Algorithms for propagation-aware underwater ranging and localization

by

Elizaveta Dubrovinskaya

in partial fulfillment of the requirements for the degree of Doctor in

Telematic Engineering

Universidad Carlos III de Madrid

Advisor: Paolo Casari Tutor: Albert Banchs

February 2021

Algorithms for propagation-aware underwater ranging and localization

Prepared by:
Elizaveta Dubrovinskaya, IMDEA Networks Institute, Universidad Carlos III de Madrid
contact: lisa.d@imdea.org

Under the advice of:
Paolo Casari, University of Trento, Italy
Albert Banchs Roca, Telematic Engineering Department, Universidad Carlos III de Madrid, and IMDEA Networks Institute

This work has been supported by:

Acknowledgements

The first page of this thesis is devoted to words of gratitude for those who made this Ph.D. journey possible. The first and foremost gratitude to my supervisor Dr. Paolo Casari knows no boundaries. He believed in me to allow me to work independently while providing all the possible support and was by my side when needed. His high standards both in research and communication have become my constant beacon in my personal and professional growth. His assistance in every aspect of research work was absolutely invaluable.

Another thank you should definitely go to my tutor Prof. Albert Banchs Roca for all his help with the paperwork from the very first days of my Ph.D.

I would also like to say thanks to Dr. Ivor Nissen for his introduction to the world of underwater acoustic localization research, and for always keeping an eye on my progress.

I am grateful to Dr. Diamant for welcoming and hosting me at Haifa University and enduring me during numerous sea trials. All the field experience I got during my trips to Israel is still one of the most unforgettable go through during my Ph.D.

Furthermore, I would like to say many thanks to Dr. Konstantin Kebkal and Evologics GmbH colleagues for hosting me at their premises in Berlin and for all the invaluable help in organizing lab and field experiments. Special thanks to Veronika Kebkal who pushed everyone to make it possible (and fun!)

Also, I would like to thank Dr. Fraser Dalgleish who hosted me at Florida Atlantic University and for sharing experimental data even after I left.

A huge bunch of thanks goes to all my colleagues at IMDEA Networks for all the coffee break talks, dinner drinks, board games battles, and barbeque parties. Special thanks go to the "ground-floor" team for all the evening gossip times and fun we had.

Last but not least thank you goes to my wonderful family. Thanks to my parents for teaching me to value knowledge and for giving me everything I need. And the most important thanks - to my husband Sergey Samodov, who was solving all my life problems, listened to my nagging during the hardest days saying, "this too shall pass", waiting for me to come back. I couldn't have done this journey without you.

Published and Submitted Content

This thesis is based on the following published papers:

[1] Elizaveta Dubrovinskaya, Ivor Nissen, Paolo Casari. On the accuracy of passive multipath-aided underwater range estimation. Published in *2016 IEEE Third Underwater Communications and Networking Conference (UComms)*, 18-20 September 2016, Lerici, Italy.

- This work is fully included and its content is reported in Chapter 3.
- The author's role in this work is focused on the design, implementation and simulations of the concepts proposed in the paper.

[2] Elizaveta Dubrovinskaya, Roee Diamant, Paolo Casari. Anchorless underwater acoustic localization. Published in *2017 14th Workshop on Positioning, Navigation and Communications (WPNC)*, 23-24 October 2017, Bremen, Germany. http://eprints.networks.imdea.org/1675/1/anchorless-underwater-acoustic.pdf

- This work is partially included and its content is partially reported in Chapter 4.
- The author's role in this work is focused on the design, implementation and simulations with regarding of the concepts proposed in the paper.

[3] Elizaveta Dubrovinskaya, Fraser Dalgleish, Bing Ouyang, Paolo Casari. Underwater LiDAR signal processing for enhanced detection and localization of marine life. Published in *2018 OCEANS-MTS/IEEE Kobe Techno-Oceans (OTO)*, May 2018 Kobe, Japan.

- This work is fully included and its content is reported in Appendix B.
- The author's role in this work is focused on the design, implementation, simulations and experimental data processing with regard to the concepts proposed in the paper.

[4] Dubrovinskaya, Elizaveta, Paolo Casari. Underwater Direction of Arrival Estimation using Wideband Arrays of Opportunity. Published in *OCEANS 2019-Marseille*, June 2019, France, Marseille.

- This work is partially included and its content is partially reported in Chapter 5.

- The author's role in this work is focused on the design, implementation and simulations of the concepts proposed in the paper.

[5] Dubrovinskaya, Elizaveta, Roee Diamant, Paolo Casari. Bathymetry-aided underwater acoustic localization using a single passive receiver. Published in *The Journal of the Acoustical Society of America, S.I. on Acoustic Localization*, 31 December 2019.

- This work is fully included and its content is reported in Chapter 4.
- The author's role in this work is focused on the design, implementation, simulations and experimental data processing with regarding of the concepts proposed in the paper.

[6] Elizaveta Dubrovinskaya, Veronika Kebkal, Oleksiy Kebkal, Konstantin Kebkal, Paolo Casari. Underwater Localization via Wideband Direction-of-Arrival Estimation Using Acoustic Arrays of Arbitrary Shape. Published in *Multidisciplinary Digital Publishing Institute Sensors*, July 2020.

- This work is fully included and its content is reported in Chapter 5.
- The author's role in this work is focused on the design, implementation and simulations with regarding of the concepts proposed in the paper.

[7] Paolo Casari, Filippo Campagnaro, **Elizaveta Dubrovinskaya**, Roberto Francescon, Amir Dagan, Shlomo Dahan, Michele Zorzi, Roee Diamant. ASUNA: A Topology Data Set for Underwater Network Emulation. Published in *IEEE Journal of Oceanic Engineering*, 18 March 2020.

- This work is partially included and its content is partially reported in Appendix A.
- The author's role in this work is focused on the experimental data processing and simulations/emulations with regard to the concepts proposed in the paper.

Additionally, the following paper is currently under submission

Elizaveta Dubrovinskaya, Amer Shaddad, Shlomo Dahan, Roee Diamant, Paolo Casari. Validation of Localization via Wideband Acoustic Arrays for Underwater Fauna Monitoring at Sea. Submitted to *IEEE/MTS OCEANS*, 2021.

- This work is included and its contents are partially reported in Chapter 6.
- The author's role in this work is focused on the design, implementation and extension of a detection, localization and tracking algorithm for underwater fauna, including experimental data processing.

Project Acknowledgements

Part of the research in this thesis has been supported by the EU H2020 program under project SYMBIOSIS (G.A. no. 773753).

Abstract

While oceans occupy most of our planet, their exploration and conservation are one of the crucial research problems of modern time. Underwater localization stands among the key issues on the way to the proper inspection and monitoring of this significant part of our world. In this thesis, we investigate and tackle different challenges related to underwater ranging and localization. In particular, we focus on algorithms that consider underwater acoustic channel properties. This group of algorithms utilizes additional information about the environment and its impact on acoustic signal propagation, in order to improve the accuracy of location estimates, or to achieve a reduced complexity, or a reduced amount of resources (e.g., anchor nodes) compared to traditional algorithms.

First, we tackle the problem of passive range estimation using the differences in the times of arrival of multipath replicas of a transmitted acoustic signal. This is a cost- and energy- effective algorithm that can be used for the localization of autonomous underwater vehicles (AUVs), and utilizes information about signal propagation. We study the accuracy of this method in the simplified case of constant sound speed profile (SSP) and compare it to a more realistic case with various non-constant SSP. We also propose an auxiliary quantity called effective sound speed. This quantity, when modeling acoustic propagation via ray models, takes into account the difference between rectilinear and non-rectilinear sound ray paths. According to our evaluation, this offers improved range estimation results with respect to standard algorithms that consider the actual value of the speed of sound.

We then propose an algorithm suitable for the non-invasive tracking of AUVs or vocalizing marine animals, using only a single receiver. This algorithm evaluates the underwater acoustic channel impulse response differences induced by a diverse sea bottom profile, and proposes a computationally- and energy-efficient solution for passive localization.

Finally, we propose another algorithm to solve the issue of 3D acoustic localization and tracking of marine fauna. To reach the expected degree of accuracy, more sensors are often required than are available in typical commercial off-the-shelf (COTS) phased arrays found, e.g., in ultra short baseline (USBL) systems. Direct combination of multiple COTS arrays may be constrained by array body elements, and lead to breaking the

optimal array element spacing, or the desired array layout. Thus, the application of state-of-the-art direction of arrival (DoA) estimation algorithms may not be possible. We propose a solution for passive 3D localization and tracking using a wideband acoustic array of arbitrary shape, and validate the algorithm in multiple experiments, involving both active and passive targets.

Table of Contents

List of Figures

List of Acronyms

AoA angle of arrival

AUV autonomous underwater vehicle

BER bit error rate

CCDF complementary cumulative distribution function

CDF cumulative distribution function

CIR channel impulse response

COTS commercial off-the-shelf

CTD conductivity-temperature-depth

DoA direction of arrival

ECEF earth-centered, earth-fixed

ETA estimated time of arrival

FAU Florida Atlantic University

FFT fast Fourier transform

GCS geographical coordinate system

GPIO general purpose input/output

GPS global positioning system

HBOI Harbour Branch Oceanographic Institute

LBL long baseline

LFM linear frequency modulation

LOS line of sight

LSE least squares estimator

MFP Matched field processing

MVDR minimum variance distortionless response

NMF Normalized matched filter

PDF probability density function

PER packet error rate

RF Radio Frequency

RMSE Root Mean Square Error

RMS Root Mean Square

RSSI received signal strength indicator

RSSI Received Signal Strength Indicator

RSS received signal strength

SBL short baseline

SDM Software defined modem

SNR Signal-to-Noise Ratio

SSP sound speed profile

TDoA time difference of arrival

TL transmission loss

ToA time of arrival

ToA Time-of-Arrival

ToF time of flight

ULA uniform linear array

UMSLI Unobtrusive Multistatic Serial LiDAR Imager

URA uniform rectangular array

USBL ultra short baseline

UWSN underwater sensor networks

WSN wireless sensor networks

Part I

Introduction

"Ocean is more ancient than the mountains, and freighted with the memories and the dreams of Time."

H.P. Lovecraft (1890 – 1937)

1 Introduction

The oceans occupy more than two-thirds of our planet, and one of today's priorities is learning how to preserve it. Collecting more data about the ocean can help us strike a balance between human activities and the conservation of marine life, or to prevent the devastating effects of earthquakes and floods. Besides, the modern world can carry various threats such as terrorist attacks in ports or on ships, which can be prevented by a timely detection and localization of the possible threat sources. Despite the importance of studying the oceans, at the moment we know only a portion of what we would like to know. The underwater environment is extremely difficult to study: the challenges related to human diving, the high complexity of special equipment for immersion, and, perhaps most importantly, the intricacy of underwater communications with which we could collect data from distributed sensors or navigate remote devices such as autonomous underwater vehicles (AUVs) still prevent humans from fully interacting with oceanic environments. Cabling at the bottom of the ocean is very costly and challenging, so when possible, wireless connections would be preferred.

Today we are so used to fast wireless Internet and easy satellite navigation, that it can be hard to imagine the world without it. However, these systems are built based on high-frequency radio waves, that hardly propagate in salty (hence electrically conductive) ocean waters. It strongly affects the way communication and navigation systems are built under the surface of the ocean. The creation of cheaper and faster deployment methods for various distributed underwater systems is limited by the need to localize underwater devices, as most algorithms employed in terrestrial radio networks cannot be directly applied in underwater networks.

Currently, most commercial off-the-shelf (COTS) underwater ranging and localization systems are based on acoustic signals. Depending on the frequency, these signals can propagate from meters to thousands of kilometers.

In this thesis, various algorithms for underwater ranging and localization are presented. First, we study the scenario with a single passive receiver and active source. We investigate the accuracy of a range estimation algorithm based on the time-of-arrival

differences between multipath replicas of a signal transmitted by an anchor node at a known location to a silent node that needs to localize itself. Then, we proceed with an algorithm to localize a source of acoustic signals using only a single receiver and the difference in the underwater channel induced by a non-homogeneous ocean bottom profile (or "bathymetry"). Further, we investigate another scenario with an active transmitter and array of hydrophones. This task was set within the framework of the SYMBIOSIS project. The SYMBIOSIS platform is an optical-acoustic system for biodiversity monitoring. In this context, we developed a system for the 3D acoustic localization of marine fauna based on acoustic arrays consisting of several sub-arrays combined together into an arbitrary layout.

1.1. Challenges

There exist multiple reasons for the necessity of ranging and localization underwater: some of these research challenges are listed below.

- *Slow and non-linear propagation.* The speed of sound in water is usually 1450-1550m/s [12], which is about five orders of magnitude slower than the radio wave propagation speed in air. Moreover, the speed of sound is changing over depth and distance. It creates a non-homogeneous environment where the sound propagates non-linearly. Measuring the local instantaneous speed of sound can be possible with the help of extra sensors.

- *Dynamic underwater environment.* Properties of the underwater channel change depending on the weather, surface state, underwater currents with different salinity and speed, and many other characteristics. Every time a transmission is made, multiple reflections from environmental features, like waves, rocks or even water bubbles can create a very different channel.

- *Difficulties with fine synchronization.* Many algorithms coming from traditional terrestrial wireless networks require a fine synchronization between nodes or array parts. However, in an underwater scenario, due to the long deployment time of wireless remote nodes and strong changes in pressure and temperature, even a good clock suffers from drift. Slow non-linear signal propagation induces the complexity of synchronization algorithms.

- *Energy and size constraints* Most of the nodes in an underwater network (including AUV) lack proper power supply, and battery life has to be saved whenever possible. That includes limitations on signal emission, extra sensors or transmitters, computational power, etc.

- *Accessibility of experimental work* Sea trials and deployments are difficult, long, and expensive. When an experiment is organized, weather and conditions on board might not allow performing all planned activities.

In this work we try to address most of these challenges and propose a group of algorithms for underwater ranging and localization that consider properties of underwater acoustic channel including non-linear propagation and changing nature of environment.

1.2. Contributions

The main contributions of this work have been presented in the following venues. Three publications were published in journals: one publication in a special issue of *The Journal of the Acoustical Society of America*, a second one in the *MDPI Sensors* journal's *Special Issue on the Internet of Underwater Things*, and a third one was published in the *IEEE Journal of Oceanic Engineering*. Additionally, three publications in conferences, top-known in the field of underwater localization: two of them in *IEEE/MTS OCEANS*, one publication in *UComms*. One more publication was recognized with the "Best Paper Award" at the *IEEE WPNC* workshop. In more detail:

Contribution 1. *Algorithms for underwater bathymetry-aided passive range estimation and localization*

We have investigated the algorithms that improve the accuracy of ranging by using additional information about the underwater acoustic propagation environment, and studied the effects of imprecise environmental knowledge on the accuracy of underwater localization and range measurements.

- **Elizaveta Dubrovinskaya**, Ivor Nissen, Paolo Casari. On the accuracy of passive multipath-aided underwater range estimation *[UComms 2016]*.

- **Elizaveta Dubrovinskaya**, Roee Diamant, Paolo Casari. Anchorless underwater acoustic localization *[IEEE WPNC 2017]*.

- **Elizaveta Dubrovinskaya**, Roee Diamant, Paolo Casari. Bathymetry-aided underwater acoustic localization using a single passive receiver *[The Journal of the Acoustical Society of America 146 (6), 4774-4789, S.I. on Acoustic Localization]*.

Contribution 2. *Underwater Direction of Arrival Estimation using Wideband Arrays of Opportunity*

We present a scheme to estimate the direction of arrival of acoustic signals reflected by underwater targets using wideband hydrophone arrays of opportunity. Such arrays may be obtained by arranging together multiple smaller sub-arrays that were originally designed to work independently. The array that results is often affected by spatial ambiguity due to improper array element spacing. Our proposed scheme solves this issue by fusing direction-of-arrival information with coarse multilateration outputs. This makes it possible to eliminate most of the ambiguity, and yields accurate direction-of-arrival estimates. Our simulation results show that our scheme achieves satisfactory direction of arrival estimates.

- **Elizaveta Dubrovinskaya**, Paolo Casari. Underwater Direction of Arrival Estimation using Wideband Arrays of Opportunity *[IEEE/MTS OCEANS 2019-Marseille].*

- **Elizaveta Dubrovinskaya**, Veronika Kebkal, Oleksiy Kebkal,Konstantin Kebkal, and Paolo Casari. Underwater Localization via Wideband Direction-of-Arrival Estimation using Arrays of Arbitrary Shape *[MDPI Sensors, S.I. on the Internet of Underwater Things].*

Additional contributions.

- Paolo Casari, Filippo Campangaro, **Elizaveta Dubrovinskaya**, Roberto Francescon, Amir Dagan, Shlomo Dahan, Michele Zorzi, Roee Diamant. ASUNA: A topology data set for underwater network emulation *[IEEE Journal of Oceanic Engineering].*
We report the details of ASUNA, a freely shared data set for underwater network emulation. In this work we analyze logs from previously performed multiple sea experiments to provide a benchmark database of time-varying network topologies. This enables researchers to run underwater communications and networking tests by "replaying" the conditions of the experiments included in our ASUNA database, instead of facing the costly operations required for sea trials.

- **Elizaveta Dubrovinskaya**, Fraser Dalgleish, Bing Ouyang, and Paolo Casari. Underwater LiDAR signal processing for enhanced detection and localization of marine life *[2018 OCEANS-MTS/IEEE Kobe Techno-Oceans (OTO)].* This work presents an algorithm designed for underwater Light Detection And Ranging (LiDAR) signal processing, and experimental work related to its validation.

1.3. Outline of the thesis

The rest of the thesis is organized in chapters that delineate the aforementioned contributions.

First, Chapter 2, in the next Part, presents background on underwater localization.

Further, work on multipath-aided underwater localization is presented in Part III, where we study the accuracy of multipath-based ranging (Chapter 3), and propose our approach for model-based localization (Chapter 4).

Part IV describes an algorithm for underwater localization based on direction of arrival (DoA) using complex arrays of arbitrary shapes (Chapter 5), and shows the application of this technique in several experiments (Chapter 6).

Chapter 7 draws conclusions and final remarks related to the work carried out in the thesis.

Part II

Underwater Acoustic localization: background and state of the art

"If you cause your ship to stop and place the head of a long tube in the water and place the outer extremity to your ear, you will hear ships at a great distance from you."

Leonardo da Vinci, 1490

2 Background on underwater acoustic localization

At the dawn of life, hearing was developed among the first organs of perception. Fossil bony fishes got their ability to hear about 350-400 million years ago [13]. Indeed, sound is one of the main sources of information underwater: light can only penetrate from a few dozens meters of clear water down to a few centimeters in turbid conditions. Salty water is a conductive medium, therefore radio waves dampen down after few centimeters, and up to a few meters away from the transmitter, depending on frequency. Instead, acoustic waves are able to propagate in the water up to hundreds kilometers.

Sound propagates as a pressure wave. As such, it propagates in water faster and for longer distances than in air, because the particles of a denser medium are closer to one another, and can transfer the energy of such a wave easier. However sound is still significantly slower than radio waves in air – a means that is widely used in conventional terrestrial communications. The difference between the propagation speed of sound in water and radio in air in speed propagation is considerable: about 1500 m/s vs about $3 \cdot 10^8$ m/s, respectively. That makes many traditional "in-air" approaches for localization inapplicable under the surface of ocean.

Initially, specific underwater localization techniques were probably inspired by biosonars used by toothed whales (e.g., dolphins). These animals developed such ability about 35 million years ago [14]. Similar mechanisms were independently developed by bats in air. They produce a specific sound and then listen to its reflections with two ears. These animals can estimate the distance from the echoing object by measuring the time that passed from the moment they emitted the sound to the moment they heard it back. Additionally, they can hear the difference of signal strength and time difference between two ears to estimate the azimuthal angle of arrival of the reflecting surface relative to their own body.

2.1. Historical approaches

Learning how sound propagates in water has always attracted scientists and researchers. From the beginning of navigation, mankind faced a need to measure the

depth of water bodies, in order to prevent ships from running aground. While this task was performed manually (with a special person to measure depth with a stick or a lead-line called lot), a large number of ships still got eventually damaged by unexpected underwater obstacles.

It is authentically known that, already in the fourth century BC, Aristotle knew that sound can be heard in water as it does in air [15, p. 6]. About two thousand years later, in the 15th century, Leonardo da Vinci noted that with sound helped estimate the distance from a lightning event by calculating the time between the lightning itself and the ensuing thunder sound. Furthermore, he mentioned the ability of sound to propagate underwater. Then, over the course of 3 centuries, there was a strong development of the mathematical apparatus that makes it possible to describe the propagation of sound in fluids. The law of refraction, also known as Snell-Descartes law and formulated during the 17th century, is one of the key laws used for modern hydroacoustic channel modelling based on ray tracing [16, p. 66]: A first documented experiment to measure the speed of sound in water was performed in 1826 in lake of Geneva by Jean-Daniel Colladon and his colleague Charles Sturm, during their research work on the compressibility of liquids (see fig. 2.1). The measured value of 1435 m/s was proved to be relatively accurate by later experiments, having in mind that the scientists did not have access to any electrical instrumentation [17].

These and other theories were further developed by Lord Rayleigh in the book "Theory of sound" (1877). Among the first scientific works specifically on hydroacoustics there is a paper "Sound signals" (1882) by F.F. Petrushevsky, where the author summarized the studies of the laws of sound propagation in sea water. He also noticed that sound can propagate over longer distances in shallow waters due to reflections from the bottom and surface of the water.

The beginning of the 20th century was a golden age of hydroacoustics. Advances in piezoelectric instrumentation allowed creation of various remotely controlled sound transmitting and receiving devices, often called projectors and hydrophones respectively. At first, they were mostly used for communication between ships (e.g., via morse codes) and were mostly replaced with radio communications after their invention.

After RMS Titanic catastrophe in 1912, many scientists were inspired to seek a robust solution to prevent collision with icebergs and other underwater objects. Moreover, during first World War many countries had a lot of interest in underwater warfare and submarine detection and localization. At this time multiple scientists invented and tested various sound-based devices independently at about the same time, including British meteorologist Lewis Fry Richardson, German physicist Alexander Behm, and French physicist Paul Langevin, together with Russian emigrant electrical engineer Konstantin Chilovsky.

First sonar-like devices were based on bulky and unreliable piezo-electric transducers.

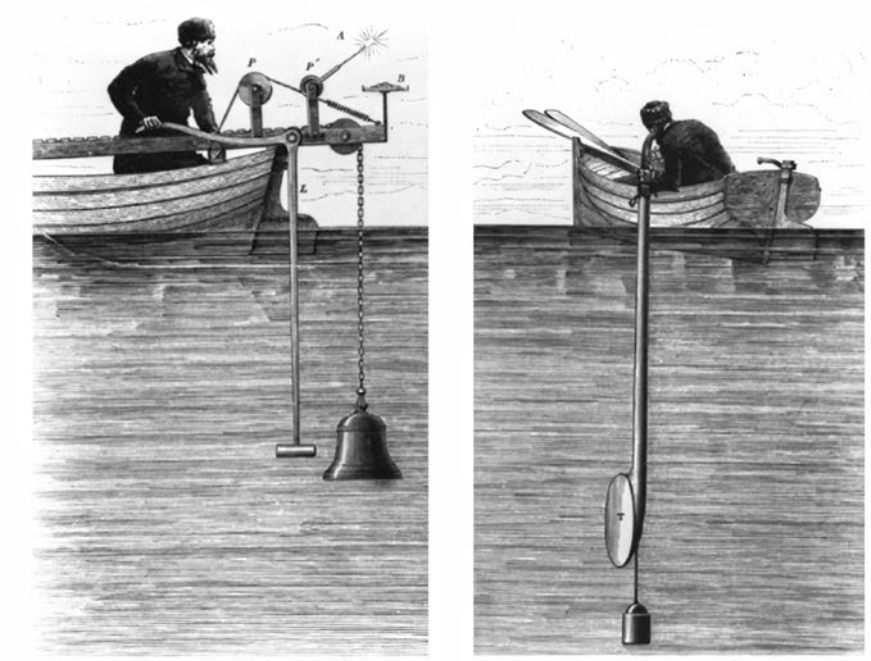


Figure 2.1: Historical experiment: measurement of sound speed in Geneva lake in 1826 [8]. One boat was carrying a remotely controlled submerged bell, that was activated together with a rocket. The second boat, 13487 m away, was carrying a person with a submerged ear trumpet. The measured time between the rocket priming and sound signal noted by the second boat was about 9.4s.

Throughout the 20th century, the field was rapidly evolving together with the development of electromagnetic instrumentation. The term sonar was first introduced by American engineers in 1930s. It was an acronym for "SOund NAvigation and Ranging" and was inspired by a similar existing acronym RADAR ("RAdio Detection and Ranging"). The British name asdic for hydrolocation systems stems from the need to classify developments related to the ultrasound devices. Therefore, "superson-ic" turned into ASD-ic, where ASD stated for "Anti-Submarine Detection" [15]. Nowadays various echolots, sonars and acoustic localization systems are used in almost any discipline related to underwater environment. The possibilities offered by the use of sound for object localization under the surface of the oceans are very attractive. However, it is crucial to face the special characteristics of this environment being a media for sound propagation.

2.2. Modeling underwater acoustic channel properties

2.2.1. Sound speed

Sound is a vibration wave that propagates through an elastic medium, such as air or water. The speed of this propagation (the wavefront's longitudinal motion) depends on the medium: in general, the stiffer and the more dense is the media, the faster sound travels. In fluids this relation is described by Newton-Laplace equation:

$$c = \sqrt{\frac{K}{\rho}} \tag{2.1}$$

where c is sound speed in fluids, K is the bulk modulus (defined as the coefficient of stiffness to volumetric deformation), and ρ is density. When it comes to real sea environments, both parameters are typically variable throughout the watercolumn and change with the environment. Measuring these parameters directly can be challenging. There exist several empirical formulas to estimate the relationship between sound speed and parameters that can be measured, e.g. [18] or [19]. The formulas are fairly complex, reaching standard errors of 0.05 m/s and 0.036 m/s respectively. However for many applications, the high precision of these empirical formulas is neglected by the low resolution of collected data, for which simpler formulas can already work sufficiently well.

One of the commonly used empirical formulas is the following [12, p. 1]:

$$c = 1449.2 + 4.6T - 0.055T^2 + 0.00029T^3 + (1.34 - 0.010T)(S - 35) + 0.016z \tag{2.2}$$

where T is temperature in degrees Celsius, S is salinity in parts per thousand, and z is depth in meters. These parameters are often measured using special devices called conductivity-temperature-depth (CTD) profilers. Conductivity is a good measure of water salinity, and each quantity can be translated into the other using, e.g., empirical tables

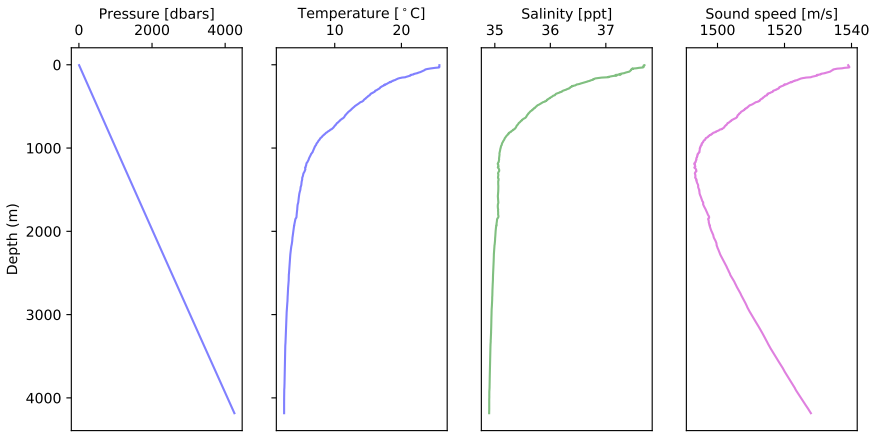


Figure 2.2: Example of SSP profile calculated using Eq.2.2 and real measured CTD profile taken at Northern Atlantic (24°29'33.7"N 39°19'43.3"W) [9]

[20].

All these parameters are variable throughout the water column, and the resulting sound speed for each measurement is called sound speed profile (SSP). Examples of such measured CTD profile taken at Northern Atlantic and calculated SSP is presented on figure 2.2.

According to refraction law (2.3), propagation of wave in non-homogeneous medium will not be linear. The path of the wave can be defined using Fermat's principle, that states that the path taken by a ray between two given points is the path that can be traversed in the least time.

$$\frac{\sin\theta_2}{\sin\theta_1} = \frac{c_2}{c_1} \tag{2.3}$$

where θ_1 and θ_2 are angles of incidence and refraction respectively, c_1 and c_2 are wave propagation speed in the 2 mediums (see figure 2.3 b). In case of linear gradient SSP sound will propagate by a circle trajectory with radius that is defined as:

$$\begin{aligned} R &= -\frac{c_2}{g} \\ g &= \frac{c_2 - c_1}{h_2 - h_1} \end{aligned} \tag{2.4}$$

where g is a slope, c_1 and c_2 are sound speed values at depth h_1 and h_2 respectively (see Figure 2.3 c). However, for many realistic SSP it is difficult to find a straightforward formula to calculate the trajectory of sound propagation. Instead, there exist various numerical algorithms for sound propagation modelling and prediction. They will be presented more in details in the section 2.2.5.

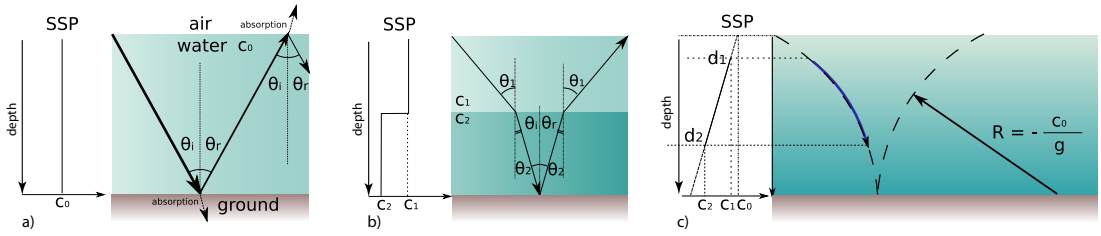


Figure 2.3: Sound propagation in water with a) straight SSP b) 2-value SSP c) constant gradient SSP

2.2.2. Variability of natural water basins

Even if it were possible to measure all instantaneous characteristics of the underwater environment with no error, it would still be difficult to predict the exact sound propagation, as there many relevant physical properties remain practically unpredictable and rapidly varying. The full description of these changes is out of scope of this thesis. However, some of them are mentioned below.

- *Currents.* Examples of such variability are underwater currents, characterized by the continuous movements of water. Each current may yield differences in salinity and temperature, affecting sound propagation. Besides, strong currents may affect bathymetry changes, and even the shoreline. Inflow from freshwater rivers or drying salty lakes create another source of significant changes in underwater acoustic channels.

- *Small-scale turbulence.* Turbulence is a chaotic water motion characterized by changes in pressure and particle velocity, induced by external or internal forces. In nature, water basins reside under the influence of diverse phenomena, including wind, temperature variation, anthropogenic motions and others. Each may create big or small scale turbulence effects.

- *Air bubbles.* Waves may create bubbles of air suspended in water. These bubbles can seriously affect acoustic channel attenuation due to thermal and scattering effects, and may resonate at different frequencies dependent on their size. Various effects of air bubbles was studied e.g. here: [21].

- *Ice* In cold areas, especially at the highest latitudes, water surface may freeze, partially or completely. Sound speed in ice is significantly higher and is usually above 3650 m/s dependent on salinity and density of ice. Therefore, frozen surfaces, as well as icebergs, affect sound propagation. An in-depth study of ice effects on sound propagation is presented in [22].

2.2.3. Ambient noise

Several factors affect the amount of acoustic noise perceived by an underwater acoustic device. The main factors are listed below [23].

- *turbulence* N_t. Water turbulence creates noise affecting mainly the lowest sections of the acoustic spectrum, typically at less than 10 Hz.

- *sea state* N_s. Wind-induced surface waves create a specific noise in the spectrum between 100 Hz and 100 kHz. The corresponding noise level typically increases with the wind speed.

- *weather* N_w. Various weather conditions may contribute to ambient noise level. For example, heavy rain will produce a noise peak at about 15 kHz.

- *thermal* N_t. High frequency transmissions may suffer from thermal noise that grows linearly with frequency and is notable for frequencies above 10kHz.

