Master of Science Thesis in Electrical Engineering
Department of Electrical Engineering, Linköping University, 2023

Blind Acoustic Feedback Cancellation for an AUV

Hampus Frick

LINKÖPING UNIVERSITY

Master of Science Thesis in Electrical Engineering

Blind Acoustic Feedback Cancellation for an AUV

Hampus Frick

LiTH-ISY-EX--23/5607--SE

Supervisor: **Daniel Bossér**
DEPARTMENT OF ELECTRICAL ENGINEERING,
Linköpings University (LiU)
Louise Fuchs
SAAB Dynamics
Per Abrahamsson
SAAB Dynamics

Examiner: **Fredrik Gustafsson**
DEPARTMENT OF ELECTRICAL ENGINEERING,
Linköpings University (LiU)

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

Abstract

SAAB has developed an autonomous underwater vehicle that can mimic a conventional submarine for military fleets to exercise anti-submarine warfare. The *AUV* actively emits amplified versions of received sonar pulses to create the illusion of being a larger object. To prevent **acoustic feedback**, the *AUV* must distinguish between the sound to be actively responded to and its emitted signal. This master thesis has examined techniques aimed at preventing the *AUV* from responding to previously emitted signals to avoid **acoustical feedback**, without **relying on prior knowledge** of either the received signal or the signal emitted by the *AUV*. The two primary types of algorithms explored for this problem include **blind source separation** and **adaptive filtering**.

The **adaptive filters** based on **Leaky Least Mean Square** and **Kalman** have shown promising results in attenuating the active response from the received signal. The **adaptive filters** utilize the fact that a certain hydrophone primarily receives the active response. This hydrophone serves as an estimate of the active response since the signal it captures is considered unknown and is to be removed.

The techniques based on **blind source separation** have utilized the recordings of three hydrophones placed at various locations of the *AUV* to separate and estimate the received signal from the one emitted by the *AUV*. The results have demonstrated that neither of the reviewed methods is suitable for implementation on the *AUV*. The hydrophones are situated at a considerable distance from each other, resulting in distinct time delays between the reception of the two signals. This is usually referred to as a **convolutive mixture**. This is commonly solved using the frequency domain to transform the **convolutive mixture** to an **instantaneous mixture**. However, the fact that the signals share the same frequency spectrum and are adjacent in time has proven highly challenging.

Keywords: acoustical feedback, adaptive filter, least leaky mean square, Kalman, blind source separation, independent component analysis, complex independent component analysis, degenerative unmixing estimation technique, convolutive mixture.

Acknowledgments

I would like to express my appreciation to my supervisors at *SAAB Dynamics*, **Louise Fuchs**, and **Per Abrahamsson**. **Louise**, for your unwavering guidance, great ideas, and invaluable feedback. **Per**, for your assistance and introducing me to the world of *AUVs* and underwater acoustics. I have truly learned a lot from both of you.

I would like to express my sincere gratitude to my supervisor at *LiU*, **Daniel Bossér**. Your counseling and invaluable support throughout this journey have been of immense value. Thank you for your dedication to making this project succeed.

Furthermore, I would also like to thank **Jennie Asp**, the supervisor of the Sonar division at *SAAB*. I sincerely appreciate your inclusive attitude and unwavering support throughout my tenure at SAAB. Last but not least, thank you **Fredrik Gustafsson** for showing interest in my thesis and examining it.

Linköping, June 2023
Hampus Frick

Contents

Notation

Notations

Notations	Meaning
E	Expected Value
H	Hermitian Transpose
T	Transpose

Abbreviations

Abbreviations	Meaning
AUV	Autonomous Underwater Vehicle
BSS	Blind Source Separation
CICA	Complex Independent Component Analysis
DCA	Dependent Component Analysis
FOI	Totalförsvarets forskningsinstitut / Swedish Defence Research Agency
ICA	Independent Component Analysis
HLMS	Hybrid Least Mean Square
LLMS	Leaky Least Mean Square
LMS	Least Mean Square
PA	Pulse Active - Active Response of the AUV
PCA	Principal Component Analysis
PE	Pulse Emitted - Pulse emitted by searching boat
RMSE	Root Mean Square Error
STFT	Short Time Fourier Transform
SNR	Signal to Noise Ratio
SVD	Singular Value Decomposition

1 Introduction

1.1 Background

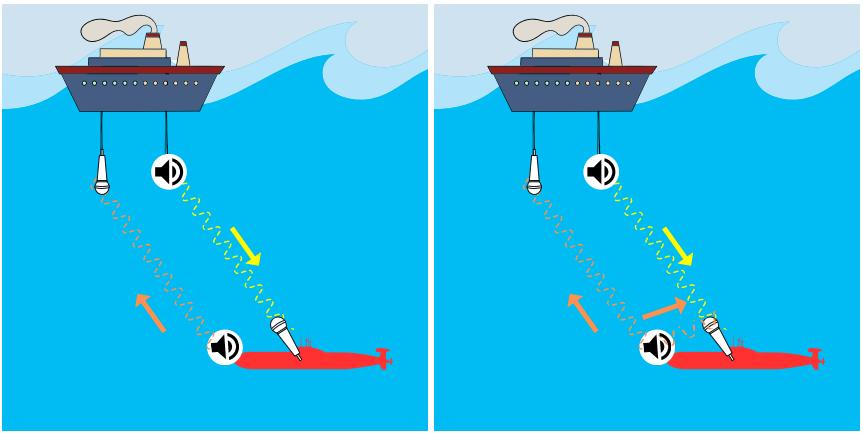
Military fleets around the world exercise anti-submarine warfare. *SAAB* has developed an *AUV* which can mimic a larger conventional military submarine. The reason for using the *AUV* instead of a more conventional military submarine is the reduced cost of performing these exercises.

A common approach to detect submerged objects is to use active sonar. Simplified, this is done by emitting sound pulses from a beacon that will reflect when hitting surrounding objects, thus; creating an echo that can be observed with a hydrophone. When measuring the time delay and strength of the emitted pulses and the received echo, one can determine information such as the distance and size of surrounding objects [20].

The *AUV* mimics the response of a larger submarine by measuring the strength of the sound received by an emitted sonar. Then, the *AUV* enhances the strength of the natural echo reflected by its frame to appear as a larger object. A simplified illustration of this can be seen in figure 1.1a.

1.2 Problem Formulation

The hydrophones of the *AUV* will receive a wide range of sounds from various sources. The two most important ones are the active sonar pulse from the searching boat and the active response emitted by the *AUV* itself, illustrated in figure 1.1b. This will result in a mixture of two highly similar signals containing an almost identical frequency spectrum, slightly separated in time, with varying amplitude. However, we lack access to the characteristics of the active response



(a) *Illustration of the AUV mimicking a conventional submarine by emitting active responses to received sonar.*

(b) *Illustration of the AUV mimicking a conventional submarine by emitting active responses to received sonar, while registering its own emitted signal.*

Figure 1.1: *Illustration of the AUV responding to a nearby ship's sonar.*

and consider it as an unknown parameter. This is because the delay between receiving a pulse and providing an active response needs to be minimized. This delay would increase if the active response were treated as a known parameter. Our objective is to isolate the signal to which a response is intended from the active response without prior knowledge.

If the *AUV* were to disregard the phenomenon of self-registration of its own emitted pulse, acoustical feedback would occur. The responses emitted by the *AUV* would not solely be based on the pulse received from the ship but would also incorporate earlier emitted pulses, resulting in meaningless noise.

The *AUV* is equipped with three hydrophones. Earlier work at *SAAB* for dealing with the problem explained above has concentrated solely on utilizing one of these hydrophones. This is because one of the hydrophones is positioned at the rear, exposed to a lower degree of ambient noise, primarily focusing on recording sonar pulses. The remaining two hydrophones are primarily used for navigation. However, this master's thesis has focused on utilizing all three hydrophones to mitigate acoustical feedback for the *AUV*; see figure 2.2 for the layout of the *AUV*.

1.3 Examined Methods

This master's thesis focuses on two primary approaches to address the issue of acoustical feedback: signal separation and adaptive filtering.

1.3.1 Signal Separation

Due to the absence of prior knowledge about the signals, a technique referred to as blind source separation - *BSS* has been investigated. These methods are primarily used to separate speech from one another, usually referred to as the *Cocktail party problem*. This concept involves extracting individual voices from one or multiple recordings in environments where multiple individuals speak, and noise is present. The idea is to utilize the three hydrophones of the *AUV* to separate the two mixed signals with *BSS*.

Several methods within the field of *BSS* have been evaluated; however, the results have not demonstrated satisfactory performance.

1.3.2 Adaptive Filtering

Considering that the active response is unknown, one approach to estimate the active response is to use the side hydrophones at the *AUV*. The side hydrophones are positioned close to the *AUV's* own emitter; thus, the active response is significantly more prominent for these recordings. Subsequently, this data can serve as a source for generating an estimation of the active response, which can be subsequently filtered out in the rear hydrophone.

The results have shown promising outcomes and have been applied to both simulated and real-world data.

1.4 Research Questions

In order to address the problem outlined in section 1.2 and derive essential conclusions from the methods discussed in section 1.3, the following research questions are presented:

- **RQ1** - Can blind source separation effectively separate the active response while preserving the receiving signal, and what types of constraints may arise during the separation process?
- **RQ2** - To what extent is it feasible to implement the reviewed blind source separation in the *AUV* for generating active responses?
- **RQ3** - Can adaptive filtering effectively attenuate the active response while preserving the receiving signal? Additionally, what constraints and challenges may arise?

- **RQ4** - To what extent is it feasible to implement the reviewed adaptive filters in the *AUV* for generating active responses?
- **RQ5** - What areas of interest could be investigated for future work and why?

2

Background of the *AUV*

This chapter outlines the signals' background and essential properties regarding the *AUV*.

2.1 Layout of the *AUV*

Figure 2.1 depicts the *AUV* seen from above with the three available hydrophones colored in red. Two are placed on the sides of the *AUV*, one pointing to starboard and the other to port, referred to as the side hydrophones. The third hydrophone is dragged 75 meters behind, called the rear hydrophone.

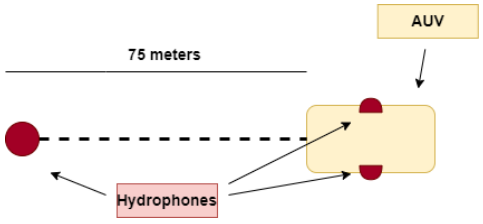


Figure 2.1: *The AUV depicted from above with its three hydrophones.*

2.2 Mixture of Signals Received by the *AUV*

Figure 2.2 depicts the *AUV* positioned in the water with a nearby ship. Noteworthy, no signals are emitted.

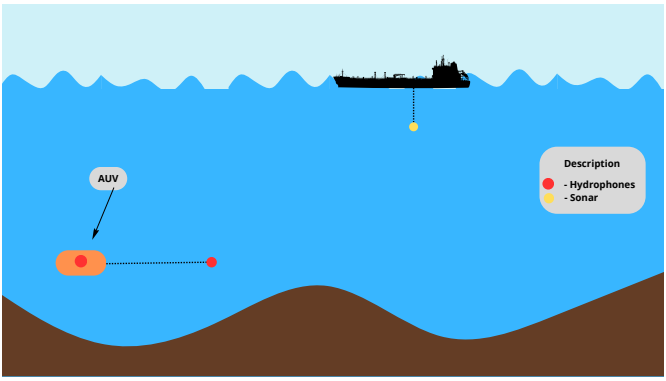


Figure 2.2: *AUV and ship with no signals.*

Figure 2.3 then illustrates when the ship is emitting sonar pulses to detect the *AUV*, note no reflections such as the ones depicted in figure 4.1 are present. The pulses PE_1 to PE_5 are chronologically emitted by the ship's sonar. Due to the travel time for sound, the emitted pulses will be received at different instances for the three hydrophones. This is portrayed in figure 2.3, as PE_1 is received at the side hydrophone at the same time as PE_2 is received at the rear.

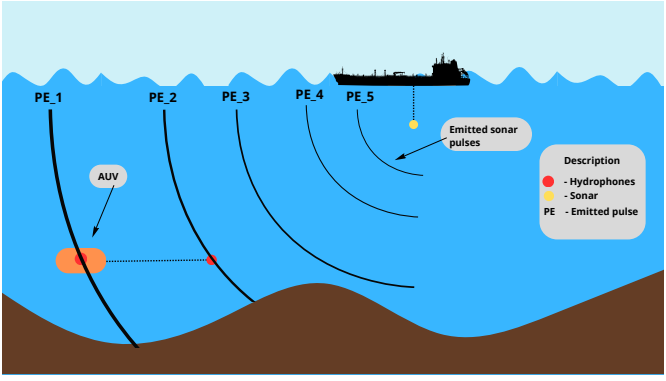


Figure 2.3: *AUV and ship with the signal PE.*

Figure 2.4 depicts when the *AUV* is actively responding to the received signals from the ship. There are two key characteristics portrayed in this figure.

Firstly, the rear hydrophone is considered the primary hydrophone due to its significantly lower noise level compared to the side hydrophones. If the active response *PA* is based purely on what is received at the rear hydrophone, acoustical feedback will occur. The main task of this master's thesis is to make the *AUV* actively reply to *PE* while not responding to *PA*.

Secondly, the sonar pulses, *PA* and *PE* will be registered at different time instances by the three hydrophones of the *AUV*. This is portrayed in figure 2.4 as in this example, PA_2 and PE_5 are registered at the rear hydrophone at the same time as PA_3 and PE_4 are at the side.

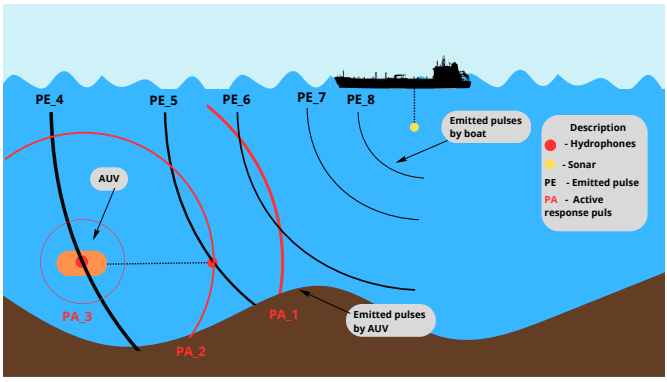


Figure 2.4: *AUV and ship with the signals PE and PA.*

2.3 Notable characteristics of the *AUV*

Following list displays some notable characteristics of the *AUV*:

- The hydrophones at the sides of the *AUV* are currently used predominately for navigation and are exposed to increased noise compared to the one at the rear.
- A 13 ms delay occurs between receiving a signal at the rear hydrophone and providing an active response.
- The rear hydrophone is positioned 75 meters behind the *AUV* creating a time delay for signals received by the rear and the hydrophones at the sides. This constitutes a maximum delay of approximately 0.05 seconds; see section 3.5 for further equations.

3

Method

3.1 Limitations

The limitations of this master thesis are presented in the list below. A few requirements for a desirable solution, such as that PA and PE are considered unknown, may act as limitations and are presented in section 3.2.

- Only one boat emitting sonar pulses PE.
- Sonar pulses limited to a maximum frequency of 20KHz for simulation.
- Sufficient depth resulting in negligible reflections.
- Sufficient signal amplitude to distinguish it from noise.

3.2 Requirements

A delay of 13ms for a signal registered at the rear hydrophone to be actively responded to occurs with the current setup; see section 2.3. This delay is predominantly due to hardware limitations such as buffer sizes and available computing power. A desirable algorithm for our problem, see section 1.2, should not increase this existing delay. To achieve this, a proposed algorithm must be computationally efficient and cautious of not using data that will saturate the buffers.

The requirement of using a computationally efficient algorithm imposes limitations on the selection of algorithms to be investigated; this is further evaluated in section 3.4. However, the limitation of buffer size imposes more arduous restrictions, with the most crucial of PA not considered known.

3.2.1 *PA* Considered Unknown

With the current delay of 13ms, several buffers of the *AUV* are working at full capacity. By accessing the active response *PA*, the existing buffers will be exhausted, and the delay of 13ms is deemed to be prolonged, yielding an unsatisfactory solution. Therefore, *PA* will be considered an unknown variable.

3.2.2 *PE* Considered Unknown

The emitted signal *PE* could be anticipated either to be a known signal. However, a system for dealing with fully known signals *PE* is already implemented in the *AUV*. Due to the reasons listed below, this system is not always in use:

- The complex process to incorporate the signals in advance within the *AUV*.
- For an actual exercise with the *AUV*, the characteristics such as frequency of *PE* may change during the exercise depending on the operator.
- The signal *PE* may be confidential.

Therefore, *PE* will be considered unknown.

