nima Riahi and Peter Gerstoft. The seismic traffic footprint: Tracking trains, aircraft, and cars seismically. *Geophysical Research Letters*, 42(8):2674–2681, April 2015. ISSN 1944-8007. doi: 10.1002/2015GL063558.

[38] Hannah Ritchie. The state of the world's elephant populations. *Our World in Data*, 2022. URL `https://ourworldindata.org/elephant-populations`. Accessed: 2024-09-06.

[39] Ralph O. Schmidt. Multiple emitter location and signal parameter estimation. *IEEE Transactions on Antennas and Propagation*, 34(3):276–280, March 1986. ISSN 1558-2221. doi: 10.1109/TAP.1986.1143830.

[40] *SM-24 Geophone Element*. Sensor Nederlands BV, 2006. URL `https://cdn.sparkfun.com/datasheets/Sensors/Accelerometers/SM-24Brochure.pdf`. Accessed: 2024-09-06.

[41] Claude E. Shannon. Communication in the Presence of Noise. *Proceedings of the IRE*, 37(1):10–21, January 1949. ISSN 0096-8390. doi: 10.1109/JRPROC.1949.232969.

[42] Shishir Kumar Singh, M Premalatha, and Gita Nair. Ellipsoidal gating for an airborne track while scan radar. In *Proceedings International Radar Conference*, RADAR-95, pages 334–339, Alexandria, VA, USA, 1995. IEEE, IEEE. doi: 10.1109/RADAR.1995.522568.

[43] Soundjay. Woman Scream 01. URL `https://www.soundjay.com/human/sounds/woman-scream-01.mp3`. Accessed: 2024-09-06.

[44] Petre Stoica, Prabhu Babu, and Jian Li. SPICE: A Sparse Covariance-Based Estimation Method for Array Processing. *IEEE Transactions on Signal Processing*, 59(2):629–638, February 2011. ISSN 1053-587X. doi: 10.1109/TSP.2010.2090525.

[45] Petre G. Stoica and Randolph Moses. *Spectral Analysis of Signals*. Pearson Prentice Hall, Upper Saddle River, NJ, 2005. ISBN 9780131139565.

[46] E Suriñach, I Vilajosana, Giorgi Khazaradze, B Biescas, G Furdada, and Joan Manuel Vilaplana. Seismic detection and characterization of landslides and other mass movements. *Natural Hazards and Earth System Sciences*, 5 (6):791–798, October 2005. ISSN 1684-9981. doi: 10.5194/nhess-5-791-2005.

[47] Alexandre Szenicer, Michael Reinwald, Ben Moseley, Tarje Nissen-Meyer, Zachary Mutinda Muteti, Sandy Oduor, Alex McDermott-Roberts, Atilim G. Baydin, and Beth Mortimer. Seismic savanna: machine learning for classifying wildlife and behaviours using ground-based vibration field recordings. *Remote Sensing in Ecology and Conservation*, 8(2):236–250, November 2021. ISSN 2056-3485. doi: 10.1002/rse2.242.

[48] *Very Low Noise, 24-Bit Analog-to-Digital Converter*. Texas Instruments, 2003. URL `https://www.ti.com/lit/ds/sbas288k/sbas288k.pdf`. Accessed: 2024-09-06.

[49] Harry L. Van Trees. *Optimum Array Processing*. John Wiley & Sons, New York, 2002. ISBN 9780471093909.

[50] Videvo. Police Siren Variou TE2027601. URL `https://www.videvo.net/sound-effect/police-siren-variou-te2027601/253680/`. Accessed: 2024-09-06.

[51] Spahr C. Webb. *Chapter 19 - Seismic Noise on Land and on the Sea Floor*, volume 81 of *International Geophysics*, pages 305–318. Academic Press, 2002. ISBN 9780124406520. doi: 10.1016/S0074-6142(02)80222-4.