- *human* N_h. A serious source of underwater noise is ship motors. According to study [24], global marine traffic is increasing noise level by 3dB per decade. A container ship can easily produce a propeller noise having an acoustic level of 180 dB re 1 μPa, measured at 1 meter of distance from the engine. In areas close to ports and marine traffic paths, noise may seriously affect underwater communications and marine life [25]. With the development of sustainable energy installations, wind turbines can become another growing source of anthropogenic underwater ambient noise.

- *animal* N_a. One of the strongest noise sources in is a tiny creature called "pistol" or snapping shrimp. They snap their claws with an extremely loud wideband pop sound up to 220dB. Other sources are various mammals including whales and dolphins: their vocalizations often lie in the same frequency range as communication and sonar systems. Some fishes can also produce various sounds during spawning season or warning predators. Fish sounds are usually low frequency and rarely exceed 1 kHz.

2.2.4. Other challenges

- *Doppler shift*

While modern terrestrial radio communications have to deal with Doppler shift at high speeds, acoustic communications are also affected by the same phenomenon at a broader scale, especially due to the much lower speed of acoustic wave propagation. This effect can be negative for many applications, especially with limited frequency spectrum. However, other applications may use this information to estimate the target speed.

- *Synchronization*

High pressure and temperature changes below ocean surfaces often cause time drifts even for high quality quartz clocks. There have been several attempts to limit these drifts using atomic clocks (e.g. [26]), however these clocks are expensive and not suitable for many applications.

- *Deployment issues*

Unlike conventional terrestrial deployments, mostly done directly, underwater deployments require special remote equipment and exact positioning of deployed nodes is not always possible. Common sources of error in positioning are deployment issues (e.g. inexact anchor placements), or node displacement/drifts caused by underwater currents.

2.2.5. Sound propagation modelling

While most of the applications require experimental testing, physical deployments may be challenging and expensive or even impossible to implement. Preliminary tests using simulation may often help test various scenarios before implementations and tests with real devices. Every simulation model has certain assumptions concerning the real-world, and the complexity of such models depends on the application. A detailed review of various models and corresponding software can be found in the book [16]. In this part, commonly used models and corresponding software will be named; some will be used further in the next chapters.

2.2.5.1. Simplified empirical modeling

In many applications, simple models that consider propagation delay, noise level, and transmission loss calculated based on empirical methods. An example of such an application could be calculations of SNR for sonar power adjustments. Similar empirical models are used in WOSS [27]. The parameters for empirical models are usually estimated using several sources of experimental data.

For many applications it is not important how exactly sound waves propagate within the watercolumn. However, it is crucial to understand the approximate intensity received sound signal and transmission loss (TL). There exist two simplified wave expansion models: spherical and cylindrical. The first one assumes that sound propagates in a limitless homogeneous environment and can be applied in case of small point sound source, deep water and relatively small distances. In this case TL can be calculated as:

$$TL(dB) = 20\log_{10}(d) \tag{2.5}$$

where d is the linear distance between sound source and sink. The second one assumes

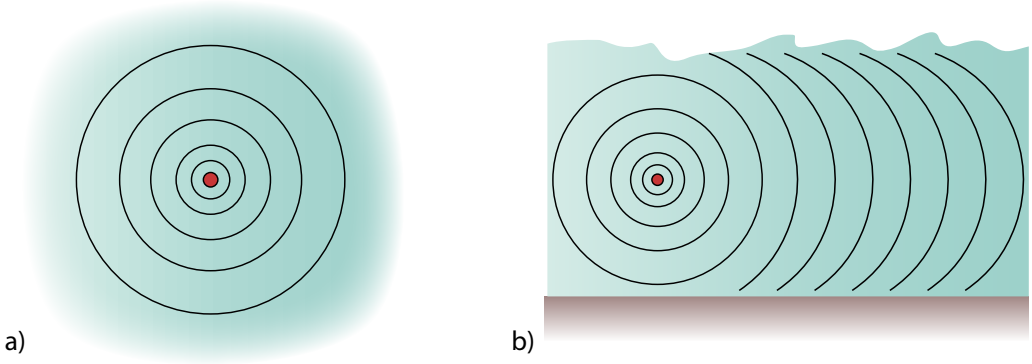


Figure 2.4: TL models: a) spherical TL, b) cylindrical TL

that the wave propagation is limited by two parallel boundaries (e.g. surface and bottom) and most of the sound signal's energy is distributed between them. It is a common approximation of sound wave expansion in shallow water. Before sound wave reaches boundaries it will be approximated as spherical, and then converted to cylindrical model of TL:

$$TL(dB) = (20\log_{10}(h) - 10\log_{10}(h)) + 10\log_{10}(d) = 10\log_{10}(h) + 10\log_{10}(d) \quad (2.6)$$

where d is a horizontal distance between sound source and sink, and h is depth [28].

- *Absorption* Previous models for the calculation of transmission loss do not consider absorption effects. While in the section 2.2.5 it will be considered more complex models, there exist a simplified formula which can be used for approximate calculations [28]:

$$TL(dB) = 10log_{10}(h) + 10log_{10}(d_{\mathrm{m}}) + a \cdot d_{\mathrm{km}} \quad (2.7)$$

where a is absorption coefficient that is usually 0.06-3.5 dB/km, and the quantities d_{m} and d_{km} refer to the distance being expressed in meters and kilometers, respectively.

- *Sound scattering* Other sources of transmission loss can be various turbulence sources that affect wave propagation. Small scale turbulence and bubbles can even completely block line of sight (LOS) or introduce extra delays into propagation time. Boundaries of water is another significant source of TL

  - *ocean surface* Uneven water surface is causing multiple reflections and partial absorption of acoustic signal.
  - *deep scattering levels* Bottom usually consist of multiple layers of

sediments with different acoustic properties. It causes various effects on sound propagation including scattering, absorption and reflection.

2.2.5.2. Stochastic modeling

Since the acoustic channel's nature is fast-changing and often unpredictable, probabilistic approaches to acoustic channel modeling may become a trade-off between the complexity of deterministic approaches and the approximations of simplified empirical models. They describe a complex acoustic channel with a random distribution function. Parameters of the stochastic model are usually taken from experimental data. However, there are examples of models that have been verified with simulation data. These models can be used in applications that do not consider exact acoustic signal propagation parameters, but rather communication parameters such as Signal-to-Noise Ratio (SNR), delay spread, and similar.

2.2.5.3. Deterministic modeling

The underwater environment can be presented as a combination of various physical phenomena, each of which can be mathematically modeled. The deterministic approach assumes that the same parameters of the model will induce the same modeled channel. Theoretically, perfect knowledge about the environment may allow an accurate deterministic model representation of real physical processes. In practice, it is extremely challenging to get ideal data about the environment, especially that of small scale and in high resolution. However, even with imperfect knowledge, it is possible to model the environment with an accuracy high enough to be used for multiple applications.

Deterministic modeling is based on various ways to solve the wave equation:

$$\frac{\partial^2 s(t)}{\partial t^2} = C^2 \nabla^2 s(t) \tag{2.8}$$

where $s(t)$ is a signal scalar function of t time, ∇^2 is Laplacian operator, and C here is a fixed non-negative coefficient.

Direct solution of partial differential equation in 3D is complex. Approximate solutions create a set of various mathematical models, whose classification is difficult: modern software solutions are using various combinations of these models. Two most common models will be mentioned here: ray-tracing approach and normal mode approach.

- *Normal mode based approach*

This approach is based on a normal mode solution of Eq.2.13:

$$s(t) = \phi e^{-i\omega t} \tag{2.9}$$

$$\phi = F(z)S(d) \tag{2.10}$$

$$\frac{\partial^2 F}{\partial z^2} + (k^2 - \xi^2)F = 0 \tag{2.11}$$

$$\frac{\partial^2 S}{\partial d^2} + \frac{\partial S}{d\partial d} + \xi^2 S = 0 \tag{2.12}$$

where ϕ is potential function, ω is a phase function, $F(z)$ is a depth function and $S(d)$ is a distance function. ξ^2 is used as a separation constant and k is a wave number. In contradistinction to ray-based approaches, this model treats signal as a wave and considers various wave effects such as dispersion, especially noticeable at low frequencies. However in order to have a good model, wave properties of boundaries have to be well known, which is unlikely to be possible, especially in shallow waters. An example of software that is using normal mode based approach is [29].

- *Ray-tracing approach*

Ray theory is based on Helmholtz equation solution:

$$\nabla^2 s + k^2 s = 0 \tag{2.13}$$

where k is the wave number. Ray-tracing approach models acoustic signal as a set of ideally narrow rays and traces their paths through given media considering its various inhomogeneities and partially reflecting boundaries. This approach is frequency independent as it models sound signal propagation as amplitude and phase functions of space coordinates:

$$s = A(x, y, z)e^{i\cdot\omega(x,y,z)} \tag{2.14}$$

where A is an amplitude function of coordinates x, y, z and ω is a phase function respectively. It considers that media properties (such as bathymetry, sound speed) are varying with coordinates but the differences are negligible within one wavelength. Since this approach is geometry-based, wave phenomena such as interference or diffraction are not considered. A common software that is using ray-tracing approach is Bellhop [30]. A significant advantage of this approach is the possibility of obtaining a relatively accurate arrival structure with considerably less computational efforts than using other methods. For this reason, the ray-based approach was preferred in this thesis.

2.3. Overview of existing approaches to underwater acoustic localization

Nowadays the need of accurate underwater localization and navigation is motivated by many scientific needs. One of the fastest developing areas is underwater robotics, which is hardly possible without proper positioning systems. Shallow sea wind electricity farms and mining platforms require exact installation and maintenance of their submerged equipment, as well as various ecological, geophysical and other sensors placed in water column or on the bottom of the sea. Ports and other sea adjacent territories are in need for an appropriate protection from naval and terrorist threats. Fishing industry and supervisory authorities would like to know better about the current state of various underwater animal populations.

To make the most of the hydroacoustics capabilities in all of these applications and to cope with the challenges of the underwater environment, multiple algorithms have been created for underwater localization. It is very difficult to cover the entire range of possible technologies within the framework of this thesis, therefore scenarios will be restricted to those that mainly rely on acoustic-based sensors. Here localization is understood as finding relative or absolute coordinates of a desired object, called target in the following. When the algorithms are run at the target, localization can be called self-localization. Localization is an important step of navigation that is focused on controlling of movement of various objects, including autonomous underwater vehicles (AUVs).

Most of the methods for underwater localization are basically not different from those in terrestrial deployments. Among the main differences are the higher probability of error, the slowness of data communication and the serious restrictions on the reception and transmission of signals due to the limited spectrum and available energy. How exactly these characteristics are treated in underwater localization will be discussed in the subsection. This section focuses on basic localization scenarios.

2.3.1. Scenarios for underwater localization

Most of scenarios for underwater localization can be described using a scheme at Figure 2.5.

At first, every scenario includes some information about the given equipment to be deployed, e.g. type of sensors and their mutual arrangement. Then it is needed to define which information can be obtained using the given equipment and how it is possible to process it. As a result, in each scenario it is required to estimate the location of the desired object in a certain way.

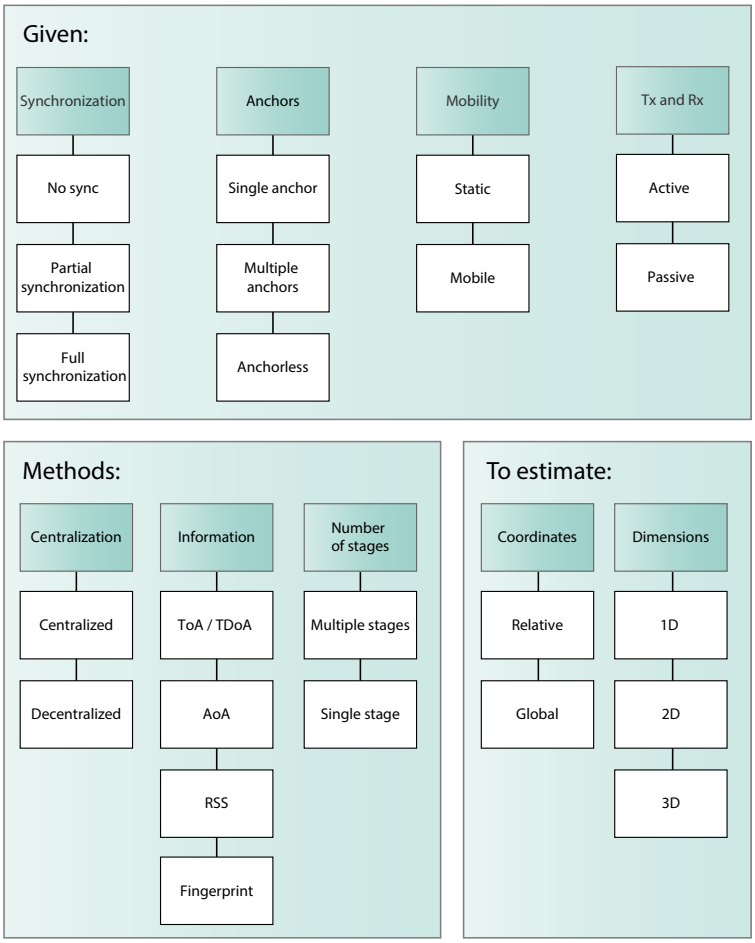


Figure 2.5: Taxonomy of underwater localization scenario, methods and outputs.

2.3.2. Scenario conditions

2.3.2.1. Synchronization

While many localization approaches in terrestrial wireless sensor networks (WSN) rely on synchronization, in underwater environment synchronization is challenging. There were many attempts to introduce algorithms for synchronization underwater [31, 32] or perform localization in presence of synchronization errors [33].

2.3.2.2. Number of anchors

The estimated location of a target can be defined through a relative position to other nodes in the underwater network, called anchor nodes. In a typical network, there could be one or more anchor nodes. A classical example of localization using multiple anchor nodes is long baseline (LBL) (studied, e.g., here [34]). There are also methods to define the location of the target relative to the environment, e.g., to the bottom profile. The

latter is sometimes referred to as anchorless localization.

2.3.2.3. Mobility

The scenario may include static and mobile nodes. A static node remains stationary within the watercolumn, on the surface or bottom. On the contrary, mobile nodes typically change their location over time.

2.3.2.4. Acoustic sensors

Each node in an acoustic sensor network shall be equipped with a sensor that can either transmit an acoustic signal (acoustic transmitter), receive it (hydrophone) or both (e.g., acoustic transceiver). In case a node can produce acoustic signals, it is called an active node, otherwise it is passive. Additionally, there are some sensors that can measure extra parameters of acoustic signal such as particles velocity [35, 36].

2.3.3. Underwater localization methods

2.3.3.1. Centralization

In some applications, all information from multiple underwater sensors can be collected at a single place, e.g., a central node. This is quite common for applications with post processing, or in networks with sensors connected by wires or air wireless communication. However in other scenarios that do not assume full communications between sensors at each stage (e.g., AUV swarms or distributed sensor networks with limited connectivity) it is possible that nodes have to make independent decisions based on collected data. In this case, the scenario can be called decentralized. An example of decentralized localization can be found in [37]. Some scenarios may assume both options for different nodes.

2.3.3.2. Number of stages

After exchanging the information prescribed by the localization scheme, it becomes possible to convert this information directly into coordinates using one the methods described above. However, due to erroneous or insufficient data, one information exchange stage may not be enough to estimate the coordinates of a target. Single-stage localization methods assume that, after each stage, there is a decision about the estimated location of the target. Multiple-stage methods can require more than one round to estimate the position and rely on intermediate estimates. While most of algorithms for underwater localization are single-stage, an example of multi-stage algorithm can be found in [38].

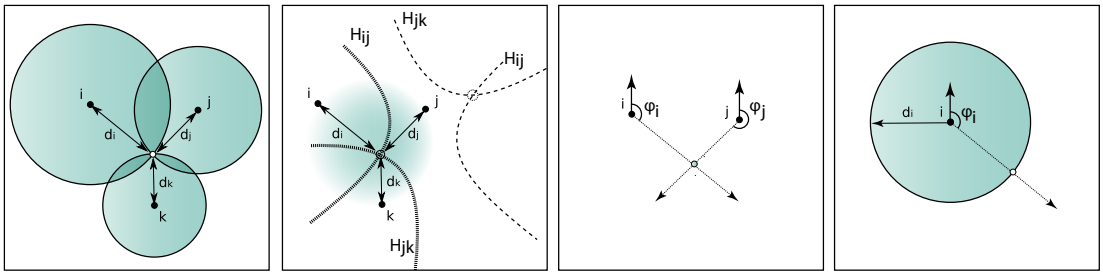


Figure 2.6: Various localization techniques a) based on Time-of-Arrival (ToA), b) based on time difference of arrival (TDoA) c) based on angle of arrival (AoA) d) combined ToA and AoA.

2.3.3.3. Information

Depending on the number of sensors and on the presence or absence of synchronization between nodes, there could be several methods to estimate the target location. Most of them are employed in terrestrial WSNs as well, however the accuracy of these methods is usually significantly lower in underwater scenarios, due to time drifts and sound speed variability in time and space. This part will cover the classic localization methods. A first set of methods based on range estimation between anchor nodes and target is presented in figure 2.7. The general idea of these methods is to measure how much time it takes for a signal to travel between nodes. These methods in general work both with anchor nodes as transmitters and as receivers, given a reciprocal channel assumption.

- *ToA*. A first baseline method works as follows. Node 1 can send a signal to a node 2, and node 2 can measure the traveling time between nodes, t_{travel} (see figure 2.7 a). It is often called ToA. In this case only one node needs to be active, however both nodes need to be synchronized. The acoustic time of flight can be converted into a range estimate by assuming constant known sound speed and negligible refraction, leading to a (quasi-) rectilinear shape of acoustic propagation paths. In this case we can compute:

$$d = c \cdot t_{travel} \tag{2.15}$$

where d is the linear distance between nodes, c is sound speed and t_{travel} is measured time between signal emission and signal reception. If there is no synchronization between two nodes, but both nodes are active, the range can be measured by round trip time (see Figure 2.7b). In this case node 1 is transmitting the acoustic signal, node 2 receives the signal, processes it and transmits back.

$$d = \frac{1}{2}(t_{travel} - t_{service}) \cdot c \tag{2.16}$$

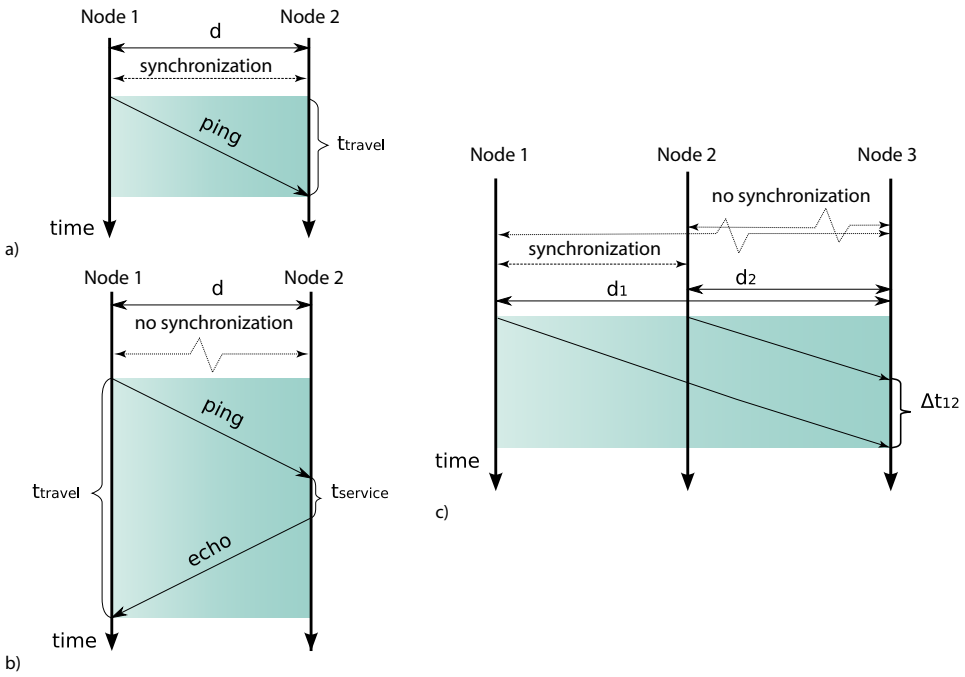


Figure 2.7: Graphical representation of various range measurement methods. ToA: a) Travel time b) Round trip travel time. c) TDoA.

- *TDoA.* Another option for indirect range estimation is to measure TDoA (see figure 2.7c). In this case, only anchor nodes need to be synchronized. At the target node, the algorithm measures TDoA. With only 1 TDoA measurement it is not possible to estimate the range between the target and the anchor nodes, but it is possible to define a locus (i.e., a geometric description of all possible locations of target node) as a hyperbolic line in 2D or a two-sheeted hyperboloid in 3D. Multiple measurements will produce a set of loci, whose intersection will identify the desired location of target node, assuming no error in measurements (see fig. 2.6b).

- *received signal strength (RSS).* This method is widely used in radio communications, but its application is limited in the acoustic underwater environment. In many situations including deep water basins, the TL often does not have a clear relationship with distance, and is affected by lack of LOS and rich multipath. However there are some examples of successful localization based on RSS in shallow water underwater sensor networks (UWSN) [39], or methods where RSS is used as an auxiliary metric to improve the accuracy of other localization methods [40].

- *AoA.* If a node has more than one associated acoustic sensor, it may be possible to estimate the AoA of the acoustic signal (see fig. 2.8). In this case, the

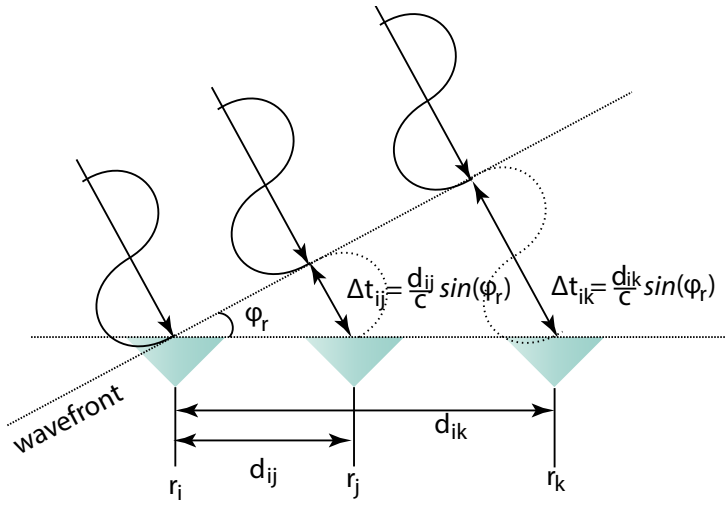


Figure 2.8: AoA estimation. Δt_{ij} is time difference of arrival between sensors r_i and r_j, d_{ij} is linear distance between corresponding sensors, c is sound speed and φ_r is angle of wavefront incidence. Similarly for sensors r_i and r_k.

phase difference between different sensors is measured. Dependent on frequency and sound speed, this information can be converted into a very fine time difference of arrival even when the latter is less than the sampling time of the acoustic device's analog-to-digital converter:

$$\Delta t_{ij} = \frac{\Delta \phi_{ij}}{2\pi f} \tag{2.17}$$

where Δt_{ij} is time difference of arrival between i and j sensors. In order to avoid that the phase difference exceeds one wave period, usually it is recommended to place nearest receivers within half a wavelength from each other. On Fig. 2.6c we show an example of localization technique based on AoA information.

- *Fingerprint or Matched field processing (MFP).* MFP or similar fingerprint-based localization are often used in terrestrial WSN, e.g. based on millimeter-wave signals [41] or acoustic signals [42]. As discussed in this chapter earlier, underwater acoustic channel properties are highly time-varying. Therefore, direct methods of MFP applications can be computationally ineffective.

- *Combined techniques* While each aforementioned technique can be used separately, often their combination is more effective and suitable for certain applications. On Fig. 2.6d shows an example of combining ToA and AoA in a scenario with a single anchor node.

2.3.3.4. Coordinates

The aforementioned algorithms help find coordinates relative to the anchor nodes. Relative coordinates can be represented by polar (elevation angle ϕ, azimuth angle θ and range ρ), cylindrical (azimuth angle θ, distance d and depth z) or Cartesian (horizontal plane x, y and depth z) coordinates. For certain applications, it is required to find an exact geographical position. In this case relative coordinates can be converted to geographical coordinate system (GCS) if such position of anchor nodes is known. A common convention for GCS is using latitude, longitude and elevation/depth coordinates or earth-centered, earth-fixed (ECEF) Cartesian coordinates.

2.3.3.5. Dimensions

Different applications require different dimensionality of localization. Range estimates enable localization only along a single dimension. Merging range information with angle of arrival information or depth estimates enables localization in two dimensions. Three-dimensional localization requires to consider three quantity estimates, e.g., range, angle of arrival and depth, or range and two angles of arrival (e.g., elevation and azimuth).

2.4. Classical deployment scenarios

While installation technologies are constantly improved, baseline localization methods can be divided into the three main deployment scenarios presented in Fig. 2.9.

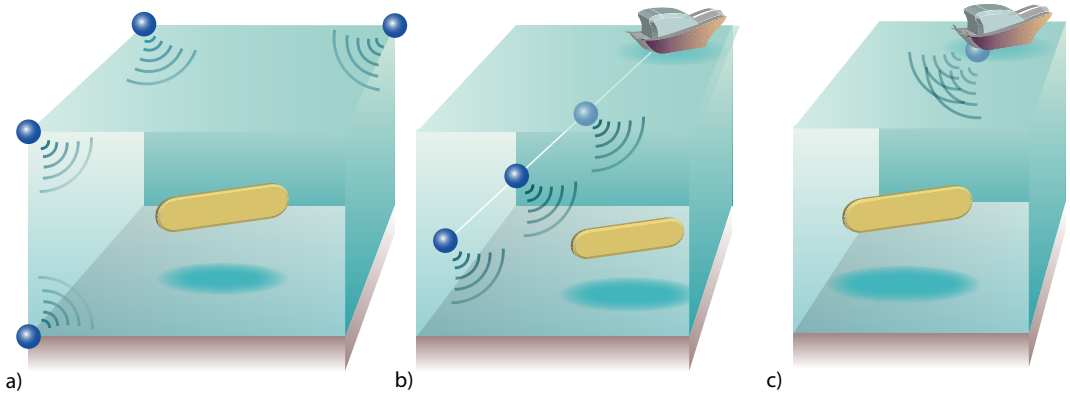


Figure 2.9: Classic deployment scenarios for underwater localization: a) LBL, b) short baseline (SBL), c) ultra short baseline (USBL)

- *LBL.* In this scenario, anchor nodes are deployed separately as surface buoys or anchored nodes in the area of interest, usually on the borders of the area (see Fig.2.9a) This scenario assumes a challenging deployment. However, if the nodes are correctly deployed, LBL schemes achieve good accuracy.

- *SBL.* The relatively small distance between sensors in this scenario allows deployment directly from a ship or pier. At the same time, the small distance between the sensors can reduce the accuracy of localization, including inaccuracies in synchronization and installation, as well as signal processing [43] (see Fig.2.9b)

- *USBL.* If the distance between the acoustic sensors is small enough to meet the condition of less than half the wavelength, then it becomes possible to estimate the direction of arrival (DoA) of incoming acoustic signals through array signal processing techniques (see Fig.2.9c). Such systems are usually easy to install on any floating craft and are used further as a sub-array to compose larger arrays in Part IV.

- *Synthetic arrays* If the target to be localized is static (e.g. it is a sea bottom feature), a similar mathematical apparatus can be used substituting multiple anchor nodes with a single moving anchor node. Examples of such synthetic array processing can be found in [44, 45].

An in-depth analysis of the state of the art related to each of the topics presented in this thesis is provided in the following chapters, tailored to their respective contents.

Part III

On multipath-aided range estimation and localization

As discussed in previous chapters, underwater acoustic localization is required by various applications of sciences related to the ocean, including marine life monitoring and conservation of endangered species. For certain applications that consider localization of sporadic targets it is important to simultaneously observe broad underwater environments. While most modern localization methods assume the deployment of multiple sensors [46], the deployment of a large number of active sonars to cover wide areas is often challenging, and may be cost- and energy-ineffective. Moreover, active acoustic equipment may introduce various changes in natural habitat of marine animals. autonomous underwater vehicles (AUVs) can be another key instrument for observation of marine life, their cost-effective localization that can be relatively easy deployed is another issue. There were several attempts to solve the issue of underwater localization using a single passive sensor, by leveraging additional information about the environment, and by modeling sound signal propagation. A careful analysis of the previous literature in this field is provided in Sec. 4.1 In this part, we extend the current state of the art by analyzing the accuracy of multipath-aided underwater range estimation and localization. In Chap. 3, the performance of acoustic underwater passive multipath-aided range estimation under imperfect sound speed profile (SSP) knowledge is discussed. In the same chapter, we also propose a method to improve the accuracy of range estimation by introducing an auxiliary variable named "effective sound speed," that takes into account the difference between rectilinear and non-rectlinear acoustic propagation paths. Chap. 4 is dedicated to another approach inspired by traditional wireless sensor networks (WSN) fingerprint-based localization. In this chapter, we utilize the dependence of the acoustic channel on the spatial diversity of the bathymetry profile. We propose an algorithm that exploits this diversity in a computationally effective way. By applying a modified Viterbi algorithm to filter out locations that were unlikely visited, we obtain the most likely path of a mobile source. We test the accuracy of our algorithm in different simulated conditions, and validate its performance using data from real sea experiments.

"For most of history, man has had to fight nature to survive: in this century he is beginning to realize that, in order to survive, he must protect it"

Jacques-Yves Cousteau

3 On the accuracy of passive multipath-aided underwater range estimation

Extended ocean investigation is mandated by several application in the conservation, ecology, resource prospection and also tactical domains. In particular, AUVs have the potential to be a key tool for this purpose [47]. However, AUV self-localization remains a challenge due to the characteristics of the underwater environment and the computational and energy constraints imposed by the AUVs themselves. A variety of different algorithms and architectures to overcome these challenges has been proposed over time (e.g. [47,48]). Still, several challenges remain to be solved. A precise and robust range estimation process, which possibly does not consume an excessive amount of computational resources on AUVs is one of the essential steps in many localization algorithms and advanced inertial navigation systems [49]. Conventional two-way range estimation requires an acoustic transmitter to be integrated on the AUV, which may sometimes be impossible due to size or energy consumption constraints. Without a transmitter available on the node to be localized, range estimation can be performed using one-way algorithms, which often require a challenging or time-consuming synchronization step [50].

A common technique for underwater localization is based on the computation of time difference of arrival (TDoA) among the signals emitted by a set of anchor nodes [51]. Anchor nodes send pings or structured signals, whose TDoA can be measured at the mobile node. The strong aspect of this solution is that it does not require time synchronization. However, an accurate translation of time differences into distances requires a detailed environmental model, especially in terms of SSP and capability to trace the history of signals heard at the receiver. Employing only some average sound speed value might lead to large ranging errors, and produce results that are valid only as a first-order approximation.

Additional information on the environment can be factored in by analyzing multipath sound propagation. For example, this concept has been applied to the localization of vocalizing mammals in [52], where the authors process the sound recorded by widely spaced receivers, under the assumption that the SSP and the bathymetry of the area are fully known. However, obtaining measurements of the SSP via, e.g., conductivity-

temperature-depth (CTD) equipment may be often impractical or exceedingly expensive. The idea of using additional information about multipath propagation for passive range estimation has been proposed in the past (e.g., [53]), and multipath-aided range estimation has been applied to motion tracking for several types of targets [52, 54, 55].

In this contribution, we discuss the performance of passive, multipath-aided range estimation between an anchor node and a second node whose location is unknown. We argue that obtaining full knowledge of the environment is often infeasible, and discuss a method to improve the accuracy of ranging by taking into account at least the effect of refraction on sound propagation. The primary advantage of our method is that it leverages on the processing of multipath arrivals at the receiver, and therefore it does not require additional equipment or CTD sensors. However, for the same reason, our approach works preferably in scenarios offering a rich multipath, e.g., shallow waters, and in the absence of broad distortions to multipath propagation such as a strongly mixed surface layer.

In the remainder of this work, we introduce the geometry of the problem and the basics of the ranging algorithm (Sec. 3.1), we analyze its accuracy in the presence of different uncertainties (Sec. 3.2) and finally draw concluding remarks in Sec. 3.3.