In the simulation environment, see section 3.5, *PE* will always have the characteristics of being a *Chirp* signal; hence, this known characteristic could be utilized. However, the results would be somewhat irrelevant because a method assuming these characteristics would be difficult to develop beyond the constructed simulation.

3.3 Examined Methods

The main focus of this master's thesis revolves around two essential methods: Blind Source Separation - *BSS* and Adaptive Filters, as discussed in Section 1.3. The sections below motivate why these were chosen.

3.3.1 Motivation of Blind Source Separation

Earlier research conducted at Totalförsvarets forskningsinstitut - *FOI* [1] used *ICA*, a common *BSS* method, for self-noise suppression applied to underwater acoustics. Their report focused on estimating a specific noise by separating it from other known signals. Notably, this master's thesis diverges by focusing on separating two highly similar signals. Meanwhile, similarities remain, such as estimating an unwanted signal by separation and several hydrophones recording the signals underwater. Therefore, *BSS* was chosen as one area to be focused on.

3.3.2 Motivation of Adaptive Filtering

A common approach when reducing the presence of noise or an unwanted signal is to construct an adaptive filter [11]. The same report as in section 3.3.1 by *FOI* [1] examined the use of adaptive filters to reduce the influence of a specific noise. Their report estimated *AR* parameters to model the unwanted noise. The idea of using adaptive filters will be the second main area to be focused on to reduce the acoustical feedback of *PA*.

3.4 Evaluation

This section outlines the evaluation methods employed in this study. To assess the simulated data, we will utilize the Root Mean Square Error - *RMSE* in conjunction with power spectra and spectrograms. However, when evaluating real data, we encounter a challenge—the absence of ground truth. Consequently, the *RMSE* cannot be used to assess real data results. Instead, more emphasis has been placed on the spectrograms.

3.4.1 Fundamental Requirement with RMSE

For the simulation, a fundamental approach for assessing the usefulness of an estimation of *PE* or *PA* involves comparing the obtained results with the signal received by the hydrophones. Due to the layout of the *AUV* the received signals for the side hydrophones will have the characteristics of $PA \gg PE$ and for the rear $PA \ll PE$. An essential requirement for separated signals is that they should resemble better to ground truth than what the recording of the hydrophones does. Namely, the separated component depicting *PA* needs to resemble more to ground truth of *PA* than what the recording of the side hydrophones does. The same applies for *PE*, but then to the rear hydrophone. The phenomenon of *PA* being more prominent at the side hydrophones and that *PE* is more prominent in the rear can be seen in the simulated data in figure 3.2 further below. The Root Mean Square Error - *RMSE* determines how much the two datasets resemble. The *RMSE* is calculated as,

$$RMSE = \sqrt{E\{(\hat{x} - x)^2\}}\,[12]. \tag{3.1}$$

3.4.2 Computational Demand

Due to the requirements in section 3.2, a proposed algorithm has to be somewhat computationally efficient. This is a pretty vague requirement with no set limits. However, a chosen algorithm must be capable of online use at the *AUV* without introducing any significant delays in emitting an active response. All algorithms investigated follow this requirement; more computationally demanding algorithms have already been disregarded.

3.4.3 Evaluation - Blind Source Separation - BSS

There are several ways to evaluate the quality *BSS*. Two common parameters for evaluation are Distortion D and Separation S, described below.

The distortion D of the jth separated output can be calculated as,

$$\mathrm{D}_j = 10\log_{10}\left(\frac{E\{(x_{(j,s_j)} - \alpha_j y_j)^2\}}{E\{(x_{(j,s_j)})^2\}}\right) \ \{\text{unit: dB}\}. \tag{3.2}$$

with,

$$\alpha_j = \frac{E\{x^2_{(j,s_j)}\}}{E\{y_j^2\}}.$$

y_j corresponds to the jth source and $x_{(j,s_j)}$ denoted the recorded signals. The quality of a separation S can be calculated as,

$$\mathrm{S}_j = 10\log_{10}\left(\frac{E\{(y_{(j,s_j)})^2\}}{E\{(\sum_{i\neq j} y_{(j,s_j)})^2\}}\right) \ \{\text{unit: dB}\}, \tag{3.3}$$

with $y_{(j,s_j)}$ as the jth output of the mixing when only the source s_i is active. Several factors can be used for evaluation, such as how the *SNR* and the number of sources affect the result [19]. However, due to the unsatisfying results in section 5.1, it has shown sufficient to compare the separated signals' frequency spectrum and appearance to ground truth. This is done predominantly with *RMSE*.

3.4.4 Evaluation - Adaptive Filtering

For simulation, *RMSE* is also used to evaluate the filtered signals to ground truth. Compared to *BSS*, spectrograms of the unfiltered and filtered signals are analyzed. In the time domain, the mixture of *PE* and *PA* poses a challenge in accurately distinguishing between the two signals. However, the used signals have the characteristics of being a *Chirp*, and the frequencies of *PE* and *PA* will not overlap at the same time. A spectrogram with sufficient accuracy will then be able to separate *PE* and *PA* from one another. The comparison of spectrograms between the unfiltered and filtered signals provides a basis for analyzing the extent to which *PE* has been preserved and *PA* has been eliminated. The method of using a spectrogram can be used for both simulated and real data. Then a performance comparison between simulated and actual data can be made.

3.5 Simulation

A successful solution will either be able to separate *PA* and *PE* from one another or attenuate *PA* from the data received by the rear hydrophone. A simulation environment was created to determine if the methods in section 4.2 and 4.3 were successful. There are several advantages of using a simulation environment prior to using actual data. Such as access to ground truth, the possibility to turn off time delay due to propagation, and the ability to adjust the noise level.

Figure 3.1 shows a block diagram of the simulation setup. All elements inside the blocks are explained in section 3.5.1 - 3.5.7. Simplified, the simulation environment is constructed such as the signal *PE* is emitted by a fictitious boat at a specific location. Then propagation time, absorption-, spreading- and diffraction loss are calculated concerning the three hydrophones of the *AUV*. The active response *PA* is then generated by the signal *PE* received at the rear hydrophone. Meanwhile, propagation time, absorption, spreading- and diffraction loss are calculated from the emitter of *PA* to the three hydrophones. Noise is also added to all receiving hydrophones.

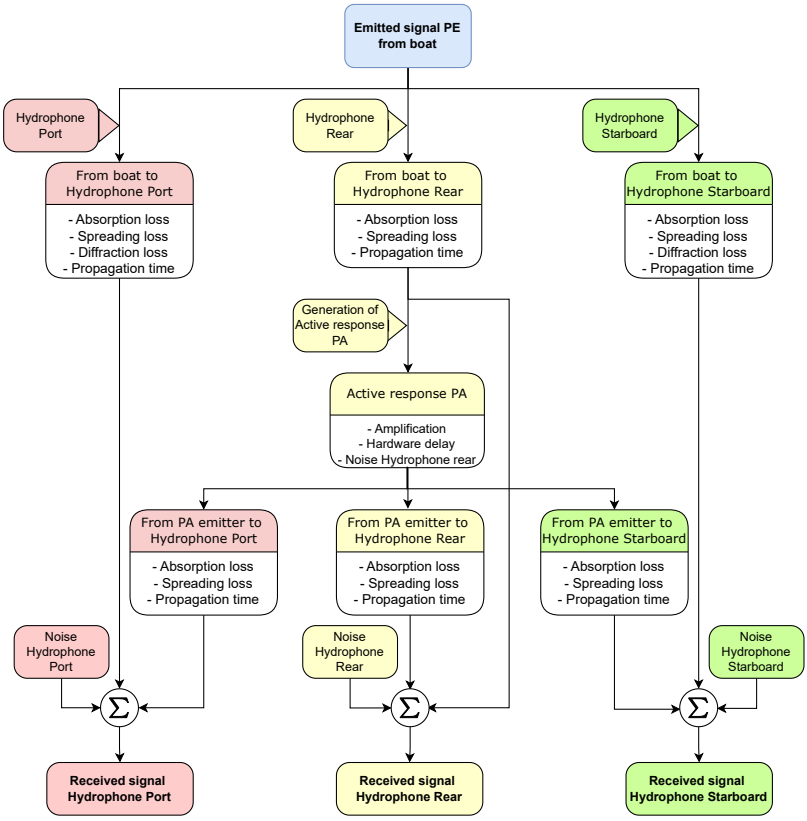


Figure 3.1: *Block diagram of the simulation.*

3.5.1 Emitted Signal PE

The emitted signal *PE* will be a pulse modulated *Chirp* signal. A cosinus and a rectangular function can express the equation of this signal as,

$$u(t) = A\cos(2\pi t(f_0 + \phi t)) \cdot \text{rect}\left(\frac{t - nT}{T_p}\right). \tag{3.4}$$

A is the amplitude of the signal, f_0 is the starting frequency in Hz with ϕ as the *Chirp* rate, n is the pulse number, T is the pulse repetition interval with T_p as its duration [17].

3.5.2 Absorption Loss

Francois and *Garrison* suggest the following formula for calculating α as the total absorption in seawater:

$$\alpha = \frac{A_1 P_1 f_1 f^2}{f_1^2 + f^2} + \frac{A_2 P_2 f_2 f^2}{f_2^2 + f^2} + A_3 P_3 f^2 \quad \left\{\text{unit: } \frac{\text{dB}}{\text{km}}\right\}. \tag{3.5}$$

Equation (3.5) is based on the notation proposed by *Fisher* and *Simmons* [7]. The first two terms describe the chemical relaxation process of boric acid and magnesium sulfate, with the final term defining the absorption from pure water. P_{1-3} are the pressure dependencies, A_{1-3} are experimentally given constant, f_{1-2} are the relaxation dependencies and f is the frequency of the sound wave [8, 9].

Ainslie and *McColm* suggest a simplified version of equation (3.5) for the following oceanographic conditions:

$$\begin{gathered} -6 < T < 35\,^{\circ}\text{C} \quad (S = 35\ ppt,\ pH = 8,\ z = 0) \\ 7.7 < pH < 8.3 \quad (T = 10\,^{\circ}\text{C},\ S = 35\ ppt,\ z = 0) \\ 5 < S < 50\ ppt \quad (T = 10\,^{\circ}\text{C},\ pH = 8,\ z = 0) \\ 0 < z < 7km \quad (T = 10\,^{\circ}\text{C},\ S = 35\ ppt,\ pH = 8)\,. \end{gathered}$$

With ppt referring to part per trillion. Due to the limitations of this master thesis, see section 3.1, and the properties of the general water environment in the Baltic Sea, the above conditions are assumed to be fulfilled [15]. Then equation (3.5) is transformed to,

$$\alpha = \frac{A_1 f_1 f^2}{f_1^2 + f^2} + \frac{A_2 f_2 f^2}{f_2^2 + f^2} + A_3 f^2 \quad \left\{\text{unit: } \frac{\text{dB}}{\text{km}}\right\}, \tag{3.6}$$

with the parameters:

$$f_1 = 0.78\left(\frac{S}{35}\right)^{1/2} e^{\left(\frac{T}{26}\right)} \ \{\text{unit: kHz}\}, \tag{3.7}$$

$$f_2 = 42e^{\left(\frac{T}{17}\right)} \text{ {unit: kHz}}, \tag{3.8}$$

$$A_1 = 0.106e^{\left(\frac{pH-0.8}{0.56}\right)} \left\{\text{unit: } \frac{\text{dB}}{\text{km} \cdot \text{kHz}}\right\}, \tag{3.9}$$

$$A_2 = 0.52\left(1 + \frac{T}{43}\right)\left(\frac{S}{35}\right)e^{\left(\frac{-z}{6}\right)} \left\{\text{unit: } \frac{\text{dB}}{\text{km} \cdot \text{kHz}}\right\}, \tag{3.10}$$

$$A_3 = 0.00049f^2e^{\left(\frac{-T}{27}+\frac{Z}{17}\right)} \left\{\text{unit: } \frac{\text{dB}}{\text{km} \cdot \text{kHz}^2}\right\}, \tag{3.11}$$

and f as the frequency of the sound wave [2].

3.5.3 Spreading Loss

Sound waves propagating in water will suffer a spreading loss related to the distance traveled. The loss is simplified to either be **Cylindrical Spreading** or **Spherical Spreading** depending on the situation. Sound waves in water can be assumed to be spherical until the wave reaches either the seafloor or surface; then, it can be approximated to be cylindrical [20]. This is how the spreading loss is constructed for this simulation.

Cylindrical Spreading

Intensity loss I for cylindrical spreading is determined by the radius r as,

$$I = I_0\left(\frac{r_0}{r}\right) \left\{\text{unit: } \frac{\text{W}}{\text{m}^2}\right\}. \tag{3.12}$$

r_0 is the reference radius with I_0 as its acoustical source level [26]. Then the spreading loss SL is calculated as:

$$SL = -10\log_{10}\left(\frac{I}{I_0}\right) = -10\log_{10}\left(\frac{r_0}{r}\right) \text{ {unit: dB}} [20]. \tag{3.13}$$

Spherical Spreading

Intensity loss I for spherical spreading is determined by the radius r as,

$$I = I_0\left(\frac{r_0^2}{r^2}\right) \left\{\text{unit: } \frac{\text{W}}{\text{m}^2}\right\}. \tag{3.14}$$

r_0 is the reference radius with I_0 as its acoustical source level [26]. Then the spreading loss SL is calculated by:

$$SL = -10\log_{10}\left(\frac{I}{I_0}\right) = -10\log_{10}\left(\frac{r_0^2}{r^2}\right) = -20\log_{10}\left(\frac{r_0}{r}\right) \text{ {unit: dB}} [20]. \tag{3.15}$$

3.5.4 Propagation time

The speed of sound in water is rather complex to determine. Factors like frequency, salinity, depth, temperature, and so forth contribute in various ways. However, an exact speed of sound is not necessary for this simulation. Therefore, the speed of sound was set to 1500 meters per second as a rough estimation [20]. This yields a propagation time T depending on distance M in meters as,

$$T = \frac{M}{1500} \quad \{\text{unit: s}\}. \tag{3.16}$$

3.5.5 Diffraction Loss

The reception of *PE* by the starboard and port hydrophones depends on the direction from which the signal is transmitted. If *PE* is sent from a ship facing the starboard side of the *AUV*, the port hydrophone will register a lower amplitude due to diffraction loss as port is shadowed. More advanced approaches to managing creeping waves have been examined in [27]. Due to their complexity, a simplified method has been developed based on *SAAB's* experiences. Considering the unit circle from the perspective of a side hydrophone, where 0 degrees is to the right of it, the signal decreases linearly from two dB to five dB between 180 - 270 degrees, and vice versa from 270 to 360 degrees, see figure 2.1.

3.5.6 Amplification & Hardware delay

PE with noise received by the rear hydrophone is amplified by 10 dB and then emitted as *PA* on top of the *AUV*. Due to the existing hardware delay of the *AUV*, *PA* will be emitted 13 ms seconds after *PE* was registered by the rear hydrophone.

3.5.7 Noise

The recordings are subject to normally distributed noise. The side hydrophones are exposed to higher noise levels than the rear hydrophones, as they are more susceptible to noise interference. Furthermore, the active response *PA* will also include noise when emitted.

3.5.8 Signals Generated by the Simulation

Figure 3.2 depicts the signals received by the three hydrophones in the simulation environment. The pulse marked with a *1* at the rear hydrophone is emitted by the fictitious ship called *PE*. Pulse *2* represents the active response signal *PA*, and box *3* indicates the moments when *PE* and *PA* overlap. For port, *4* highlights *PE*, *5* shows the presence of *PA*, and *6* represents the instances where *PE* and *PA* mix. The same applies to the starboard side.

Due to propagation time, the reception of *PE* occurs later at the port and starboard hydrophones compared to the rear. Notably, the amplitude of *PE* and *PA*

exhibits a reversed pattern for the port and starboard sides compared to the rear. This reversal is attributed to the closer positioning of the *PA* emitter to the port and starboard sides.

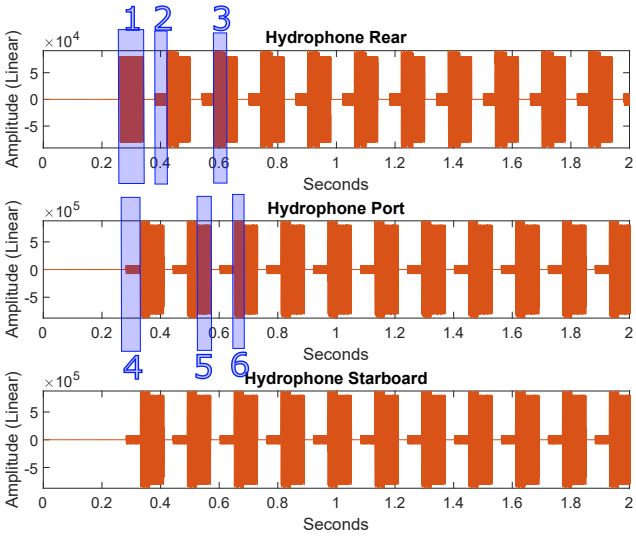


Figure 3.2: *Received signals by hydrophones in the simulation.*

3.6 Acoustic Measurement Data

The methods that exhibited the most promising results were subsequently evaluated based on real-world data. In this data, *PE* and *PA* are present, and the recording of all three hydrophones is available. The sampling rate used is 204800 Hz. However, all limitations in section 3.1 can not be guaranteed. For instance, reflections shown in figure 4.1 and 4.2 will exist due to the fundamental properties of sound in water.