[52] Jason D. Wood, Caitlin E. O'Connell-Rodwell, and Simon L. Klemperer. Using seismic sensors to detect elephants and other large mammals: a potential census technique. *Journal of Applied Ecology*, 42(3):587–594, June 2005. ISSN 1365-2664. doi: 10.1111/j.1365-2664.2005.01044.x.

[53] Gustav Zetterqvist. Direction of Arrival Estimation for Wildlife Protection - Soundfiles. Zenodo, 2024. doi: 10.5281/ZENODO.13833589.

[54] Gustav Zetterqvist, Fredrik Gustafsson, and Gustaf Hendeby. Using Received Power in Microphone Arrays to Estimate Direction of Arrival. In *ICASSP 2023 - 2023 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, Rhodes Island, Greece, June 2023. IEEE. doi: 10.1109/ICASSP49357.2023.10097197.

[55] Gustav Zetterqvist, Fredrik Gustafsson, and Gustaf Hendeby. Directional Sensitivity-Based DOA Estimation Using a Fourier Series Model. *IEEE Sensors Journal*, August 2024. Submitted, under review.

[56] Gustav Zetterqvist, Erik Wahledow, Philip Sjövik, Fredrik Gustafsson, and Gustaf Hendeby. Elephant DOA Estimation using a Geophone Network. In *2023 IEEE 26th International Conference on Information Fusion (FUSION)*, Charleston, SC, USA, June 2023. IEEE. doi: 10.23919/FUSION52260.2023.10224115.

Licentiate Theses
Division of Automatic Control
Linköping University

P. Andersson: Adaptive Forgetting through Multiple Models and Adaptive Control of Car Dynamics. Thesis No. 15, 1983.

B. Wahlberg: On Model Simplification in System Identification. Thesis No. 47, 1985.

A. Isaksson: Identification of Time Varying Systems and Applications of System Identification to Signal Processing. Thesis No. 75, 1986.

G. Malmberg: A Study of Adaptive Control Missiles. Thesis No. 76, 1986.

S. Gunnarsson: On the Mean Square Error of Transfer Function Estimates with Applications to Control. Thesis No. 90, 1986.

M. Viberg: On the Adaptive Array Problem. Thesis No. 117, 1987.

K. Ståhl: On the Frequency Domain Analysis of Nonlinear Systems. Thesis No. 137, 1988.

A. Skeppstedt: Construction of Composite Models from Large Data-Sets. Thesis No. 149, 1988.

P. A. J. Nagy: MaMiS: A Programming Environment for Numeric/Symbolic Data Processing. Thesis No. 153, 1988.

K. Forsman: Applications of Constructive Algebra to Control Problems. Thesis No. 231, 1990.

I. Klein: Planning for a Class of Sequential Control Problems. Thesis No. 234, 1990.

F. Gustafsson: Optimal Segmentation of Linear Regression Parameters. Thesis No. 246, 1990.

H. Hjalmarsson: On Estimation of Model Quality in System Identification. Thesis No. 251, 1990.

S. Andersson: Sensor Array Processing; Application to Mobile Communication Systems and Dimension Reduction. Thesis No. 255, 1990.

K. Wang Chen: Observability and Invertibility of Nonlinear Systems: A Differential Algebraic Approach. Thesis No. 282, 1991.