3.1. Range estimation algorithm

3.1.1. Scenario and assumptions

The TDoA-based algorithm considered in this work is designed to estimate the range between a single anchor node with known position (e.g., a floating global positioning system (GPS) [51] or bottom-mounted buoy) and a second node within the coverage range of the buoy's signal. We recall that the latter node is assumed to be passive: therefore, it can only receive the signal from the anchor node, but cannot transmit to it. For brevity, in the following we will refer to the anchor node as the "Tx" and to the passive node as the "Rx". We assume that the signal transmitted by the Tx undergoes multipath propagation, so that multiple arrivals of the same signal are detected and distinguished by the Rx. We also assume that the Rx knows the bottom depth and its own depth. The latter can be estimated, e.g., via a pressure sensor, but we note that in the presence of a sufficiently rich multipath it would be possible to partially compensate for this by solving for the depth of the Rx along with its range from the Tx. We also assume that the surface and bottom are flat, which is a common assumption in several shallow water scenario. In any event, the algorithm can be easily extended to accommodate for, e.g., sea bottom slopes.

At the time a TDoA measurement is taken, the Rx is considered to be static or moving at a speed that is sufficiently small not to affect the measurement. We will

now proceed by outlining the range estimation algorithm in a homogeneous (i.e., non-refractive) environment, where the SSP is constant and sound wave propagation can be assumed to take place along rectilinear trajectories. We will then extend the algorithm to the more general case of a non-constant SSP.

3.1.1.1. Constant SSP

In case the SSP can be assumed to be constant throughout the water column, the geometry of the problem is amenable to a tractable analytical solution. Call d the horizontal distance between the Tx and the Rx, and let the notation τ_{ij} indicate the TDoA between two paths having incurred i and j reflections off the sea surface or bottom. For clarity, we indicate where the reflections have occurred by appending the indices "s" and "b" for a surface and a bottom reflection, respectively, For example, τ_{02sb} indicates the TDoA between the line of sight (LOS) arrival and a second arrival which was reflected twice, first on the surface and then on the bottom. With reference to Fig. 3.1, we have the following equations:

$$d^2 = R_0^2 - (h_{Rx} - h_{Tx})^2 \tag{3.1}$$

$$d^2 = (R_0 + c\tau_{01s})^2 - (h_{Rx} + h_{Tx})^2 \tag{3.2}$$

$$d^2 = (R_0 + c\tau_{01b})^2 - (2h_B - h_{Rx} - h_{Tx})^2 \tag{3.3}$$

$$d^2 = (R_0 + c\tau_{02sb})^2 - (2h_B + h_{Rx} - h_{Tx})^2 \tag{3.4}$$

$$d^2 = (R_0 + c\tau_{02bs})^2 - (2h_B - h_{Rx} + h_{Tx})^2 \tag{3.5}$$

where R_0 is the length of the LOS path, c is the constant sound speed value and h_{Rx}, h_{Tx} and h_B are the depths of the Rx, of the Tx and of the bottom, respectively. The solution to the range estimation problem is then found by solving the system of equations above for d. We remark that Eqs. (3.1) to (3.5) require knowledge of the reflection history of the arrivals. This task is simplified by assuming that the anchor node is deployed either on the surface or on the bottom, so that reflected paths incur, respectively, at least one bottom or one surface reflection. We remark that this is not a restrictive assumption, as anchor nodes for localization systems are likely to be deployed on the surface (to take advantage of GPS fixes) or on the sea bottom at a fixed location (e.g., to avoid being damaged by passing ships). In the following, we assume that the Tx is located on the bottom (left panel in Fig. 3.1).

3.1.2. Inhomogenous environment model

In a realistic environment, the SSP is non constant throughout the water column. The refraction of the anchor's signal that ensues leads to non-ideal propagation geometries where, e.g., reflected arrivals might be detected earlier than the LOS arrival, or where some

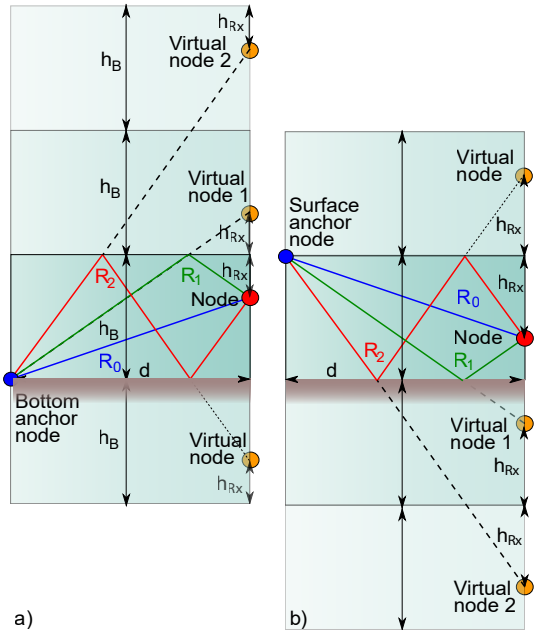


Figure 3.1: Constant SSP scenario: a) anchor node located on the sea bottom; b) anchor node located on the surface.

arrival (including the LOS one) might not exist even in the absence of barriers. The range estimation can still be carried out by fixing some approximate value for c and by finding an approximate solution to (3.1)-(3.5), e.g., via a least squares estimator (LSE). In this work, we employ the latter as a baseline reference algorithm, by assuming a fixed sound speed value at 1500m/s. However, the TDoAs between the LOS and the reflected arrivals can be employed to infer some information about the propagation of sound between from the Tx to the Rx, and thereby improve the quality of the range estimation.

In order to do so, we define the effective sound speed c^{eff} as a measure of the deviation between the rectilinear trajectory joining the Tx and the Rx and the actual trajectory followed by two rays for which we compute a TDoA. By recalling that R_i and R_j are the lengths of the trajectories of arrivals i and j under the assumption of rectilinear propagation, we define the effective sound speed as

$$c_{ij}^{eff} = (R_i - R_j)/\tau_{ij}. \tag{3.6}$$

As we will show in Sec. 3.2, the use of c^{eff} can noticeably improve the range estimation in a number of non-homogeneous environments. The main drawback related to c^{eff} is that it must be computed at a moment when the range between the Tx and the Rx is known. However, in many cases this is a feasible assumption, for example, if a ship is deploying a bottom node, the measurement can be taken just prior to releasing the node. Similarly, an AUV could start from a known location, take this measurement, and then begin its mission.

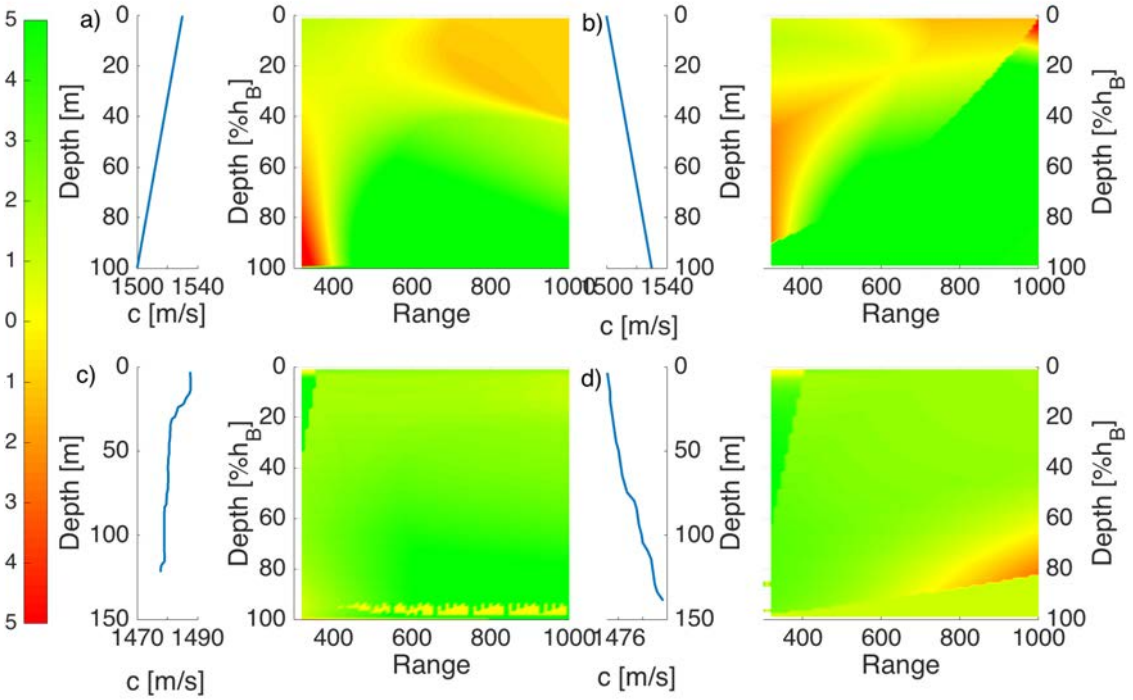


Figure 3.2: Difference (in %) between the relative error of the LSE range estimator with and without the use of c^{eff} for different SSPs: a) synthetic upward refractive SSP; b) synthetic downward refractive SSP; c) real summer SSP; d) real winter SSP. The measurement of c_{ij}^{eff} for each pair of arrivals i, j was taken once at a range of 600 m and at a depth equal to $0.5h_B$, and then used for the estimation of all other ranges.

3.2. Simulation results

In the following we will carry out our sensitivity analysis by modelling acoustic propagation via the Bellhop ray tracing software [30]. Our results were obtained using both simplified upward/downward refractive SSPs and actual SSPs taken from the world ocean database [9], in a scenario with flat bottom and surface, and mixed sand/mud bottom sediments. The output of Bellhop was preprocessed to compute TDoAs, which were then employed as an input to the range estimation algorithm. In particular, we selected three arrivals, corresponding to LOS, one reflection and two reflections, respectively. If due to specific propagation conditions, one or more of these arrivals are missing, the algorithm tries to compute an estimate using only the remaining ones, or fails otherwise.

3.2.1. Performance improvement using c^{eff}

We start by showing the range accuracy improvement achieved by employing effective sound speed values when finding a solution to Eqs. (3.1)–(3.5). Fig. 3.2 depicts the

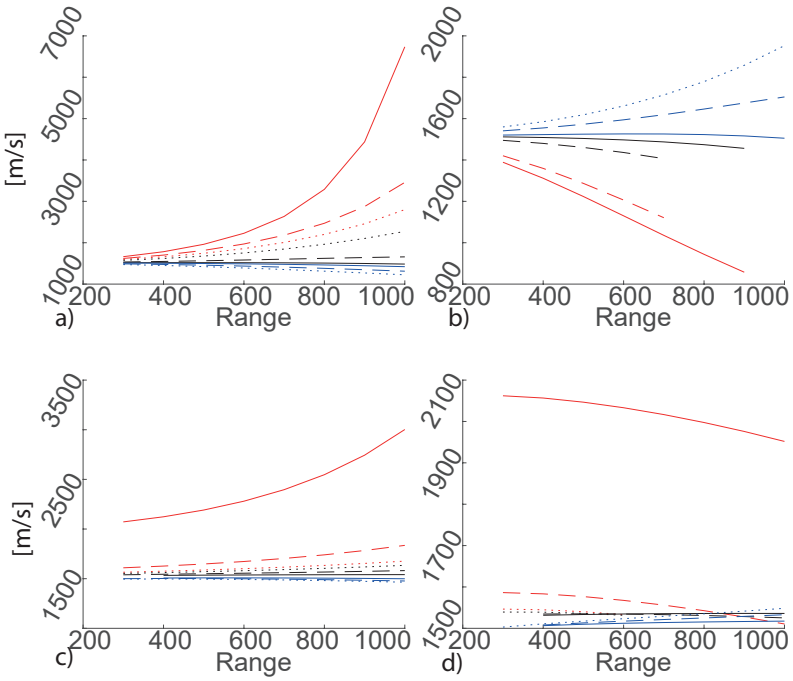


Figure 3.3: Effective sound speed values c_{01}^{eff} (red), c_{02}^{eff} (black) and c_{12}^{eff} (blue) for the same SSPs considered in Fig. 3.2, and for different depths of the Rx node (solid: $0.1h_B$; dashed: $0.5h_B$; dotted: $0.9h_B$).

difference between the relative error achieved when using c^{eff} values and the relative error achieved by a simple LSE that assumes the sound speed to be constant and equal to $c = 1500$ m/s. Green hues denote an improvement, whereas red hues a lower accuracy. In the vast majority of cases, computing c^{eff} values makes it possible to achieve more accurate results. However, in some cases the accuracy achieved is worse than with a constant sound speed value. In particular, we observe that the main factors that affect the accuracy of the estimation are the depth and range of the Rx, and the SSP.

To understand better the reason of this effect, we investigated the variation of c^{eff} with these three factors. This result is shown in Fig. 3.3. We observe that c_{01}^{eff} is the most affected by the range and depth of the Rx, especially when the latter is located close to the surface: by recalling that we consider the Tx to be located on the bottom, this effect can be explained by the fact that the LOS and the surface-reflected paths are almost superimposed. Moreover, we observed that in most cases c_{02}^{eff} and c_{12}^{eff} tend to evolve symmetrically with respect to some intermediate value that depends on the scenario. This suggested that relying on the TDoA between the corresponding arrival pairs would yield a solution that is less affected by the range and depth of the Rx node. As shown in Fig. 3.4, this improved the accuracy of the estimation considerably, as relative errors for realistic SSPs were below 1% when the node could receive all arrivals (LOS, surface-reflected and surface-bottom reflected) and was not exceedingly close to the surface or bottom.

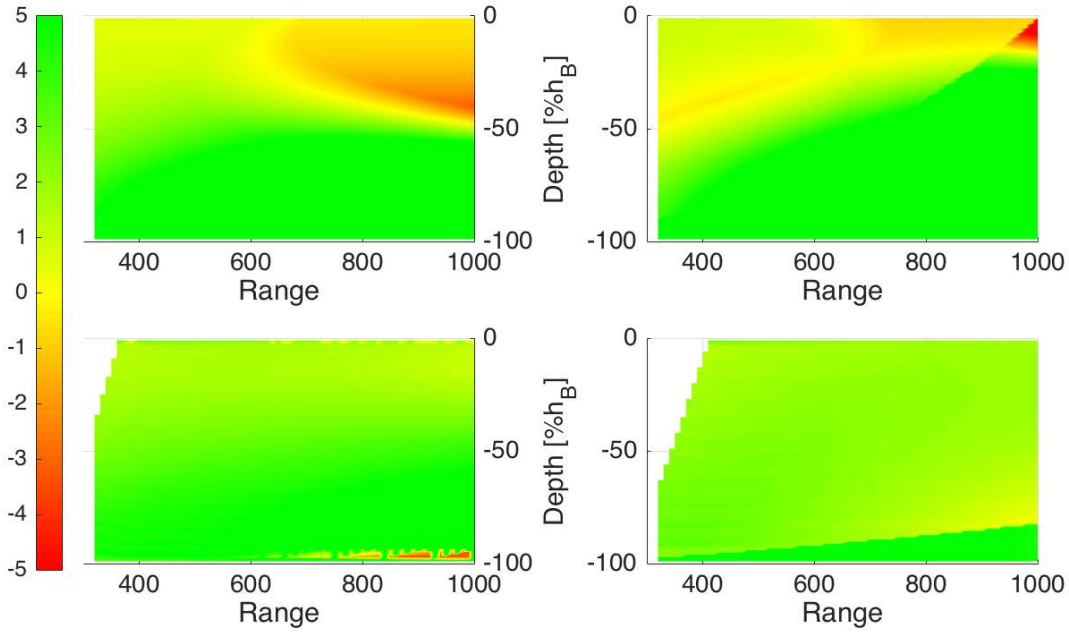


Figure 3.4: Difference (in %) between the relative error of the LSE range estimator with and without the use of c^{eff} using the 02-12 arrival pairs, for the same SSPs of Fig.3.2.

3.2.2. Sensitivity to erroneous input data

We now focus on the sensitivity of the algorithm to errors affecting the bottom depth, the Tx location the TDoA and the Rx depth estimates.

As our approach is based on TDoA measurements, a primary requirement is that these measurements are precise. According to our simulations, quantizing TDoA values down to a precision of less than 10 μs introduces significant errors in ranging estimation. The sensitivity of range estimation to errors in TDoA estimates is shown in Fig. 3.5. For this evaluation, we introduce on each TDoA value a relative error drawn uniformly at random in the interval $[0, y]$, where y is a tunable parameter. We observe that a random error up to $y = 5$ % yields a relative error of about 5% on range estimates.

We now assume that the bottom depth estimate is affected by a random relative error. Such an error affects the accuracy of the range estimation process in a way that depends strongly on the position of the Rx. This can be observed from Fig. 3.6, which represents the relationship between the error affecting h_B and the relative range estimation error. Specifically, when the Rx is not close to the surface or bottom, the effect of a 1% error affecting h_B leads to a relative range estimation error of 2 to 3%. Conversely, if the Rx is located close to the surface or bottom, TDoAs tend to decrease, causing the ranging error to increase considerably. In addition, Fig. 3.6 conveys how the ranging error varies for different Tx-Rx distances via the thickness of the line.

Finally, we consider the performance of range estimation when the value of the Rx

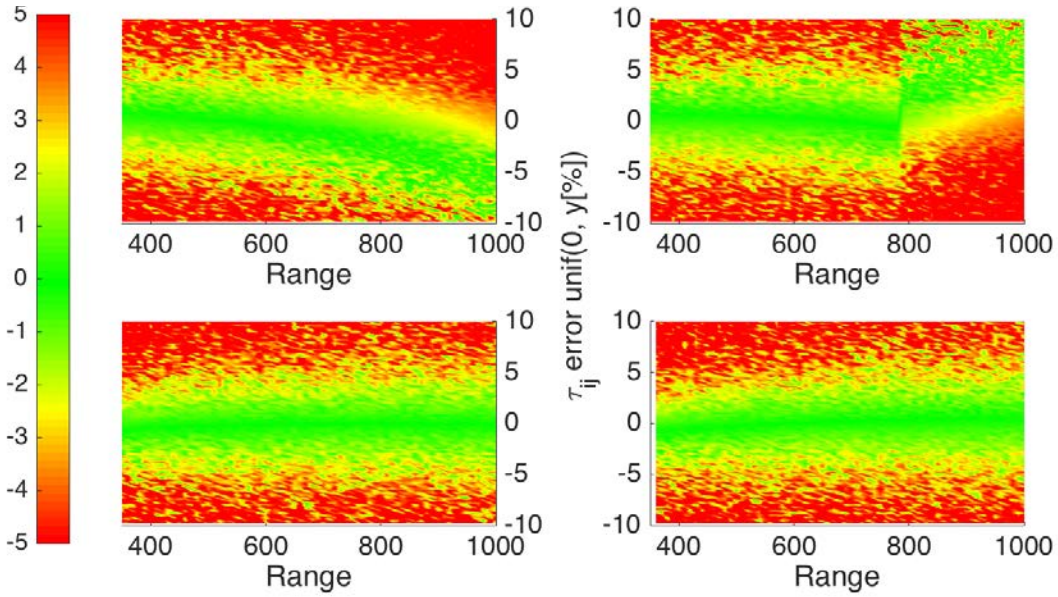


Figure 3.5: Relative range estimation error (in %) in the presence of errors in TDoA measurements for the SSPs in Fig. 3.2. The introduced relative error is drawn at random in the interval $[0, y]$, where y is the value reported on each panel's y-axis. The depth of the Rx is $0.4h_B$.

depth h_{Rx} is erroneous or not available. In this case, it is still possible to jointly estimate the range and h_{Rx}. Because in this case there exist two unknowns to be estimated, ranging can be completed successfully only when at least three arrivals are perceived by the Rx. Therefore, the error given by the algorithm strongly depends on the SSP and on the position of the Rx. In turn, the sensitivity of the algorithm to all input data increases, and errors are typically higher than in the previous cases, even if all input data is exact and despite the use of c^{eff}. Fig. 3.7 shows the relative range estimation error for the SSPs considered so far (see Fig. 3.2). We observe that due to the absence of a sufficient number of arrivals, in several cases (white areas) it is not possible to estimate the range. For realistic sound speed profiles (bottom panels), the typical relative error was about 3%, and increased up to 15% for the downward refractive profile (top left pane). In any event, the computation of c^{eff} helped also in this case, as it improved the results by 3-5% (see Fig. 3.7).

3.3. Conclusions

In this contribution, we studied the accuracy of multipath-aided range estimation based on TDoA in the presence of errors on different parameters required to compute the range estimate. We showed that summarizing the propagation history of arrivals via the computation of an effective sound speed value can significantly improve the accuracy of

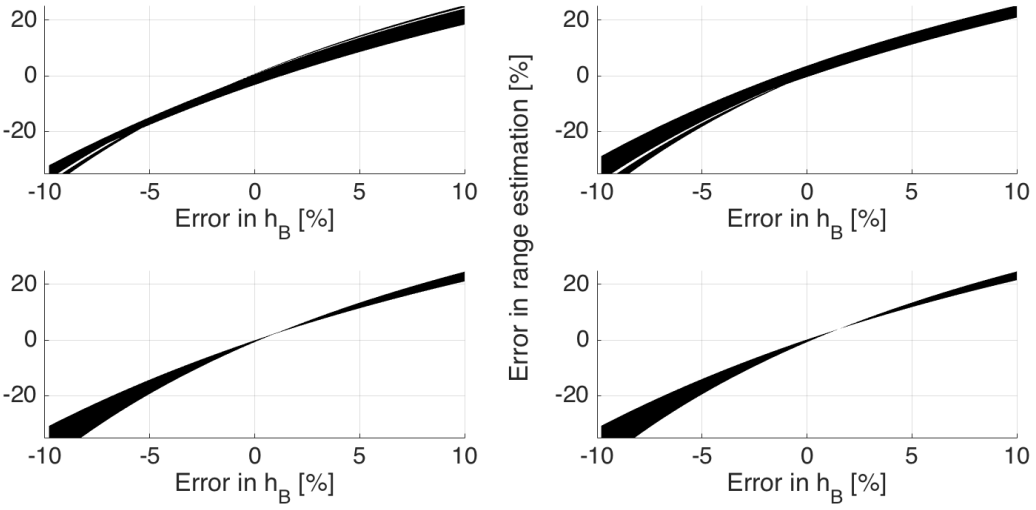


Figure 3.6: Relative range estimation error (in %) as a function of relative error affecting h_B, for the same SSPs of Fig. 3.2. The depth of the Rx is $0.4h_B$. Thicker lines represent a larger variation of the error with the range between the Tx and the Rx for the same error on h_B.

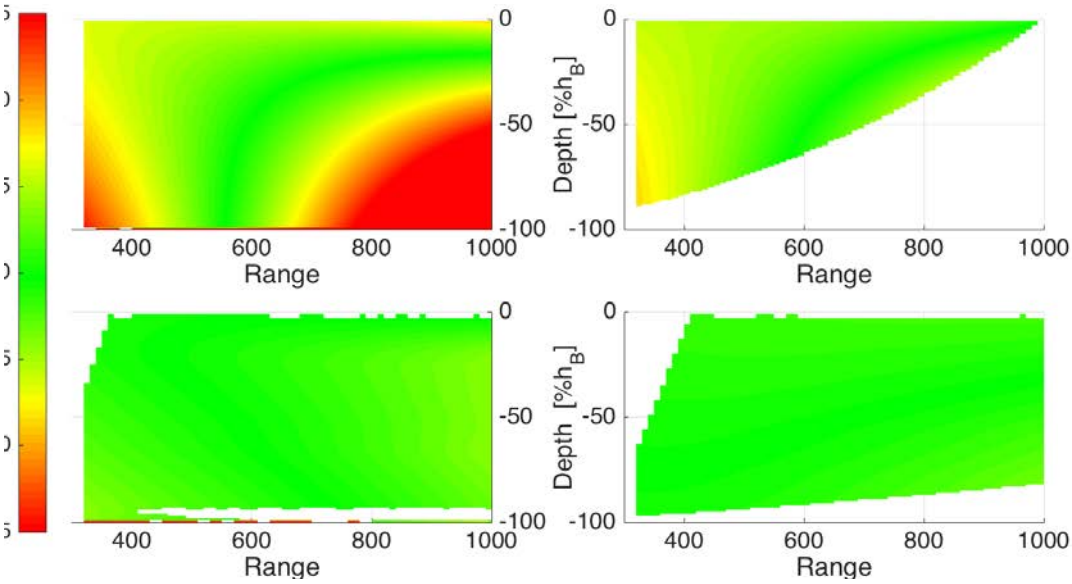


Figure 3.7: Relative range estimation error [%] in case the range and h_{Rx} are jointly estimated, for the same SSPs reported in Fig. 3.2.

the estimate. With a proper choice of arrival pairs, the relative error in range estimation in our simulations decreased to values below 1% for a number of realistic sound speed profiles. We evaluated the sensitivity of the algorithm to different uncertainties affecting the TDoA values, the bottom depth, and considered the case where both the range and the depth of the node must be jointly estimated.

In the following chapter, we discuss how the multipath propagation of underwater acoustic signals gives rise to patterns that can be learned and recognized, and that can help not only find the range between a reference node and another underwater device in its vicinity, but also infer the location of the device.

4 Underwater acoustic localization using a single passive receiver

Estimating the location of an autonomous underwater vehicle (AUV) is a required step for the operation of these devices for applications like ocean exploration, control of secure areas, and environmental monitoring. In these applications, the AUV covers large areas, and its self-navigation system may drift significantly. Localizing the AUV via non-inertial systems may greatly help reduce such drift and improve the AUV's location reckoning. Localization is typically achieved through a set of fixed receiving hydrophones spread across the AUV deployment area. Yet, due to the wide area covered by the AUV during its mission, its transmissions tend to be detected very sparsely over both space and time. This is especially the case if the AUV's mission area is very large, and would imply the (expensive) deployment of a significant amount of equipment in order to cover the intended area with a sufficient density to enable reliable multilateration estimates. Instead, in order to balance a reasonable target detection probability with long term deployment constraints and costs, the coverage of large areas is typically achieved through sparse deployments. As a result, it is often the case that the signals used to detect a target are practically received only by a single receiver. Most existing algorithms to localize submerged devices require the presence of several anchor nodes [56], or prescribe message exchanges between the device and the anchors [57]. Alternatively, range estimates from a single mobile anchor have been suggested assuming knowledge of the receiver's movement between subsequent transmissions through, e.g., acceleration measurements [50]. Yet, this also requires interaction with the device to be localized.

In this contribution, we offer a solution for the challenge of localizing a non-collaborative single AUV. As opposed to localization methods that rely on a receiving array, our method assumes only the presence of a single stationary and passive receiving element, and the knowledge of the transmitted signal (for example, the structure of the AUV pinger's signals), but does not require knowledge of the pinger's transmission times. Our approach is inspired by localization algorithms based on fingerprinting [58]: these algorithms evaluate the correlation between some significant and distinguishable channel characteristics (e.g., the power-delay profile, the number of distinguishable arrivals,

the angular spectrum of these arrivals, and so forth), and the same characteristics preliminarily measured at a number of locations, and collected together in a fingerprint database. Instead, our method hinges on the spatial diversity of the sea bottom bathymetry to match the measured channel impulse response (CIR) with a set of CIRs generated through an acoustic propagation model. To that end, we target those environments where the bathymetry and the sound speed profile (SSP) in the water column induce different channel impulse responses for different emitter-receiver location pairs. This is often the case for shallow-water environments with a diverse non-flat bathymetry, but also for deeper waters where sea bottom hills, mountains, or steep slopes may exist.

We base our method on the modeling of expected acoustic CIRs for different possible locations of an acoustic source around the moored receiver. After measuring the CIR for each received signal, we correlate it with the pre-computed modeled CIRs in order to estimate the distance, depth and bearing of the transmitter. This makes it possible to point the location of the sound-emitting AUV to the position for which the modeled CIR best fits the measured CIR. We repeat the process as the AUV moves and keeps emitting signals. The result is a sequence of location estimates whose size equals the number of detected sound emissions. These location estimates are expected to be noisy, since there may be several modeled CIRs that are significantly correlated with each measured CIR. To filter this noise, we create a trellis of possible locations, which are chosen from the output of the cross-correlation between the modeled and measured CIRs, and which satisfy a given maximum AUV speed. The final path of the AUV is obtained via an efficient trellis search process similar to the Viterbi algorithm.

Our contribution is twofold:

- A localization approach for an AUV using a single receiving element;
- An efficient method to reduce the state space resulting from the cross-correlation of modeled and measured CIRs, and thereby significantly decrease the complexity of the AUV path estimation process.

We evaluate our method through both simulations (based on real bathymetry and sound speed information) and a proof-of-concept sea trial. Our results show that the proposed approach can estimate the AUV path with an acceptable localization error.

The remainder of this chapter is organized as follows: Section 4.1 provides an account of related work; Section 4.2 details the localization algorithm; Section 4.3 presents simulation results; Section 4.4 describes our proof-of-concept sea trial; Section 4.5 concludes the chapter.

4.1. Related Work

4.1.1. Techniques for Underwater Acoustic Localization

A comprehensive survey of underwater acoustic localization is presented in [47,59], and involves techniques for range estimation, bearing estimation, or both. Typical approaches to localization include long baseline (LBL) [47] (based on trilateration, and thus requiring the interaction between the device to be localized and the anchors), short baseline (SBL), usually operated from a single vessel and ultra short baseline (USBL) systems [60], that estimate the location of the device via time of arrival (Time-of-Arrival (ToA)) and angle of arrival (AoA) measurements. As the accuracy of the angle estimation process directly depends on the stability of the equipment and is sensitive to strong multipath, range-based approaches are more typically used.

Typical underwater ranging schemes rely on ToA, time difference of arrival (TDoA) or received signal strength (RSS), which is translated into distance via an acoustic propagation model [61]. ToA measurements can be obtained by separately analyzing the reflection patterns of transmitted signals [62], which can be estimated via matched filtering or by using phase-only correlation and the kurtosis metric to mitigate channel-enhanced noise [63]. Still, ToA measurements tend to be noisy due to multipath: mistaking a non-specular multipath component for the direct path is often regarded as measurement noise [64], and can be mitigated by transmitting signals having a narrow auto-correlation [65, 66], or by averaging ToA measurements over different signals [67]. Yet, instead of considering multipath as a distortion, the wealth of multipath arrivals can be exploited in passive systems in order to improve the localization accuracy, as well as to find the range of the acoustic source [1] or to localize it with multiple receivers through a propagation model [52].

In the literature, the closest approaches to our proposed scheme target localization with less than three reference nodes, often by exploiting some form of knowledge about the environment. For example, the work in [68] introduced a model-based range-bearing localization scheme that employs two receiving hydrophones. The method identifies multipath arrivals at the hydrophones and tracks them using a particle filter. An ambiguity surface is then constructed based on the expected multipath structure (derived via a ray model) and used to determine the most likely target location. To localize a source, the work in [58] proposes to match received signals against a set of fingerprints measured by an array of receivers. The authors test the feasibility their approach in a pool, which represents a static environment where fingerprints remain sufficiently stable over time. However, systematic fingerprint measurements in uncontrolled open sea environments would be more challenging, due to the rapidly changing nature of underwater acoustic channels.

Matched-field processing, a family of array processing-based methods to estimate the parameters of the ocean waveguide based on the full field structure of acoustic signals, can also be extended to underwater localization [69]. For example, the work in [70] assumes the three-dimensional knowledge of the SSP and of the bathymetry over a 600×600 km^2 area. The area is further divided in squares of side 5 km and normal mode theory is employed to predict sound propagation for a hypothetical source located in the center of each square. The sound field replicas thus obtained are matched to the acoustic field measurements collected through a 21-element vertical array, in order to infer the most likely location of the source. Matched-field localization has been recently achieved using compressed sensing (CS), which has the advantage of providing sparse solutions to inference problems using convex optimization [71]. Specifically, the proposed approach employs CS (implemented through the basis pursuit algorithm and the Lasso path) to find the best matching between field replicas and measurements, and shows that CS reliably handles coherent sources as well. Earlier, CS was considered to localize an underwater device by means of ultrawideband radio CIR fingerprinting [72]. Here, CS is implemented using the orthogonal matching pursuit and Lasso-II algorithms. Although the method achieves good localization accuracy, it remains suitable only for very short ranges, due to the strong attenuation of RF waves in salted waters. An approach to estimate the range of a source with respect to a single receiver is presented in [73]. The authors assume that a moving source transmits signals with a period Δt while moving around a hydrophone, and determine striation patterns in the function relating the signal observation time to Δt. These these patterns are then employed to infer the velocity and range of the source, based on the assumption that the ranging operations take place in a shallow-water environment with waveguide invariant $\beta = 1$. When β is unknown, the Automatic Identification System (AIS) of nearby vessels can be opportunistically used to estimate it, by relating their received signal, intensity and frequency to their known position [74].