Figure 3.3 illustrates an actual recording obtained by the *AUV*, featuring the simultaneous presence of both *PE* and *PA*. Boxes *1* and *3* highlight the active response *PA*, as shown in the figure, where the signal exhibits regular repetition across all three hydrophones. Boxes *2* and *4* highlights the received signal *PE*, which also exhibits repetitive patterns but with different intervals compared to *PA*. The objective is to eliminate or, at the very least, reduce the presence of *PA* while preserving the *PE* signal.

Noteworthy, *PA* for this measured data is not an active response to *PE*. Instead, *PA* is a pulse emitted by the *AUV* detached from what is received. In this measurement, *PA* and *PE* represent two distinct types of signals.

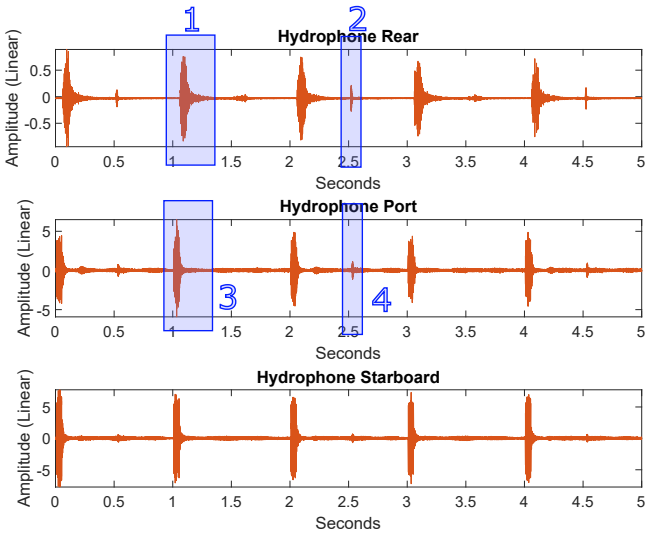


Figure 3.3: *Real data recorded by the hydrophones of the AUV*

4

Theory

4.1 General

This section describes the phenomenon of reflections of sound in water and defines Cross-Correlation, which is used for almost all algorithms. Reflections have not been implemented in the simulation environment; nonetheless, this phenomenon is of great importance for the acoustic measurement data and will be discussed in chapter 6.

4.1.1 Reflections of Sound Waves Underwater

How sound waves, i.e., sonar, propagate and behave in water depends on several factors. Sound waves are not only reflected by objects but also by seafloor and sea surface, depicted in figure 4.1.

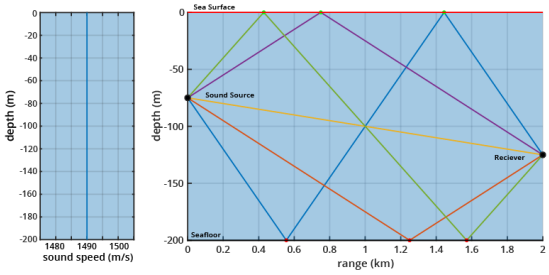


Figure 4.1: *Reflection of a sound source beneath the sea surface. Used with permission, © Discovery of Sound in the Sea (dosits.org), University of Rhode Island. 2020 [25].*

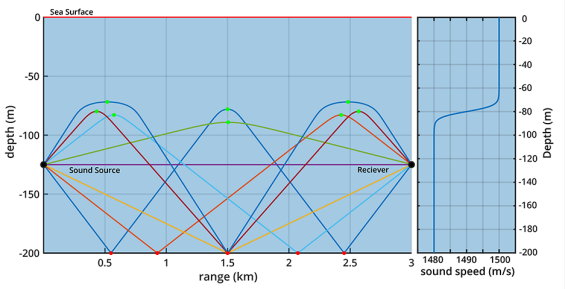


Figure 4.2: *Reflection of a sound source beneath the sea surface with temperature differences. Used with permission, © Discovery of Sound in the Sea (dosits.org), University of Rhode Island. 2020 [25].*

Sound waves also reflect at parts of the ocean with different temperatures; hence, the most common example is during the summer when the sun has heated the top layer of the water. Then a reflection occurs further down where the shift of temperature occurs; this is portrayed in figure 4.2. The sound speed also depends on temperature, which is depicted to the right in figure 4.2 [20].

4.1.2 Cross-Correlation

Cross-Correlation is a commonly used technique for measuring the level of correlation between two series. With two series $x(i)$ and $y(i)$ where i = 0,1,2...N-1, the degree of correlation $R_{xy}(d)$ at delay d is calculated as,

$$R_{xy}(d) = \frac{\sum_i((x(i) - E\{x\}) * (y(i-d) - E\{y\}))}{\sqrt{\sum_i(x(i) - E\{x\})^2}\sqrt{\sum_i(y(i-d) - E\{y\})^2}} \text{ [18].} \tag{4.1}$$

4.2 Blind Source Separation - BSS

Blind Source Separation is a big group of various algorithms to separate a mixture of multiple components into single ones blindly. The most common problem, referred to as the *Cocktail party problem*, is separating individual speeches in a noisy environment with several people speaking simultaneously. Instead of separating speech mixtures, the theory below will be focused on how to separate *PA* from *PE* received by the hydrophones of the *AUV*.

4.2.1 Overview

The general linear model of a *BSS* problem in the time-domain can be described as,

$$\mathbf{X(t)} = \mathbf{AS(t)} + \mathbf{N}. \tag{4.2}$$

Were, **X** denotes a vector of measured signals, **A** is an unknown mixing matrix, **S** is a vector of the unknown source signals and **N** is noise. Then the procedure is to estimate a demixing matrix by the measured signals, yielding an estimation of the source signals $\hat{\mathbf{S}}$ [21].

4.2.2 Pre-Processing

Principle Component Analysis - PCA

PCA can be used as a dimension reduction with less redundancy if the given dataset is a multivariate measurement of high dimension. This can be used as a prepossessing step before, for instance, an *ICA* [4].

The dataset is structured in a way that organizes each measurement as a row in $\mathbf{X}_{i,j}$. Here, i represents the number of individual measurements, while j denotes the number of samples for each measurement. The first thing is to calculate the row-wise mean as,

$$\bar{x}_j = \frac{1}{n}\sum_{i=1}^{n}\mathbf{X}_{i,j}.$$

Then a mean matrix is calculated as,

$$\bar{\mathbf{X}} = \begin{bmatrix}1\\ \vdots \\ 1\end{bmatrix}\bar{x}_j.$$

The covariance matrix **C** is next calculated with mean-subtracted **B** as,

$$\begin{aligned}\mathbf{B} &= \mathbf{X} - \bar{\mathbf{X}},\\ \mathbf{C} &= \frac{1}{n-1}\mathbf{B}^T\mathbf{B}.\end{aligned} \tag{4.3}$$

The first principal component is then given by $\mathbf{u}_1$ as,

$$\mathbf{u}_1 = \underset{\|\mathbf{u}_1\|=1}{\arg\max}\, \mathbf{u}_1^T\mathbf{B}^T\mathbf{B}\mathbf{u}_1\ . \tag{4.4}$$

Noticeably, if a *PCA* is performed on complex-valued data, then the hermitian transpose has to be used for equation (4.3) and (4.4) [4].

Whitening

An important pre-processing step before doing, for instance, an *ICA* is to whiten the data. A zero-mean vector $\mathbf{z} = (z, ..., z_n)^T$ is described as white when the covariance of **z** is the unit matrix.

For a vector $\mathbf{x}$ with n elements that is to be whitened, the linear transformation $\mathbf{V}$ is calculated so $\mathbf{z}$ is its white transformation, shown here,

$$\mathbf{z} = \mathbf{V}\mathbf{x}. \tag{4.5}$$

Firstly, the covariance $\mathbf{C}$ of $\mathbf{x}$ is calculated as done in equation (4.3) and denoted below as $\mathbf{C}_x$. Then let $\mathbf{E}$ be the matrix whose columns are the unit-norm eigenvectors of $\mathbf{C}_x$. While $\mathbf{D}$ is the diagonal matrix of the eigenvalues of $\mathbf{C}_x$. The whitening transform is then given by,

$$\mathbf{V} = \mathbf{D}^{\frac{-1}{2}}\mathbf{E}^T. \tag{4.6}$$

Further reading is refereed to [13].

4.2.3 Independent Component Analysis - Maximization of nongaussianity - ICA

Description

ICA is a method to discover independent components of a mixture. This is done by transforming the observed features into a new feature space, such that each new feature space is statistically independent. While the original feature space has as much as possible in common with the new [13].

There are several versions and techniques for estimating independent components of a mixture displayed in equation (4.2). The one described here is based on the Maximization of Nongaussianity using Negentropy [13].

Preprocess

Before the fixed-point algorithm is performed, the data has to be whitened; see section 4.2.2 [13]. Depending on the quality of the data, *PCA*, see section 4.2.2, may prove beneficial to the final result [4].

Fixed-point algorithm

$\mathbf{W}$ below is an estimation of $\mathbf{A}^{-1}$ in equation (4.2), which is used to retrieve the estimated sources $\hat{\mathbf{S}}$. Initially, $\mathbf{W}$ is decided by a guess. The fixed-point algorithm for separating independent sources by Maximization of Nongaussianity using Negentropy is,

$$\mathbf{W}^+ = E\{\mathbf{x}g(\mathbf{W}^T\mathbf{x})\} - E\{g'(\mathbf{W}^T\mathbf{x})\}\mathbf{W} \tag{4.7}$$

with a symmetric orthogonalization of $\mathbf{W}^+$ as,

$$\mathbf{W}_{new} = (\mathbf{W}^{+}(\mathbf{W}^{+})^{T})^{\frac{-1}{2}}\mathbf{W}^{+}. \tag{4.8}$$

g and g' are defined as:

$$g(y) = \tanh(ay)\,,$$
$$g'(y) = a(1 - \tanh^2(ay))\,.$$

$\mathbf{x}$ is the pre-processed data, where $1 \leq a \leq 2$ as a suitable constant proposed by *Aapo*. $\mathbf{W}$ in equation (4.7) is the estimated demixing matrix, initially chosen randomly. $\mathbf{W}$ is then updated with equation (4.7) to $\mathbf{W}^{+}$. Next, the symmetric orthogonalization of $\mathbf{W}^{+}$ is calculated with equation (4.8) yielding $\mathbf{W}_{new}$ [13]. Depending on the chosen convergence criterion β in equation (4.10), $\mathbf{W}_{new}$ is either determined to be the final demixing matrix, or $\mathbf{W}$ is set to $\mathbf{W}_{new}$ in equation (4.7) and the fixed-point algorithm is repeated [13]. If the convergence criterion is fulfilled, $\mathbf{W}_{new}$ is used as an estimation of $\mathbf{A}^{-1}$ in equation (4.2) to retrieve the estimated sources $\hat{\mathbf{S}}$.

As the algorithm progresses, the estimated demixing matrix will converge as follows as it improves,

$$\frac{1}{N}\sum_{i=1}^{N}\mathbf{Q}(i) \rightarrow 1\,,$$

with $\mathbf{Q}(i)$ defined as,

$$\mathbf{Q}(i) = |\mathbf{W}_{new}^{T} \cdot \mathbf{W}|\,. \tag{4.9}$$

N is the number of components searched and i is the row of the matrix. A threshold β is then chosen to determine when the algorithm has finished. The threshold is commonly constructed with equation (4.9) as,

$$\beta > \frac{1}{N}\sum_{i=1}^{N}(\mathbf{Q}(i) - 1). \tag{4.10}$$

The smaller value decided for β yields a more accurate solution; however, further calculations are required [13].

4.2.4 Complex Independent Component Analysis - CICA

Description

Simplified, the data is transformed with a Short-time Fourier transform, *STFT* with a complex-valued *ICA* applied to each frequency bin. Fundamentally, the goal and procedure are similar to the *ICA* in section 4.2.3.

A convolutive mixture (a mixture containing time delays) is considered instantaneous (no time delays present) in the frequency domain. This is because convolution in time becomes multiplication in frequency [5, 22].

Frequency Domain

An *STFT* is applied at the measured signals $\mathbf{X}$ in equation (4.2) as,

$$\hat{\mathbf{X}}(\omega, t_s) = \sum_t \mathrm{e}^{-i\omega t}\mathbf{X}(t)w(t - t_s). \tag{4.11}$$

The frequency ω and t_s as the window position are defined as:

$$\omega = 0,\ \frac{1}{N}2\pi,\ \frac{2}{N}2\pi,\ \ldots,\ \frac{N-1}{N}2\pi,$$
$$t_s = 0,\ \Delta T,\ 2\Delta T,\ \ldots,$$

where N is the number of points in the Fourier transform, w is the chosen window function, and ΔT is the shifting time for the window. Equation (4.11) applied at (4.2), without noise $\mathbf{N}$, yields,

$$\hat{\mathbf{X}}(\omega, t_s) = \hat{\mathbf{A}}(\omega)\hat{\mathbf{S}}(\omega, t_s). \tag{4.12}$$

Equation (4.12) illustrates how the measured signals are related to the sourced signals in the time-frequency domain [16]. Noteworthy, $\hat{\mathbf{A}}(\omega)$ is still a stationary mixing matrix, but regarding frequency [16, 29].

Preprocess of data

First of all, the data given by $\hat{\mathbf{X}}(\omega, t_s)$ has to be whitened described in section 4.2.2 [13]. Depending on the data, a *PCA* may prove beneficial, see section 4.2.2 [4].

Fixed-point algorithm

$\mathbf{W}$ below is an estimation $\hat{\mathbf{A}}(\omega)^{-1}$ in equation (4.12), which is used to retrieve the estimated sources $\hat{\mathbf{S}}(\omega, t_s)$. The complex-valued fixed-point algorithm for separating independent sources is defined as,

$$\mathbf{W}^+ = E\{\mathbf{x}(\mathbf{W}^H\mathbf{x})g(|\mathbf{W}^H\mathbf{x}|^2)\} - E\{g(|\mathbf{W}^H\mathbf{x}|^2) + |\mathbf{W}^H\mathbf{x}|^2 g'(|\mathbf{W}^H\mathbf{x}|^2)\}\mathbf{W}, \tag{4.13}$$

with a symmetric orthogonalization of $\mathbf{W}^+$ as,

$$\mathbf{W}_{new} = \mathbf{W}^+((\mathbf{W}^+)^H\mathbf{W}^+)^{\frac{-1}{2}}. \tag{4.14}$$

g and g' are defined as:

$$g(y) = \frac{1}{a + y},$$

$$g'(y) = -\frac{1}{(a+y)^2}.$$

$\mathbf{x}$ is the pre-processed data, a is a constant which *Aapo* proposes to be ≈ 0.1 and, H is the hermitian transpose. $\mathbf{W}$ in equation (4.13) is the estimated demixing matrix, initially chosen randomly. $\mathbf{W}$ is then updated with equation (4.13) to $\mathbf{W}^+$. Next, the symmetric orthogonalization of $\mathbf{W}^+$ is calculated with equation (4.14) yielding $\mathbf{W}_{new}$ [3, 14]. Depending on the chosen convergence criterion β, $\mathbf{W}_{new}$ is either determined to be the final demixing matrix, or $\mathbf{W}$ is set to $\mathbf{W}_{new}$ in equation (4.13) and the fixed-point algorithm is repeated. β is used and determined in the same way as for the noncomplex *ICA*; however, due to the complex numbers, the hermitian transpose is used [13]. When the criterion is fulfilled, $\mathbf{W}_{new}$ is used as an estimation of $\hat{\mathbf{A}}(\omega)^{-1}$ in equation (4.12) to retrieve $\hat{\mathbf{S}}(\omega, t_s)$.

4.2.5 Degenerative Unmixing Estimation Technique - DUET

Description

The *DUET* method creates *STFT* of two data sets. Then a ratio between these two is calculated to estimate the *Relative delay* and *Relative attenuation*; see sections below for definition. This ratio is presented in a histogram, with each source hopefully displaying its peak. Then the goal is to determine the time-frequency masks that separate each source from the data. This is done by determining the closets distance for each time-frequency point to the location of the peaks for the *Relative delay* and *Relative attenuation* [22].