J. Sjöberg: Regularization Issues in Neural Network Models of Dynamical Systems. Thesis No. 366, 1993.

P. Pucar: Segmentation of Laser Range Radar Images Using Hidden Markov Field Models. Thesis No. 403, 1993.

H. Fortell: Volterra and Algebraic Approaches to the Zero Dynamics. Thesis No. 438, 1994.

T. McKelvey: On State-Space Models in System Identification. Thesis No. 447, 1994.

T. Andersson: Concepts and Algorithms for Non-Linear System Identifiability. Thesis No. 448, 1994.

P. Lindskog: Algorithms and Tools for System Identification Using Prior Knowledge. Thesis No. 456, 1994.

J. Plantin: Algebraic Methods for Verification and Control of Discrete Event Dynamic Systems. Thesis No. 501, 1995.

J. Gunnarsson: On Modeling of Discrete Event Dynamic Systems, Using Symbolic Algebraic Methods. Thesis No. 502, 1995.

A. Ericsson: Fast Power Control to Counteract Rayleigh Fading in Cellular Radio Systems. Thesis No. 527, 1995.

M. Jirstrand: Algebraic Methods for Modeling and Design in Control. Thesis No. 540, 1996.

K. Edström: Simulation of Mode Switching Systems Using Switched Bond Graphs. Thesis No. 586, 1996.

J. Palmqvist: On Integrity Monitoring of Integrated Navigation Systems. Thesis No. 600, 1997.

A. Stenman: Just-in-Time Models with Applications to Dynamical Systems. Thesis No. 601, 1997.

M. Andersson: Experimental Design and Updating of Finite Element Models. Thesis No. 611, 1997.

U. Forssell: Properties and Usage of Closed-Loop Identification Methods. Thesis No. 641, 1997.

M. Larsson: On Modeling and Diagnosis of Discrete Event Dynamic systems. Thesis No. 648, 1997.

N. Bergman: Bayesian Inference in Terrain Navigation. Thesis No. 649, 1997.

V. Einarsson: On Verification of Switched Systems Using Abstractions. Thesis No. 705, 1998.

J. Blom, F. Gunnarsson: Power Control in Cellular Radio Systems. Thesis No. 706, 1998.

P. Spångéus: Hybrid Control using LP and LMI methods – Some Applications. Thesis No. 724, 1998.

M. Norrlöf: On Analysis and Implementation of Iterative Learning Control. Thesis No. 727, 1998.

A. Hagenblad: Aspects of the Identification of Wiener Models. Thesis No. 793, 1999.

F. Tjärnström: Quality Estimation of Approximate Models. Thesis No. 810, 2000.

C. Carlsson: Vehicle Size and Orientation Estimation Using Geometric Fitting. Thesis No. 840, 2000.

J. Löfberg: Linear Model Predictive Control: Stability and Robustness. Thesis No. 866, 2001.

O. Härkegård: Flight Control Design Using Backstepping. Thesis No. 875, 2001.

J. Elbornsson: Equalization of Distortion in A/D Converters. Thesis No. 883, 2001.

J. Roll: Robust Verification and Identification of Piecewise Affine Systems. Thesis No. 899, 2001.

I. Lind: Regressor Selection in System Identification using ANOVA. Thesis No. 921, 2001.

R. Karlsson: Simulation Based Methods for Target Tracking. Thesis No. 930, 2002.

P.-J. Nordlund: Sequential Monte Carlo Filters and Integrated Navigation. Thesis No. 945, 2002.