The presence of an array of transmitters is assumed in [75], where the authors pre-compute the CIR from each transmitter to at all points of a grid that finely covers the water column along a given bearing. The location of a receiver is estimated by comparing the CIRs measured by the receiver against pre-computed CIRs. The system finally employs the determined location to tune transmit beamforming. In [76], an AUV is located by fusing AUV heading and velocity information from some external sensor with acoustic phase information. The phase is measured from a batch of signals transmitted by a fixed projector of known location and received by a single hydrophone at the AUV.

4.1.2. Differences with respect to Indoor Localization

While fingerprinting is an established localization technique for terrestrial radio networks [77, 78], one of its key assumptions is that radio measurements are repeatable and slowly varying in space [79], so that a device can actually afford to compute several

statistics of a received radio signal and fuse them into a fingerprint vector [80]. Conversely, the underwater acoustic channel tends to be much more dynamic, with several arrivals coming from multiple reflection over the surface, bottom and volume scatterers. Moreover, the spatial coherence of the underwater channel is very limited, and a transmitter could experience very different channels when communicating to a static receiver from different locations. Similar uses of ray tracing to aid indoor localization (e.g., see [81]) typically do not experience these issue, as they can rely on more stable radio channels. Filtering multiple sequential measurements through the Viterbi algorithm [82] or other techniques (such as probability maps reproducing the expected movement of mobile devices [83] or conditional random fields [84]) makes it possible to eliminate this uncertainty. However, the number of possible indoor positions to be matched by a terrestrial radio fingerprinting algorithm is usually very limited, yielding a state space of tractable size. On the contrary, in our underwater approach the location of the target could be anywhere around the location of the single receiver, yielding an order-of-10^7 state space size. This calls for methods to reduce the complexity of trellis exploration. We also remark that direction-of-arrival fingerprinting-based localization has been reconsidered in the field of millimeter wave communications (e.g., see [41, 85]), where however the devices can leverage large arrays to reliably decouple propagation paths in the received angular spectra. This is in contrast with our assumption of using a single receiving element, and remains very different from the rich CIRs usually measured in underwater communications.

4.1.3. Summary

The literature that most closely relates to our contribution is summarized in Table 4.1, where we report the requirements, description, and shortcomings of each approach. From this comparison, it becomes clear that the most prominent contribution of our approach is the localization of a moving AUV in 3D using a single receiver (and assuming only a single transmitter at the AUV). While our approach has some aspects in common with matched field processing and fingerprinting, it remains unique in that it reduces the ambiguity of the matching between measured and pre-computed CIRs through a trellis search approach similar to the Viterbi algorithm, rather than resorting to fusing information from multiple transmitters or receivers. Moreover, we only process acoustic data, and do not require any external sensors to support the localization process.

With respect to the preliminary work in [2], the algorithm presented in this chapter is much less sensitive to trellises that are not fully connected and to imperfect estimates of the initial AUV location; in addition, we include a performance verification through a sea experiment, and compare against benchmark approaches both in the simulations and in the sea experiment.

Table 4.1: Summary of the most relevant related work

Ref.	Approach	Requirements	Details	Shortcomings
[68] (2015)	Multipath tracking	– Two receivers – Known environmental parameters	– Compare multipath vs. ray model – Particle filter extracts arrivals – Ambiguity surface search	– Assumes isovelocity profile
[58] (2009)	Fingerprinting	– Fingerprint database – Broadband signal	– Database of modeled CIRs – Pattern matching of CIR measurements at different frequencies	– Multiple receivers – Maintenance of fingerprint database in ocean environments
[70] (1990)	Matched field processing	– Hydrophone array – Known environmental parameters	– Acoustic field replica computation – Gridded virtual source positioning – ML or Bartlett processing	– Requires multiple receivers to decrease ambiguity
[72] (2014)	Radio UWB fingerprinting	– UWB radio modeling to pre-compute field dictionary – Multiple antennas	– UWB fingerprinting – CS solution via orthogonal matching pursuit and Lasso-II	– Multiple antennas – Limited to short-range localization
[71] (2017)	Compressed sensing for matched field processing	– Hydrophone array – Known environmental parameters	– Acoustic field replica computation –CS solution via basis pursuit and Lasso	– Multiple receivers
[73] (2012)	Range estimation	– Single receiver – Known waveguide invariant	– Source velocity computation – Identification of point closest to receiver	– Range-only
[75] (2018)	Fingerprinting	– Known environmental parameters	– CIR computation over a fine 2D vertical grid – Matching with measurements from multiple transmitters	– Multiple projectors – Fixed-bearing localization
[76] (2014)	Acoustics-aided inertial tracking	– External bearing/speed sensor	– Extract phase from train of sine waves – Solves inverse problem to determine AUV location	– Requires accurate bearing/speed measurements

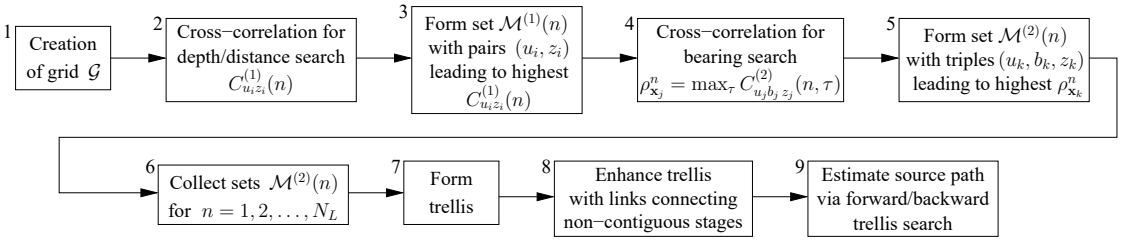
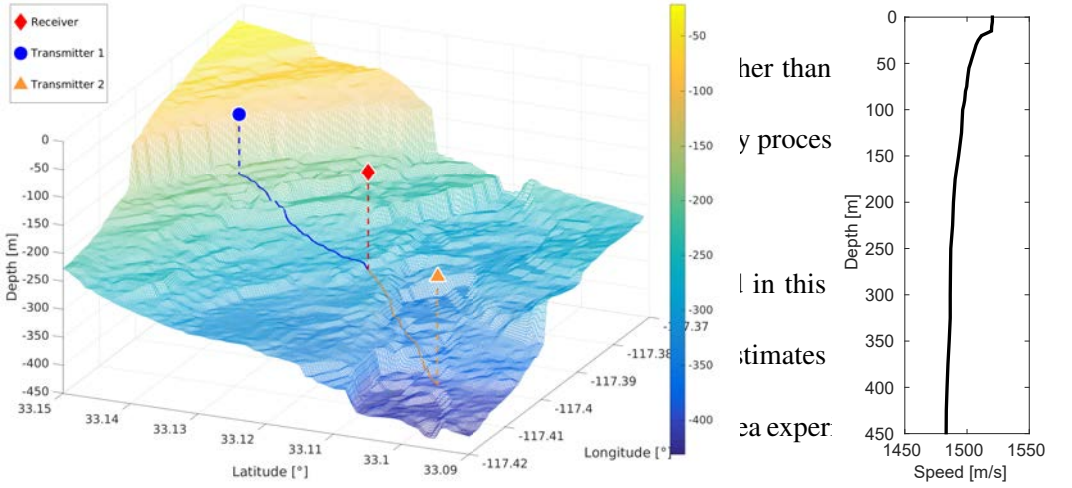


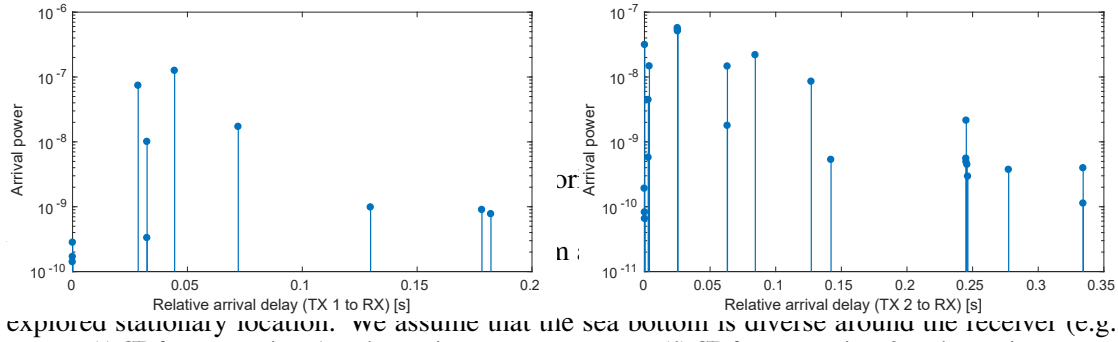
Figure 4.1: Block diagram of the AUV location and path estimation algorithm.

FIG. 1. Block diagram of the AUV location and path estimation algorithm.



(a) Map of the San Diego bay area, showing a variable bathymetry, the location of the receiver and of two transmitters, and the sea bottom profiles between each transmitter and the receiver.

(b) Example of SSP measured in the area.



(c) CIR from transmitter 1 to the receiver.

(d) CIR from transmitter 2 to the receiver.

Figure 4.2: Illustration of our single-receiver localization method. When the environment is sufficiently diverse (a), the CIRs differ significantly across different locations (c), (d). This can be leveraged for localization.

4.2. Algorithm Description

4.2.1. Key Idea

We summarize the key idea behind our algorithm with the help of the flow chart in Fig. 4.1. We operate the AUV localization algorithm from a single receiver deployed

at a known and well-explored stationary location. We assume that the sea bottom is diverse around the receiver (e.g., see Fig. 4.2), leading to a spatially-dependent CIR, which we exploit in order to estimate the location of the AUV via a fingerprinting-based location system. Since such a system requires to measure a three-dimensional database of fingerprints (which is not feasible in underwater scenarios due to the resource- and time-intensiveness of underwater acoustic measurements), we resort to a database of modeled CIRs instead. Such database is pre-computed via a numerical sound propagation model, such as the Bellhop ray tracing simulator (see Ch. 3 in [86] and [30]).

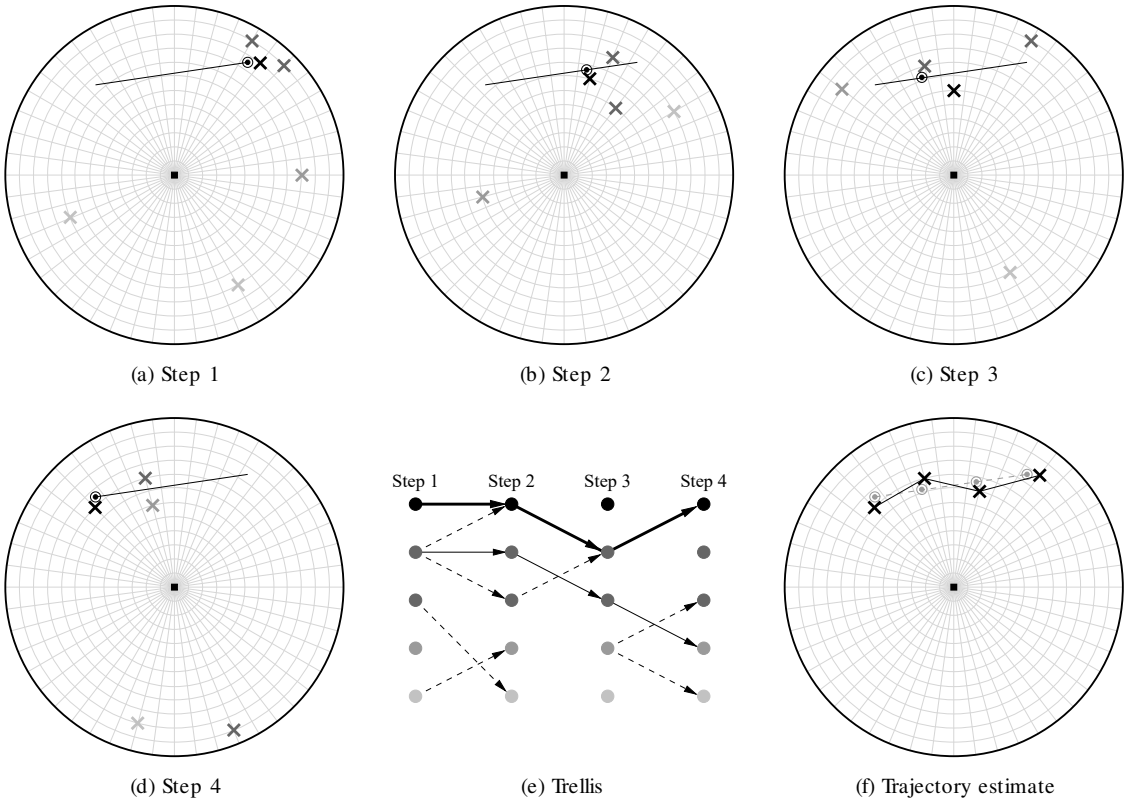
Whenever an acoustic signal is received from the AUV, we estimate the CIR of the corresponding acoustic channel and correlate it with our database. In order to reduce the complexity of this step, we first correlate the CIR with specular and surface-reflected arrivals from the modeled CIRs: this excludes bearing-dependent bottom arrivals, and allows us to retrieve a set of possible values for the AUV's depth and distance. We then compute one further round of cross-correlations, this time with the whole channel impulse response (thus including bottom reflections), for the selected depths and distances, and for every bearing value. The result is a number of possible AUV locations. We repeat the process for several subsequent acoustic signals emitted from the AUV, which may correspond to the same location, or to different locations in case the AUV is moving. Finally, we apply an efficient, low-complexity tracking mechanism in order to filter all matching locations found, and to obtain a source trajectory estimate.

Fig. 4.3 presents an example of the output of four subsequent location estimates. Each of panels (a) through (d) shows a map of the scenario. Our single receiver is shown as a centrally located square, whereas the AUV that moves along the trajectory represented as a black line. At each of the positions marked by two concentric circles, the AUV emits a signal that is employed by the receiver to compute location estimates as explained above. In panels (a)-(d), these location estimates are represented as grey crosses, where a darker grey shade indicates a higher confidence. The algorithm outputs multiple estimates for each AUV location, each with a different levels of confidence (higher confidence is represented using a darker grey shade in panels (a)–(d). Note that the the point of highest confidence may not be the closest to the actual AUV location.

To rule out spurious estimates, we order the computed locations into a trellis (Fig. 4.3), and run a forward-backward path search procedure similar to the Viterbi algorithm. In this case, the black path in Fig. 4.3e is selected, corresponding to the trajectory shown in Fig. 4.3f.

4.2.2. Preliminary Assumptions and Setup

The first step to localize the AUV is to detect is periodic pinger signals. We assume that no prior information is available about the location, the instantaneous speed, or the trajectory of the AUV, and that the AUV does not collaborate to the localization



(a) Step 1 (b) Step 2 (c) Step 3

(d) Step 4 (e) Trellis (f) Trajectory estimate

Figure 4.3: High-level illustration of the key idea behind our single-receiver localization process. Panels (a)–(d) show a sound source moving along a straight trajectory. At four locations, the source emits a signal. The receiver (located at the center of the area) measures the CIR and compares it against a database of modeled channel responses. This translates into the location estimates indicated by the crosses, where a darker grey shade indicates higher confidence. A trellis search algorithm (e) is then applied to find the most likely source path (panel (f)).

process. Hence, a solution based on updating the parameters of a dynamic model for the AUV through filtering is not an option in our scenario. We only assume that the emitted signal's waveform is either known, or can be reliably estimated, such that the channel impulse response can be evaluated. By this, we take into account received multipath, but ignore interference. Hence, our method is geared into the localization of a single source. We assume that an initial survey has been carried out in order to measure the bathymetry of the area surrounding the moored receiver with a fine resolution. The 1-meter resolution obtained by a 400-kHz multibeam sonar (see our experimental results in Section 4.4) is more than sufficient in this respect. We further require periodic direct or indirect measurements of the local SSP .

The area explored to localize the AUV is limited by the coverage of the bathymetry measurements, by the reception capabilities of the receiver, and by constraints on the

emitter's source level. This yields a bounded depth range between $z^{\mathrm{s}}_{\min}$ and $z^{\mathrm{s}}_{\max}$. We further assume the AUV is moving at an absolute maximum speed of $v^{\mathrm{s}}_{\max}$, known to the receiver. This leads to an expectation on the maximum distance traveled by the AUV between two subsequent emissions. We note that the knowledge of the AUV's maximum speed is not strictly required, but the availability of this information improves the performance and greatly reduces the complexity of our method. At different locations, indexed by $n = 1, \ldots, N_L$, the source emits acoustic signals that are detected by the receiver along with each significant multipath arrival. The locations are expressed in terms of a cylindrical coordinate system as $\mathbf{x}^{\mathrm{s}}_n = (u^{\mathrm{s}}_n, b^{\mathrm{s}}_n, z^{\mathrm{s}}_n)$ where, at location index n, $u^{\mathrm{s}}_n \in [0, u_{\max}]$ is the great-circle distance in meters between the receiver and the source, $b^{\mathrm{s}}_n \in [0°, 360°)$ is the bearing of the AUV with respect to the receiver (i.e., the angle at which the receiver sees the source, measured clockwise from due north) and $z^{\mathrm{s}}_n \in [z^{\mathrm{s}}_{\min}, z^{\mathrm{s}}_{\max}]$. We define the AUV's path as the ordered source location sequence $\{\mathbf{x}^{\mathrm{s}}_1, \ldots, \mathbf{x}^{\mathrm{s}}_{N_L}\}$.

The database of modeled CIRs set up by the receiver is computed at all points of a cylindrical grid designed to span the ranges $\mathcal{U} = \{\delta_u, 2\delta_u, \ldots, u_{\max}\}$, the bearing angles $\mathcal{B} = \{\delta_b, 2\delta_b, \ldots 360°\}$, and the depth values $\mathcal{Z} = \{z^{\mathrm{s}}_{\min}, z^{\mathrm{s}}_{\min} + \delta z, \ldots, z^{\mathrm{s}}_{\max}\}$. The set of grid points is then defined as $\mathcal{G} = \mathcal{U} \times \mathcal{B} \times \mathcal{Z}$, where we denote $\mathbf{g}_{u_i b_i z_i} \in \mathcal{G}$ as the ith grid point, $i = 1, \ldots, |\mathcal{G}|$. This corresponds to the first box in Fig. 4.1.

4.2.3. AUV Location Estimation

For each grid point $\mathbf{g}_{u_i b_i z_i}$, the receiver models the expected CIR using a propagation model. For this purpose, we employ the Bellhop ray tracing software (see Ch. 3 in [86] and [30]). Bellhop is an established solution to numerically solve pressure wave propagation equations by taking into account boundary conditions. In particular, Bellhop can factor in, among others: the SSP at multiple points throughout the water body section that joins the transmitter to the receiver; the relevant bathymetry in the area, including abrupt changes; the shape of surface waves; and the geo-acoustic properties of the sea bottom sediments. Bellhop has been used to model acoustic channels in different communication contexts, and served as the basis for more complex models (e.g., see [87], [88]). In our context, Bellhop yields accurate time-of-arrival information for each acoustic path, and sufficiently accurate complex amplitude information, so that the outcome of correlation operations can be trusted. We will show that Bellhop offers sufficiently reliable CIR modeling in a sea trial in Section 4.4.

The output of Bellhop includes a list of expected multipath arrivals, along with their amplitude, phase, delay, and reception angle. Moreover, for each arrival, Bellhop reports the list of bottom and surface reflections it incurred. This information is employed to construct two modeled responses, namely a partial CIR $h^{(1)}_{u_i z_i}(t)$, containing only the

specular and surface-reflected arrivals,[1] and the complete CIR $h^{(2)}_{u_i b_i z_i}(t)$. As the specular and surface-reflected arrivals are practically independent of the bearing of the AUV relative to the receiver, and rather depend only on the SSP , on u_i, and on z_i, the subscript b_i has been dropped in $h^{(1)}_{u_i z_i}(t)$.

From the modeled CIRs, the receiver obtains two separate fingerprints, $h^{(1)}_{u_i z_i}$ and $h^{(2)}_{u_i b_i z_i}$. When the source is at location $\mathbf{x}_n$, its emitted signal is received as

$$r_n(t) = \hat{h}_{u_n b_n z_n}(t) \otimes s(t) + \nu(t)\,, \tag{4.1}$$

where $\hat{h}_{u_n b_n z_n}(t)$ is the CIR estimated from a received signal, $s(t)$ is the emitted signal waveform, $\nu(t)$ is the ambient noise, and $\otimes$ denotes convolution. The receiver then computes

$$f^{(1)}_{u_i z_i} = h^{(1)}_{u_i z_i}(t) \otimes s(t) \tag{4.2a}$$

$$f^{(2)}_{u_j b_j z_j} = h^{(2)}_{u_j b_j z_j}(t) \otimes s(t)\,, \tag{4.2b}$$

and matches $r_n(t)$ against the fingerprints $f^{(1)}_{u_i z_i}$ and $f^{(2)}_{u_j b_j z_j}$ corresponding to the grid points in $\mathcal{G}$ as follows.

For each point $(u_i z_i)$ in the grid, we compute the normalized correlation

$$C^{(1)}_{u_i z_i}(n) = \frac{\displaystyle\int_0^{+\infty} r_n(t)\, f^{(1)}_{u_i z_i}(t-\tau)\,\mathrm{d}t}{\left(\displaystyle\int_0^{T} r_n(t)^2 \mathrm{d}t \int_0^{+\infty} f^{(1)}_{u_i z_i}(t)^2 \mathrm{d}t\right)^{1/2}}\,, \tag{4.3}$$

where T is the signal's duration, and τ is the time epoch corresponding to the correlation maximum. Note that in (4.3) we perform a normalized correlation to reduce sensitivity to CIRs characterized by different power attenuation. Define $\mathcal{M}^{(1)}(n)$ as the set of all pairs (u_j, z_j) corresponding to those $C^{(1)}_{u_i z_i}(n)$ that exceed a certain threshold Θ_D, $\forall\,(u_i, z_i) \in \mathcal{G}$, where we set Θ_D, using the analysis in [89]. We remark that we do not limit set $\mathcal{M}^{(1)}(n)$ to contain just the coordinates of the single grid point yielding the maximum correlation. In fact, at this point, the estimation of the correct distance and depth may be hindered by the lack of, e.g., the specular arrival, which can occur in the presence of SSP patterns with a sufficiently steep gradient and for a sufficiently large distance between the AUV and the receiver (e.g., see the example on page 46 of Bellhop's manual [30]). Including a number of possible matching locations is more robust against such errors. The above steps correspond to boxes 2 and 3 in Fig. 4.1.

For each $(u_j, z_j) \in \mathcal{M}^{(1)}(n)$, and $\forall\, b \in \mathcal{B}$, we compute the following normalized

[1]Note that this information is part of Bellhop's standard output data, and that we do not need to track any propagation history analysis for any of the components of the measured CIR.

correlations:

$$C^{(2)}_{u_j b_j z_j}(n,\tau) = \frac{\displaystyle\int_0^{+\infty} r_n(t)\, f^{(2)}_{u_j b_j z_j}(t-\tau)\,\mathrm{d}t}{\left(\displaystyle\int_0^{T} r_n(t)^2 \mathrm{d}t \int_0^{+\infty} f^{(2)}_{u_j b_j z_j}(t)^2 \mathrm{d}t\right)^{1/2}} \,. \tag{4.4}$$

Call

$$\rho^n_{\mathbf{x}_j} \;=\; \max_{\tau}\, C^{(2)}_{u_j b_j z_j}(n,\tau)\,, \tag{4.5}$$

and define $\mathcal{M}^{(2)}(n)$ as the set of all triples $\mathbf{p}_k = (u_k, b_k, z_k)$ corresponding to the $R^{(2)}$ highest values of $\rho^n_{\mathbf{p}_k}$ $\forall\,(u_j, z_j) \in \mathcal{M}^{(2)}(n)$ and $\forall\, b \in \mathcal{B}$, where $R^{(2)}$ is a user-defined parameter (in our performance evaluation, we set $R^{(2)} = 70$). The above steps correspond to boxes 4 and 5 in Fig. 4.1.

In ideal conditions, e.g., with an extremely dense grid $\mathcal{G}$, in the absence of noise, and with perfect environmental information, it would be enough to limit set $\mathcal{M}^{(2)}(n)$ to the coordinates of the point $\mathbf{p}_k = (u_k, b_k, z_k)$ for which $\rho^n_{\mathbf{p}_k}$ is highest. However, in any practical scenario, the grid point closest to the actual position of the AUV might not yield the highest correlation due to noise, outdated environmental information, or a combination of both. In this perspective, it is convenient to set $R^{(2)}$ to some large value. On the other hand, it is computationally infeasible to have an exceedingly large set $\mathcal{G}$. For this reason, we reduce the complexity of the search whenever possible by limiting the location search area through a bound on the distance between the AUV and the receiver. For example, if the source level is known, this bound can be obtained based on a Received Signal Strength Indicator (RSSI) as in [40]. Furthermore, in the following we present a filtering scheme that reduces the complexity of path estimation.

4.2.4. AUV Path Estimation

After determining the possible matching locations $\mathcal{M}^{(2)}(n)$ for $n = 1, \ldots, N_L$, we proceed to find the most likely sequence of AUV's locations among all possible options using a path estimation algorithm. Without prior information about the AUV motion pattern, we avoid assuming a dynamic model solved by filtering, but rather work on a trellis such as the one shown in Fig. 4.4. The trellis has N_L stages, one for each transmission received from the AUV. In each stage, different nodes represent different estimated locations, so that the first stage of the trellis represents all location estimates for the first detected signal from the AUV (set $\mathcal{M}^{(2)}(1)$), the second stage contains the estimates in set $\mathcal{M}^{(2)}(2)$, and so forth until the last stage, which contains the estimates in $\mathcal{M}^{(2)}(N_L)$. We assign a confidence index to each node in the trellis (the value of the normalized cross-correlation between the modeled and measured channels, see (4.5)), and organize them into a $R^{(2)} \times N_L$ matrix $\mathbf{T}$ (boxes 6 and 7 in Fig. 4.1). Both the nodes in the ith trellis stage and the entries in the ith column of $\mathbf{T}$ are sorted in order of decreasing

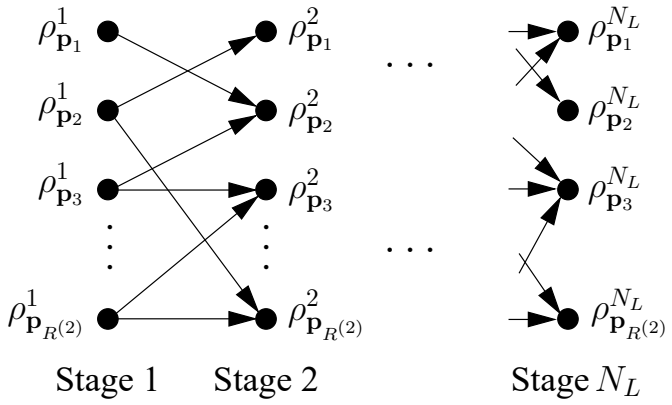


Figure 4.4: Example of trellis employed by the tracking algorithm for the source path estimation. Each node represents a location estimate. Trellis links exist only among locations that are closer than the maximum distance $d_{\max}$ covered by the AUV when traveling at full speed between subsequent signal transmissions.

confidence, i.e., $[\mathbf{T}]_{1,i} = \rho^i_{\mathbf{p}_1}$, $[\mathbf{T}]_{2,i} = \rho^i_{\mathbf{p}_2}$, $[\mathbf{T}]_{R^{(2)},i} = \rho^i_{\mathbf{p}_{R^{(2)}}}$, and

$$\rho^i_{\mathbf{p}_1} > \rho^i_{\mathbf{p}_2} > \cdots > \rho^i_{\mathbf{p}_{R^{(2)}}} \,. \tag{4.6}$$

4.2.4.1. Setting the Path Weights

The objective of path estimation is to find the best sequence of nodes across consecutive trellis stages. To that end, a link exists between an entry in stage n and an entry in stage $n+1$ if the locations represented by these nodes are closer than the maximum distance the AUV could cover when traveling at full speed $v^{\mathrm{s}}_{\max}$ between the nth and the $(n+1)$th signal detections (recall that the maximum absolute speed is assumed to be known). Formally, call $e_{\ell_n \ell_{n+1}}$ the edge that connects the ℓ_nth node at stage n in the trellis (entry in column n of $\mathbf{T}$) and the ℓ_{n+1}th entry at column $n+1$. Call $A(e_{\ell_n \ell_{n+1}}) = \mathbf{p}_{\ell_n}$ and $S(e_{\ell_n \ell_{n+1}}) = \mathbf{p}_{\ell_{n+1}}$ the ancestor and the successor of edge $e_{\ell_n \ell_{n+1}}$, respectively. Define the edge weight as

$$\sigma(e_{\ell_n \ell_{n+1}}) = \begin{cases} 1\,, & \text{if } d(\mathbf{p}_{\ell_n}, \mathbf{p}_{\ell_{n+1}}) \le d_{\max} \\ \dfrac{d_{\max}}{d(\mathbf{p}_{\ell_n}, \mathbf{p}_{\ell_{n+1}})}\,, & \text{if } d_{\max} < d(\mathbf{p}_{\ell_n}, \mathbf{p}_{\ell_{n+1}}) \le 1.5\, d_{\max} \\ 0\,, & \text{otherwise} \end{cases} \tag{4.7}$$

where $d(\mathbf{x}, \mathbf{y}) = \|\mathbf{x} - \mathbf{y}\|_2$ is the Euclidean distance between locations $\mathbf{x}$ and $\mathbf{y}$, t_n and t_{n+1} are the reception epochs of the nth and $(n+1)$th detected signals, respectively, and $d_{\max} = v^{\mathrm{s}}_{\max}(t_{n+1} - t_n)$ is the maximum distance that the AUV could have traveled between time epochs t_n and t_{n+1}. Only edges with non-zero weights are considered for

path estimation. To form a continuous path, we require connected edges. In particular, if for edge $e_{\ell_n \ell_{n+1}}$ it occurs that its ancestor $\mathbf{p}_{\ell_n}$ is not successor of any edge $e_{\ell_{n-1}\ell_n}$, or that its successor $\mathbf{p}_{\ell_{n+1}}$ is not ancestor of any edge $e_{\ell_{n+1},\ell_{n+2}}$, then the weight of edge $e_{\ell_n \ell_{n+1}}$ is set as zero, and the edge is removed from the trellis.

We remark the similarities between the path estimation algorithm and the Viterbi algorithm for tracking within a trellis (see also [90]). While the Viterbi algorithm would yield the optimal solution, it would include all grid points in $\mathcal{G}$ in each stage of the trellis. This would require $|\mathcal{G}|$ entries in each column of $\mathbf{T}$, which would compound to a huge state space and imply an exceedingly high computational complexity, especially if $|\mathcal{G}|$ is very large. In addition, solving through the Viterbi algorithm would require an estimation for the emission and transition probabilities, which involves some hard assumptions on the CIR and noise models. Instead, our version relies on confidence indices, and makes it possible to trim the state space according to physical movement constraints. This leads to a significant performance improvement and to a feasible path estimation complexity.

4.2.4.2. Finding the Best Path

Let $\mathcal{E}(n) = \{e_{\ell_n \ell_{n+1}}\}$ be the set of edges that link a node in stage n of the trellis to a node in stage $n+1$, and use (4.7) to define the following metric for each edge

$$\lambda(e_{\ell_n \ell_{n+1}}) = \rho^n_{\mathbf{p}_{\ell_n}} \rho^{n+1}_{\mathbf{p}_{\ell_{n+1}}} \sigma(e_{\ell_n \ell_{n+1}}) \,, \tag{4.8}$$

where the confidence indices are taken from $\mathbf{T}$. Define a generic path on the trellis as

$$\Psi = \{e_1, \ldots, e_{N_L}\} \,, \tag{4.9}$$

where e_i is a shorthand for $e_{\ell_i \ell_{i+1}} \in \mathcal{E}(i)$, and all edges are such that $S(e_i) = A(e_{i+1})$, $i = 1, \ldots, N_L - 1$. Define the overall path metric as

$$\Lambda(\Psi) = \frac{\prod_{i=1}^{N_L - 1} \lambda(e_i)}{\prod_{i=1}^{N_L - 2} \rho^i_{S(e_i)}} \,, \tag{4.10}$$

i.e., as the product of the confidence metrics for all edges that belong to the path, divided by the confidence of intermediate nodes in order to avoid accounting for them twice. The path estimate is finally found as

$$\widehat{\Psi} = \arg\max_{\Psi} \Lambda(\Psi) \,, \tag{4.11}$$

and we indicate the sequence of locations traversed by $\widehat{\Psi}$ as $\{\hat{\mathbf{x}}_1, \hat{\mathbf{x}}_2, \ldots, \hat{\mathbf{x}}_{N_L}\}$.