Algorithm

Two sources depicted as in equation (4.2), without noise, with an *STFT* applied as in equation (4.11) can be portrayed as,

$$\begin{bmatrix} \hat{x}_1(\tau,\omega) \\ \hat{x}_2(\tau,\omega) \end{bmatrix} = \begin{bmatrix} 1 & \dots & 1 \\ a_1 e^{-i\omega\delta_1} & \dots & a_N e^{-i\omega\delta_N} \end{bmatrix} \begin{bmatrix} \hat{s_1}(\tau,\omega) \\ \vdots \\ \hat{s_N}(\tau,\omega) \end{bmatrix}. \tag{4.15}$$

With a as the relative attenuation factor between the two sources and δ as the corresponding delay [34]. For signals exhibiting W-disjoint orthogonality, where only one source is active for every (τ,ω), then equation (4.15) with j as the index of the active source at (τ,ω) can be described as,

$$\begin{bmatrix} \hat{x}_1(\tau,\omega) \\ \hat{x}_2(\tau,\omega) \end{bmatrix} = \begin{bmatrix} 1 \\ a_j e^{-i\omega\delta_j} \end{bmatrix} \hat{s_j}(\tau,\omega) \quad \forall(\tau,\omega) \in \Omega_j, \tag{4.16}$$

$$\Omega_j := \{(\tau,\omega) : \hat{s_j}(\tau,\omega) \neq 0\}. \tag{4.17}$$

For simplicity when creating the histogram, the relative attenuation a_j is transformed to the symmetric attenuation, α_j as,

$$\alpha_j := a_j - \frac{1}{a_j}.$$

The local symmetric attenuation and the delay can then be estimated as [29],

$$\tilde{\alpha}(\tau,\omega) := \left|\frac{\hat{x}_2(\tau,\omega)}{\hat{x}_1(\tau,\omega)}\right| - \left|\frac{\hat{x}_1(\tau,\omega)}{\hat{x}_2(\tau,\omega)}\right|, \tag{4.18}$$

$$\tilde{\delta}(\tau,\omega) := -\frac{1}{\omega}\angle\frac{\hat{x}_2(\tau,\omega)}{\hat{x}_1(\tau,\omega)}[29]. \tag{4.19}$$

Motivated by [29], the *MLE* gives the estimation as below, see [29, 34] for further reading,

$$\tilde{\alpha}_j = \frac{\iint_{(\tau,\omega)\epsilon\Omega_j} |\hat{x}_1(\tau,\omega)\hat{x}_2(\tau,\omega)|^p \omega^q \tilde{\alpha}(\tau,\omega) d\tau d\omega}{\iint_{(\tau,\omega)\epsilon\Omega_j} |\hat{x}_1(\tau,\omega)\hat{x}_2(\tau,\omega)|^p \omega^q d\tau d\omega} \tag{4.20}$$

$$\tilde{\delta}_j = \frac{\iint_{(\tau,\omega)\epsilon\Omega_j} |\hat{x}_1(\tau,\omega)\hat{x}_2(\tau,\omega)|^p \omega^q \tilde{\delta}(\tau,\omega) d\tau d\omega}{\iint_{(\tau,\omega)\epsilon\Omega_j} |\hat{x}_1(\tau,\omega)\hat{x}_2(\tau,\omega)|^p \omega^q d\tau d\omega}. \tag{4.21}$$

With p and q as scalars chosen by the user. For further explanation on how these are chosen see [29]. A histogram H is then constructed as,

$$H(\alpha,\tau) = \iint\limits_{(\tau,\Omega)\in I(\alpha,\delta)} |\hat{x}_1(\tau,\omega)\hat{x}_2(\tau,\omega)|^p \omega^q d\tau d\omega \tag{4.22}$$

$$I(\alpha,\delta) := \{(\tau,\omega) : |\hat{\alpha}(\tau,\omega) - \alpha| < \Delta_\alpha,\ |\hat{\delta}(\omega,\tau) - \delta| < \Delta_\delta\,. \tag{4.23}$$

The histogram peak centers are then desirably positioned at $\tilde{\alpha}_j, \tilde{\delta}_j$. These time-frequency peaks of the histogram are then assigned with,

$$J(\tau,\omega) := \arg\min_{\mathbf{k}} \frac{|\tilde{a}_k e^{-i\tilde{\delta}_k\omega}\hat{x}_1(\tau,\omega) - \hat{x}_2(\tau,\omega)|^2}{1+\tilde{a}_k^2}. \tag{4.24}$$

The symmetric attenuation is then transformed back to attenuation via,

$$\tilde{a}_j = \frac{\tilde{\alpha}_j + \sqrt{\tilde{\alpha}_j^2 + 4}}{2}. \tag{4.25}$$

Latly, the mixing parameter estimate $\tilde{M}$ are used to estimate the searched components $\tilde{\hat{s}}_j$ as,

$$\tilde{M}_j(\tau,\omega) := \begin{cases} 1 & J(\tau,\omega) = j \\ 0 & \text{otherwise} \end{cases} \tag{4.26}$$

$$\tilde{\hat{s}}_j(\tau,\omega) = \tilde{M}_j(\tau,\omega)\frac{\hat{x}_1(\tau,\omega) + \tilde{a}_j e^{-\tilde{\delta}_j\omega}\hat{x}_2(\tau,\omega)}{1+\tilde{a}_j^2} \quad [29,\ 34]. \tag{4.27}$$

Noteworthy, this solution does not calculate a demixing matrix $\mathbf{A}^{-1}$ as the *BSS* problem described in equation (4.2). Instead, the sources are estimated by masks and given by equation (4.27).

4.2.6 Additional Employed Methods

During the research and evaluation of *BSS* methods, the *Least Dependent Component Analysis Based on Mutual Information* [31] and the *AMUSE* [32] algorithms were implemented. Neither of the results for these methods will be presented because of the inferior results. The reasons for these unsatisfactory results are the same as for *ICA*; see the discussion in section 6.1.1.

4.2.7 Important Aspects

Signal Properties

BSS is to a large extent focused on the separation of a mixture of speech [21]. Vital properties of speech are:

- Each speech signal typically has a unique temporal structure over short time frames.
- Speech signals are quasi-stationary for small time durations ($\approx$ 10 ms), but non-stationary over longer periods.
- Speech typically has a fundamental frequency between 85 to 255 Hz. [22]

Independence - ICA

ICA algorithms are usually restrained by three requirements, these are:

- A maximum of one source is Gaussian distributed. An *ICA* algorithm cannot separate several Gaussian distributed components.
- The sources are independent of one another, not only uncorrelated.
- The matrix $\boldsymbol{A}$ in equation (4.2) is full-rank [13].

There are methods referred to as Dependent Component Analysis - *DCA* as a workaround dealing with dependent components [28]. Section 6.1.1 discusses why these are not further evaluated.

4.3 Adaptive Filtering

PA and *PE* are considered unknown, making it difficult to determine what to be adaptively filtered out. The idea is that the hydrophones at the sides of the *AUV*, positioned close to the emitter of *PA*, will predominately receive the signal *PA*. Compared to the rear hydrophone positioned further away from the *PA* emitter. Then this data can be used to determine what to be filtered out in the rear hydrophone.

The investigated adaptive filters will model *PA* as linear regression. The linear regression is defined by,

$$y(t) = \varphi^T(t)\theta(t) + e(t). \tag{4.28}$$

φ is the regression vector containing the sampled audio from one of the side hydrophones. Due to the layout of the *AUV*, see section 2.1, this data is assumed to have the characteristics of $PA \gg PE$. θ is the parameter vector, e is noise and y is the sampled signal of *PA* in the rear hydrophone.

$\hat{\theta}$ are then estimated by the adaptive filters described in the following sections. Figure 4.3 portrays the setup with an adaptive filter with the hydrophones of the *AUV*. $\hat{y}$ is then presumed to be somewhat similar to *PA*, yielding an estimation denoted $\hat{PE}$.

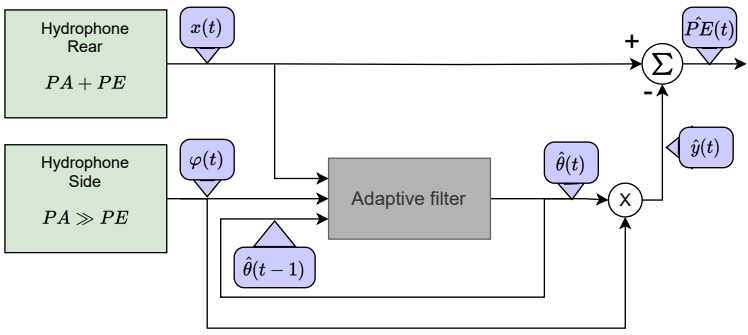


Figure 4.3: *Block diagram of an adaptive filter estimation of* $\hat{\theta}$

However, due to the layout of the *AUV*, see section 2.1, the hydrophones will capture signals with varying time delays. This has been accounted for by using Cross-Correlation, see section 4.1.2, yielding a system diagram with the estimation $\hat{t}$ in figure 4.4, with d denoting the delay.

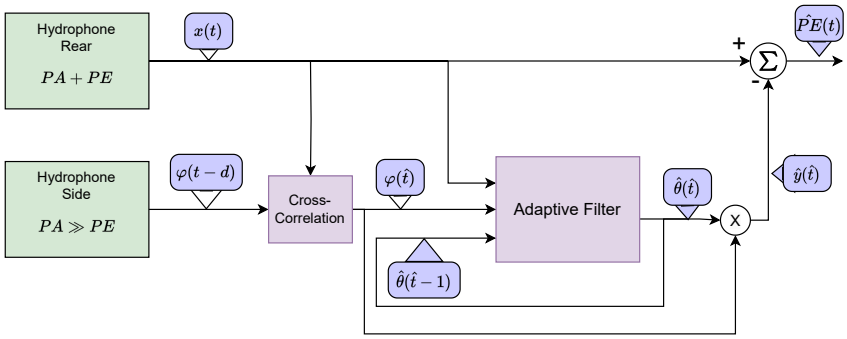


Figure 4.4: *Block diagram of an adaptive filter estimation of $\hat{\theta}$ and $\hat{t}$*

4.3.1 Leaky Least Mean Square - LLMS

Least mean Square - *LMS*, filters are a class of gradient descent algorithms of adaptive filters that estimates a transfer function. Simplified, an *LMS* algorithm minimizes the expected value of the squared prediction error. A regular *LMS* algorithm estimating the parameter $\hat{\theta}$ in equation (4.28) can be depicted as,

$$\hat{\theta} = \hat{\theta}(t-1) + \mu\varphi(t)(x(t) - \varphi^T(t)\hat{\theta}(t-1)). \tag{4.29}$$

With μ as step-size, φ as the noise source, x as the signal source [11, 33].

Due to the solution's requirements, a fast convergence rate is important. Therefore, its enhancement Leaky Least Mean Square - *LLMS* is further investigated instead of a regular *LMS*. This is due to its faster convergence rate [10, 30]. The estimation of $\hat{\theta}$ using *LLMS* is calculated as,

$$\hat{\theta} = (1-\gamma)\hat{\theta}(t-1) + \mu\varphi(t)(x(t) - \varphi^T(t)\hat{\theta}(t-1)). \tag{4.30}$$

Where $0 < \gamma \ll 1$ is the *leaky* coefficient chosen by the user [6]. This coefficient helps to regularize the solution towards zero [11]. The step-size parameter μ determines the degree of responsiveness of the filtration process to new values, with larger values leading to a more aggressive filtration approach [10].

4.3.2 Kalman

Another class of adaptive filters is based on the Kalman theory. Compared to the *LMS* filter, Kalman filters interpret the linear regression as a state space model. For the linear regression seen in equation (4.28), $\hat{\theta}$ is estimated with,

$$\hat{\theta} = \hat{\theta}(t-1) + K(t)(x(t) - \varphi^T(t)\hat{\theta}(t-1)), \tag{4.31}$$

$$K(t) = \frac{P(t-1)\varphi(t)}{R(t) + \varphi^T(t)P(t-1)\varphi(t)}, \tag{4.32}$$

$$P(t) = P(t-1) - \frac{P(t-1)\varphi(t)\varphi^T(t)P(t-1)}{R(t) + \varphi^T(t)P(t-1)\varphi(t)} + Q(t). \tag{4.33}$$

$R(t)$ is the covariance of $e(t)$ in equation (4.28). $Q(t)$ is the covariance of $w(t)$ which is defined as,

$$\theta(t+1) = \theta(t) + w(t). \tag{4.34}$$

These two covariances are frequently unknown, and the user is often restrained to guess [11, 23]. One interpretation of the parameter $Q(t)$ pertains to its role in determining the level of confidence attributed to the observations and subsequently influencing the aggressiveness of the filtration process [23].

4.3.3 Bayesian Information Criterion - BIC

The Bayesian Information Criterion - *BIC* estimates a suitable amount of parameters for $\hat{\theta}$. A candidate model M_k using *BIC* can be calculated as,

$$M_k = -2\ln L(\hat{\theta}_k|x) + k\ln(n), \tag{4.35}$$

with L as the maximum likelihood of the parameter values $\hat{\theta}$ at order k given the observed data x and n is the number of data points [24]. Further reading of *BIC* is directed to [24].

5
Results

This chapter will present the findings obtained from the methods investigated as described in chapter 4. The chapter is structured into three parts, with the first presenting the results of the *BSS* techniques and the second part portraying the results of the adaptive filters applied to simulated data. Finally, the third part presents the results of the adaptive filters applied to real-world data.

5.1 Blind Source Separation - BSS

To briefly summarize the results of *ICA*, the results did not prove themselves satisfactory. For the simulated data depicted in figure 5.1, *ICA* in combination with *PCA* and Cross-Correlation have not proven better than only using the recorded signal by the hydrophones as described in section 5.1.2. Therefore, these methods will not be tested for real-world data due to unsatisfactory results.

When propagation time was disregarded in the simulation environment, the results for *ICA* improved drastically; see figure 5.4a. Hence, the time delays between signals have been concluded to be the reason for the above unsatisfying results.

The method based on *DUET* was then investigated as a solution to the problems given by time delays due to its ability to utilize time delay as a source of information. However, the results were unsatisfactory; therefore, *DUET* was only applied to simulated data.

The method based on *CICA* was developed as a potential solution to address the challenges arising from time delays. This approach utilizes the frequency domain to convert the convolutive mixture into an instantaneous one (a linear

combination), as explained in Section 4.2.4. However, the obtained results were inconclusive, and this method was solely applied to simulated data.

5.1.1 Simulated Data

Figure 5.1 depicts the signals received by the three hydrophones. This data will be used for all *BSS* methods except when *ICA* was analyzed without time delays. The pulse marked with a *1* at the rear hydrophone is emitted by the fictitious ship called *PE*. Pulse *2* is the active response *PA*, and box *3* shows when *PE* and *PA* mix. For port, *4* highlights *PE*, *5* shows *PA* and *6* when *PE* and *PA* mix, with the same for starboard. Due to propagation time, *PE* is received later for port and starboard than the rear. The amplitude of *PE* and *PA* are somewhat the opposite of port and starboard compared to rear. This is because the emitter of *PA* is positioned closer to port and starboard. Figure 5.2 shows the ground truth of the pulses *PE* and *PA*. The goal of the *ICA* algorithm is to separate *PA* and *PE* from one another and resemble the two of them with ground truth as much as possible.

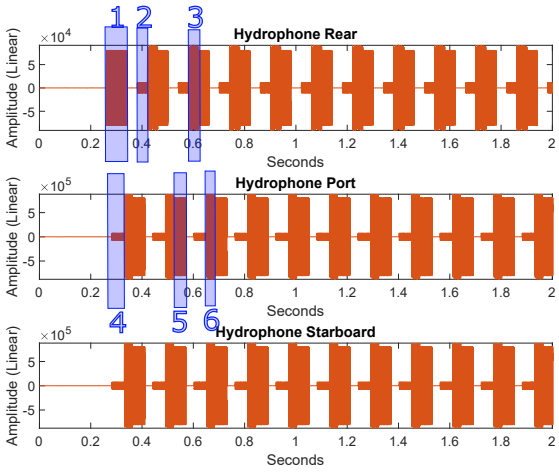


Figure 5.1: *Received signals by hydrophones. With time delay.*

5.1.2 Evaluation

This section only repeats how the evaluation will be done for the *BSS* methods below; this is further explained in section 3.4. Due to the layout of the *AUV*, the received signals for the side hydrophones will have the characteristics of $PA \gg PE$ and for the rear $PA \ll PE$. An essential requirement for separated signals is that they should resemble better to ground truth than what the recording of the hydrophones does. Namely, the separated component depicting *PA* needs to resemble more to ground truth of *PA* than what the recording of the side hydrophones

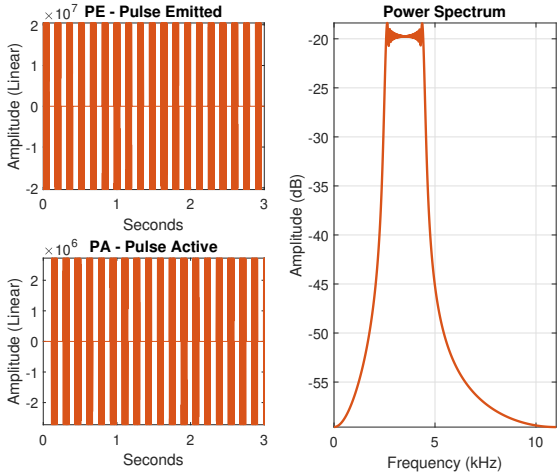


Figure 5.2: *Ground Truth of PE and PA in figure 5.1. The two signals share the same Power Spectrum when normalized.*

does. The same applies for *PE*, but then to the rear hydrophone. This can be seen in figure 5.1, as box *1* showing *PE* do have a significantly larger amplitude than *PA* shown by box *2* for the rear hydrophone. The opposite can be observed for the side hydrophones. As box *4* highlights *PE*, which is much lower in amplitude than *PA* shown in box *5*. To determine if a separation is better than using only the recordings of the hydrophones, *RMSE* is used, and the results are outlined in appendix A.