M. Östring: Identification, Diagnosis, and Control of a Flexible Robot Arm. Thesis No. 948, 2002.

C. Olsson: Active Engine Vibration Isolation using Feedback Control. Thesis No. 968, 2002.

J. Jansson: Tracking and Decision Making for Automotive Collision Avoidance. Thesis No. 965, 2002.

N. Persson: Event Based Sampling with Application to Spectral Estimation. Thesis No. 981, 2002.

D. Lindgren: Subspace Selection Techniques for Classification Problems. Thesis No. 995, 2002.

E. Geijer Lundin: Uplink Load in CDMA Cellular Systems. Thesis No. 1045, 2003.

M. Enqvist: Some Results on Linear Models of Nonlinear Systems. Thesis No. 1046, 2003.

T. Schön: On Computational Methods for Nonlinear Estimation. Thesis No. 1047, 2003.

F. Gunnarsson: On Modeling and Control of Network Queue Dynamics. Thesis No. 1048, 2003.

S. Björklund: A Survey and Comparison of Time-Delay Estimation Methods in Linear Systems. Thesis No. 1061, 2003.

M. Gerdin: Parameter Estimation in Linear Descriptor Systems. Thesis No. 1085, 2004.

A. Eidehall: An Automotive Lane Guidance System. Thesis No. 1122, 2004.

E. Wernholt: On Multivariable and Nonlinear Identification of Industrial Robots. Thesis No. 1131, 2004.

J. Gillberg: Methods for Frequency Domain Estimation of Continuous-Time Models. Thesis No. 1133, 2004.

G. Hendeby: Fundamental Estimation and Detection Limits in Linear Non-Gaussian Systems. Thesis No. 1199, 2005.

D. Axehill: Applications of Integer Quadratic Programming in Control and Communication. Thesis No. 1218, 2005.

J. Sjöberg: Some Results On Optimal Control for Nonlinear Descriptor Systems. Thesis No. 1227, 2006.

D. Törnqvist: Statistical Fault Detection with Applications to IMU Disturbances. Thesis No. 1258, 2006.

H. Tidefelt: Structural algorithms and perturbations in differential-algebraic equations. Thesis No. 1318, 2007.

S. Moberg: On Modeling and Control of Flexible Manipulators. Thesis No. 1336, 2007.

J. Wallén: On Kinematic Modelling and Iterative Learning Control of Industrial Robots. Thesis No. 1343, 2008.

J. Harju Johansson: A Structure Utilizing Inexact Primal-Dual Interior-Point Method for Analysis of Linear Differential Inclusions. Thesis No. 1367, 2008.

J. D. Hol: Pose Estimation and Calibration Algorithms for Vision and Inertial Sensors. Thesis No. 1370, 2008.

H. Ohlsson: Regression on Manifolds with Implications for System Identification. Thesis No. 1382, 2008.

D. Ankelhed: On low order controller synthesis using rational constraints. Thesis No. 1398, 2009.

P. Skoglar: Planning Methods for Aerial Exploration and Ground Target Tracking. Thesis No. 1420, 2009.

C. Lundquist: Automotive Sensor Fusion for Situation Awareness. Thesis No. 1422, 2009.

C. Lyzell: Initialization Methods for System Identification. Thesis No. 1426, 2009.

R. Falkeborn: Structure exploitation in semidefinite programming for control. Thesis No. 1430, 2010.

D. Petersson: Nonlinear Optimization Approaches to $\mathcal{H}_2$-Norm Based LPV Modelling and Control. Thesis No. 1453, 2010.