As a means of measuring the discrepancy between the true and the estimated sequence of AUV's locations, we consider the Root Mean Square (RMS) point-wise distance between corresponding points of the true and estimated paths. Formally,

$$\varepsilon^{d}_{\widehat{\Psi}} = \left(\frac{1}{N_L} \sum_{n=1}^{N_L} d(\hat{\mathbf{x}}_n, \mathbf{x}^{\mathrm{s}}_n)^2 \right)^{1/2} . \tag{4.12}$$

We also convey the source bearing estimation effectiveness of our approach via the bearing error

$$\varepsilon^{a}_{\widehat{\Psi}} = \frac{1}{N_L} \sum_{n=1}^{N_L} |\hat{b}_n - b^{\mathrm{s}}_n| \, , \tag{4.13}$$

where $d(\cdot, \cdot)$ denotes the distance between two points in the cylindric coordinate system.

4.2.4.3. Refinement

In this section, we present two refinements to the above algorithm. The first refinement relates to the possible case that there exists no edge with a non-zero weight connecting two trellis stages n and $n+1$. This would lead to a partitioning of the trellis. We correct for these cases by allowing stage $n-1$ to directly connect to stage $n+1$. Specifically, the corresponding edge $e_{\ell_{n-1}\ell_{n+1}}$ will have a weight equal to

$$\lambda(e_{\ell_{n-1}\ell_{n+1}}) = \rho^{n-1}_{\mathbf{p}_{\ell_{n-1}}} \rho^{n+1}_{\mathbf{p}_{\ell_{n+1}}} \sigma(e_{\ell_{n-1}\ell_{n+1}}) \, , \tag{4.14}$$

where $\sigma(\cdot)$ is the same as in (4.7).

The above recovery mechanism is further enhanced to handle cases of broader trellis partitioning due to bursts of errors. These bursts are caused by strong noise from, e.g., a nearby vessel or waves, or due to erroneous bathymetry information at some locations. The result of such bursts are sets of short paths for which the maximization in (4.11) is not optimal, i.e., the problem becomes non-convex. Considering this case, we increase the number of paths in Ψ through our second refinement procedure as follows.

We start by observing that, from the perspective of path finding, we can calculate paths by taking sets of estimated locations either in order they occur in time, or by reversing this order. In other words, the trellis stages in Fig. 4.4 and the corresponding columns in $\mathbf{T}$ can be flipped, such that the first contains location estimates in $\mathcal{M}^{(2)}(N_L)$, the second contains the estimates in $\mathcal{M}^{(2)}(N_L - 1)$, and so forth until the last column, which contains the samples in $\mathcal{M}^{(2)}(1)$. Call Ψ_F a forward path on the trellis traversing locations $\{\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_{N_L}\}$, and call Ψ_B a backward path computed on the reversed trellis, traversing locations $\{\mathbf{y}_{N_L}, \mathbf{y}_{N_L-1}, \ldots, \mathbf{y}_{N_1}\}$. If $\widehat{\Psi}_F$ and $\widehat{\Psi}_B$ are the best forward and backward paths according to (4.11), respectively, we set the final path estimate $\widehat{\Psi} = \widehat{\Psi}_F$ if $\Lambda(\widehat{\Psi}_F) > \Lambda(\widehat{\Psi}_B)$, and $\widehat{\Psi} = \widehat{\Psi}_B$ otherwise. In case of significant interruptions in the trellis structure, the above scheme increases the probability to find the correct path. The

scheme is also beneficial if the estimate of the initial location on the forward path is incorrect, making the path search diverge to a mostly wrong sequence of locations. In case of a well connected trellis, instead, the scheme is likely to find the same path twice, with no effect on the accuracy of the algorithm.

The complexity of the algorithm relates to the number of correlation operations and to the trellis search. For each received source signal, the algorithm computes $\kappa|\mathcal{G}| + \mathcal{O}(|\mathcal{M}^{(1)}(n)|) \approx \kappa|\mathcal{G}|$ correlations in order to extract the possible position estimates in set $\mathcal{M}^{(2)}(n)$, where κ is a proportionality factor that account for the search space reduction enabled, e.g., by RSSI bounding considerations as mentioned in Section 4.2.3. For a signal of bandwidth–time duration product BT, the complexity of each normalized cross correlation is $\mathcal{O}\left(B^2\right)$. With $\mathcal{O}(N_L|\mathcal{M}^{(2)}(n)|)$ operations for the trellis search, the overall complexity is $\mathcal{O}\left(N_L|\mathcal{M}^{(2)}(n)| + |\mathcal{G}|B^2\right)$. Comparing this with the complexity of the Viterbi algorithm, i.e., $\mathcal{O}\left(N_L|\mathcal{G}|^2\right)$ (see also [82]), a significant complexity reduction exists.

4.2.5. Discussion

Our method considers the practical case of observing an unknown target. This target can move in any dynamic pattern and even irregularly. Hence, we avoid evaluating its position through filtering, and rather follow a trellis search approach over the confidence indices. This also means that the path found from all feasible solutions is the one with maximum overall confidence index, and thus isolated positions associated to a high confidence value will not be chosen. This is appropriate, since we are looking for a systematic solution, rather then an individual match. Our solution for the trellis search takes a suboptimal approach by taking into account sets of only two nodes. This has the drawback that a single node in the trellis may have a higher impact than it should. Yet, without prior knowledge of the target and to keep the calculations feasible we avoid other solutions in the form of, e.g., dynamic programming. Further, we note that the accuracy of our method depends on the quality of the channel estimation process, which improves with the bandwidth of the emitted signal.

For channel modeling, we use the bathymetry and the sound speed profile. Without up-to-date information about instantaneous sea conditions, we avoid a time-varying propagation model and use instead a static model. Instead, the time-variation of the channel is taken into account by the AUV's motion, both by calculating different channels for different locations, and by using the maximum velocity $v^{\mathrm{s}}_{\max}$. This parameter trades off complexity with performance, as higher values for the maximum speed corresponds to additional possible paths in the considered trellis. Another significant assumption is the ability to estimate the channel from the received signals. Clearly the performance of our approach depends on the accuracy of such estimation. While channel estimation is beyond the scope of this work, possible techniques for such an estimation can be rake

receivers [91], blind source separation [92], or cyclo-stationary analysis [93], to name a few options.

4.3. Simulation Results

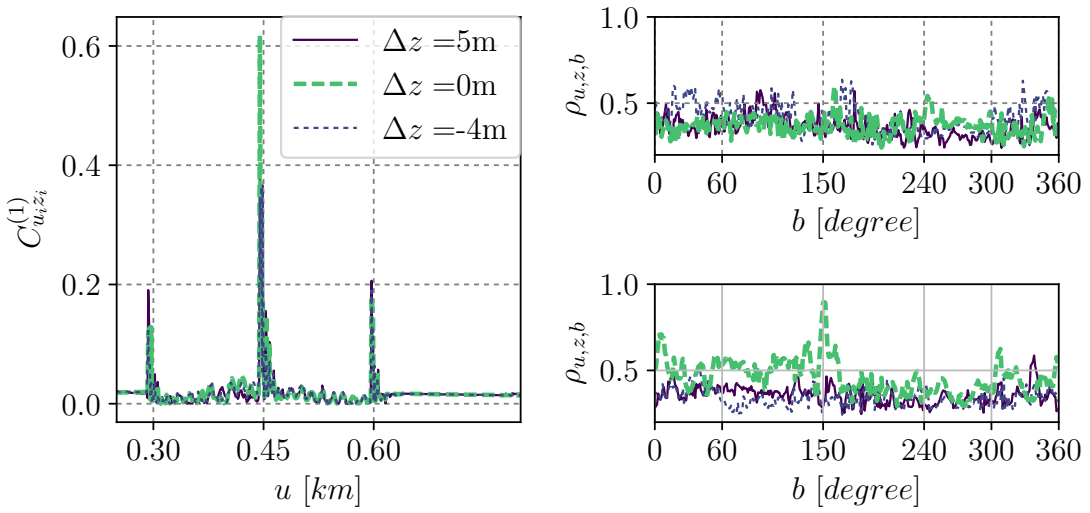
4.3.1. Scenario and Parameters

For our simulations, we consider a portion of the San Diego bay area, off the coast of US's southern California, which is a well-explored area. We place the receiver at the coordinates $[32.9390°N, 117.2816°W]$. We take the area's bathymetry data from the US Coastal Relief model [94] (revealing that the average depth in the area is about 50 m), and employ an SSP sample taken at the observed area. The SSP has a downward-refractive shape, typical of shallow Californian waters during warm seasons, as depicted in Fig. 4.2b. We assume that the water surface is flat.

In our simulations, we deploy both the receiver and the source at depths of 10 m. Still, we remark that the receiver is not aware of the source's depth. The simulation starts by deploying the emitting source at random in the area at a range of 500 m from the receiver. The source then chooses a bearing uniformly at random and moves along the corresponding direction with constant speed chosen at random for the time required to carry out 10 transmissions. The locations $\mathbf{x}_n^{\mathrm{s}}$ and $\mathbf{x}_{n+1}^{\mathrm{s}}$, where two subsequent emissions take place, are chosen uniformly at random such that $d(\mathbf{x}_n^{\mathrm{s}}, \mathbf{x}_{n+1}^{\mathrm{s}}) \leq d_{\max}$, and we set $d_{\max} = 50$ m.

The fingerprint grid pre-computed by the receiver spans a total range $u_{\max} = 1.5$ km around the receiver, with a resolution of 1 m. The whole azimuthal plane is considered, with a resolution of 1°, and the CIRs are computed for all depths between 5 m and 15 m, also with a resolution of 1 m. This choice leads to a total of about 6 million points in set $\mathcal{G}$, and emphasizes the need for our path finding algorithm, as it has much lower complexity than the regular Viterbi algorithm.

The signal transmitted by the source, $s(t)$, is chosen to be a linear chirp signal of duration 100 ms and bandwidth of 10 kHz, centered at a carrier frequency of 12 kHz. Based on these signal parameters and using the analysis in [89], for the computation of (4.3) and the formation of set $\mathcal{M}^{(1)}(n)$, we choose $\Theta_D = 0.1 \ \forall\, n$. For each emission from a given source-receiver location pairs, the channel impulse response is computed through Bellhop [30], using as parameters the SSP and the available bathymetry samples along the direction from the source to the receiver. The ambient noise at the receiver is modeled as an additive white Gaussian process, whose power is tuned so as to achieve a prescribed signal-to-noise ratio (SNR).



(a) Distance-depth estimation

(b) Bearing estimation for a wrong (top) and correct distance (bottom). Key: see Fig. 4a.

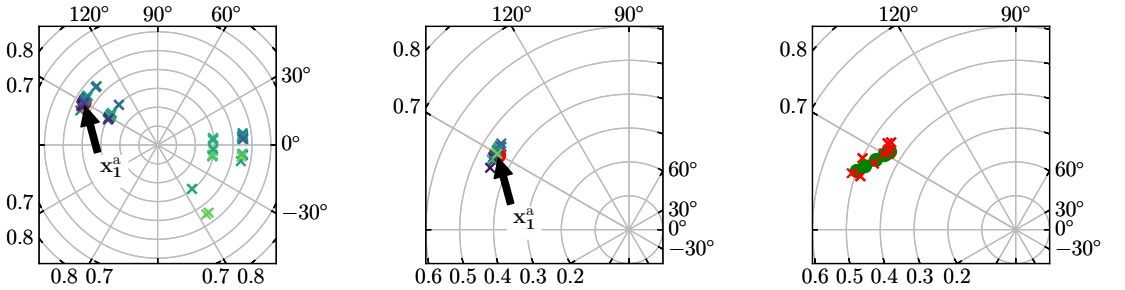
Figure 4.5: Example of correlation values for $\mathbf{x}^{\mathrm{s}} = [446\,\mathrm{m}, 150°, 10\,\mathrm{m}]$ at an SNR of 30 dB, for different value of the offset Δz between the actual depth and the tested depth.

4.3.2. Examples

A sample result from (4.3) is shown in Fig. 4.5. We observe a clear peak suggesting that the source is located at a distance of approximately 450 m from the receiver, at a depth of 10 m. This is due to the presence of all expected specular and surface-reflected arrivals in the received signal. If, e.g., the specular arrival were missing, the correlation peak at 450 m would not be as high. This is why we consider all three significant peaks, including those at about 300 m and 600 m, and for all depths where such peaks exceed Θ_D.

To populate set $\mathcal{M}^{(2)}(n)$, we set $R^{(2)} = 70$. A sample computation of (4.4) for some range-depth pairs in $\mathcal{M}^{(1)}(n)$ is shown in Fig. 4.5. While in this particular case a peak stands out corresponding to the correct bearing of about 150°, often such a favorable result does not occur. The chosen value of $R^{(2)}$ makes it possible to considerably increase the probability that the actual bearing is included in $\mathcal{M}^{(2)}(n)$, while keeping the computational effort controlled.

An example of the output of the algorithm for a single step n and a whole path is provided in Fig. 4.6. Each panel shows a view of the area around the receiver, which is located at the center of panel 4.6a or towards the bottom-right corner in panels 4.6b and 4.6c. Fig. 4.6a shows the location estimates for the first signal emission, obtained from the set $\mathcal{M}^{(2)}(1)$ as described above; Fig. 4.6b shows the location estimates after filtering out those that are not part of any edge set $\mathcal{E}(n-1)$ or $\mathcal{E}(n)$, $\forall\, n$; Fig. 4.6c shows the final path estimate compared to the actual path of the source. For the numerical



(a) Location estimates after cross-correlation

(b) Filtered estimates

(c) Whole path estimate (red crosses) vs. actual path (green dots)

Figure 4.6: Example of cross-correlation results for model-based range-bearing localization (distances in [km]). Our method successfully rules out spurious locations and achieves an accuracy of $\varepsilon^d \approx 160$ m and $\varepsilon^a \approx= 4.4°$.

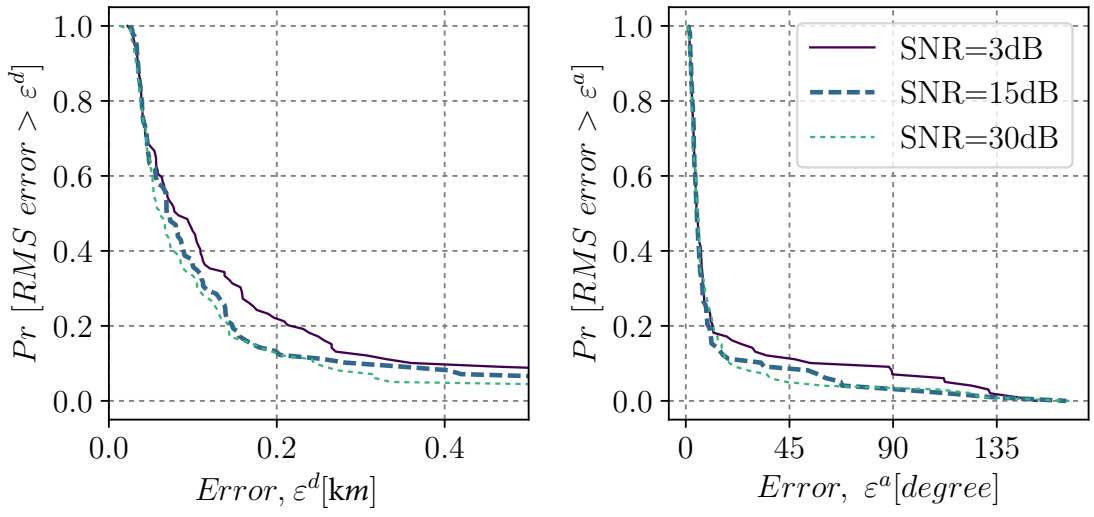
results discussed in the following, we consider a Monte-Carlo set of 100 source paths, which corresponds to 1000 signal emissions.

4.3.3. Localization Accuracy Under Varying SNR

We start our performance evaluation by running our algorithm in the presence of exact environmental data under different SNR values. The complementary cumulative distribution functions (CCDFs) of the Root Mean Square Error (RMSE) affecting the distance and bearing estimates are shown in Figs. 4.7a and 4.7b, respectively. Thanks to the perfect knowledge of both the bathymetry and the SSP in the observed area, neither result shows a significant dependence on the SNR, even after decreasing it to as low as 3 dB, which tends to make additional peaks appear in the correlation outputs. We observe that the average RMSE varies from about 120 m for an SNR of 30 dB, up to about 170 m for an SNR of 3 dB, with a median error around 80 m, which is satisfactory given the grid resolution employed and the use of a single receiving element. The bearing estimation results show even higher accuracy, with a mean estimation error $\varepsilon_a < 20°$ even for an SNR of 3 dB, and a median error of less than 10°.

4.3.4. Localization Accuracy Under Imperfect Bathymetry Data

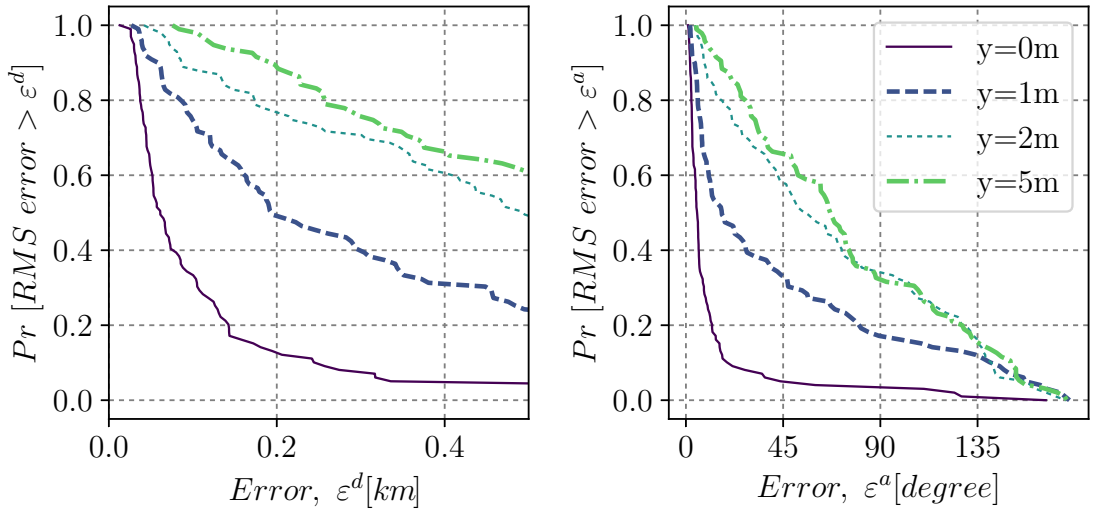
The above simulation results show accurate localization for different SNR levels. However, the results are obtained assuming perfect bathymetry and sound speed profile knowledge. In our setting, the receiver is an anchored station, e.g., a marine observatory, and thus we argue that accurate sound speed measurements are possible and do often exist in such marine observatories (e.g., see [95]). Still, while fine-gridded bathymetry mapping can be made around the observatory, small errors and outdated measurements in the resulting depth map may exist. We now explore the sensitivity of our localization method to imperfect bathymetry information.



(a) CCDF of the distance RMSE

(b) CCDF of the bearing RMSE

Figure 4.7: Accuracy of the path estimation algorithm in the presence of exact environmental data, for different values of the Signal-to-Noise Ratio (SNR). Even for low values of the SNR the arrival structure in the CIRs does not change considerably, and has no significant effects on performance.



(a) CCDF of the distance RMSE

(b) CCDF of the bearing RMSE

Figure 4.8: Accuracy of the path estimation algorithm in the presence of imperfect bathymetry data. Erroneous bathymetry significantly affects the algorithm's performance. For limited errors ($y = 1$ m) the results are still viable for several applications.

In the following analysis, to each true bathymetry sample we add an offset drawn uniformly at random in the interval $[-y, y]$, where y (in m) is a tunable parameter. We collect a Monte-Carlo set of 100 source paths and compute the CCDFs of the RMSE for both the distance and the bearing. The results are shown in Figs. 4.8. We observe

that, as expected, mismatched bathymetry data worsens the path estimation performance. However, for a limited offset on bathymetry samples, up to $y = 1$ m, the median RMS distance error remains below 200 m (or 6% of the total observed area), which is still a reasonably good result given the presence of a single receiving element. Instead, an error of up to $y = 5$ m yields comparatively worse performance. However, we remark that this is an extreme case, as such an error amounts to about 10% of the average sea bottom depth in the area, and current sea bottom mapping systems typically ensure sub-meter bathymetry measurements for depths of less than 200 m (e.g., this is the case for Kongsberg Maritime's 400 kHz EM 2040 multibeam sonar system we use in our sea experiment).

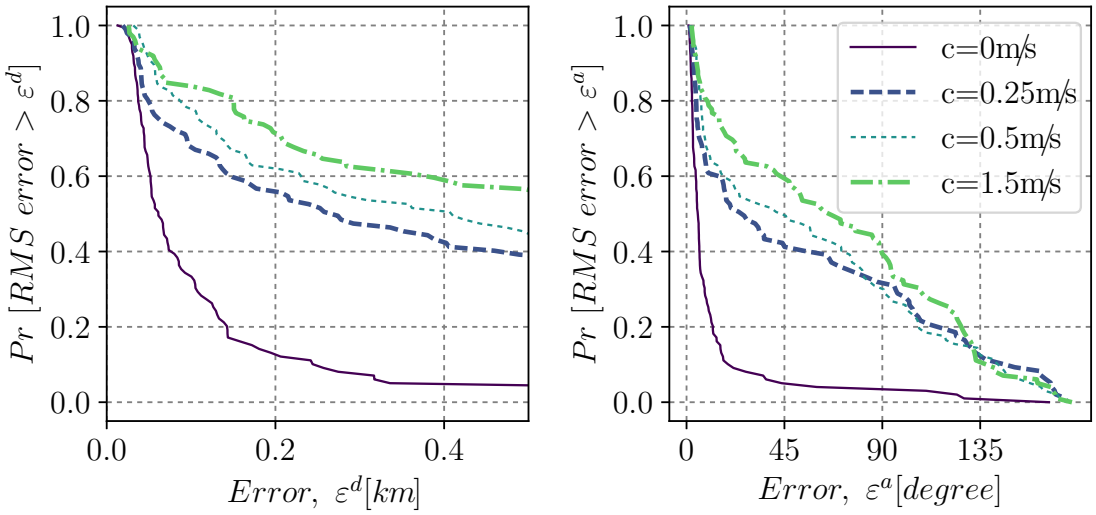
Similar conclusions as for the distance-based sensitivity of the algorithm can be drawn also for the bearing estimation error. Fig. 4.8 shows that for $y = 1$ m, the increase in the median bearing estimation error is roughly 20°, and increases to roughly 55° for $y = 5$ m. This result emphasizes the need for accurate bathymetry information. Still, we argue that even such rough localization estimates can be instrumental for some applications. For example, security or environmental monitoring systems, where even a rough estimate can trigger a more accurate investigation by human personnel or more complex detection mechanisms; or fauna and habitat monitoring applications, where it is often sufficient to find the approximate path of a vocalizing animal.

4.3.5. Localization Accuracy Under Imperfect SSP Data

In order to evaluate the impact of imperfect SSP data on the performance of our algorithm, we add an offset drawn uniformly at random in the interval $[-c, c]$ to each true SSP sample, and carry out Monte-Carlo simulations for different value of c.

The CCDFs of the distance and bearing RMSE are provided in in Figs. 4.9a and 4.9b, respectively. While the chosen values for c preserve the general downward-refractive properties of the SSP , even a small value tends to cause significant changes in the structure of multipath arrivals. For $c = 0.25$ m/s, we already observe a median distance error of about 250 m and a median bearing error of about 30°. It could be argued that these values are still practical for rough localization applications, where the only need is to know whether the AUV is practically following a desired trajectory or is falling significantly off track. As expected, increasing c tends to reduce both the distance and the bearing estimation accuracy. This emphasizes the need to maintain SSP estimates updated at the receiver, and to recompute the CIRs in the grid $\mathcal{G}$ in the presence of significant changes.

We finally remark that, besides bathymetry and SSP, high sea states may induce significant surface waves that would also contribute to modifying surface-reflected multipath components of the modeled and measured CIRs. Since it is not feasible to create different modeled CIR sets $\mathcal{G}$ for many realizations of the surface waves and for



(a) CCDF of the distance RMSE

(b) CCDF of the bearing RMSE

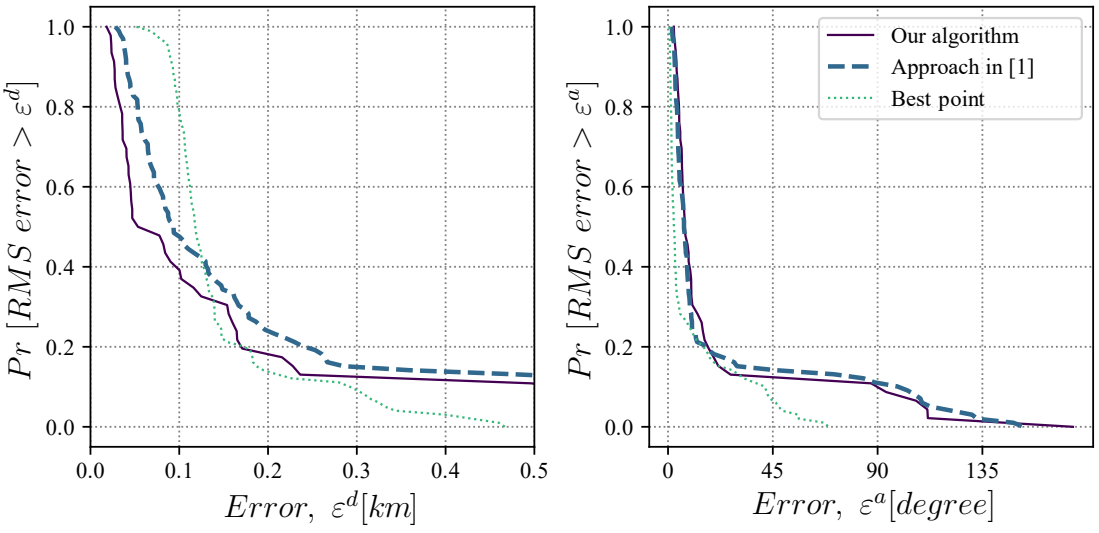
Figure 4.9: Accuracy of the path estimation algorithm in the presence of imperfect SSP data. Increasing deviations from the actual SSP tend to significantly change the multipath arrival structure. For limited deviations, the median localization and angle error remain acceptable.

different sea states, in this case it would be appropriate to skip the correlation-based depth/distance estimation that results in sets $\mathcal{M}^{(1)}(n)$. Instead, it would be possible to populate $\mathcal{M}^{(1)}(n)$ with all pairs of depth and distance values that satisfy RSSI bounds, and then proceed with the computation of the cross-correlations that lead to set $\mathcal{M}^{(2)}(n)$.

4.3.6. Comparison against benchmark localization schemes

We conclude our evaluation with a comparison among our algorithm, its preliminary version in [2], and a benchmark scheme that, for every location index n corresponding to a signal received by the buoy, chooses the most likely source location as the grid point in $\mathcal{M}^{(2)}(n)$ yielding the largest correlation value (dubbed "best point" in the following). This is akin to a classical fingerprint-based localization approach, where the fingerprint is defined as the value of (4.4). Fig. 4.10 shows the CCDFs of the distance and bearing RMSE for all above approaches carrying out all operations listed in Section 4.2, including the forward-backward refinement of Section 4.2.4.3.

The results confirm the expectation that our approach achieves a lower estimation error. This is also due to the forward-backward search refinement, which reduces the chance that a comparatively low correlation value in the first point of the source's trajectory hampers the correct estimation of the whole path. Specifically, the median distance RMSE decreases from about 100 to about 50 m, in the presence of a comparable angle RMSE. Although the best point scheme provides a good estimation of the bearing



(a) CCDF of the distance RMSE

(b) CCDF of the bearing RMSE

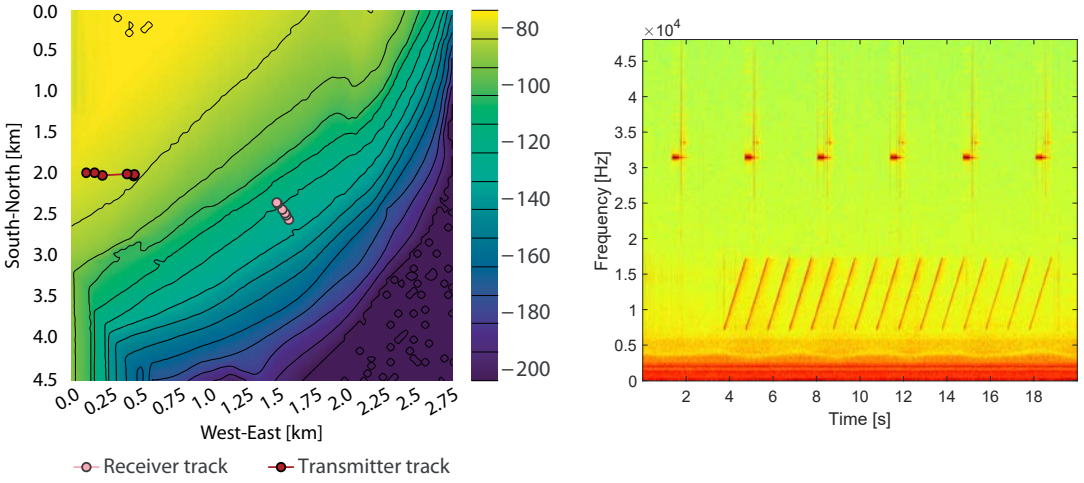
Figure 4.10: Comparison among different location estimation schemes: our algorithm, the preliminary version of our approach in [2], and the best point benchmark (corresponding to selecting the location that yields the highest cross-correlation value).

in these simulations, its distance error is still very significant, with 80% of the errors being greater than 100 m. Moreover, while the tail of the error distribution for the best point scheme is better than for the algorithms in [2] and in this contribution, these tails already correspond to significant errors (e.g., $>$ 300 m in terms of distance and $>$ 45° in terms of bearing).

4.4. Experimental results

4.4.1. Experiment Setup

In the previous section, we explored the performance of our localization scheme in simulations. Since these simulations rely heavily on a numerical acoustic propagation model, we now complete our analysis and show the performance achieved by our algorithm in a sea trial. The experiment was carried out in February 2017 in northern Israel (coordinates 33°01'57.0"N 34°55'41.2"E), in waters with a maximum depth of 140 m. The measured sound speed was 1529 m/s with a water temperature of 21° Celsius at the sea surface, and 1521 m/s with a water temperature of 17° Celsius at the sea bottom. The sound speed gradient between the surface and bottom was approximately constant. the current at the water surface was roughly 0.5 knot, the wave height was roughly 40 cm, and the sea bottom was sandy. A 5 m-resolution bathymetry was collected using a Kongsberg EM 2040 400 kHz multibeam sonar. The bathymetry of the explored area is shown in Fig. 4.11a, and included a steep slope ranging from 60 m to 140 m. The top-left side of



(a) Contour bathymetry map of the sea experiment area.

(b) Spectrogram of the received signals.

Figure 4.11: Setup of the sea experiment carried out in Mediterranean Sea waters near Haifa, Israel, in Feb. 2017.

the figure shows artificial data as no measurements were collected in that region.