5.1.3 Independent Component Analysis - Maximization of nongaussianity - ICA

This section covers the result given by *ICA*. *ICA* is first applied to the data in figure 5.1 when time delays are present. Then to further distinguish the reason for these results, *ICA* is applied to data without time delays seen in figure 5.4.

ICA applied to simulated data with time delays

Time delays due to propagation time between the hydrophones have proven more challenging than anticipated; this is further reviewed in section 6.1.1. To overcome this problem, Cross-Correlation was used to shift *PE* in phase between the hydrophones. Then *PCA* was used to extract the correlation between the shifted data yielding a new feature space presumed similar to when no time delays were present. *ICA* was then performed yielding the components shown in figure 5.3.

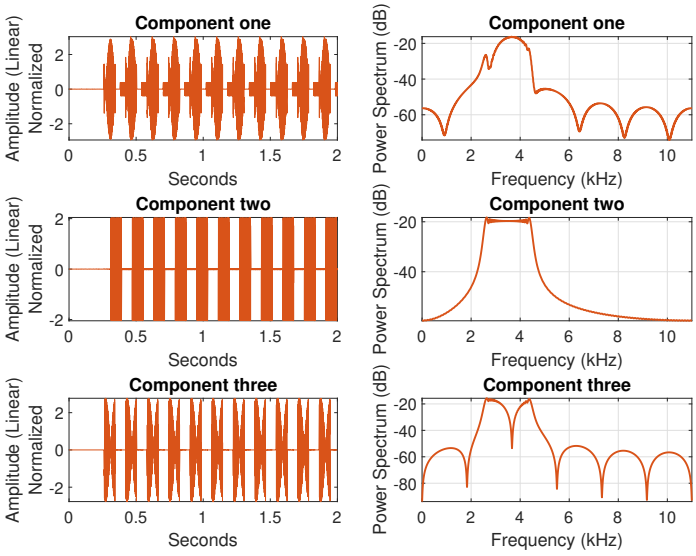
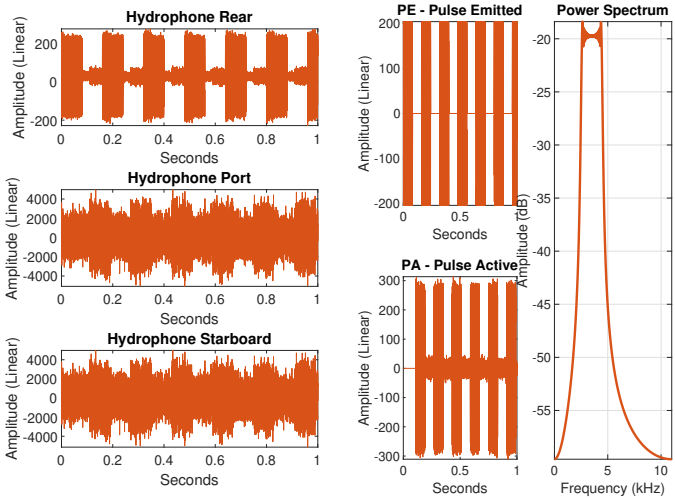


Figure 5.3: *Resulting components and their power spectrum by the data in figure 5.1 when Cross-correlation, PCA, and ICA are performed.*

The resulting *RMSE* of the components shown in figure 5.3 are outlined in table A.2. The goal is to archive a better *RMSE* than table A.1. However, neither component in 5.3 showed a lower *RMSE* than what the recorded data by the rear hydrophone did compared to ground truth in figure 5.2. Therefore, this method can be disregarded because it is more advantageous to use the recording by the rear hydrophone solely.

ICA applied to simulated data without time delays

When no propagation time is present, namely when there are no time delays between the received signals, the results improve drastically. A scenario like this is not realistic; however, attempts were made to emulate it using Cross Correlation; as seen in the preceding section, these attempts were unsuccessful. Figure 5.4a depicts the received signals with excessive noise when no time delays due to propagation time are present. The noise is presumed to be somewhat exaggerated compared to what is to be expected in reality. Figure 5.4b depicts the ground truth of these signals. The reason for *PA* consisting of noise is because the response emitted from the *AUV* also consists of a response to noise.



(a) *Received signals by hydrophones. With excessive noise but no time delay.*

(b) *Ground truth of PE and PA in figure 5.4a. PE and PA share the same Power Spectrum when normalized.*

Figure 5.4: *Simulated signals without propagation time.*

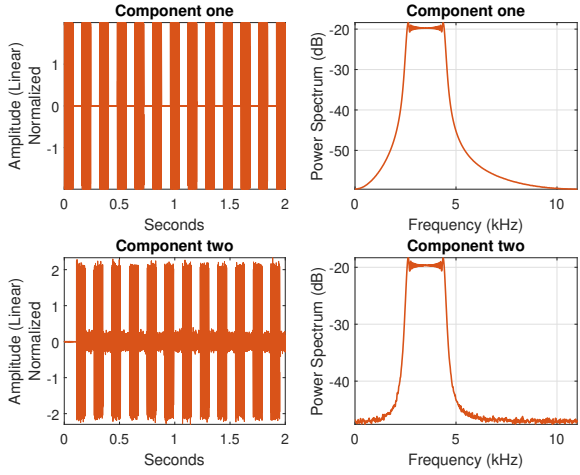


Figure 5.5: *Resulting components of PCA and ICA by the data in figure 5.4a.*

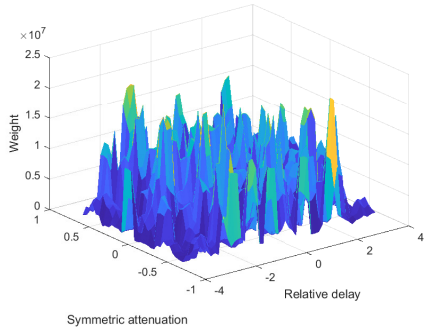
The resulting components of the data shown in figure 5.4a are displayed in figure 5.5. The *RMSE*, listed in table A.2, shows a much better estimation of both *PE*

and *PA* generated from the separation than solely using the recorded data by the hydrophones. Thus, when no time delays are present, the separation is rather successful, further discussed in section 6.1.1.

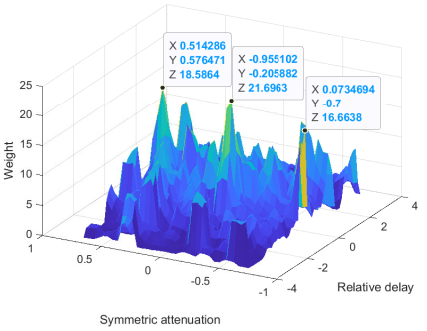
5.1.4 Degenerative Unmixing Estimation Technique - DUET

Compared to *ICA* in section 5.1.3, the *DUET* method does not calculate a demixing matrix; instead, it separates components by using masks. The first step is to create a histogram based on the components' relative delay and attenuation. The data displayed in figure 5.1 earlier is used to evaluate this method; however, due to the structure of *DUET*, only the rear and starboard hydrophones are used; see section 4.2.5 for the theory.

Figure 5.6a displays the histogram of the relative delay and attenuation constructed by the data mentioned above. When constructing a histogram like this, the histogram peak centers are desirably positioned clearly, and optimally the peaks are located at the searched sources, see equation (4.24). However, this is not true for 5.6a. Then Cross-Correlation is used to get *PE* in phase between the channels so the relative delay, see equation (4.21), of this source, is close to zero. The data is also normalized to counter the influence of *PA* received with an increased amplitude of 20dB. This results in the histogram in figure 5.6b.



(a) *Histogram of the data received by the hydrophones at the rear and starboard shown in figure 5.1.*



(b) *Histogram of the Cross-Correlated and normalized data received by the hydrophones at the rear and starboard shown in figure 5.1.*

Figure 5.6: *Histogram of the Relative delay and Symmetric attenuation.*

The choosing of the peaks portraying the estimated sources $\hat{\mathbf{s}}$ in equation (4.24) is a pretty arbitrary task. However, the three most apparent peaks are pointed out in figure 5.6b. The peak furthest to the left will be referred to as component one, the peak in the middle will be referred to as component two, and so forth from left to right.

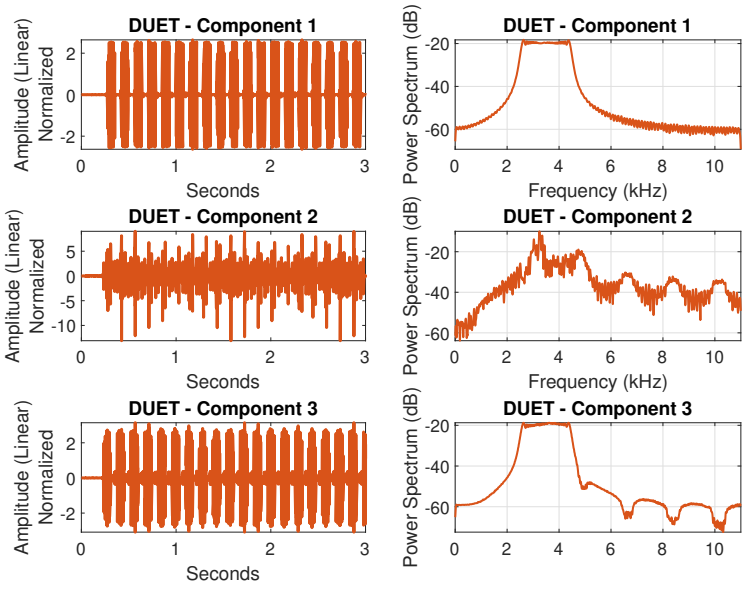


Figure 5.7: *Resulting components given by the chosen parameters shown in figure 5.6.*

The resulting components are displayed in figure 5.7. The components' resulting *RMSE* compared to ground truth is larger than ground truth compared to the signals received by the hydrophones; see table A.3. Hence, the separation gives inferior results estimating *PE* and *PA* compared to solely using the recorded signals.

5.1.5 Complex Independent Component Analysis - CICA

The Complex Independent Component Analysis *CICA* explained in section 4.2.4 will here be performed on the data shown in figure 5.1. The idea is to use an *STFT* to overcome the problem of time delays. This is because a convolutive mixture in the time domain transforms into an instantaneous in the frequency domain, namely, a linear combination. Figure 5.8 displays the three components of an *CICA* used with a complex-valued *PCA*.

Table A.4 shows the results of the *RMSE* of the components in figure 5.8 to ground truth of *PE* and *PA*. Neither shows a better result than only using the hydrophone recorded signals to estimate *PE* or *PA*. Thus, using solely the recordings of the hydrophones is more favorable than using the separated components to estimate *PE* or *PA*.

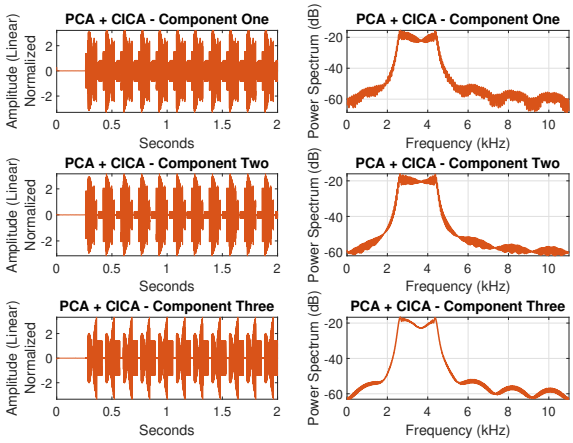


Figure 5.8: *Resulting components using CICA and PCA with an STFT of the data displayed in figure 5.1.*

5.2 Adaptive Filtering - Simulation

As earlier described, the received data of the rear hydrophone is considered to consist of $PE \gg PA$, and for port and starboard, the opposite $PE \ll PA$ is to be assumed. The idea of the adaptive filters is to use the side hydrophones as an estimation of *PA*, which is to be attenuated in the rear hydrophone.

Due to the promising results, it is not necessary to analyze these methods without the presence of time delays as done for *BSS*. Furthermore, the investigated adaptive filters will be evaluated using real data in the following section.

In this section, the evaluation measurements will not be repeated as in the case of *BSS*. For a detailed explanation of the evaluation methods, see section 3.4.1 for further clarification.

5.2.1 Simulated Data

Figure 5.9 displays the simulated data that will be used for the methods below. The data is almost identical to the one for *BSS* shown in figure 5.1. However, the position of the fictitious ship is not exactly the same, yielding a slight difference when the pulses reach the hydrophones. Therefore, the *RMSE* values concerning the ground truth cannot be directly compared to the adaptive filters with previous *BSS* methods.

In figure 5.9, the pulse marked with a *1* at the rear hydrophone is the pulse emitted by the fictitious ship called *PE*. Pulse *2* is the active response *PA*, and

box *3* shows when *PE* and *PA* mix. For port, *4* highlights *PE* and *5 PA*, with the same for starboard. Due to propagation time, *PE* is received later for port and starboard than the rear. The amplitude of *PE* and *PA* are somewhat the opposite of port and starboard compared to rear. This is because the emitter of *PA* is positioned closer to port and starboard. The goal is to decrease as much as possible of *PA* in the data received by the rear hydrophone; hence, the box denoted *2*.

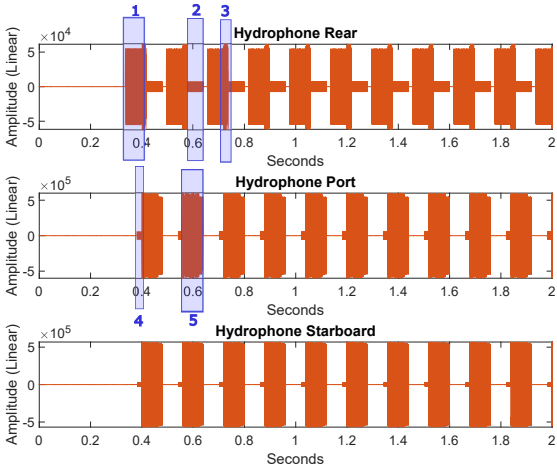


Figure 5.9: *Received signals by hydrophones. With time delay.*

5.2.2 Adaptive Filter based on Kalman

An adaptive filter based on the Kalman filter is used in this section. The theory is explained in section 4.3 with a visual chart of the setup displayed in figure 4.4. Figure 5.9 illustrates the data being investigated below.

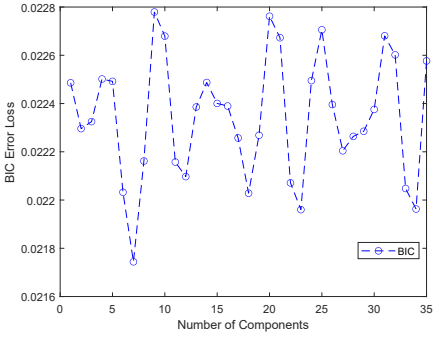


Figure 5.10: *BIC curve of the adaptive Kalman filter for the data in 5.9*

Figure 5.10 displays the *BIC* curve for the Kalman filter with the data in figure 5.9. A slightly lower error score is shown at seven components. Therefore, seven components are chosen for $\hat{\theta}$.

Figure 5.11 depicts the data registered by the rear hydrophone on the first row, as shown in figure 5.9. The second row shows the Kalman-filtered signal. The filtered signal is when the data received by the starboard is used as a noise generator of what is to be filtered out in the rear. The goal is the reduce *PA*, the box named *2*. Box *5* shows a significant decrease in amplitude, implying a reduced presence of *PA*. *PE* seem to be preserved by looking at box *1* and *4*. However, box *6*, which contains a mixture of *PA* and *PE*, might be excessively damped when comparing it to box *3* and *1*.