Z. Sjanic: Navigation and SAR Auto-focusing in a Sensor Fusion Framework. Thesis No. 1464, 2011.

K. Granström: Loop detection and extended target tracking using laser data. Thesis No. 1465, 2011.

J. Callmer: Topics in Localization and Mapping. Thesis No. 1489, 2011.

F. Lindsten: Rao-Blackwellised particle methods for inference and identification. Thesis No. 1480, 2011.

M. Skoglund: Visual Inertial Navigation and Calibration. Thesis No. 1500, 2011.

S. Khoshfetrat Pakazad: Topics in Robustness Analysis. Thesis No. 1512, 2011.

P. Axelsson: On Sensor Fusion Applied to Industrial Manipulators. Thesis No. 1511, 2011.

A. Carvalho Bittencourt: On Modeling and Diagnosis of Friction and Wear in Industrial Robots. Thesis No. 1516, 2012.

P. Rosander: Averaging level control in the presence of frequent inlet flow upsets. Thesis No. 1527, 2012.

N. Wahlström: Localization using Magnetometers and Light Sensors. Thesis No. 1581, 2013.
R. Larsson: System Identification of Flight Mechanical Characteristics. Thesis No. 1599, 2013.
Y. Jung: Estimation of Inverse Models Applied to Power Amplifier Predistortion. Thesis No. 1605, 2013.
M. Syldatk: On Calibration of Ground Sensor Networks. Thesis No. 1611, 2013.
M. Roth: Kalman Filters for Nonlinear Systems and Heavy-Tailed Noise. Thesis No. 1613, 2013.
D. Simon: Model Predictive Control in Flight Control Design — Stability and Reference Tracking. Thesis No. 1642, 2014.
J. Dahlin: Sequential Monte Carlo for inference in nonlinear state space models. Thesis No. 1652, 2014.
M. Kok: Probabilistic modeling for positioning applications using inertial sensors. Thesis No. 1656, 2014.
J. Linder: Graybox Modelling of Ships Using Indirect Input Measurements. Thesis No. 1681, 2014.
G. Mathai: Direction of Arrival Estimation of Wideband Acoustic Wavefields in a Passive Sensing Environment. Thesis No. 1721, 2015.
I. Nielsen: On Structure Exploiting Numerical Algorithms for Model Predictive Control. Thesis No. 1727, 2015.
C. Veibäck: Tracking of Animals Using Airborne Cameras. Thesis No. 1761, 2016.
N. Evestedt: Sampling Based Motion Planning for Heavy Duty Autonomous Vehicles. Thesis No. 1762, 2016.
H. Nyqvist: On Pose Estimation in Room-Scaled Environments. Thesis No. 1765, 2016.
Y. Zhao: Position Estimation in Uncertain Radio Environments and Trajectory Learning. Thesis No. 1772, 2017.
P. Kasebzadeh: Parameter Estimation for Mobile Positioning Applications. Thesis No. 1786, 2017.
K. Radnosrati: On Timing-Based Localization in Cellular Radio Networks. Thesis No. 1808, 2018.
G. Lindmark: Methods and Algorithms for Control Input Placement in Complex Networks. Thesis No. 1814, 2018.
M. Lindfors: Frequency Tracking for Speed Estimation. Thesis No. 1815, 2018.
D. Ho: Some results on closed-loop identification of quadcopters. Thesis No. 1826, 2018.
O. Ljungqvist: On motion planning and control for truck and trailer systems. Thesis No. 1832, 2019.
P. Boström-Rost: On Informative Path Planning for Tracking and Surveillance. Thesis No. 1838, 2019.
K. Bergman: On Motion Planning Using Numerical Optimal Control. Thesis No. 1843, 2019.
M. Klingspor: Low-rank optimization in system identification. Thesis No. 1855, 2019.
A. Bergström: Timing-Based Localization using Multipath Information. Thesis No. 1867, 2019.
F. Ljungberg: Estimation of Nonlinear Greybox Models for Marine Applications. Thesis No. 1880, 2020.
E. Hedberg: Control, Models and Industrial Manipulators. Thesis No. 1894, 2020.
R. Forsling: Decentralized Estimation Using Conservative Information Extraction. Thesis No. 1897, 2020.

D. Arnström: On Complexity Certification of Active-Set QP Methods with Applications to Linear MPC. Thesis No. 1901, 2021.
M. Malmström: Uncertainties in Neural Networks: A System Identification Approach. Thesis No. 1902, 2021.
K. Nielsen: Robust LIDAR-Based Localization in Underground Mines. Thesis No. 1906, 2021.
H. Haghshenas: Time-Optimal Cooperative Path Tracking for Multi-Robot Systems. Thesis No. 1915, 2021.
A. Kullberg: On Joint State Estimation and Model Learning using Gaussian Process Approximations. Thesis No. 1917, 2021.
J. Nordlöf: On Landmark Densities in Minimum-Uncertainty Motion Planning. Thesis No. 1927, 2022.
S. A. Zimmermann: Data-driven Modeling of Robotic Manipulators—Efficiency Aspects. Thesis No. 1963, 2023.
S. Shoja: On Complexity Certification of Branch-and-Bound Methods for MILP and MIQP with Applications to Hybrid MPC. Thesis No. 1967, 2023.
A. Hellander: On Optimal Integrated Task and Motion Planning with Applications to Tractor-Trailers. Thesis No. 1981, 2023.
J. Wilroth: Exploring Auditory Attention Using EEG. Thesis No. 1993, 2024.
C. Huang: On Indoor Localization Using Magnetic Field-Aided Inertial Navigation Systems. Thesis No. 2003, 2024.

FACULTY OF SCIENCE AND ENGINEERING

Linköping Studies in Science and Technology.
Licentiate Thesis No. 2006, 2024

Department of Electrical Engineering
Linköping University
SE-581 83 Linköping, Sweden

www.liu.se