The experiment included an 80-feet long vessel, RV EDEN, and a 13 feet rubber boat dragging a floating buoy from which an acoustic emitter was deployed, see Fig. 4.12. The rubber boat represented the opportunistic sounds source, and the RV EDEN represented the single receiver. During the transmissions, the distance between the vessels was roughly 1200 m. The transmissions from the rubber boat included a sequence of 15 linear chirps at the frequency range of 7 kHz to 17 kHz, each of duration of 1 s. Transmissions were made with the EvoLogics S2C R 7/17W underwater acoustic software defined modem at a source level of 170 dB re 1μPa @1m. Receptions at the RV EDEN were made through the custom uRadar recorder, whose receive sensitivity at the transmissions' frequency range is about 190 dBV re 1μPa. Both the transmitter and the receiver were deployed at depth of 10 m. A time-frequency spectrogram of the received signals is shown in Fig. 4.11b. Besides the transmitted chirp signals, we observe the signals of the RV EDEN's own echo-sounder. To mitigate the ambient noise as well as the signals of the echo-sounder, we filtered each chirp signal. Synchronization was performed using a normalized matched filter [89].

4.4.2. Results

We start from Figs. 4.13 and 4.15, which detail the results of the comparison between the modeled CIRs and the measured one for nine consecutive received signals. The results are shown as a polar map centered around the location of the RV EDEN, where each contour line represents a distance of 300 m from the vessel. In Fig. 4.13, each small grey

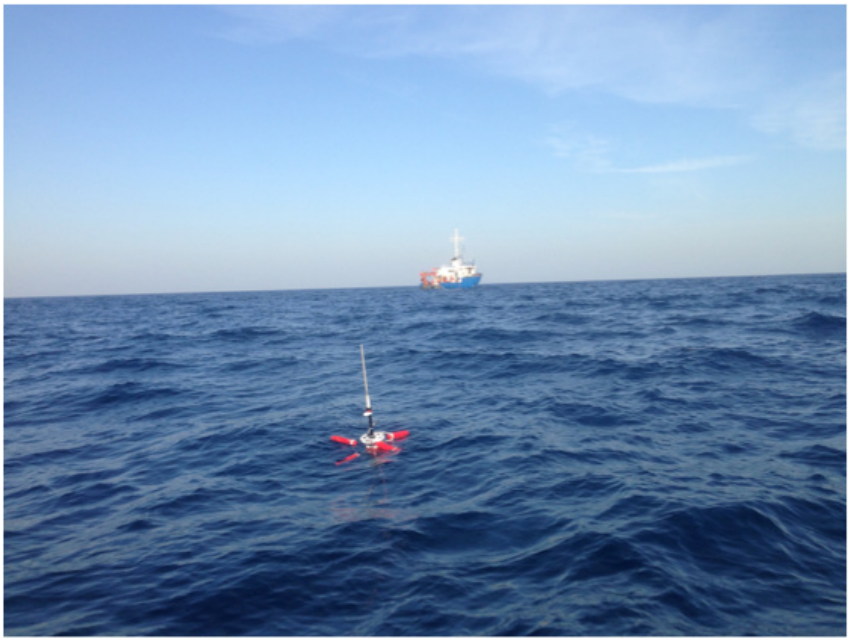


Figure 4.12: Picture of the buoy and ship from which the transmitter and receiver were deployed, respectively.

cross represents a possible position for the source, i.e., a comparison output that passed the detection threshold Θ_D (see Section 4.2.3) and was included in set $\mathcal{M}^{(2)}(i)$, $i = 1, \ldots, 9$. The thicker green cross marks the true location of the source. We observe that, for each of the nine received signals, many possible locations are obtained as a result of the cross-correlation operations carried out by our method. Fig. 4.14 shows one comparison between a CIR measured from a received signal and the CIR template constructed starting from Bellhop's output (light blue) and corresponding to a location close to the true location of the source. We observe that although the channel model is imperfect, all significant peaks in the measured CIR are well represented, leading to a good overall matching. However, other locations also lead to a similarly strong matching, resulting in several location estimates being significantly far from the source, and collectively resembling a random cloud of possible source locations (the small grey crosses). The best point algorithm (purple triangle), that points to the location yielding the maximum correlation for each signal, suffers from significant errors in three cases out of nine. These results support the simulation outcomes, showing that even when the bathymetry is fully known, relying only on the spatial diversity of the channel impulse response yields significant residual uncertainty. Processing the outcome of the best point algorithm through a Kalman filter does not yield significantly better results, even if the filter is fed with the actual velocity of the AUV (in contrast with our approach, that only requires to know the maximum AUV velocity, $v^{\mathrm{s}}_{\mathrm{max}}$). The corresponding location estimates are shown in Fig. 4.14 as grey squares. Conversely, our algorithm (red triangles) exploits the trellis search to achieve

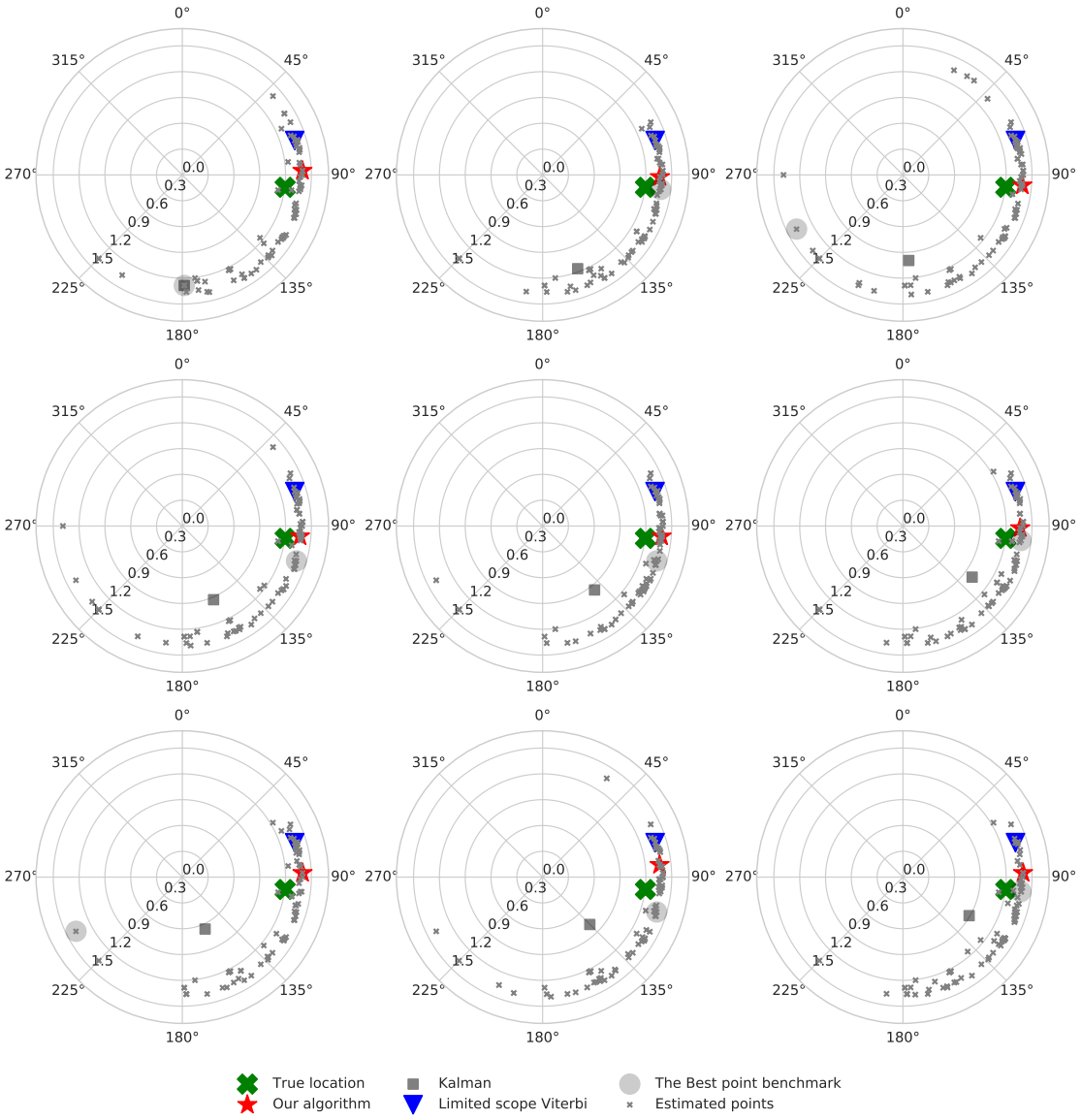


Figure 4.13: Sequence of location estimates for nine subsequent transmissions from a drifting source in the sea experiment, showing the true location of the source (green cross), the estimate of our algorithm (red triangles), the best point benchmark estimates (purple triangles), the Kalman filter results (grey squares), and the limited-scope Viterbi estimates (blue triangles).

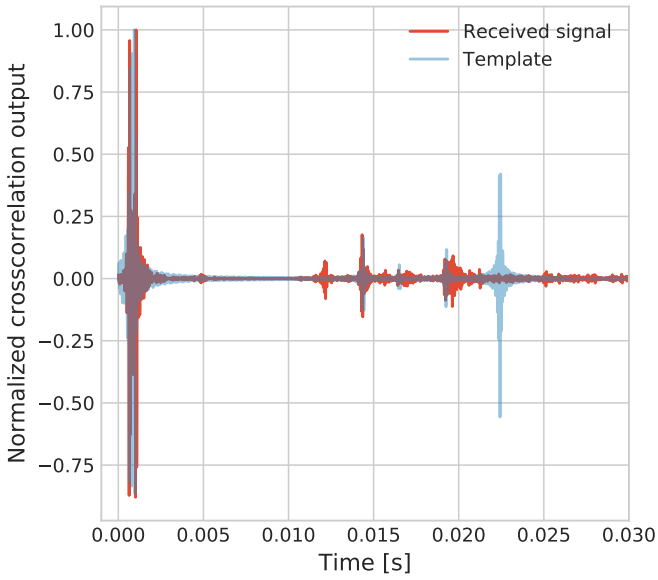


Figure 4.14: Comparison between a template CIR obtained from Bellhop (light blue) and the CIR measured from a signal received during the sea experiment.

a more precise estimation and removes outliers, resulting in a much smaller localization error.

We also compare the above results against those of the Viterbi algorithm. Given the size of the state space, in order to be able to run the algorithm we artificially reduce the search scope to a 90° sector centered on the true bearing of the source, and to the distances ranging from 1000 to 1400 m. While this gives a clear advantage to the Viterbi algorithm, it is a necessary step to allow the search space to fit in its data structures. The Viterbi results are shown as blue triangles in Fig. 4.13. We observe that the algorithm correctly predicts the fact that the source is static, but achieves a slightly worse location error despite the limitation of the search scope. This outcome is due to systematic, non-Gaussian errors incurred when modeling real underwater propagation using an acoustic propagation model under imperfect information (e.g., in this case, the resolution of the bathymetry and SSP data).

We summarize the path estimation results for the comparison outputs in Fig. 4.13 are shown 4.15 with the same color coding as above. In this case, the green crosses represent the average location of the drifting RV EDEN throughout the experiment. The results show a nice match between the estimated location and the true one for our algorithm and the Viterbi algorithm, and for only a subset of the location estimates of the best point algorithm. The Kalman filter results also show that this method is very sensitive to incorrect estimates of the initial location of the AUV.

The estimated locations predicted by our approach also correctly follow the drifting direction of the source boat. The total localization error for our algorithm is between

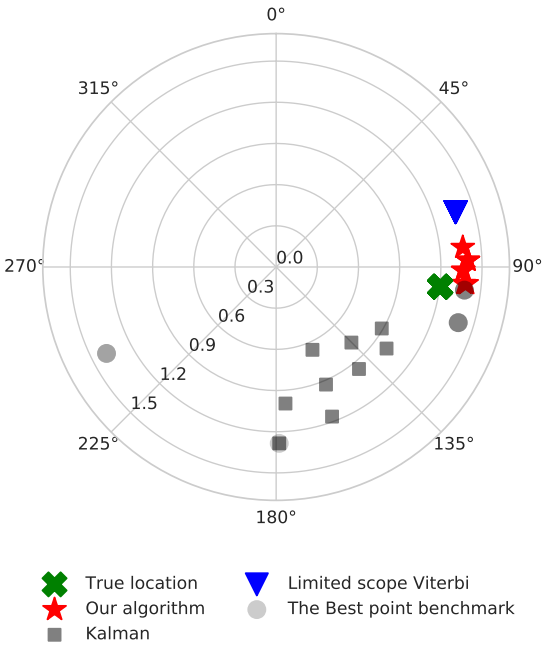


Figure 4.15: Final drifting source path estimate in the sea experiment, showing the true location of the source (green cross), the estimate of our algorithm (red triangles), the best point benchmark estimates (purple triangles), the Kalman filter results (grey squares), and the limited-scope Viterbi estimates (blue triangles). The total localization error for our algorithm is between 174 m (5.8%) and 330 m (11%), with a bearing error between 2 and 12 degrees.

174 m and 330 m, with a bearing error between 2 and 12 degrees. While these errors may seem large, we argue that for the task of localizing an AUV in a long term mission, this is still acceptable. This is because, first, after a few hours especially in deep water, the self-navigation system of the AUV completely drifts and thus any localization solution of limited expected error will benefit the operation [96], and second, compared to the typical detection range of roughly 5 km for the AUV's pinger (e.g., [97]), the above reported localization error as in our experiment is still a good result. Given that this result was obtained using only one receiver in real sea conditions, it demonstrates well the applicability of our suggested localization method.

4.5. Conclusions

In this work, we presented a novel approach for the acoustic localization of a non-cooperating AUV emitting acoustic signals. Our approach relies on a single passive and stationary receiving element and on the modeling of acoustic propagation given knowledge of the bathymetry, sound speed profile and bottom sediments in the deployment area. The method is based on the comparison of a channel impulse response evaluated from a

received acoustic signal, against a database of channel impulse response fingerprints. As the latter are modeled instead of measured, we require no periodic channel fingerprint acquisition in the area around the receiver. To filter noise, locations that show a good match between the measured and the modeled channels are arranged into a trellis. A location path is then estimated while limiting transitions between the trellis nodes according to an assumed maximum velocity for the AUV. Our approach makes it possible to estimate the path traveled by the AUV with low complexity and with high accuracy. Such accuracy decreases (but still remains sufficient for a variety of applications) if the receiver holds outdated environmental data. A proof-of-concept sea experiment demonstrates the applicability of our method to real sea conditions with a localization error as low as 5.8%, which is a remarkably good accuracy given the use of a single stationary receiver and the realistic, imperfect bathymetry and sound speed profile measurements.

Part IV

Direction of arrival estimation and localization through 3D acoustic arrays of arbitrary shape

In search of a balance between localization accuracy and the cost and complexity of deployment, there could be situations when, for a given accuracy, it is necessary to use multiple pre-assembled arrays. Fusing information from multiple acoustic sensors can have a positive effect on the reliability of the results. Despite this, often such design may not meet the condition of the distance between the elements at half the minimum wavelength due to design considerations such as cabling, batteries or other construction issues. In this case, the ability to use information from all the sensors is limited due to the spatial ambiguity caused by the improper spacing of array elements (see Sec.2.3.3.3). In the context of the SYMBIOSIS project, we worked on a localization algorithm for pelagic fish species using a software-defined version of a commercial ultra short baseline (USBL) array of five elements. Several such sub-arrays were joined together to form a larger array with better localization capabilities. In Chap. 5, we propose an algorithm for 3D wideband direction of arrival (DoA) estimation for such "opportunistically" joined arrays. We provide an in-depth analysis of its performance in various simulated conditions and validate the results in a preliminary experiment performed in a lake in Germany. Timely software maintenance and proper assembly of the complex final design was challenged by the travel difficulties caused by the well-known events of 2020. In this regard, changes were made to the algorithm that allowed for additional functions and worked in the absence of key elements. The changes and the results of final experiments are described in Chap. 6.

"What is a scientist after all? It is a curious man looking through a keyhole, the keyhole of nature, trying to know what's going on"

Jacques-Yves Cousteau

5 Underwater Localization via Wideband Direction-of-Arrival Estimation Using Acoustic Arrays of Arbitrary Shape

Underwater sensing and remote telemetry tasks produce the most valuable results when they can clearly geo-locate sensed data values. This is especially important when sensing the presence of acoustic signals coming, e.g., from wildlife or man-made devices: in these cases, estimating the location of the acoustic source and tracking it over time typically yields significantly valuable information, and may likely be the ultimate task of the sensing process. Common solutions for this challenging task include image- and video-based monitoring [98], LiDAR systems [3], as well as acoustic systems [47].

In underwater scenarios, accurate localization typically requires acoustic arrays. These pieces of equipment encompass multiple hydrophones or acoustic transceivers. This enables spatial filtering to increase directivity towards specific directions, and makes it possible to estimate the DoA of a signal. In several cases, underwater acoustic arrays are sizeable, and may require complex handling for deployment at sea.

When possible, the shape of an acoustic array is designed to fit the need of some application. For example, the side-scan sonar of an autonomous underwater vehicle (AUV) is usually a 2D matrix of acoustic elements [99, 100] designed to cover a given aperture with a given resolution, expressed in terms of the beamwidth of the main lobe of the array's beam patterns. Other relevant examples include linear arrays for DoA estimation [101] and three-dimensional general-purpose scanners such as tetrahedral and pyramidal arrays [102].

If arrays are modularly designed, it is typically possible to improve the spatial scanning performance by joining several arrays into a more complex layout. For example, several linear arrays could be joined into a 2D matrix, or into a 3D cylindrical configuration that makes it possible to scan a water volume and discriminate among different directions of arrival. Tesei et al. [103] showed that fusing the information from multiple arrays improves localization and ranging, even if these arrays are located far apart, and receive uncorrelated acoustic signals.

In this chapter, we are interested in "acoustic arrays of opportunity." As opposed to

arrays specifically designed for a given task, acoustic arrays of opportunity are typically composed of multiple sub-arrays originally designed to work independently, where each sub-array may contain one or more hydrophones or acoustic transceivers. Sometimes, sub-arrays come as standalone units: an opportunistic array would merge and co-operate multiple such units.

The main challenge related to acoustic arrays of opportunity is that each sub-array may have a physical design or may present mounting constraints that prevent the array of opportunity from having the optimal structure for a given task. For example, we may have to ensure some minimum spacing among the sub-arrays in order to preserve connectors, or to avoid that power and data cables bend in excess of their specifications. Additionally, the sub-arrays may have pre-defined shapes, and it is typically unfeasible to reconfigure these shapes into other layouts.

There are at least two important consequences to the above constraints. First, it may be impossible to construct typical array topologies such as uniform linear arrays (ULAs), uniform rectangular arrays (URAs), or cylindrical arrays [104]. Second, the resulting layout may force a larger-than-optimal spacing among closest array elements, e.g., larger than $\lambda/2$, where λ is the wavelength corresponding to the maximum operational acoustic frequency of the array's hydrophones. An improper spacing of the array elements causes spatial ambiguities in beamforming and DoA estimation operations. For example, multiple, equally strong lobes in the opportunistic array's beam pattern, or equivalently, it may become impossible to distinguish among multiple equally likely DoA estimates. Additional physical characteristics of the array elements, such as a non-omnidirectional radiation pattern, may be insufficient to remove such spatial ambiguities [105]. When the purpose of the array is localization through multilateration, larger-than-$\lambda/2$ spacing may also lead to significant errors [106].

The above issues have an even larger impact when employing wideband 3D DoA estimation algorithms. Notably, most wideband algorithms [105] work with predefined array shapes, or are limited to 2D, to specific signals, or to a known number of targets [107]. In some cases, the preferred solution is to directly employ particle velocity sensors [108–110].

In this work, we propose a wideband DoA estimation scheme based on the delay-sum algorithm. Our scheme that works with arbitrary array layouts, where both the 3D arrangement of the array elements and the spacing among them are potentially irregular or arbitrary. This fits well our assumption that the array opportunistically merges independent subsystems. To remove spatial ambiguity, we compute rough target location estimates via multilateration, using time difference of arrival (TDoA) measurements from the array elements. We then restrict the DoA search space to an area around the target, using multilateration estimates as side information. This rules out or at least dampens ambiguous directivity peaks. When the approach is successful, we accrue the additional

advantage that peaks affected by spatial ambiguity are narrower [105], and thus yield a more accurate DoA estimate. We observe that such an approach works even if the array elements are not sufficiently spaced to achieve high resolution [103], and therefore multilateration estimates are not extremely accurate.

We evaluate our approach by running emulations and by performing a lake experiment. Emulations provide a controlled environment that simplifies acoustic propagation and detection, but recreates realistic acoustic background conditions by using real clutter noise recordings from an in-water experiment. The lake experiment, instead, involves real hardware and realistic water conditions, including small movements due to currents and waves. In both cases, our results show that our algorithm effectively estimates DoAs and 3D target locations. Using the extra freedom allowed by emulation, we also show that merging together realistic pyramidal arrays (such as those found in off-the-shelf equipment, e.g., [111]) yields better DoA estimation performance than typical cylindrical arrays having $\lambda/2$ element spacing.

In summary, our approach yields the following advantages: (i) it provides a framework to merge together smaller arrays into a larger "array of opportunity" to achieve better DoA estimation accuracy; (ii) it provides a method to rule out the ambiguity that may result from the suboptimal spacing of the array elements; (iii) it works with wideband signals and arbitrary array topologies; (iv) it yields good performance in emulated sea environments as well as in a proof-of-concept experiment. In particular, emulation results show that, at a Signal-to-Noise Ratio (SNR) of -20 dB, a 15-element array of opportunity achieves lower average and median localization error (27 m and 12 m, respectively) than a 30-element array with proper $\lambda/2$ element spacing (33 m and 15 m, respectively). In a proof-of-concept lake experiment, we additionally show that our algorithm achieves a 90th-percentile DoA estimation error of 4° and a 90th-percentile total location error of 5 m when applied to a real 10-element array of opportunity. This realistic performance evaluation substantially extends the preliminary simulation results in our previous work [4]. Moreover, in this chapter, we included additional related literature, extended the presentation and explanation of our approach with examples of intermediate steps, and provided evidence that the components of our algorithm that complement wideband DoA estimation can run in real time on an embedded platform.

The remainder of this chapter is organized as follows. In Section 5.1, we survey relevant related work. In Section 5.2, we describe our DoA estimation method. We introduce materials and methods for our performance evaluation in Section 5.3. Sections 5.4 and 5.5 cover the evaluation of our proposed scheme via emulation and via a lake experiment, respectively. Finally, we discuss our results in Section 5.6, and draw concluding remarks in Section 5.7.

5.1. Related Work

The engineering of array processing schemes for underwater detection and communication spans several disciplines, from sonar systems to communications and underwater target detection with either passive or active arrays [112]. Recent advances in this fields include the application of different estimation or signal processing techniques to classical beamforming algorithms, with the objective of improving their accuracy and decrease their complexity. For example, using a particle filter to estimate the DoA of an acoustic source improves the performance of Bartlett and conventional beamformers [113]. Real data from the SwellEx'96 sea experiment validate the findings of the study. Chen et al. [114] improve the performance of a blind DoA estimation algorithm from the literature [115] by exploiting partial knowledge on the structure of the signal transmitted by an acoustic source. With the aim to reconstruct the DoA of a wideband underwater signal, Tang et al. [101] use sparse signal representation and provide further methods to eliminate the aliasing originating from the over-completeness of the measurement dictionary. The authors prove the effectiveness of their algorithm using a uniform linear array to detect the breathing sounds of divers equipped with closed-circuit rebreathers.

Van Kleunen et al. [116] consider a blind node integrating a 4-element linear array for DoA estimation, and mix DoA with time of flight (ToF) information related to the signals that the blind node receives from synchronous reference nodes. On a similar vein, Guo et al. [117] employ a linear array to localize a node emitting acoustic signals, by leveraging the multipath components appearing at the receiver. Weighed subspace fitting helps avoid the explicit estimation of the DoA for each multipath arrival. Tesei et al. [103] discuss sound source localization in 3D using either one or two tetrahedral arrays deployed at different locations. Despite synchronous sampling in the two systems, their algorithm does not process the arrays jointly, as the distance between the arrays decorrelates the received signals.

Multiple works applied compressive sensing and sparse reconstruction techniques to underwater array processing. For example, Song et al. [118] use compressive beamforming to estimate the DoA of an underwater acoustic source via a forward-looking sonar, and validate the system using field experiment data. Sparse reconstruction [104] enables the estimation of the DoA of sound emitted by underwater vessels. The authors carry out an experiment using a passive towed linear array sonar that showcases the performance of their algorithm. Two-dimensional continuous compressive sensing is used in [119] to estimate the complete set of measurements of a URA starting from a sparser array. The resulting estimates are then employed to impute the missing measurements and compute the DoA of a signal impinging on the array. Coherent signal subspace processing and compressive sensing are jointly considered for wideband

DoA estimation in [120]. Compared to the conventional minimum variance distortionless response (MVDR) beamformer, the proposed method yields higher resolution.

Acoustic vector sensors, also known as particle velocity sensors, provide a first estimate of the direction of arrival of an underwater signal. Owing to this, several works rely on vector sensors for DoA estimation [121–123]. With a focus on computationally efficient DoA estimation, Bereketli et al. [110] employ an acoustic vector sensor to estimate the DoA of an impinging signal in a shallow water scenario, where strong multipath echoes degrade the quality of DoA estimates. Vector sensors can also be arranged into arrays, and coherently processing the sensors' signals improves the resolution of underwater DoA estimation for wideband coherent sources [124].

Wideband beamforming recently spurred significant interest in the broadband terrestrial radio communication domain [125–127]: as underwater acoustic systems are typically wideband, similar techiques find applications for underwater acoustic detection and communications as well. Liu and Weiss [105] extensively cover classical approaches and recent research results for wideband array processing with applications to signal enhancement and DoA estimation. Multichannel processing through diversity combining and optimal beamforming is the focus of the work in [128], which targets the reception of high-speed underwater acoustic communication signals. The authors show that beamforming enables the design of a significantly simpler receiver, which can coherently extract multipath signal energy in a sea experiment. A similar approach [129] employs mono-pulse processing to cancel incoherent multipath components that would interfere with the receiver in a reverberating shallow water acoustic communications scenario. Bayesian methods (BMs) are also applied to the estimation of the DoA of wideband linear frequency modulation (LFM) signals using a uniform linear array [130]. The authors resort to the fractional Fourier transform to extract the wanted signal from a reverberating background and improve the operating signal-to-noise ratio.

Typically, the development of signal processing algorithms for underwater acoustic arrays assumes a simple array topology, for which the steering vectors and array manifolds can be computed in close-form. Often, linear arrays are used [104, 113, 114, 116–118] or rectangular arrays [119]. A study involving 3D, 4-element tetrahedral arrays is provided in [103]. Unlike the above literature, in this work, we propose a wideband DoA estimation algorithm that works on any 3D underwater array layout. Thus, our approach encompasses imperfect array design occurrences, or the opportunistic combination of multiple arrays into a larger structure to seek better spatial performance. A key realistic assumption of our work is that array sensors may not be properly spaced at a distance of $\lambda/2$ from one another, possibly resulting in spatial ambiguity that must be compensated for.

5.2. Wideband DoA Estimation Algorithm

We now proceed to introduce and explain our wideband DoA estimation method. We start by providing the key idea behind the algorithm in Section 5.2.1, and continue with the details in Section 5.2.2.

5.2.1. Key Idea

We assume to operate an array of known topology, but whose elements are not necessarily arranged to obey the $\lambda/2$ spacing constraint. The array elements are co-located, and the array control electronics synchronously retrieve acoustic samples from all elements. The task of the array is to detect the DoA of signals with a known structure, either emitted by AUVs and other man-made equipment, or emitted by a projector co-located with the array and reflected back by the target.

Our algorithm mitigates spatial ambiguity via side information in the form of a rough location estimate derived from TDoA-based multilateration. This information helps filter the output of a wideband delay-sum DoA estimation algorithm and thus rules out most of the ambiguous DoA estimates.

5.2.2. Algorithm Description

Call $f(t)$ the signal that the array seeks the DoA of. While this can be any signal, if the array is co-located with a projector and listens to reflections of the projector's signal, a typical solution is to employ linear chirps spanning the frequency interval from $f_{\min}$ to $f_{\max}$ over a time interval of duration T. Such chirp would have the form $f(t) = \cos\left(2\pi \frac{f_{\max}-f_{\min}}{2T} t^2 + f_{\min} t\right)$. Let $s_n(t)$ be the real-valued signal received by the nth array element.
the $s_n(t$

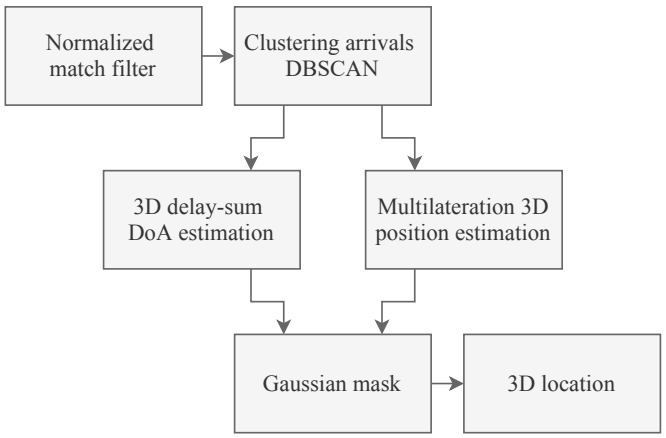


Figure 5.1: Flow diagram of the DoA estimation and localization algorithm.

the output of the NMF is expressed as

$$R_n(\tau) = \frac{\int_0^{+\infty} f(t)\, s_n(t+\tau)\, \mathrm{d}t}{\left(\int_0^T f^2(t)\, \mathrm{d}t \int_0^{+\infty} s_n^2(t)\, \mathrm{d}t \right)^{1/2}} . \tag{5.1}$$

We search relevant peaks in $R_n(\tau)$ via a sliding window method. In more detail, we consider a window of length T aligned with the beginning of $R_n(\tau)$, and take the highest peak in the window; then, we slide the window, take again the highest peak, and repeat the process until we cover the whole of $R_n(\tau)$. This filtering step eliminates secondary peaks that are never the tallest in any window. Call $\mathcal{P}_n$ the set of peaks that survived filtering, where the features that fully define a peak $p \in \mathcal{P}_n$ are its time of occurrence t, its amplitude a and the hydrophone n that detects it, i.e., $p = (t, a, n) \in \mathcal{P}_n$.

We proceed by applying the DBSCAN algorithm [131] over the whole set of peaks

$$\mathcal{P} = \mathcal{P}_1 \cup \cdots \cup \mathcal{P}_n . \tag{5.2}$$

DBSCAN approximates the function

$$\mathcal{C} = \mathcal{D}(\mathcal{P}) \tag{5.3}$$

that returns all subsets of arrivals $C \in \mathcal{C}$, such that each subset C contains groups of detections that correspond to the same target. We choose DBSCAN because it works based on point density, which we found to be a very good indicator of target detection. In fact, when a signal from a target insonifies the array, no two NMF peaks related to this detection should be farther in time than the maximum propagation delay between any two array elements. This allows us to define the density of the NMF peaks in time across multiple channels. Moreover, DBSCAN executes very fast [132, 133] on our NMF time series (typically in less than 1 ms) and the algorithm does not need prior information about the number of points that are part of a cluster, or the number of clusters in the dataset. As a result, DBSCAN is suitable for multiple target detection in scenarios with multiple targets.

We configured DBSCAN to seek arrivals detected in at least 70% of the array elements, and spaced in time no more than the maximum propagation delay between any two elements. This makes it possible to discard peaks that are not detected reliably by all elements, or that are separated by a large time delay, enough to suspect that they may correspond to different emissions from the environment, or to reflections from different targets. We remark that the threshold on the number of array elements that should detect target-related peaks configures a trade-off between the probability of missing a detection, and the probability of wrongly including a detection that pertains to a different target.

In fact, some array elements may be shadowed by other sensors, cables, or structural components of the array, and therefore a cluster C may contain peaks from only a subset of the array elements. Additionally, if a cluster contains multiple arrivals within t_{max} on the same channel, we have an option to filter the arrivals depending on the output of the normalized matched filter.

Using real data from a lake experiment, we show an example of clustering result for peaks collected by a synchronously-sampled 10-element array in Figure 5.2. The array is configured to seek linear chirp signals of duration 10 ms spanning the acoustic frequencies from 7 to 17 kHz. The light-blue time series in the background of Figure 5.2 is the output of the NMF for one of the ten acoustic channels, namely channel 1, depicted over time (measured in seconds). The sampling frequency is 62.5 kHz. For this channel, the peak extraction algorithm described above filters the peaks marked as a large blue circle. The same algorithm, applied to the NMF output of the remaining nine channels (whose time series are omitted for clarity), leads to the peaks marked as small, purple circles. Altogether, these peaks form set $\mathcal{P}$. DBSCAN processes the peak set to detect the target. Out of all peaks, DBSCAN singles out the orange-colored ones as being likely associated with the target, due to their density and their appearance in several acoustic channels. As confirmed by the vertical orange line, these peaks correctly align with the ground truth of the target's position, which corresponds to a time of arrival of about 0.16 s.