The values of $R(t)$ and $Q(t)$, which the user chooses, can be considered to be a ratio between each other; therefore, $R(t)$ was selected to be one. At the same time, $Q(t)$ has been tried for several values. A higher value of the parameter $Q(t)$ leads to increased damping of the *PA* signal in box *5* while also causing a more attenuated signal when *PA* and *PE* are mixed in container *6*. Conversely, lower values of $Q(t)$ produce opposite outcomes, resulting in less damping. After evaluating the results and considering the tradeoff, it has been determined that the most favorable compromise is achieved with a $Q(t)$ value of 0.01.

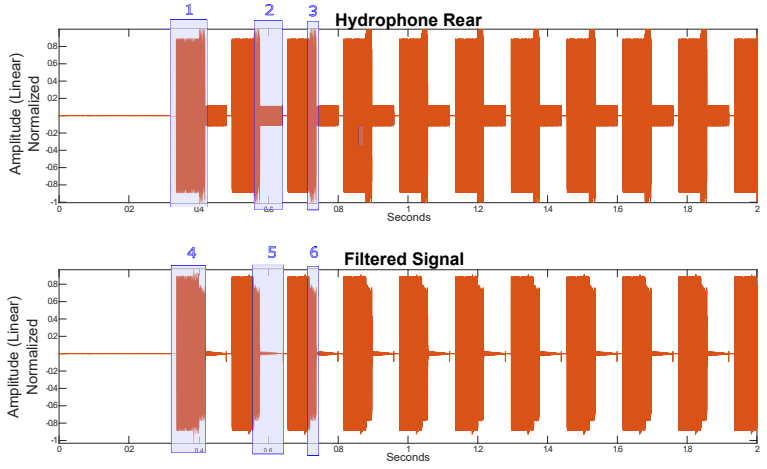
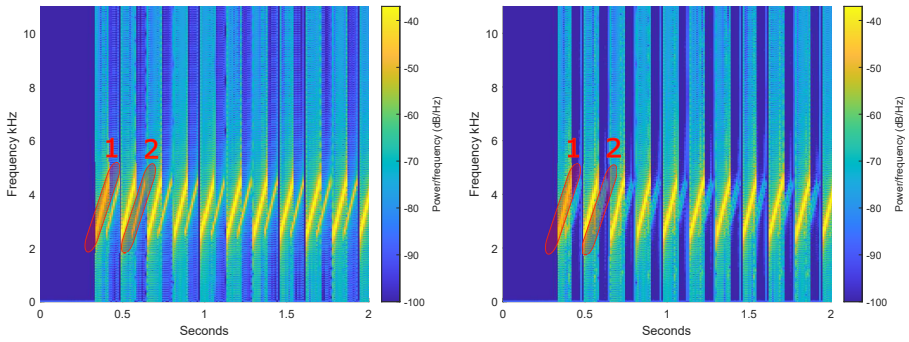


Figure 5.11: *The recorded data of the rear hydrophone in figure 5.9 and the filtered signal with Kalman*

Table A.6 shows the *RMSE* of the filtered signal in figure 5.11 to ground truth while table A.5 shows the *RMSE* for received signal by the rear hydrophone compared to ground truth. The filtered signal shows a lower *RMSE* than only using the unfiltered version of the rear hydrophone, implying successful filtering has occurred.



(a) *Spectrogram of received signal by the rear hydrophone in figure 5.11.*

(b) *Spectrogram of filtered signal in figure 5.11.*

Figure 5.12: *Spectrograms of the signals in figure 5.11.*

Figure 5.12 displays the spectrogram of the signals in figure 5.11. 5.12a depicts the data for the rear hydrophone with box *1* displaying *PE* and *2* showing *PA*. Optimally the filtered signal should preserve *1* but attenuate *2*. Figure 5.12b displays the filtered signal with box *1* as *PE* and box *2* as the remaining *PA*. *PE* has lost around three dB in the end where *PE* and *PA* has been mixed; however, *PA* has decreased by 25 dB.

5.2.3 Adaptive Filtering based on Leaky Least Mean Square - LLMS

The *BIC* curve shown in figure 5.10 for Kalman is almost identical to the one for *LLMS*. Therefore, this graph is not presented, and the same number of parameters in $\hat{\theta}$ are chosen, namely seven.

Figure 5.13 depicts the data registered by the rear hydrophone on the first row, as shown in figure 5.9. The second row shows the *LLMS* filtered signal. The filtered signal is when the data received by the starboard is used as a noise generator of what is to be filtered out in the rear. The goal is the reduce *PA*, the box named *2*. Box *5* shows a significant decrease in the amplitude, implying a reduced presence of *PA*. *PE* seem to be preserved by looking at box *1* and *4*. However, compared to the Kalman filtering in figure 5.11, the mixture of *PE* and *PA* is not as damped by looking at box *6* compared to *3* and *1*, which may be favorable.

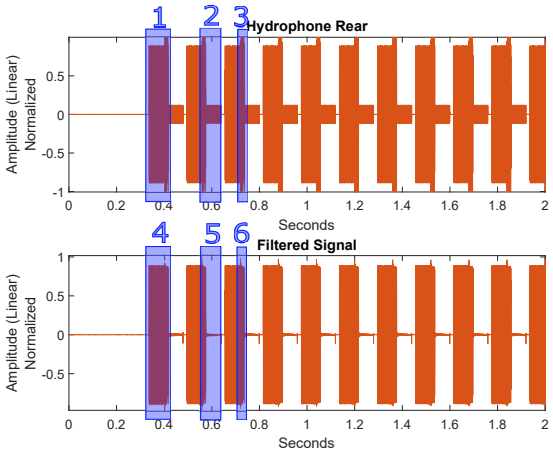
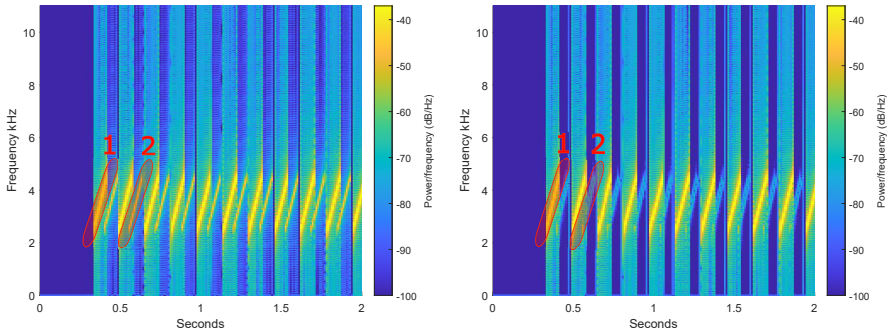


Figure 5.13: *The recorded data of the rear hydrophone in figure 5.9 and the filtered signal with LLMS.*

The leaky parameter γ in equation (4.30) was examined and set to be 0.0001. A higher value of γ yielded inferior results overall. Conversely, the step-size parameter μ proved to be more noteworthy. Increasing μ resulted in a more aggressive filtration approach, while lower values had the opposite effect. Similar to $Q(t)$ in the Kalman filter, μ can be regarded as a parameter influencing the responsiveness of the filtration process to new values and determining its aggressiveness [10]. However, a larger value of μ did not have any major impact on the energy when *PE* and *PA* were mixed, as observed in box *6*. Compared to $Q(t)$, which decreased the energy of the mixing to a larger extent when set to a high value. The value of μ was set to 0.1.

Table A.7 shows the *RMSE* of the filtered signal in figure 5.13 to ground truth while table A.5 shows the *RMSE* for received signal by the rear hydrophone compared to ground truth. The filtered signal shows a lower *RMSE* than only using the unfiltered version of the rear hydrophone, implying successful filtering has occurred. Moreover, the *RMSE* is lower for the *LLMS* filtering than the earlier Kalman.

Figure 5.14 displays the spectrogram of the two signals in figure 5.13. 5.14a depicts the data for the rear hydrophone with box *1* displaying *PE* and *2* showing *PA*. Optimally the filtered signal should preserve *1* but attenuate *2*. Figure 5.14b displays the filtered signal with box *1* as *PE* and box *2* as the remaining *PA*. *PE* has lost around one dB in the end where *PE* and *PA* have been mixed; however, *PA* has decreased by 35 dB. These numbers can be argued to be better than what Kalman provided earlier.



(a) *Spectrogram of received signal by the rear hydrophone in figure 5.13.*

(b) *Spectrogram of filtered signal in figure 5.13.*

Figure 5.14: *Spectrograms of the signals in figure 5.13.*

5.3 Adaptive Filtering - Acoustic Measurement Data

The adaptive filters applied to actual data have demonstrated favorable results, shown in this section.

5.3.1 Real Data

Figure 5.15 depicts a real recording done by the *AUV* with both *PE* and *PA* present. Box *1* and *3* highlight the active response *PA*, as seen in the figure, the signal repeats at regular intervals for all three hydrophones. Box *2* and *4* highlight the received signal *PE*, which also repeats itself but not with the same interval as for *PA*. The goal is to eliminate or at least reduce *PA* at the same time as *PE* is preserved.

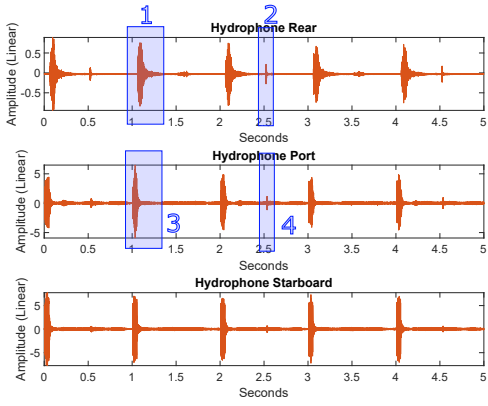
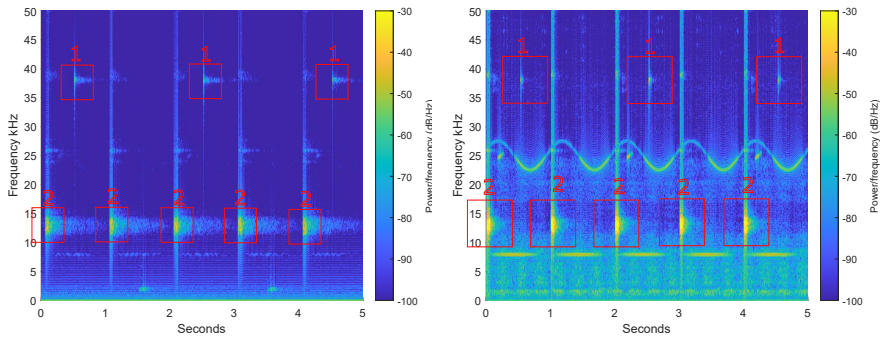


Figure 5.15: *Real data recorded by the hydrophones of the AUV*

Noteworthy, *PA* for this measured data is not an active response to *PE*. Instead, *PA* is a pulse emitted by the *AUV* detached from what is received. In this measurement, *PA* and *PE* represent two distinct types of signals. Meanwhile, *PA* is still emitted by the *AUV*, and *PE* is received pulse from another sonar.

Figure 5.16 displays the spectrograms for starboard and rear hydrophones seen in figure 5.15. The decision to use the starboard side for the adaptive filter was based on the observation that the amplitude of the *PA* is slightly more significant than the port side. However, since all signals have been normalized, port side has yielded nearly indistinguishable outcomes. Figure 5.16a displays the spectrogram for the rear hydrophone with the boxes marked *1* highlighting *PE*, while *2* encircles *PA*. The same signals are identified consistently for starboard in figure 5.16b. It can be seen that starboard hydrophone is subjected to higher noise levels, both from the water at lower frequencies and periodically occurring noise presumed to be of mechanical origin. Reflections of sound underwater can explain the tail following all pulses; see figure 4.1 and 4.2.



(a) *Spectrogram of received signal by the rear hydrophone in figure 5.15.*

(b) *Spectrogram of received signal by the starboard hydrophone in figure 5.15.*

Figure 5.16: *Spectrograms showing rear and starboard hydrophones in figure 5.15. Notable, PA is not an active response to PE.*

5.3.2 Kalman

This section will cover the results of the adaptive Kalman filter described in section 4.3 applied to the real-world data. The number of components for $\hat{\theta}$ is decided to be 15. This is due to the low error score shown in figure 5.17. It is hard to distinguish, but the error score is somewhat lower for 15 than 7.

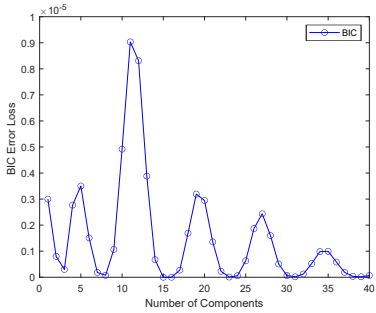


Figure 5.17: *BIC curve of the adaptive Kalman filter for the data in figure 5.15*

Figure 5.18 displays the data received by the rear hydrophone in figure 5.15 and the final filtered version. Box *1* and *3* show *PA*, which we want to reduce as much as possible. It can be observed that the filtered signal exhibits a decreased amplitude of *PA* compared to the recording done by the rear hydrophone. Meanwhile, boxes *2* and *4* highlight the preservation of *PE*, which is our desired outcome, and this may indeed be the case.

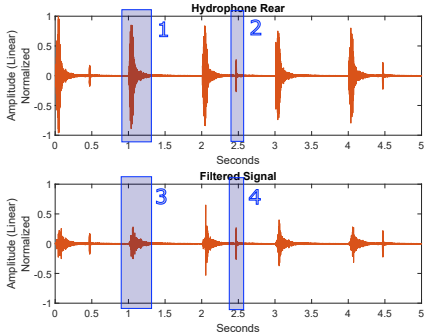
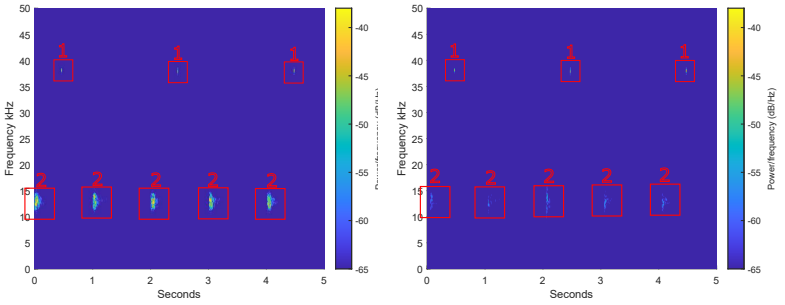


Figure 5.18: *The recorded data of the rear hydrophone in figure 5.15 and the filtered signal with Kalman*

The spectrograms of the signals portrayed in figure 5.18 are displayed in figure 5.19. The boxes denoted *1* highlights *PE* while *2* shows *PA*. The power scale on these spectrograms differs from previous ones because the frequencies require higher energy levels to become visible. Figure 5.19a presents the spectrogram of the signal received by the rear hydrophone, as depicted in figure 5.16a. Upon comparison with figure 5.19b, illustrating the spectrogram of the filtered counterpart, it appears as the energy of *PA* in box *2* has notably decreased. Additionally, it seems that *PE* in box *1* has not changed significantly. Figure 5.20 presents a magnified view of the two distinct signals depicted in figure 5.18, individually examined in the two paragraphs below.

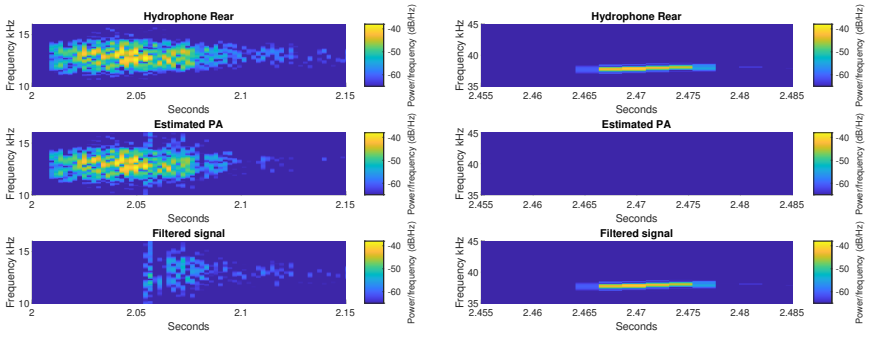


(a) Spectrogram of received signal by the rear hydrophone in figure 5.18.

(b) Spectrogram of filtered signal in figure 5.18.

Figure 5.19: *Spectrograms of the signals in figure 5.18.*

Figure 5.20a depicts the spectrogram of a single *PA* pulse received at the rear, the estimation of *PA* to be removed and the resultant filtered version. The energy of the *PA* pulse has reduced by approximately 35 dB across all frequencies at the onset of the signal. However, some residual energy persists toward the end of the pulse. This phenomenon can potentially be attributed to reflections, as illustrated in figure 4.1, and will be further investigated in Section 6.2.2. In general, there has been a significant dissipation in *PA*, which is a desirable behavior.