Once DBSCAN has detected the target, we are ready to estimate the DoA of the target's signal. To do so, we now consider a set of elevation angles Θ and azimuth angles Φ, and scan the power received by the array along every direction identified by a pair (θ, ϕ), for $\theta \in \Theta$ and $\phi \in \Phi$. We remark that we could set $\Theta = (0, \pi)$ and $\Phi = (-\pi, \pi)$ in order to cover all possible DoAs in 3D space, or rather restrict Θ and Φ to a smaller domain, in case some prior information is available.

For each cluster C, we measure the energy perceived by the array along different directions through a wideband delay-sum algorithm [105]. We stress that using wideband algorithms matches the possibly wideband signals detected or projected by the array. For example, the equipment in [111] operates across the bandwidth from 7 kHz to 17 kHz, which aligns well with the definition of wideband equipment. In any event, our approach also works with narrowband signals.

We implement the algorithm in the frequency domain. Specifically, we cut the output of the matched filter that covers the arrivals in cluster C for each hydrophone, and then apply an fast Fourier transform (FFT) to this signal chunk. We then apply a different, frequency-dependent phase shift vector to each frequency bin in order to steer the array towards the direction (θ, ϕ). Finally, we convert back to the time domain via an inverse FFT, and sum up the resulting outputs across all hydrophones.

By carrying out this operation for several 3D DoAs (θ, ϕ), we obtain a map $\alpha(\theta, \phi)$ of

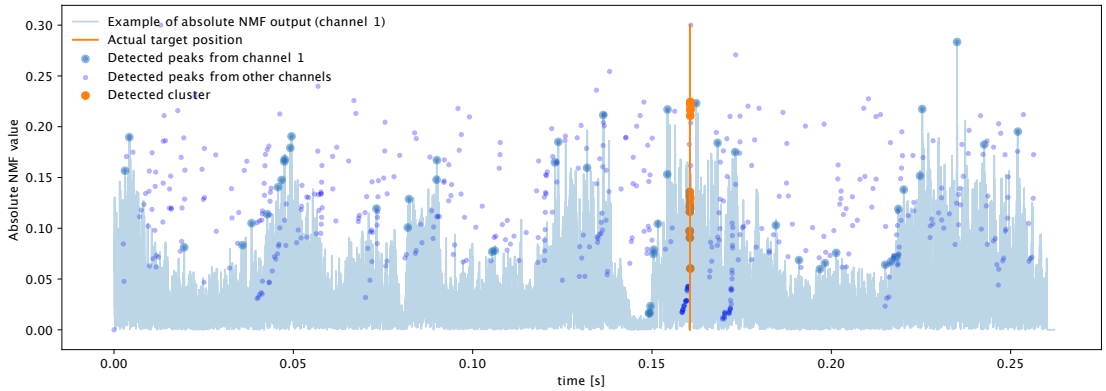


Figure 5.2: Example of successful DBSCAN clustering for peaks collected by a 10-element array. The light-blue time series is the output of the NMF for channel 1. Large blue circles represent peak detections for this NMF time series. Smaller dark-purple peaks represent peak detections from the remaining nine channels. DBSCAN correctly detects a cluster of target related arrivals around 0.16 s (vertical orange line). Data from a real lake experiment.

the power received over all scanning directions specified by sets Θ and Φ. As we consider opportunistic arrays where the elements may be spaced more than $\lambda/2$, the delay-sum map may be affected by ambiguities, hence it may indicate the reception of a significant amount of power from directions different than the true DoA of the target.

The key to recuse such ambiguity is to mask the above DoA map by roughly localizing the target in 3D space via a multilateration algorithm. We achieve this through TDoA measurements carried out across the array elements. Call

$$\mathbf{u} = [x\ y\ z]^{\mathsf{T}} \quad \text{and} \quad \mathbf{u}_n = [x_n\ y_n\ z_n]^{\mathsf{T}} \tag{5.4}$$

the Cartesian coordinates of the target and those of hydrophone n, respectively. Furthermore, call t_0 the time of occurrence of the earliest peak in cluster C. Without loss of generality, assign index 0 to the hydrophone that receives this arrival. Finally, call c the sound speed near the array. We assume that such speed is known, typically via local equipment such as a sound velocity profiler or a CTD sensor.

For each peak $p = (t, a, n) \in C$, the corresponding multilateration equation is

$$x \cdot X_n + y \cdot Y_n + z \cdot Z_n + D_n = 0\,, \tag{5.5}$$

where

$$X_n = \frac{2x_n}{c\,t} - \frac{2x_0}{c\,t_0}\,, \quad Y_n = \frac{2y_n}{c\,t} - \frac{2y_0}{c\,t_0}\,, \quad Z_n = \frac{2z_n}{c\,t} - \frac{2z_0}{c\,t_0}\,, \tag{5.6}$$

and

$$D_n = c\,(t - t_0) - \frac{x_n^2 + y_n^2 + z_n^2}{c\,t} + \frac{x_0^2 + y_0^2 + z_0^2}{c\,t_0}\,. \tag{5.7}$$

Collecting one equation such as (5.5) for every peak in cluster C results in an over-determined system of equations, which we solve through Moore–Penrose's pseudo-inverse. The result is a rough estimate of the target location $\mathbf{u}^\star = [x^\star \; y^\star \; z^\star]^\mathsf{T}$, which we convert to polar coordinates to yield the estimated location of the target, namely

$$\bar{\mathbf{u}} = [\bar{r} \; \bar{\theta} \; \bar{\phi}]^\mathsf{T} . \tag{5.8}$$

We exploit the above estimate to define a masking function having the shape of a truncated bi-variate Gaussian kernel

$$m(\theta, \phi) = \min \left\{ 1; \; \frac{1}{2\pi\sigma_\theta\sigma_\phi} \, \mathrm{e}^{-\frac{(\theta-\bar{\theta})^2}{2\sigma_\theta^2}} \, \mathrm{e}^{-\frac{(\phi-\bar{\phi})^2}{2\sigma_\phi^2}} \right\}, \tag{5.9}$$

where $\sigma_\theta = \pi/8$ and $\sigma_\phi = \pi/4$. Using $m(\theta, \phi)$, we mask the output of the wideband delay-sum beamformer in order to mitigate (and typically fully remove) ambiguities. Finally, we set the estimated DoA for the received signal as

$$\hat{\theta}, \hat{\phi} = \arg\max_{\theta,\phi} \alpha(\theta, \phi) \, m(\theta, \phi) . \tag{5.10}$$

Figure 5.3 (left panel) provides an example of the delay-sum output. Specifically, for a number of 3D DoAs characterized by a pair of angles ϕ (azimuthal angle, x-axis) and θ (elevation angle, y-axis), we steer the array towards each DoA using wideband delay-sum beamforming, and depict the normalized amount of power at the output of the array. Yellow hues correspond to a strong signal, green hues to a signal of intermediate power,

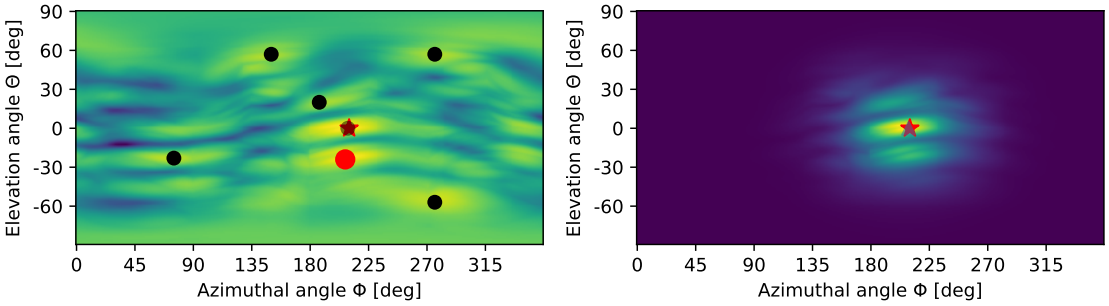


Figure 5.3: Intensity map at the output of the wideband delay-sum beamformer without (**left**) and with (**right**) TDoA multilateration-based masking. The latter mitigates the ambiguity and makes it possible to correctly estimate the location of the target (red star), while ruling out the strongest peak (red dot) which would correspond to a wrong target location. Yellow hues denote a stronger signal.

with ambiguity would lead to a wrong estimate (red circle) of the target's actual DoA (red star).

In the right panel of Figure 5.3, instead, we apply the multilateration-based mask. As explained above, we employ TDoA information extracted from the peaks that DBSCAN recognizes as being part of the same target detection (formally, cluster C above). These peaks come from different acoustic elements within the array: computing the TDoA values for these peaks makes it possible to roughly localize the target via multilateration, and to construct the mask in Equation (5.9). This filters out most of the ambiguity and points to the correct DoA (red star). We note that some local maxima still remain even after applying the mask (corresponding to the green hues above and below the starred peak). However, these peaks are now sufficiently mitigated, and do not impede a correct DoA estimation.

As a final step, we fuse the estimated DoA with ranging information and pass it on as a valid location only if the position of the target remains within the boundaries of the water column.

5.3. Materials and Methods

In this section, we summarize the methods and materials used for the performance evaluation of our proposed DoA estimation algorithm. After a short account of common assumptions in the two evaluation setups, we present our emulation framework in Section 5.3.2 and our experiment framework in Section 5.3.3. Sections 5.4 and 5.5 follow up with the results of the corresponding performance evaluations.

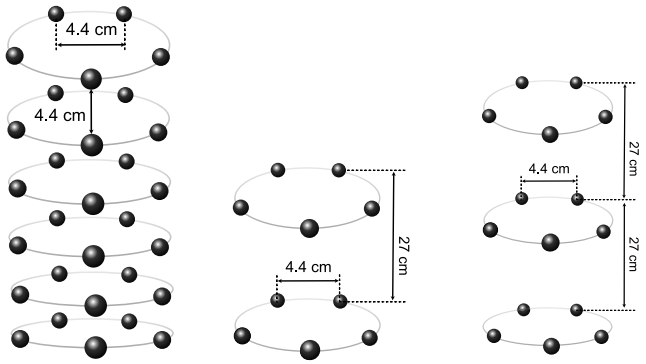
5.3.1. Common Setup and Parameter Configurations

In the following, we assume that our array of opportunity seeks linear chirp signals in order to localize a nearby target. Each chirp has duration $T = 10$ ms and spans the acoustic band from $f_{\min} = 7$ kHz to $f_{\max} = 17$ kHz. We also assume that the array can synchronously sample its elements, and can store the corresponding acoustic samples for immediate or offline processing.

5.3.2. Emulation

The key idea of the emulation is to employ measurements of noise and acoustic clutter from a lake experiment, in order to achieve a more realistic representation of the signal received by the array elements.

Specifically, we consider an experiment performed in the Werbellin freshwater lake, located 60 km north of Berlin, Germany. During this experiment, we acquired several underwater acoustic recordings containing environmental noise and clutter. We subdivide



(a) Array 1 (b) Array 2 (c) Array 3 (d) Array 4 (e) Array 5

Figure 5.4: Array topologies considered in this paper.

these recordings in chunks of 16 ms (equal to 1000 samples at a sampling frequency of 62.5 kHz) and normalize each chunk so that the standard deviation of the noise is equal to 1 throughout all chunks. Finally, we create a Monte Carlo set of emulated noise recordings, where in each recording we randomly shuffle the order of the 16-ms noise chunks.

Our emulation framework consists of a software written in Python. Here, we assume that the array of opportunity is located at a depth of 10 m within an isovelocity water body with flat surface and bottom having maximum depth of 100 m. (The lake experiment in Section 5.5 serves to test our algorithm in realistic propagation conditions, with a stratified medium and non-flat bottom.) For the emulation scenario, we assume that the array insonifies the underwater environment by transmitting the chirp signal, and listens to reflections from the environment. We thus emulate a received signal by propagating the chirp to the target, and back to the array. In particular, we shift the phase of the chirp as a function of the location of the source and of the position of each acoustic array element. We then scale a noise sequence from the Monte Carlo set to yield a desired SNR level, and superimpose the received signal to the noise. Finally, we apply our DoA estimation algorithm to the signal. We repeat the experiment for 270 different locations of the target, chosen to represent all array lookout directions. For each location, we repeat the estimation for ten different underwater noise realizations.

We consider five different array layouts, as illustrated in Figure 5.4:

- Array 1 is composed of two 5-element pyramidal arrays having a base side length of 10 cm and an height of 7.07 cm. The sub-arrays are stacked at a distance of 27 cm, and the bottom one is rotated by 45°. This is typical in the case in which each 5-element pyramidal array actually comes as a separate unit, whose connector mounting and cable bending constraints prevents placing the units closer than a given maximum distance;

- Array 2 is similar to array 1 but is composed of three pyramidal arrays stacked at a distance of 27 cm. In this case, the second array is rotated by 30° and the third by 60°;

- Array 3 is a cylindrical array composed of 6 circular sub-arrays of 5 elements each (the same number of elements as in the pyramidal arrays of Array 1 and Array 2). The distance between closest elements along the same ring and across different rings is 4.4 cm;

- Array 4 is composed of two circular sub-arrays of radius 3.5 cm, placed at a distance of 27 cm from each other. Each sub-array embeds 5 elements. The elements are equally spaced along the ring and closest elements are 4.4 cm apart;

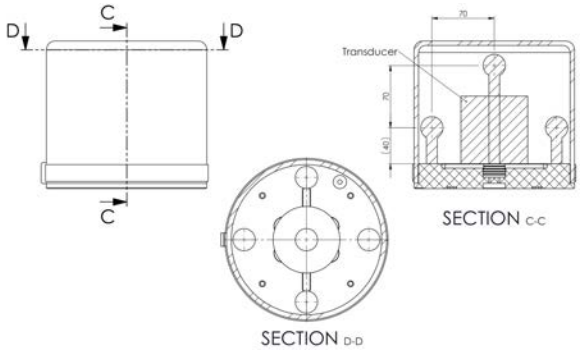
- Array 5 is similar to array 4 but is composed of three rather than two rings.

We chose the topologies of acoustic arrays 1 and 2 above as they resemble closely the arrays of opportunity attached to the underwater fauna detection platform described in Section 5.3.3.1, and used for the lake experiment in Section 5.5. Array 3 is a typical cylindrical array. Its shape enables spatial scanning along both the azimuthal plane and the elevation plane; moreover, the spacing between closest array elements is less than or equal to $\lambda/2$ up to a frequency of 17 kHz. Finally, arrays 4 and 5 are also cylindrical arrays, but the distance between subsequent circular 5-element sub-arrays is 27 cm, in order to emulate the performance of Array 3 in case it had the same number of elements and mounting constraints as arrays 1 and 2.

We remark that assuming a sound speed of 1500 m/s, a distance of 4.4 cm corresponds to $\lambda/2$ spacing up to a frequency of $\approx$17 kHz. Because array 3 is designed with proper $\lambda/2$ spacing throughout the whole bandwidth of the chirp signal, we do not apply the $m(\theta, \phi)$ mask to the wideband delay-sum output in this case.

The ambiguity of DoA estimation originates primarily from the 27-cm spacing between subsequent pyramidal arrays, as such distance is about 7 times the appropriate $\lambda/2$ spacing of 4.4 cm. In Sections 5.4 and 5.5, we show that our algorithms still obtain meaningful location estimates even when using the equipment in the above configuration.

(a) Acoustic array part of the SYMBIOSIS platform (courtesy of EvoLogics GmbH).



(b) Sagittal and top sections of EvoLogics's SDM-USBL unit (courtesy of EvoLogics GmbH).

Figure 5.5: (**a**) rendering of part of the upper portion of SYMBIOSIS platform, showing the acoustic array of opportunity employed in our experiment (two cylindrical SDM-USBL units, facing right); (**b**) internal configuration of an SDM-USBL unit. Each sphere denotes a receiving acoustic element (5 in total, arranged into a pentahedral, square-base pyramid). The unit includes a transducer (the large cylindrical element in the sagittal C-C section), not used in our setting.

5.3.3. Lake Experiment

5.3.3.1. Equipment and Software

The acoustic array we deployed for this experiment is a part of the opto-acoustic system built in the scope of SYMBIOSIS project [134] for the non-invasive monitoring of coastal and deep waters. The purpose of the platform is to detect, localize, and monitor fish stock from different target pelagic fish species, using a chain of acoustic and optical detection systems and algorithms. Figure 5.5a shows a rendering of the acoustic components in the upper portion of the SYMBIOSIS instrumentation. The top and bottom cylinders host control hardware and batteries to operate the platform, whereas the two cylindrical pieces of equipment facing right constitute the acoustic array considered in our experiment. Each cylinder contains a software-defined USBL with modem capabilities (SDM-USBL). Figure 5.5b shows the internal geometry of the unit via one sagittal and one longitudinal section. Each SDM-USBL consists of a modem transducer (at the geometric center of the USBL grid) and five receive hydrophones that surround it. The hydrophones form a pentahedral, square-base pyramid having side length of 10 cm, with one hydrophone per vertex. A set of commands enables the control of each SDM-USBL. Relevant capabilities include: (i) storing a user-defined signal with a duration of 1024 samples (two bytes per sample) with a sampling frequency of 62.5 kHz; (ii) setting the unit into a listening mode, where each hydrophone digitizes its received sound signal synchronously and stores the corresponding samples into a buffer holding 51,200 samples per channel; the sampling rate is user-defined, although the default rate of 62.5 kHz

perfectly suits our deployment; (iii) reading acoustic data from the buffers; and (iv) transmitting the stored signal through the central transducer.

Each SDM-USBL can operate in an "active" or "passive" mode. By triggering the "active" mode, the SDM-USBL sends the stored user-defined signal and then stores 51,200 samples per acoustic channel; in the "passive" mode, instead, the unit does not transmit any signal, but rather starts recording immediately. A sync-in signal allows us to trigger the two units at the same time, thus sampling synchronously from both of them (10 channels in total).

While each stand-alone USBL would natively work as a fully capable localization device, the SDM-USBL option and the external sync-in signals implemented by EvoLogics in the context of SYMBIOSIS disable the USBL firmware, and rather make it possible to collect acoustic samples synchronously from all channels. Therefore, each unit can double as an acoustic array, and the synchronous use of multiple SDM-USBL units effectively results in an array of opportunity, whose arrangement makes it equivalent to array 1 in Section 5.4.

The setup of the SYMBIOSIS platform includes an NVidia Jetson TX2 board, mainly used to control the optical components and run image recognition algorithms, which are outside the scope of this work. In our setting, we performed all signal processing steps offline. Through one of the Jetson's general purpose input/output (GPIO) pins, we issued the sync-in signal periodically in order to synchronously start recording from all of the 10 array hydrophones, while at the same time transmitting from the active target. This resulted in one acoustic time series of 0.7 s per acoustic channel. We measured the accuracy of the sync-in signal to be on the order of 200 ns, which is sufficiently accurate, given the sampling frequency of 62.5 kHz.

With reference to the scheme in Figure 5.6, we control the equipment from the laptop ashore (see also Figure 5.7b) using custom Python software. Once we issue the sync-in signal, the Jetson board collects data in real time, and uploads them to the laptop ashore. The software then processes the acoustic time series through the procedure of Section 5.2 to obtain DoA estimates.

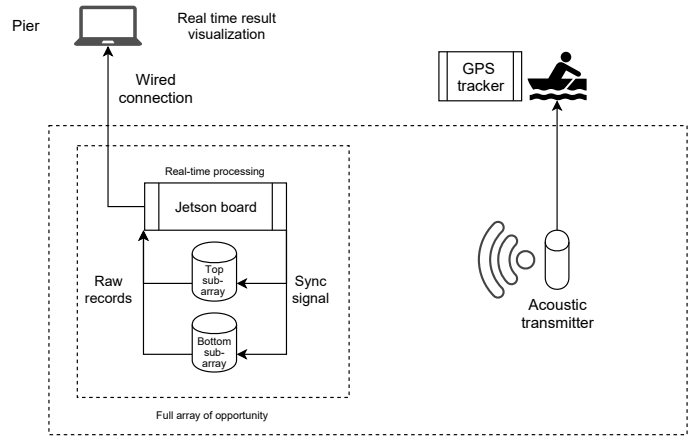
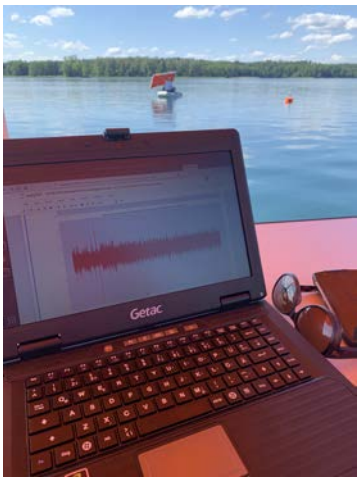


Figure 5.6: Conceptual organization of the experiment.

(a) SYMBIOSIS acoustic array on the jetty of the Werbellin lake marina.



(b) Real-time data acquisition and offline processing on the jetty.

Figure 5.7: Photos of the deployment: (**a**) acoustic array of the SYMBIOSIS platform on the jetty, before deployment; (**b**) ongoing experiment, showing a snapshot of a captured signal on the laptop's screen.

5.3.3.2. Experiment Setup

We performed our experiment on 13 June 2019. The weather remained mostly sunny throughout the day, with little wind. The water temperature ranged from 19.4 °C at the bottom, up to 21.5 °C near the lake's surface.

Figure 5.8 sketches the deployment configuration: we lowered the SYMBIOSIS unit

in the water near the jetty of the Werbellin lake marina (Figure 5.7a), and placed it on the lake bottom at a depth of 7.5 m, so that the two units that constitute the array of opportunity remained submerged at a depth of 6.7 m and 7.1 m, respectively. The red arrow in Figure 5.8 denotes the location of the equipment and the reference (i.e., 0°) direction of the acoustic array. Additionally, we deployed a small motorboat carrying an active target, namely a software-defined modem emitting linear chirp signals with a duration of 10 ms in the 7–17 kHz band. The target transmitted one such signal every 2 s. The objective of this setup is to mimic the behavior of autonomous underwater vehicles that issue heartbeat signals at fixed intervals, in order to signal their presence and operational status. The task of the array is thus to estimate and track the bearing of the chirp source. In our experiment, the depth of the active target is assumed to be known, and is fixed to approximately 8 m. This is coherent with, e.g., the detection of underwater vehicles or similar equipment, which typically embed accurate depth sensors, and can communicate the corresponding data. An operator paddled the boat with the target towards the acoustic array and slightly to the side throughout the experiment, as also seen from the ground-truth trajectory of the target (the solid blue line in Figure 5.8), and the reference orientation of the array (the red arrow in the same figure).

Figure 5.8: Geographical map of the experiment site near the Werbellin lake marina, Germany. The red arrow on the jetty represents the location and the reference (i.e., 0°) direction of the acoustic array; the blue line and arrow represent the trajectory and movement direction of the target throughout the experiment.

5.4. Emulation-Based Performance Evaluation

We now evaluate the performance of our algorithm in an emulated underwater environment, according to the emulation setup described in Section 5.3.2.

Considering a target located at a distance of 100 m and a receive SNR of −20 dB, Figure 5.9 shows the cumulative distribution function (CDF) of the azimuthal angle

estimation error (left panel), of the depth estimation error (central panel) and of the global location error (right panel) for the five array layouts introduced above. The most interesting result is that array 2 (opportunistic 3×5-element pyramidal arrays) achieves at least the same global location accuracy as array 3 (6×5-element rings with $\lambda/2$ spacing at 17 kHz), and rather shows better median and 80th percentile results, even if array 3 has 30 elements, and array 2 has only 15. We can conclude that our technique effectively coalesces the receptions of all elements in an array of opportunity, and yields good location performance by filtering out spatial ambiguities.

In more detail, we observe that array 3 achieves the best azimuth estimates as expected: e.g., it outperforms arrays 1 and 2, owing to the larger number of elements placed along the azimuthal plane. It also yields a lower maximum depth estimation error, but the rest of the CDF is better for the other arrays, which are vertically longer and favor the multilateration approach. Array 2 also achieves a marginally but noticeably lower azimuth error than array 1 because its 5-element pyramidal arrays are rotated by 30° and 60°, respectively, which yields a better discrimination capability over the azimuthal plane.

The results suggest that arrays with elements at different heights typically perform better than shorter arrays. For example, arrays 2 and 5 (having elements at 5 and 4 different heights, respectively) outperform arrays 1 (four vertical elements) and 4 (three vertical elements). Array 2 also achieves the best 90th-percentile error (12.5 m). The good azimuth and depth estimation accuracy of array 2 makes it a very good replacement of array 3. Arrays 1, 3, and 4 show a larger maximum error, but the performance up to the third quartile is comparable with that of array 3: this is remarkable, considering that arrays 1, 3, and 4 have only 10 or 15 elements.

We analyze the performance of arrays 1 and 2 in more depth in Figures 5.10 and 5.11, respectively. Here, we consider three different SNR levels of 0, -10 and -20 dB. For array 1, we observe that the accuracy decreases significantly only for the SNR level of -20 dB, for which the azimuthal angle and the depth estimation errors concur to yield a large 90th-percentile error of about 130 m. The case is different for array 2: with respect to array 1, the better multilateration performance yields comparably accurate azimuthal angle estimates for all SNR values and slightly better depth estimates even at an SNR of -20 dB. Altogether, these improvements drive the global location error to a 90th percentile of about 60 m, which is significantly lower than array 1's.

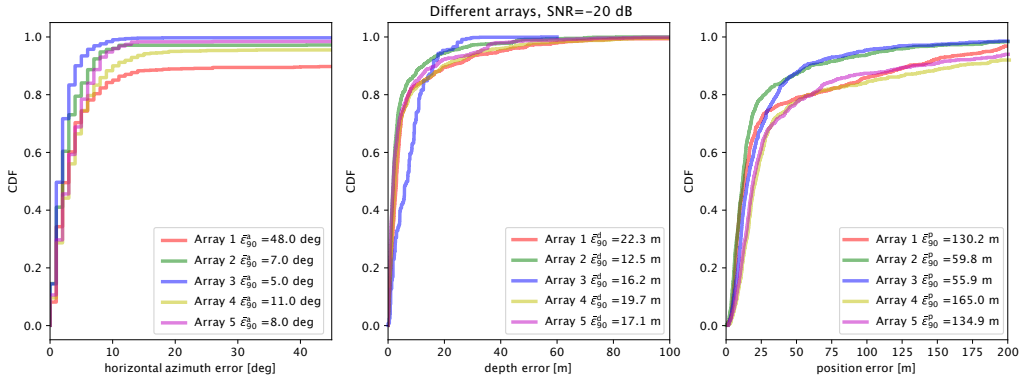


Figure 5.9: Localization error results for arrays 1 to 5, at an SNR of −20 dB: azimuthal angle error (**left**); depth error (**center**); and total location error (**right**).

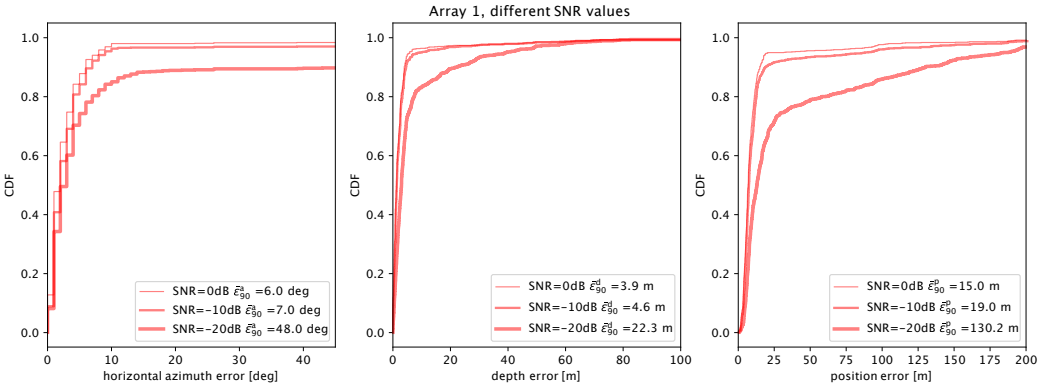


Figure 5.10: Localization error results for array 1 at different SNRs of 0, −10 and −20 dB: azimuthal angle error (**left**); depth error (**center**); and total location error (**right**).

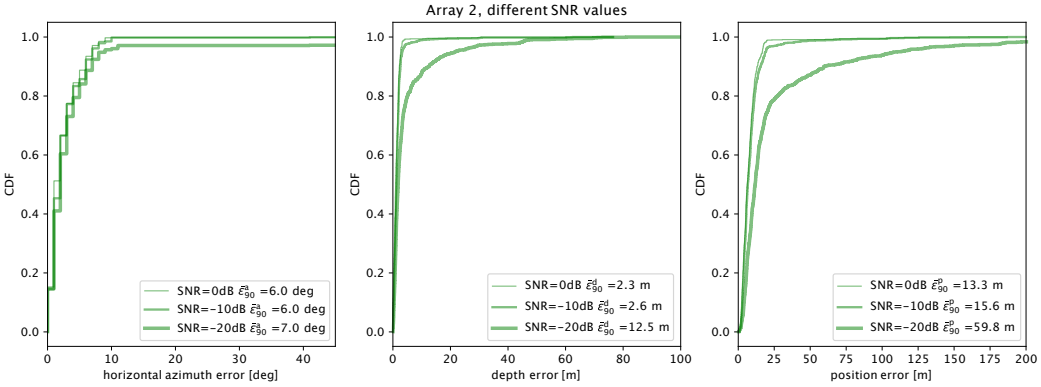


Figure 5.11: Localization error results for array 2 at different SNRs of 0, −10 and −20 dB: azimuthal angle error (**left**); depth error (**center**); and total location error (**right**).

From the above results, we conclude that the triple pyramidal structure of array 2 yields the best trade-off between azimuthal and elevation angle estimation capabilities among the tested array layouts, provided that side information is available to correct the ambiguity arising from the spacing larger than $\lambda/2$ (such as our TDoA-based mask). More broadly, we also conclude that our algorithm is a promising solution to achieve satisfactory array performance when multiple smaller sub-arrays are opportunistically combined into a larger array.

5.5. Lake Experiment Results

In this section, we present the results of the lake experiment whose configuration and setup are described in Section 5.3.3.

We start with Figure 5.12, which shows azimuthal angle estimates for different array processing techniques applied to the acoustic array of opportunity that is part of the SYMBIOSIS platform. We recall that we assume the depth of the target to be fixed and equal to 8 m throughout the experiment. We apply our algorithm to different parts of the array, and specifically: (i) only to the top sub-array (light blue triangles facing up), (ii) only to the bottom sub-array (dark-grey triangles facing down), and (iii) to the full array of opportunity (purple diamonds). Grey "+" markers show the angle estimates yielded by multilateration. Throughout the experiment, the target transmits 20 signals, numbered from 0 to 19 in Figure 5.12. Each marker thus indicates the azimuthal angle of arrival estimate (x-axis coordinate) for each transmission and for the corresponding array. We also depict a solid blue line to denote the ground truth of the target's angle of arrival as inferred from the GPS tracker's readings (cf. Figure 5.6).

We observe that our method estimates the angle of arrival from the target very well. On the one hand, it avoids the sometimes largely erroneous estimates that would be computed by using only either of the two pyramidal sub-arrays (e.g., the erroneous estimates around 50 and 210–240 degrees). On the other hand, our technique fuses information from the two sub-arrays, achieving a more accurate estimate, even when each sub-array would already be satisfactorily accurate. For example, the latter is the case of the sets of readings corresponding to transmissions 0, 1, 18, and 19 in Figure 5.12. Notably, resorting only to multilateration would not yield accurate results, as observed from the several markers located off the ground truth line. Still, multilateration is a good source of side information for ambiguous peak removal, and helps discriminate among different equivalent peaks even when the estimate is slightly off, as for transmissions 3, 5, 16, and 17. Only for transmission 10 does the inaccurate multilateration estimate offset the DoA estimate.