(a) Spectrogram of an active response to be filtered out. Box 2 in figure 5.19

(b) Spectrogram of an received pulse not be filtered out. Box 1 in figure 5.19

Figure 5.20: *Spectrograms of the active response PA and received signal PE.*

Meanwhile, *PE* is to be preserved as much as possible. Figure 5.20b depicts the spectrogram of a single *PE* pulse received at the rear, along with the estimation of *PA* to be eliminated and the resulting filtered version. The estimation of *PA* does not exhibit any noticeable energy to be removed in the presence of *PE*. The filtered version further supports this observation, as it retains the same energy level of *PE* as the recorded signal from the rear hydrophone. Such behavior is considered highly desirable.

5.3.3 Leaky Least Mean Square - LLMS

This section will present the findings of the implemented adaptive *LLMS* filter, as outlined in section 4.3 when applied to real-world data. The *BIC* graph for the Kalman filter shown in figure 5.17 is almost identical to the one for *LLMS*; therefore, the number of components in $\hat{\theta}$ are also decided to be 15.

Figure 5.21 presents the obtained data from the rear hydrophone, as depicted in Figure 5.15, along with the filtered version. Boxes *1* and *3* correspond to PA, which we aim to minimize. It is shown that the filtered signal exhibits a reduced amplitude of *PA* in comparison to the recorded signal obtained from the rear hydrophone. However, maybe not as much as the Kalman filter succeeded. Boxes *2* and *4* highlight *PE*, which is intended to be preserved, which seems to be the case.

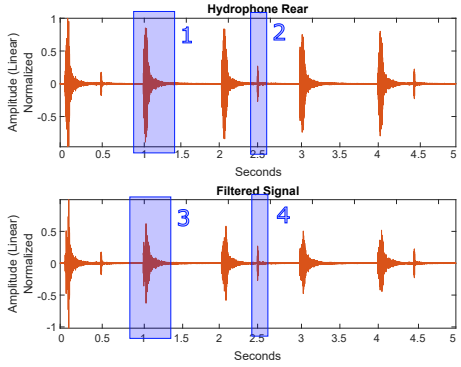
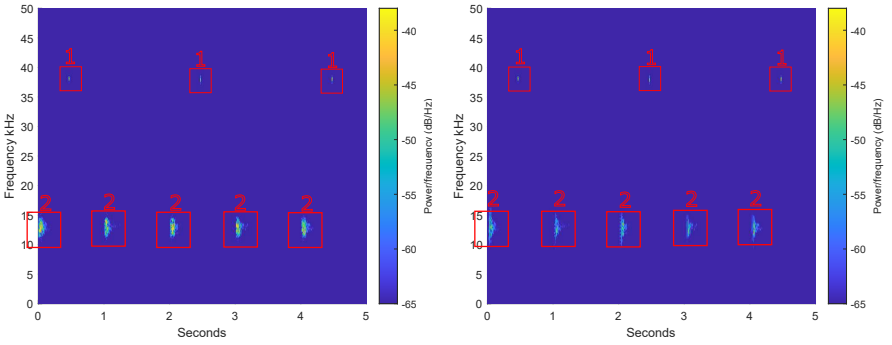


Figure 5.21: *The recorded data of the rear hydrophone in figure 5.15 and the filtered signal with LLMS*



(a) *Spectrogram of received signal by the rear hydrophone in figure 5.21.*

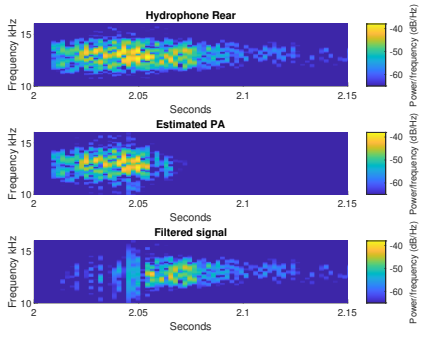
(b) *Spectrogram of filtered signal in figure 5.21.*

Figure 5.22: *Spectrograms of the signals in figure 5.21.*

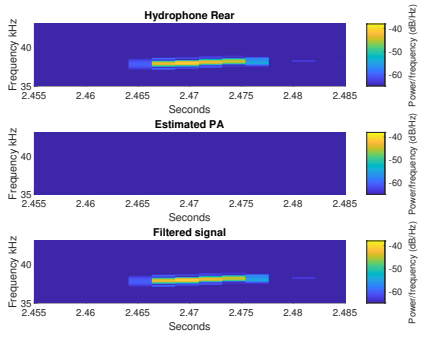
Figure 5.22 illustrates the spectrograms of the signals depicted in Figure 5.21. Box *1* highlights *PE*, while box *2* shows *PA*. Moreover, figure 5.22a displays the spectrogram of the signal received by the rear hydrophone, similar to figure 5.16a, but with a different amplitude scale. Upon comparison with figure 5.22b, which showcases the spectrogram of the filtered version, *PA* in box *2* appears to contain less energy. While *PE* in box *1* looks preserved. These two signals are individually examined in the following two paragraphs.

Figure 5.23a illustrates the spectrogram of a single pulse of *PA* captured at the rear. It also includes the estimated version of *PA* to be excluded and the resulting filtered version. The beginning half of *PA* is reduced by around 30 dB across all frequencies; however, the second part remains largely unaltered. One reason for this can be due to reflections of sound underwater portrayed in 4.1; this is further evaluated in section 6.2.2. Nonetheless, this can be seen as a poorer result than the earlier Kalman-filtered signal.

Furthermore, *PE* is to be preserved. Figure 5.23b presents the spectrogram of a single pulse of *PE* captured by the rear hydrophone, including the estimation of *PA* to be eliminated and the resulting filtered version. The estimation of *PA* does not display any energy to be removed where *PE* is present. The filtered signal supports this observation, as the energy is sustained when compared to the recorded signal by the rear hydrophone, which is a desirable behavior.



(a) *Spectrogram of an active response to be filtered out.*



(b) *Spectrogram of an received pulse not be filtered out*

Figure 5.23: *Spectrograms of one active response PA and one received signal PE.*

6

Discussion

6.1 Blind Source Separation - BSS

This section will discuss the three reviewed blind source estimation techniques for this master thesis.

6.1.1 Independent Component Analysis - Maximization of nongaussianity - ICA

When no time delays are present, the *ICA* offers a pretty robust and accurate estimation of *PE* and *PA*, see section 5.1.3. The mixing and demixing process can be explained by,

$$\mathbf{X(t)} = \mathbf{AS(t)} + \mathbf{N}. \tag{6.1}$$

Were **X** denotes a vector of measured signals, **A** is an unknown mixing matrix, **S** is a vector of the unknown source signals and **N** is noise. For a simplification of the problem, equation (6.1) can be described without noise, as below,

$$\begin{bmatrix} X_1(t) \\ X_2(t) \\ X_3(t) \end{bmatrix} = \begin{bmatrix} a_{1,1} & a_{1,2} & a_{1,3} \\ a_{2,1} & a_{2,2} & a_{2,3} \\ a_{3,1} & a_{3,2} & a_{3,3} \end{bmatrix} \cdot \begin{bmatrix} s_1(t) \\ s_2(t) \\ s_3(t) \end{bmatrix}. \tag{6.2}$$

The regular *ICA* then gives an estimation of the sources by calculating the estimation of $\hat{a}$ as,

$$\begin{bmatrix} \hat{a}_{1,1} & \hat{a}_{1,2} & \hat{a}_{1,3} \\ \hat{a}_{2,1} & \hat{a}_{2,2} & \hat{a}_{2,3} \\ \hat{a}_{3,1} & \hat{a}_{3,2} & \hat{a}_{3,3} \end{bmatrix}^{-1} \cdot \begin{bmatrix} X_1 \\ X_2 \\ X_3 \end{bmatrix} = \begin{bmatrix} \hat{s_1}(t) \\ \hat{s_2}(t) \\ \hat{s_3}(t) \end{bmatrix}. \tag{6.3}$$

The above equation, when no time delays are present, is usually referred to as an instantaneous mixture. However, when time delays are present, equation (6.2) is transformed to,

$$\begin{bmatrix} X_1 \\ X_2 \\ X_3 \end{bmatrix} = \begin{bmatrix} a_{1,1} & a_{1,2} & a_{1,3} \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} \cdot \begin{bmatrix} s_1(t-\tau_1) \\ s_2(t-\tau_2) \\ s_3(t-\tau_3) \end{bmatrix} + ...$$

$$... + \begin{bmatrix} 0 & 0 & 0 \\ a_{2,1} & a_{2,2} & a_{2,3} \\ 0 & 0 & 0 \end{bmatrix} \cdot \begin{bmatrix} s_1(t-\tau_4) \\ s_2(t-\tau_5) \\ s_3(t-\tau_6) \end{bmatrix} + \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ a_{3,1} & a_{3,2} & a_{3,3} \end{bmatrix} \cdot \begin{bmatrix} s_1(t-\tau_7) \\ s_2(t-\tau_8) \\ s_3(t-\tau_9) \end{bmatrix}. \quad (6.4)$$

Were τ represents the time delays for when the channels receive the different signals. Mixtures depending on τ are usually referred to as a convolutive mixture. Equation (6.4) is a time-varying mixing matrix (convolutive) and can not be rewritten as equation (6.2) (instantaneous). Trying to demix equation (6.4) as done in (6.3) has proven itself treacherous. The results have been unusable when time delays are present, even though the delay can be a maximum of 0.05 seconds. This can be explained by the characteristics of *Chirp* signals compared to speech for which *ICA* is commonly used. In section 4.2.7, three essential properties of speech are outlined, these are:

- Each speech signal typically has a unique temporal structure over short time frames.
- Speech signals are quasi-stationary for small time durations ($\approx$ 10 ms), but non-stationary over longer periods.
- Speech typically has a fundamental frequency between 85 to 255 Hz.

Neither of these characteristics can be applied for the used *Chirp* signal. Due to their fundamental characteristics, the *Chirp* signals can not be considered quasi-stationary for small time durations. The frequencies of the used *Chirp* signal is much greater than anticipated for speech and do not have a unique temporal structure over short time frames. This results in a mixture sensitive to time delays.

One solution to the problem portrayed in equation (6.4) is to use Cross-Correlation to make one source be in phase by all three hydrophones. Hence, creating an instantaneous mixture of a convolutive one. The idea is to make equation (6.4) more similar to (6.1) by using Cross-Correlation and *PCA*. The new feature space given by *PCA* would then result in a mixture more or less as equation (6.1); evidently, due to the results, this was not the case. This is because the convolutive mixture has still not transformed into an instantaneous one, with the unsatisfying results reflecting this.

There are a couple of requirements when using an *ICA*, with the most critical being: *The sources are independent of one another, not only uncorrelated.* One

could argue that *PE* and *PA* do not fulfill this requirement because *PA* is based on *PE*. Then other techniques, usually referred to as Dependent Components Analysis - *DCA* could be used. However, this was disregarded early in the development due to the positive results of *ICA* when no time delays were present. The good result showed that *PA* being based on *PE* did not impose a problem regarding this restriction. Furthermore, when using the frequency domain, problems of *PE* and *PA* containing the same frequency spectrum imposed issues, which are further reviewed below.

6.1.2 Degenerative Unmixing Estimation Technique - DUET

The idea of using the *DUET* method was to utilize the information given by the time delays because an essential part of this method is to construct a histogram based on sources' relative delay and attenuation. However, the results given by this method were not deemed satisfactory.

DUET is intended for speech. One important characteristic of speech is that people tend to speak with a slight variation of frequency between each other, or at least people do not speak with an exactly overlapping frequency. In our case, *PE* and *PA* consist of the exact same frequency. For several *STFT* windows, *PE* and *PA* will appear as the same signal. One requirement that has to be fulfilled for equation (4.16) to be true is that only one source can be active at every (τ, ω); however, this will not be true due to the overlapping frequency. One may think the *STFT* windows can be constructed to ensure this does not happen. This has been tried, but due to the proximity of *PE* and *PA*, this problem has been deemed impossible.

When researching this method, *DUET* has successfully separated mixtures of speech, music, and such, as long as the sources are somewhat separated in either frequency or enough in time. But, for a mixture of two identical sources adjacent to one another, the separation deteriorates.

6.1.3 Complex Independent Component Analysis - CICA

There are two essential benefits of using the frequency domain when separating sources. First, a convolutive mixture in the time domain (time delays) transforms into an instantaneous one in the frequency domain (linear combination). The second one is that echoes of sound waves are not seen as individual components; instead, they only provide a contribution in amplitude for the main signal in the frequency domain.

However, the results of the *CICA*, see section 5.1.5, are deemed inadequate. The reason for this can be traced back to the two properties listed above. The problem that the mixture is convolutive in the time domain is solved by using the *CICA*. Meanwhile, a new complication arises; hence, when using *CICA*, echoes of the signals are perceived as only a contribution in amplitude. *PE* and *PA*, we intend

to separate exhibits essentially identical frequency spectra, thereby appearing as indistinguishable signals in the frequency domain. *PE* and *PA* are closely situated in the time domain, so either one could potentially appear as an echo of the other. As for *DUET*, one could argue that the size of the *STFT* windows could be adjusted to solve this. However, this has been investigated, and *PE* and *PA* are too adjacent to one another to get this to work properly. Therefore, *CICA* is not considered a potential algorithm for the task investigated.

During the investigation of this method, it proved helpful in convolutive mixtures for speech and music.

6.2 Adaptive Filtering

The resulting estimation of *PE* given by the two adaptive filters proved superior to only using the recorded signal by the rear hydrophone for the simulated data. The filtered signal preserved the majority *PE* at the same time as *PA* was greatly attenuated. Which was the intended behavior.

Due to the satisfying results for the simulated data, the adaptive filters were evaluated using real-world data. The result was still deemed advantageous. However, some uncertainties arose due to a remaining portion of *PA*, further elaborated in section 6.2.2. In general, the adaptive filters proved to be useful.

6.2.1 Simulation

Both filters yielded desirable results for the simulated data. For the selected parameters, the *LLMS* algorithm successfully attenuated *PA* to a similar extent as the Kalman filter while simultaneously preserving a greater portion of *PE*, which is considered to be a good result. For the simulated data, *LLMS* is to be considered somewhat more advantageous than Kalman.

For the *LLMS*, the number of components in $\hat{\theta}$, the step-size μ and the leaky-coefficient γ are chosen by the user. The results remain relatively stable when altering the number of components for $\hat{\theta}$, as long as a reasonable quantity is selected. However, the computational time increases significantly once the number exceeds around 30. The selection of the leaky coefficient γ can be somewhat treacherous, as choosing a value that is not sufficiently small can lead to a highly distorted filtered version. Conversely, when γ is appropriately small, its impact on the filtering outcome is somewhat minimal. Moreover, μ is of greater significance when calibrating the filter. μ can be perceived as a parameter that controls the degree of aggressiveness exhibited by the filtering process. For the presented results, an increase in μ would lead to further attenuation of *PA*. However, this would come at the expense of excessively reducing the mixture of *PE* and *PA*, resulting in an undesirable reduction of *PE*. Furthermore, there exists a limitation on the maximum value that can be chosen for μ, as selecting an excessively large

value renders worthless results.

Regarding the adaptive filter based on Kalman, the parameters chosen by the user is the number of components in $\hat{\theta}$ and the value of Q. The same concept as for *LLMS* is true for Kalman regarding choosing the components for $\hat{\theta}$. Q can be interpreted as a parameter determining the level of confidence attributed to the observations and subsequently influencing the aggressiveness of the filtration process, similar to μ for *LLMS*. The results for the Kalman filters differ from the *LLMS*. The resulting filtration of the Kalman filter does attenuate the part when *PE* and *PA* are mixed somewhat excessively. This can be solved by lowering Q; however, at the cost of an increased presence of *PA*, which is undesirable. There is a trade-off for the Kalman filtration of how valuable it is to attenuate *PA* compared to the preservation of *PE*.

6.2.2 Acoustic Measurement Data

There is one major difference between the real data investigated compared to the simulated one. For the real-world data, *PA* is not directly a response to *PE*. Instead, *PA* is only a pulse emitted by the *AUV* regardless of the received pulse *PE*. This simplifies the process of attenuating *PA* while preserving *PE*. Furthermore, the adaptive filter based on Kalman showed better results than what *LLMS* did, which is the opposite of the simulation.

Due to the fact it is real-world data and ground truth is unknown, one important phenomenon needs to be considered. Sound waves under water reflect at the sea surface and seafloor, depicted in figure 4.1 and 4.2. These reflective waves dissipate energy, but if they are combined with non-reflective waves emitted later, they can amplify each other. In figure 5.22, the location exhibiting the highest energy within the recorded signal is situated at the midpoint of the pulse, approximately at 2.05 seconds. This occurrence can be attributed to a combination of reflective and non-reflective waves.

The adaptive filters use the side hydrophones as an estimation of *PA* to be attenuated in the rear. This is because the received signals at the side hydrophones are presumed to have the characteristics of $PA \gg PE$; however, this assumption will not be valid for all reflective waves of *PA*. Therefore, achieving a filtered signal without the presence of reflected sound waves is more arduous. This can explain why the filtered signal given by the *LLMS* seems to contain the later part of the *PA* pulse.

Meanwhile, the Kalman filtered signal appears to have attenuated the later part of the *PA* pulse more than what the *LLMS* filtration succeeded with; hence, the segment presumed to contain reflective and non-reflective sound waves. One possible explanation for this can be attributed to the parameter Q, which governs the aggressiveness of the filtration process. In the case of real data, a relatively high value is assigned to the parameter Q, thereby facilitating a more aggressive solu-

tion that effectively filters out the presumed reflective waves. The μ parameter of the *LLMS* acts similarly to Q, but when this number is increased further than what has been done, the results become unrecognizable. The exact reason for this is not fully comprehended, but selecting a larger value of μ than the used one leads to deteriorated results.

For this real-world data, when *PE* and *PA* does not directly mix, due to the fact that *PA* is unrelated to *PE*, an aggressive solution is more advantageous. Then, an adaptive filter based on Kalman has shown more favorable due to the possibility of easily increasing the parameter Q. Meanwhile, the results somewhat deteriorated when μ was increased for the *LLMS* to create further aggressive filtration. If *PA* would instead be a response of *PE* as the simulated data, then a less aggressive solution might not be as suitable due to the importance of preserving *PE*. Resulting in the *LLMS* algorithm being more advantageous, as shown for the simulated data.

However, it should be noted that the side hydrophones are subject to a significant level of ambient noise, particularly at five kHz and below. While a high pass filter can effectively address this issue for higher frequency signals, additional considerations are necessary for signals of lower frequencies. The active response must contain more energy than the ambient noise for the signals around that frequency to be distinguishable. If the active response is at five kHz or below and exhibits less energy than the ambient noise, the side hydrophones may not work properly as a generator of what is to be filtered in the rear hydrophone.

7

Conclusion

This chapter presents the conclusions derived from the research questions investigated in this master's thesis, based on the conducted work.

- **RQ1** - Can blind source separation effectively separate the active response while preserving the receiving signal, and what types of constraints may arise during the separation process?

The revised Blind Source Separation techniques can not separate the active response while preserving the receiving signal. This is due to the time delay between the two signals received by the three hydrophones. The only feasible approach, regarding the requirements of the *AUV*, is to address this issue using the frequency domain; nevertheless, it leads to the signals being perceived as a single entity due to their identical frequency spectra.

- **RQ2** - To what extent is it feasible to implement the reviewed blind source separation in the *AUV* for generating active responses?

Considering the unsatisfactory outcomes observed with the simulated data, it is not feasible to implement either of these techniques in the *AUV*.

- **RQ3** - Can adaptive filtering effectively attenuate the active response while preserving the receiving signal? Additionally, what constraints and challenges may arise?

The two adaptive filters investigated can attenuate the active response while preserving the receiving signal. The *LLMS* and Kalman filters exhibit individual advantages and disadvantages regarding attenuation, preservation, and aggressiveness. However, neither filter demonstrates a clear superiority over the other.

One significant constraint that requires consideration is that the active response is assumed to be prominent at the side hydrophones. However, as pointed out in the discussion, this assumption may not be the case if the active response consists of low-frequency components and lacks a sufficient energy level.

- **RQ4** - To what extent is it feasible to implement the reviewed adaptive filters in the *AUV* for generating active responses?

Considering the promising results, they can be feasible to implement in the *AUV*. However, the active response needs to be prominent enough at the side hydrophone to give adequate results.

- **RQ5** - What areas of interest could be investigated for future work and why?

A reasonably accurate estimation of the active response is crucial for the successful functioning of adaptive filters. One optimal solution would be to design the hardware within the *AUV* so that the emitted active response would be considered known. Alternatively, enhancing the data obtained from the side hydrophones could serve as a simpler approach to estimating the active response. As the side hydrophones are positioned less than half a meter apart, resulting in minimal time delays, employing Blind Source Separation in the time domain with Cross-Correlation might be feasible for extracting the active response from the side hydrophones. During early investigations of Blind Source Separation, this was presumed to be the case; however, it has not been extensively researched. This could be a potential solution to create an estimation of the active response to be filtered out in the rear.

Currently, Cross-Correlation is utilized to determine the delay between the registration of the active response at the side hydrophones and the rear. However, this delay will change when the *AUV* turns. Cross-Correlation could then be used to estimate this delay continuously. Alternatively, a combination of Cross-Correlation and a matching filter presents another viable solution to detect alterations in the time delay promptly.

For the researched adaptive filter, μ in *LLMS* and Q for Kalman are predetermined. An alternative approach worth exploring involves determining these parameters based on the received data to adjust the filtering aggressiveness accordingly. Preliminary attempts were made during the early stages of investigating adaptive filters; however, the results obtained were not satisfactory. Furthermore, this could be further investigated to archive better filtration.

Appendix

A

Results for Root Mean Square Error

A.1 Blind Source Separation

A.1.1 RMSE of the signals received by the hydrophones

Table A.1: *RMSE of the data received data by the hydrophones in figure 5.1 with ground truth shown in figure 5.2.*

Signals {Time domain} {Linear}	**RMSE**
Ground truth shown in figure 5.2 with the data received by the hydrophones in figure 5.1	
Ground truth of *PE* & Received signal Hydrophone Rear	0.131
Ground truth of *PA* & Received signal Hydrophone Starboard	0.116
Ground truth of *PA* & Received signal Hydrophone Port	0.081

A.1.2 RMSE Independent Component Analysis

Table A.2: *RMSE of the components given by ICA in section 5.1.3 with ground truth.*

Signals {Time domain} {Linear}	**RMSE**
Ground truth shown in figure 5.2 with the components in figure 5.3:	
Ground truth of *PE* & Component One	0.811
Ground truth of *PE* & Component Two	1.415
Ground truth of *PE* & Component Three	0.736
Ground truth of *PA* & Component One	1.419
Ground truth of *PA* & Component Two	0.014
Ground truth of *PA* & Component Three	1.417
Ground truth shown in figure 5.4b with the components in figure 5.5:	
Ground truth of *PE* & Component Two	0.001
Ground truth of *PA* & Component One	0.002

A.1.3 RMSE Degenerative Unmixing Estimation Technique

Table A.3: *RMSE of the components given by DUET in section 5.1.4 with ground truth.*

Signals {Time domain} {linear}	**RMSE**
Ground truth shown in figure 5.2 with the components in figure 5.7:	
Ground truth of *PE* & Component One	1.415
Ground truth of *PE* & Component Two	1.349
Ground truth of *PE* & Component Three	0.245
Ground truth of *PA* & Component One	0.516
Ground truth of *PA* & Component Two	1.405
Ground truth of *PA* & Component Three	1.419

A.1.4 RMSE Complex Independent Component Analysis

Table A.4: *RMSE of the components given by CICA in section 5.1.5 with ground truth.*

Signals {Time domain} {linear}	**RMSE**
Ground truth shown in figure 5.2 with the components in figure 5.8:	
Ground truth of *PE* & Component One	0.911
Ground truth of *PE* & Component Two	1.097
Ground truth of *PE* & Component Three	1.071
Ground truth of *PA* & Component One	1.340
Ground truth of *PA* & Component Two	1.302
Ground truth of *PA* & Component Three	1.111

A.2 Adaptive Filters

A.2.1 RMSE of the signals received by the hydrophones

Table A.5: *RMSE of the data received data by the rear hydrophone in figure 5.9*

Signals {Time domain} {Linear}	**RMSE**
Ground truth and the data received by the rear hydrophone in 5.9:	
Ground truth of *PE* & Hydrophone rear	0.093

A.2.2 RMSE Kalman

Table A.6: *RMSE of the signals given by the Kalman filter in section 5.2.2 with ground truth.*

Signals {Time domain} {Linear}	**RMSE**
Ground truth with the filtered signal shown in figure 5.11:	
Ground truth of *PE* & Filtered Signal	0.050

A.2.3 RMSE Leaky Least Mean Square

Table A.7: *RMSE of the signals given by the LLMS filter in section 5.2.3 with ground truth.*

Signals {Time domain} {Linear}	**RMSE**
Ground truth with the filtered signal shown in figure 5.13:	
Ground truth of *PE* & Filtered Signal	0.042

Bibliography

[1] A.Cederholm and M.Jönsson. Self-noise cancellation methods applied to acoustic underwater sensors. Oct 2008. Report number: FOI-R–2573–SE.

[2] M.A. Ainslie and J.G. McColm. A simplified formula for viscous and chemical absorption in sea water. *The Journal of the Acoustical Society of America*, 103-6:1671–1672, 1997. doi: 10.1121/1.421258.

[3] E. Bingham and A. Hyvärinen. Ica of complex valued signals: a fast and robust deflationary algorithm. In *Proceedings of the IEEE-INNS-ENNS International Joint Conference on Neural Networks. IJCNN 2000. Neural Computing: New Challenges and Perspectives for the New Millennium*, pages 357–362 vol.3, 2000. doi: 10.1109/IJCNN.2000.861330.

[4] S.L Brunton and J.N Kutz. *Data-Driven Science and Engineering: Machine Learning, Dynamical Systems, and Control*. Cambridge University Press, USA, 1st edition, Feb 2019. doi: 10.1017/9781108380690.

[5] J. Cho, J. Choi, and C.D. Yoo. Underdetermined convolutive blind source separation using a novel mixing matrix estimation and mmse-based source estimation. In *2011 IEEE International Workshop on Machine Learning for Signal Processing*, pages 1–6, Sept 2011. doi: 10.1109/MLSP.2011.6064629.

[6] S. Dixit and D. Nagaria. Lms adaptive filters for noise cancellation: A review. *International Journal of Electrical and Computer Engineering (IJECE)*, 7:2520, 10 2017. doi: 10.11591/ijece.v7i5.pp2520-2529.

[7] F.H. Fisher and V.P. Simmons. Sound absorption in seawater. *The Journal of the Acoustical Society of America*, 62:558–564, 1977. doi: 10.1121/1.381574.

[8] R.E. Francois and G.R. Garrison. Sound absorption based on ocean measurements: Part 1:pure water and magnesium sulfate contributions. *The Journal of the Acoustical Society of America*, 72-3:896–907, 1982. doi: 10.1121/1.388170.

[9] R.E. Francois and G.R. Garrison. Sound absorption based on ocean measurements: Part 2:boric acid contribution and equation for total absorption. *The Journal of the Acoustical Society of America*, 72-6:1879–1890, 1982. doi: 10.1121/1.388673.

[10] D.K. Gupta, V.K. Gupta, M. Chandra, A N. Mishra, and P.K. Srivastava. Hardware co-simulation of adaptive noise cancellation system using lms and leaky lms algorithms. In *2019 4th International Conference on Internet of Things: Smart Innovation and Usages (IoT-SIU)*, pages 1–6, 2019. doi: 10.1109/IoT-SIU.2019.8777658.

[11] F. Gustafsson, L. Ljung, and M. Millnert. *Signal Processing*. Studentlitteratur, Lund, 2010.

[12] T. O. Hodson. Root-mean-square error (rmse) or mean absolute error (mae): when to use them or not. *Geoscientific Model Development*, 15(14):5481–5487, 2022. doi: 10.5194/gmd-15-5481-2022.

[13] A. Hyvärinen, J. Karhunen, and E.Oja. *Independent Component Analysis*. John Wiley & Sons, Hoboken, New Jersey, USA, March 2001.

[14] A. Hyvärinen, P. Ramkumar, L.Parkkonen, and R.Hari. Independent component analysis of short-time fourier transforms for spontaneous eeg/meg analysis. *NeuroImage*, 49(1):257–271, Jan 2010. doi: 10.1016/j.neuroimage.2009.08.028.

[15] F. Jakobsen. The major inflow to the baltic sea during january 1993. *Journal of Marine Systems*, 6(3):227–240, 1995. doi: 10.1016/0924-7963(94)00025-7.

[16] S. Jayaraman, G. Sitaraman, and R. Seshadri. Blind source separation of acoustic mixtures using time-frequency domain independent component analysis. In *Proceedings of the 9th International Conference on Neural Information Processing, 2002. ICONIP '02.*, pages 1383–1387 vol.3, 2002. doi: 10.1109/ICONIP.2002.1202847.

[17] S. M. Kay. *Fundamentals of Statistical Signal Processing, VOL.1: Estimation Theory*. Prentice Hall, Hoboken, New Jersey, USA, 1997.

[18] A.F. Kohn. Autocorrelation and cross-correlation methods. In Metin Akay, editor, *Wiley Encyclopedia of Biomedical Engineering*, pages 260–283. John Wiley & Sons, Hoboken, New Jersey, USA, 2006.

[19] S. Kurita, H. Saruwatari, S. Kajita, K. Takeda, and F. Itakura. Evaluation of blind signal separation method using directivity pattern under reverberant conditions. In *2000 IEEE International Conference on Acoustics, Speech, and Signal Processing. Proceedings (Cat. No.00CH37100)*, volume 5, pages 3140–3143 vol.5, 2000. doi: 10.1109/ICASSP.2000.861203.

[20] L.Bjørnø. *Applied Underwater Acoustics*. Elsevier, Amsterdam Netherlands, 2017. doi: 10.1016/B978-0-12-811240-3.01001-8.

[21] P. Madhab, R. Rajib, B.Joyanta, and B. Milton. Blind source separation: A review and analysis. In *2013 International Conference Oriental COCOSDA held jointly with 2013 Conference on Asian Spoken Language Research and Evaluation (O-COCOSDA/CASLRE)*, pages 1–5, Nov 2013.

[22] S. Makino, T. Lee, and H. Sawada. *Blind Speech Separation*. Springer, Dodrecht, Netherlands, Sept 2007. doi: 10.1007/978-1-4020-6479-1.

[23] D.P. Mandic, S. Kanna, and A.G Constantinides. On the intrinsic relationship between the least mean square and kalman filters [lecture notes]. *IEEE Signal Processing Magazine*, 32(6):117–122, 2015. doi: 10.1109/MSP.2015.2461733.

[24] A.A. Neath. and J.E. Cavanaugh. The bayesian information criterion: background, derivation, and applications. *WIREs Computational Statistics*, 4(2): 199–203, 2012. doi: 10.1002/wics.199.

[25] DOSITS Discovery of Sound in the Sea. How does sound travel in shallow water? University of Rhode Island Graduate School of Oceanography, . URL `https://dosits.org/science/advanced-topics/shallow-water-propagation/`.

[26] DOSITS Discovery of Sound in the Sea. Cylindrical vs. spherical spreading. University of Rhode Island Graduate School of Oceanography, . URL `https://dosits.org/science/advanced-topics/cylindrical-vs-spherical-spreading/`.

[27] A.D. Pierce. *Acoustics - An introduction to its physical principles and applications*. Acoustical Society of America through the American Institute of Physics, 1989.

[28] L. Rui, L. Hongwei, and W.Fasong. Dependent component analysis: Concepts and main algorithms. *Journal of Computers*, 5(4), April 2010. doi: 10.4304/jcp.5.4.589-597.

[29] R. Scott. The duet blind source seperation algorithm*. In *Blind Speech Separation*, chapter 8. Springer, Dodrecht, Netherlands, Sept 2007. doi: 10.1007/978-1-4020-6479-1.

[30] R. Sinha, A. Choubey, and S.K. Mahto. A novel and efficient hybrid least mean square (hlms) adaptive algorithm for system identification. In *2015 SAI Intelligent Systems Conference (IntelliSys)*, pages 894–897, 2015. doi: 10.1109/IntelliSys.2015.7361249.

[31] H. Stögbauer, A. Kraskov, S. Astakhov, and P. Grassberger. Least dependent component analysis based on mutual information. *Physical review. E, Statistical, nonlinear, and soft matter physics*, 70, 01 2004. doi: 10.1103/PhysRevE.70.066123.

[32] L. Tong, V.C. Soon, Y.F. Huang, and R. Liu. Amuse: a new blind identification algorithm. In *IEEE International Symposium on Circuits and Systems*, pages 1784–1787 vol.3, 1990. doi: 10.1109/ISCAS.1990.111981.

[33] B. Widrow and S.D. Stearns. *Adaptive Signal Processing*. Prentice Hall, EngleWood Hills, NJ, 1985.

[34] O. Yilmaz and R. Scott. Blind separation of speech mixtures via time-frequency masking. *IEEE Transactions on Signal Processing*, 52(7):1830–1847, July 2004. doi: 10.1109/TSP.2004.828896.