We summarize the statistics of the experimental results in Figure 5.13, which shows the CDF of the azimuth estimation error (left panel) and of the total location error

(right panel) for all configurations considered in Figure 5.12. Operating the full array of opportunity with 10 elements consistently yields the most accurate results. In particular, the 90th percentile of the azimuth error (also shown in the legend for clarity) is only 4°, as opposed to 16° when operating only the top array, 106° when operating only the bottom array, and 135° when resorting to pure multilateration in order to localize the target. The only large outlier for the full array of opportunity corresponds to transmission 10's wrong multilateration estimate, which induces a wrong masking of the delay-sum estimator's output and offsets the angle estimate.

As expected, accurate azimuthal angle estimates translate into more accurate total location error estimates. In our experiment, we achieved a 90th percentile of the location error of 5 m using the full array of opportunity. This value is 5 to 16 times less than the error yielded by operating only the top and bottom sub-arrays, respectively, and 12 times less than the error yielded by multilateration.

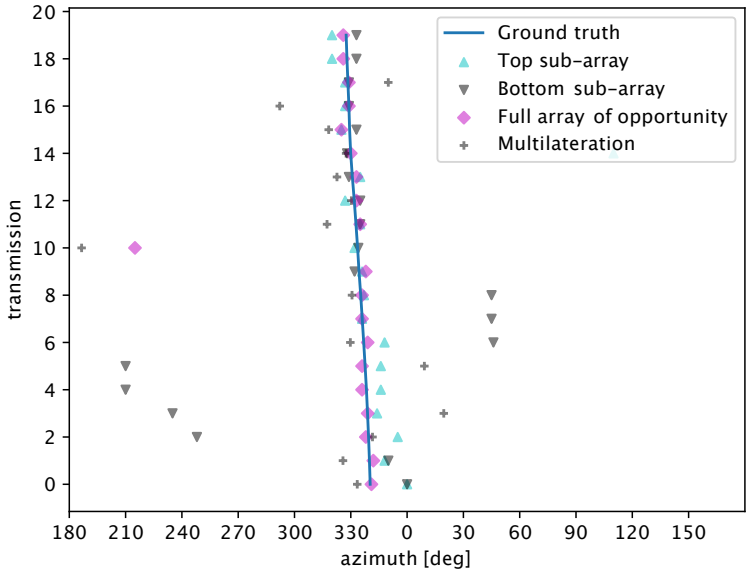


Figure 5.12: Results of the target localization experiment in the Werbellin lake using the SYMBIOSIS array of opportunity (cf. Section 5.3.3.1). While moving, the active target transmits every 2 s for 20 times. Each marker represents the azimuthal angle of arrival estimate (x-axis coordinate) for each transmission and for the corresponding array (light blue triangles: top sub-array; dark-grey triangles: bottom sub-array; purple diamonds: full array of opportunity). Grey "+" markers show the azimuth estimate yielded by multilateration. Our algorithm enables the opportunistic use of two pyramidal arrays, and makes it possible to improve the azimuth estimation accuracy with respect to using a single sub-array or multilateration per se.

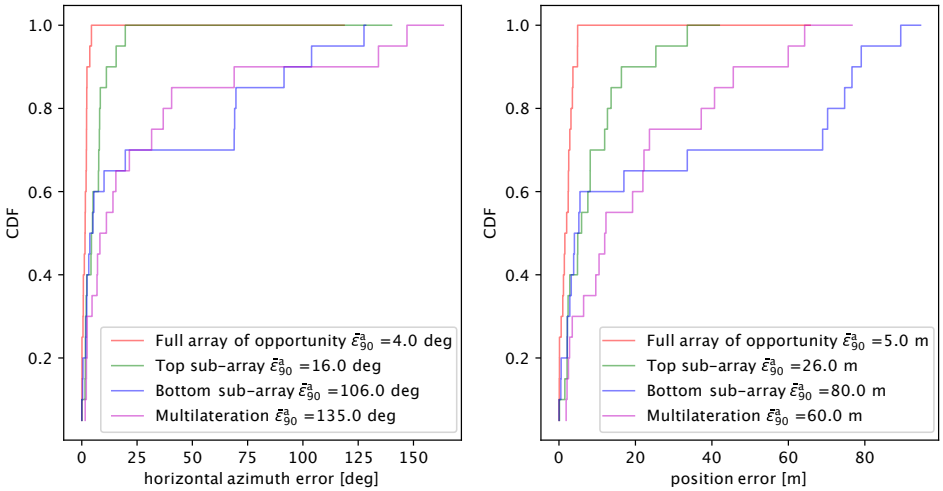


Figure 5.13: CDF of the azimuthal angle estimation error (**left**) and of the total location error (**right**) achieved in the lake experiment, showing the performance of our method as applied to different portions of the array of opportunity, as well as multilateration.

5.6. Summary of Results and Discussion

The results presented in Sections 5.4 and 5.5 show that our solution is a viable underwater observation and telemetry approach. Our key proposition revolves around the concept of acoustic array of opportunity: our algorithm is designed to exploit multiple acoustic sub-arrays or sets of acoustic receivers operating together, and rules out spatial ambiguity issues by leveraging side information from rough location estimates.

The most important implication is that an array of opportunity achieves similar performance as other standard arrays having a larger number of receiving elements and a properly designed topology. This is shown in Section 5.4, where a 15-element array of opportunity is shown to outperform or yield equivalent results as a 30-element cylindrical array. Besides this advantage, our technique may obtain extra value from underwater reception equipment that would otherwise be impossible to merge into an array, yielding economical benefits.

Our experiment results also show that relying only on multilateration would not yield results as good as our algorithm (e.g., our scheme yields a 90th-percentile location error of 5 m, whereas multilateration yields 60 m). Our approach, therefore, can harvest the value of multilateration as side information to improve the accuracy of wideband DoA estimation.

As a final remark, our scheme embeds a wideband delay-sum DoA estimator, which lower-bounds our complexity and running time. However, all additional components, such as DBSCAN and the multilateration step, complete their execution in a negligibly small time. For example, the Julia [135] implementations of the bandpass filter, normalized

matched filter, clustering, and multilateration steps complete in approximately 3.7 ms, 10.8 ms, 0.6 ms, and 0.05 ms, respectively, when run on the CPU of the Jetson TX-2 board employed in our experiments. We note that the Jetson's CPU is much slower than the CPU of the laptop we used to process acoustic receptions.

5.7. Conclusions

In this chapter, we presented a wideband DoA estimation algorithm for arrays of opportunity that coalesce smaller sub-arrays into a larger array, possibly not respecting optimal spacing constraints. We proposed to solve the spatial ambiguity issues that affect such arrays by augmenting a delay-sum DoA estimation algorithm with side information from multilateration.

Our results show that the proposed scheme yields low 3D DoA estimation errors and therefore good localization results. We test our algorithm both emulating real signal reception with the help of actual clutter and noise recordings and in a lake experiment using real underwater arrays in a realistic setting. In both cases, our results show that our algorithm is robust and achieves consistently good estimation performance, often requiring a lower number of elements than typical array topologies with proper $\lambda/2$ spacing.

Future work along the lines of this contribution includes a real-time implementation and test of our algorithm on embedded computers. We also plan sea experiments with different types of targets as well as with larger arrays.

6 Localization and tracking of pelagic fish: the SYMBIOSIS sea trials

6.1. Introduction

In this chapter, we describe a relevant application of the algorithm presented in Chap. 5, namely the localization and tracking of pelagic fish. An opportunity to test the algorithm in this specific setting came in the context of the experimental campaigns carried out under the umbrella of the EU H2020 SYMBIOSIS project [134]. This project targets the long-term, single-location monitoring of underwater environments, with the objective of detecting and measuring the presence, density, and variety of marine fauna species. The SYMBIOSIS platform is a hybrid optical-acoustic system: its design revolves around a chain of progressively more accurate detection, localization, tracking and image acquisition steps. The main stages are as follows:

- *Coarse detection*: the platform emits a 10-ms narrowband signal and processes it through a neural network, attempting to understand whether or not there exists some possible target in the surroundings.

- *Accurate detection*: the platform transmits a sequence of 20 linear chirp signals, in the bandwidth from 7 to 17 kHz, where each chirp signal has a duration of 10 ms, and transmissions happen every 1.7 s. This phase confirms the presence of a target with greater accuracy than the coarse detection step, using a combination of clustering, neural network, and signal processing algorithms.

- *Localization and tracking*: in parallel to the accurate detection step, we run our localization and tracking algorithm. For the specific case of the SYMBIOSIS platform, we adapt our algorithm as described in Section 6.2. The overarching objective of this step and of our contribution to SYMBIOSIS is to enable the prediction of fish trajectories in the vicinity of the platform. As the algorithm infers that one or more specimens are getting closer to the platform, it signals the platform controller to start image and video acquisitions from the underwater

cameras attached to the platform.

- *Optical acquisition and processing*: if detections are successful, the system records and processes images and videos: recorded images and frames are processed to detect the presence of fish specimens, extract bounding boxes that identify the specimens' position within the image or frame, and classify them using neural networks.

Several experiments were organized throughout the duration of the SYMBIOSIS project: an early-stage experiment is described in Chap. 5; subsequent experiments carried out in the context of real longer-term deployments of the whole SYMBIOSIS platform are described in this chapter. These experiments made it possible for us to test the capabilities of our localization and tracking algorithm in relevant scenarios.

Before delving into the results of the SYMBIOSIS experiments, we would like to thank the SYMBIOSIS project partners and collaborators that made the realization of these experiments possible. We especially thank the University of Haifa personnel, that materially carried out the largest majority of the scheduled activities at sea, and in particular the project coordinator, Dr. Roee Diamant.

6.2. Summary of the main extensions for the localization algorithm

The 3D localization of underwater marine fauna requires to accurately estimate the range and bearing underwater objects. The SYMBIOSIS platform enabled this capability by providing recorded acoustic information from a total of ten hydrophones, grouped into two units of five hydrophones each. Each unit is a commercial off-the-shelf (COTS) ultra short baseline (USBL) pyramidal array having a square base with 10 cm of side length, and a height of 7.07 cm. The manufacturer of the USBL arrays and SYMBIOSIS project partner EvoLogics GmbH reconfigured each unit to disable pre-programmed USBL functionalities, and just output recorded acoustic samples.

In order to provide a full solution for the detection, localization and tracking of underwater fauna, we extended and updated the algorithm presented in Chapter 5. The extended version of the algorithm includes the detection of possible targets, the removal of stationary acoustic arrivals (that typically result from reflections of transmitted signals off environment features and platform parts), and the tracking of detected targets. Therefore, the algorithm has to consider historical information from different acoustic records taken over time.

To cover an area with about 500 m radius around the platform, in each cycle the system records 0.7 s of acoustic data. Additionally, it takes about 1 s to save it and pass it to the algorithm, thus each record cycle takes about 1.7 s. Given the time constraint,

working on a very large acoustic dataset covering a long time span would imply that output location information will be likely outdated. However, the modified algorithm requires a sufficiently large number of acoustic records in order to achieve the desired level of accuracy. In order to find a good tradeoff between the above constraints, it was decided to process data with batches of 20 consecutive records. In the following, we detail how we extended and modified the localization algorithm presented in Chap. 5.

- *Target detection*

As discussed in Sec. 5, the basic detection step of the localization algorithm works by clustering peaks from a Normalized matched filter (NMF) output from multiple channels, where we recall that each acoustic channel is the output of a different array hydrophone (see Fig. 5.2). Each target detection is recorded into a database of target detections. Due to the multipath nature of the underwater acoustic channel and to the multiple reflections from environmental features, the algorithm may output multiple detections even when there is no fish around the platform. These detections correspond to stationary reflections from the water surface, or from parts of the platform itself.

- *Stationary arrival removal*

Stationary arrivals are expected to be very stable and consistent over time. In order to remove such detections, we propose a conservative stationary arrival definition: if a detection is present in all 20 records and the speed of the moving target is below 0.01m/s, we consider the object as a stationary one and remove it from further processing. In order to do so, we present the target detections information as a 3D space matrix (range of detected arrival, time of record and total NMF value of all peaks clustered in previous step). Then, we perform another clustering step based on DBSCAN, where we set the algorithm's cluster radius parameter to fit the speed limit, and we impose that the minimum number of points in each cluster should be equal to the minimum number of records in which the detection is present. Our stationary arrival definition is inspired by several experimental results. For example, in Fig. 6.1 we present two seemingly stationary arrivals which were, instead, slowly moving targets. Results like this confirm that the choice of a maximum speed of 0.01 m/s is appropriate, as larger speed would mistakenly label slowly moving targets as stationary.

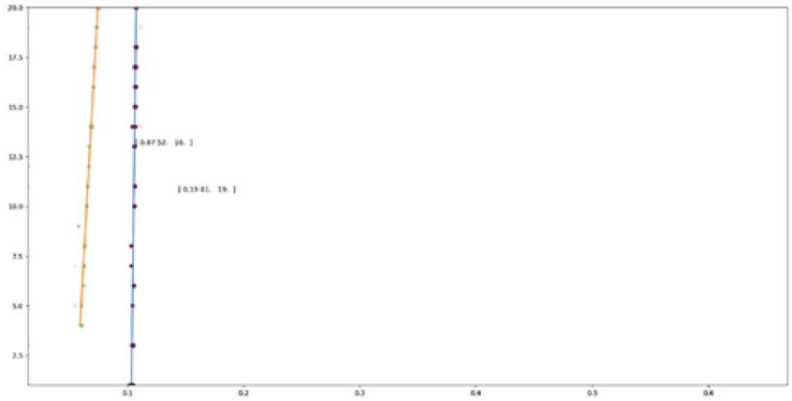


Figure 6.1: Detections related to two seemingly stationary targets over a window of 20 transmissions.

- *Path tracking* After stationary arrivals are removed, there are still multiple detections that do not fit into our requirements (e.g., single random reflections from bubbles, or targets moving too fast, such as motor boats). In order to filter these targets out, and leave only targets with speed that fit to desired parameters, one more clustering step is performed.

Once the target is tracked, we run the localization algorithm. As an output, the algorithm provides information whether any target with desired parameters was detected. If true, it calculates the average speed, the 3D coordinates of the target of interest (using the direction of arrival (DoA)-based approach in Chap. 5) and estimates the time when the target may approach the platform and become visible to underwater cameras. The outcome of the algorithm serves as a trigger to start a camera recording at the estimated time of arrival (ETA).

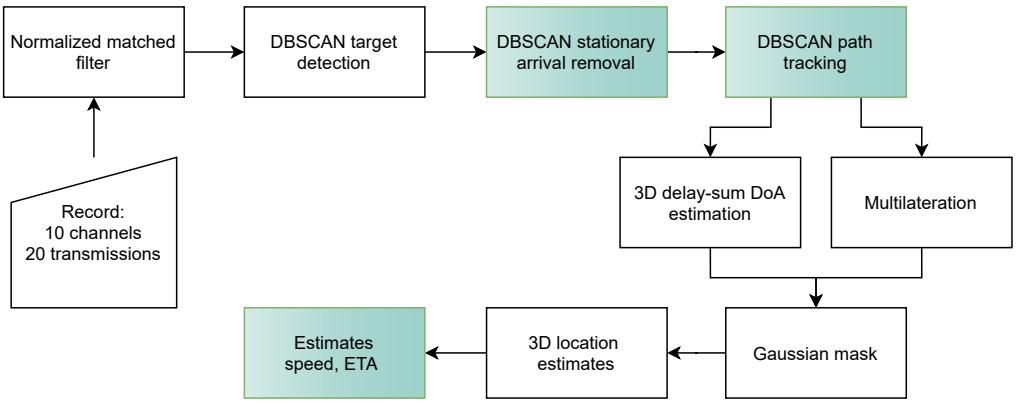
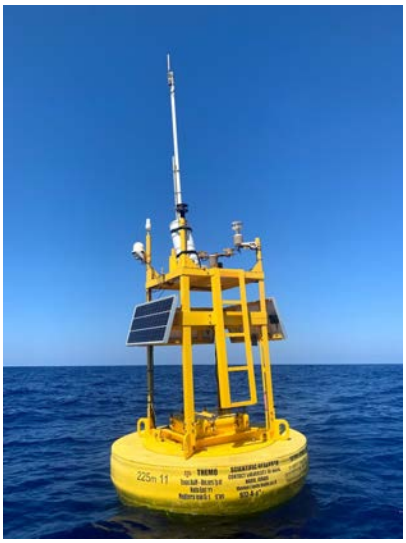


Figure 6.2: Extended version of 3D localization algorithm with tracking. The new parts of algorithm with respect to the one presented in Chap. 5 are colored in green.

6.3. Experiments at the deep-water THEMO mooring

THEMO is a marine observatory that was designed and installed in the eastern Mediterranean sea, in the context of a collaboration between the Texas A&M University, USA, and the University of Haifa, Israel. The observatory consists of two moorings: the shallow THEMO mooring, located in an area with about 120 m of depth, comparatively close to the Israeli coastal zone, and the deep THEMO mooring, located about 50 km away from the shore, where the sea depth reaches about 1500 m. The full description of the THEMO moorings can be found in [95]. During the development stage of the SYMBIOSIS project, several experiments were performed near THEMO deployments.



(a)

(b)

(c)

Figure 6.3: THEMO SYMBIOSIS deployment: a) THEMO marine observatory b) SYMBIOSIS platform with 2 USBL units during the deployment process c) Scheme of SYMBIOSIS's acoustic array with three USBL (courtesy of EvoLogics GmbH).

Our experiment took place near the deep THEMO mooring [136]. Here, SYMBIOSIS personnel used a boat to release a rehabilitated turtle. The acoustic part of the SYMBIOSIS platform that consisted of 2 USBL elements with 5 hydrophones each. An underwater transmitter sent a chirp signal every 0.11 s, and each USBL unit record one acoustic file per channel. Each file contains 20 segments of 0.7s each, with a small technical pause in between segments and a slightly bigger pause between file records, required to save file on a storage unit. This technical feature explains the uneven time spacing between transmissions and causes interruption in detected path segments. We synchronize different recordings using the first arrival as a reference.

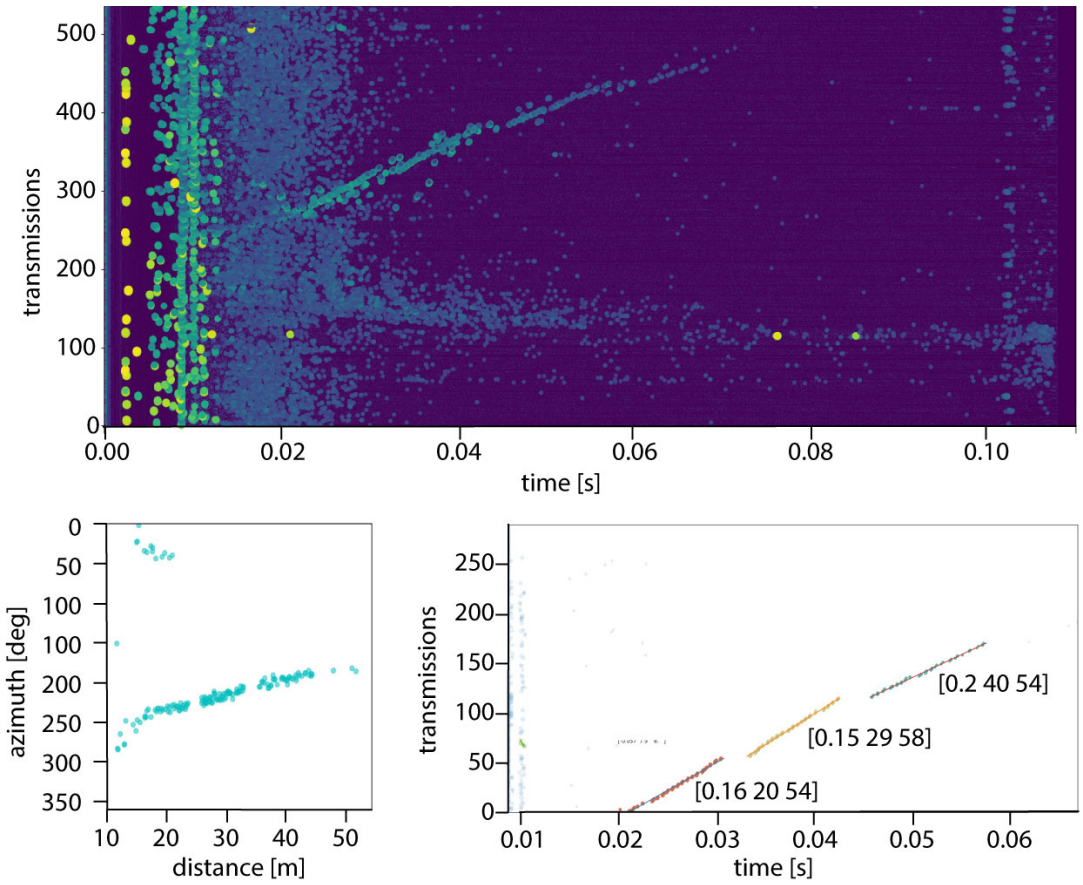


Figure 6.4: Deep THEMO dataset. a) NMF output with relevant peaks b) Azimuthal plane localization c) path tracking

Fig. 6.4a shows a comprehensive view of all relevant normalized matched filter peaks from both USBLs. Each peak represents a reflection from environment or target registered at one of the channels of receiver array. The x-axis represents the peak observation time relative to the start of a given transmission, whereas the y-axis reports the transmission count for each signal sent by the SYMBIOSIS platform (earliest transmissions at the bottom).

The strongest peaks are clearly earliest in the figure, and related to strong reflectors in the environment. Yet, a cluster of peaks starting just before transmission 300 suggests a target getting farther from the ship (as inferred from the increasing observation time of the corresponding peaks). This is compatible with the turtle's release, as the animal swam linearly away from the releasing ship. We note that the trajectory of the animal is composed of two sets of points, one showing a linear movement, and a second set composed of detected target reflections around such a linear trajectory. This is due to some elements of the SYMBIOSIS array being shadowed by other construction components of the array,

and showing a slight jitter in the sampling times. Still, the matched filter peaks are very well connected and compatible with a trajectory of a moving target.

While our algorithm can operate with multiple sub-arrays joined together, one of the key assumptions for this is that element sampling is synchronous. In SYMBIOSIS, this was realized through a sync-in signal sent by the central embedded system to all software-defined sub-arrays. Unfortunately, this signal experienced a malfunctioning during the experiment. Therefore, we employed the rest of the data from the deep THEMO site to also refine the capabilities of the localization algorithm. We tuned the parameters of the detection steps (and in particular, of the clustering step) in order to localize targets independently from each software-defined USBL array mounted on the SYMBIOSIS platform. In the worst case, this would allow us to use a single USBL unit to localize targets by choosing the one with the best output in terms of signal-to-noise ratio.

The deep THEMO deployment represented a perfect opportunity to verify and tune the capabilities of the algorithm to discriminate between static and slowly moving targets. Fig. 6.4c represents automatic isolation of relevant target reflections and target tracking. We observe a stationary arrival detected (blue dots at about 0.01 s) that is probably a reflection from parts of the deployed platform. It was successfully removed from further processing by our algorithm. For each trajectory segment, our algorithm reports three values corresponding to the movement speed, to the average distance throughout the detection, and to the number of transmissions in which the target was identified. We show such values close to each trajectory segment in the figure for clarity. Fig. 6.4b shows the result of fine localization algorithm in azimuthal plane. The algorithm shows consistent and slowly changing bearing angle estimates as the turtle swims away.

6.4. Eilat deployment

The Eilat deployment was performed during November 2020 in the Israeli Red sea, and lasted about 2 weeks. The deployment involved the whole SYMBIOSIS platform, including surface Radio Frequency (RF) communication units to enable a radio link to the platform from a shore location. The platform was deployed at about 5 m of depth, and 500 m away from the shore. The area where the platform was deployed is an extremely challenging acoustic environment. In contrast with the deep THEMO deployment, where reflections from bottom are barely noticeable, the relatively small depth of the Eilat deployment (below 30 m), and the many coral reefs on the bottom of the Red sea create a very rich multipath acoustic channel. Reflections from environmental features are usually stronger than those from small fish, which makes the detection task particularly complex. Moreover, fish detections are rare in this area at this time of the year, as confirmed by biologists collaborating with the SYMBIOSIS project.

For this deployment, we analyse a part of the dataset that consist of about 1 working

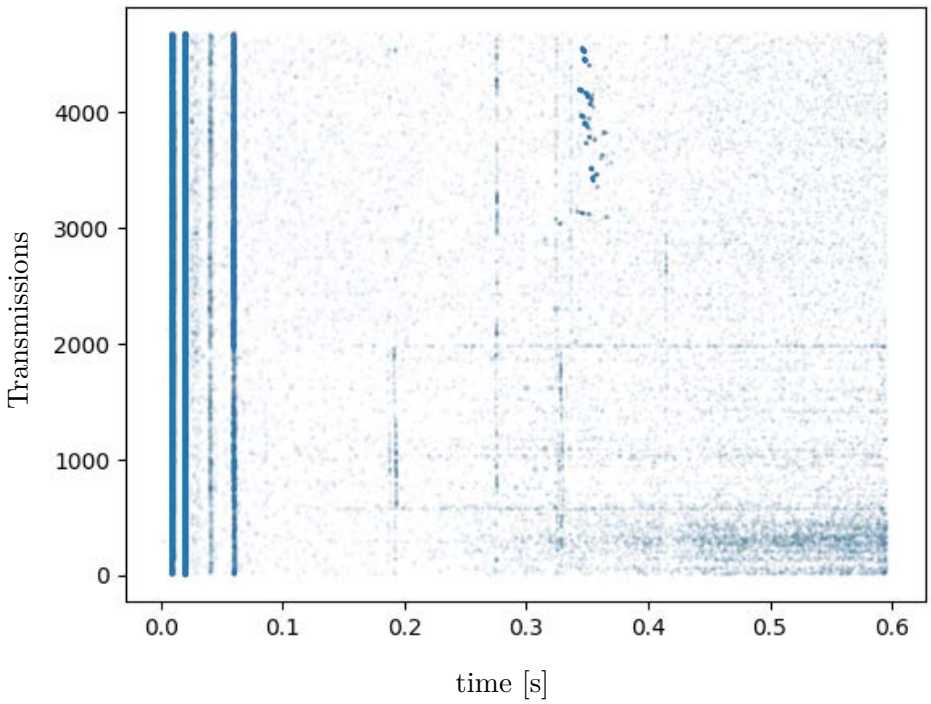


Figure 6.5: Eilat dataset. NMF peaks from over 4600 transmissions. Each peak is represented with a light blue transparent dot. The opaqueness of the dots represents the density of NMF peaks.

day of records. The relevant NMF peaks are presented in Fig. 6.5. The opaqueness of the dots represents the density of peaks in the area of interest. This picture confirms our hypothesis that the peak density is a good indicator of the presence of a reflecting object. We remark that several attempts to tune the sensitivity of the platform were performed initially, as shown by the greater level of noise at the beginning of the data set, as well as by the stationary arrivals (that appear as vertical lines in the figure), whose intensity increases and decreases between transmissions 800 and 4500. We will now focus on the last 1650 transmissions.

Fig. 6.6 shows a closer look at the waterfall of acoustic sample values for one acoustic channel of one software-defined USBL unit. In the figure, green hues represent the noise level, whereas yellow hues convey a stronger signal. The figure shows once more that array processing is paramount for underwater target detection even before accurate localization: the signal corresponding to the target is buried in noise, it remains invisible to manual inspection, and needs to be matched with additional acoustic channel outputs in order to improve the signal-to-noise ratio. In fact, simple signal thresholding would not enable reliable target detections. The Fig. 6.7 shows results from our proposed algorithm. After applying all clustering steps, the algorithm could detect multiple target (depicted in orange) and stationary arrivals (depicted in blue). We note that some arrivals did not fit the conservative conditions of stationary arrival definition, however most of them are

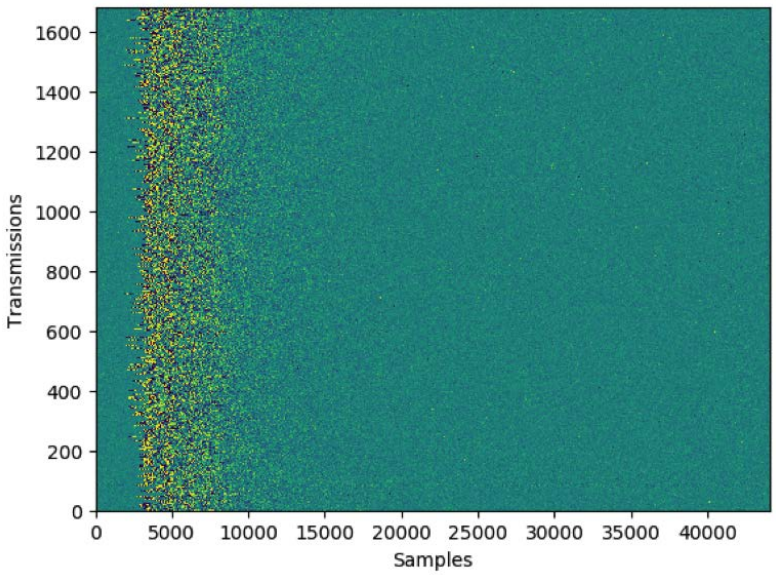


Figure 6.6: Eilat dataset. Waterfall matrix that represents raw acoustic data from one of the USBL channels for transmissions 3000-4600.

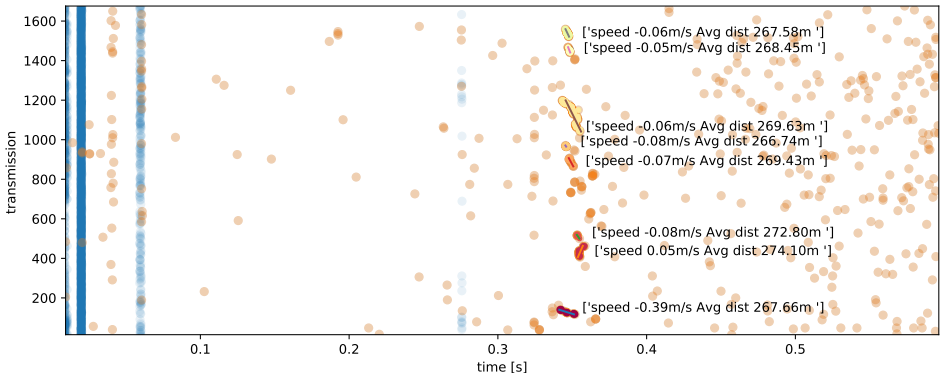
removed from further processing. Then the algorithm tracked several segments of path of some moving target.

As a final summary, Fig. 6.8 provides the statistical distribution of target detection features throughout the Eilat deployment. The three panels show the distribution of the azimuthal angles of arrival of all target detections (Fig. 6.8a), the distribution of the speed in m/s for each path tracked (Fig. 6.8b), and the distribution of the distance of the targets found (Fig. 6.8c). In line with previous results, we observe that there was a cluster of detections at about 100°, and a second cluster at 330°. Considering only the paths tracked, most of the detected targets were moving slowly, at less than 0.1 m/s. Most targets were detected at a distance of more than 250 m from the platform.

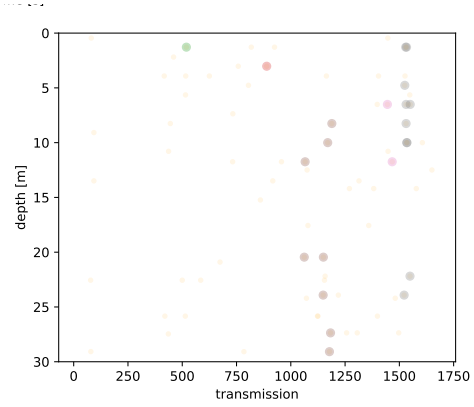
6.5. Conclusions

In this chapter, we summarized our work on the validation of the localization and tracking algorithm of Chap. 5 in real sea environments, where the task of the algorithm is to detect and track marine fauna specimens, in order to establish when fishes would come sufficiently close to an underwater platform. The algorithm forecasts when fish targets approach the platform and triggers image acquisitions from the platform's cameras.

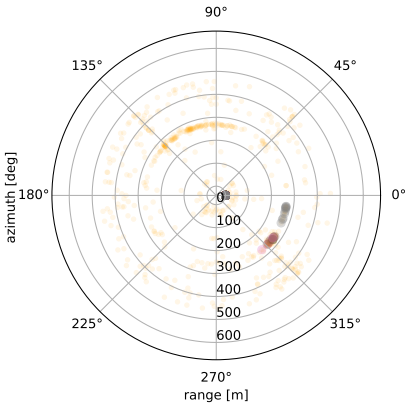
- To seize the opportunity of these deployments and validate the localization algorithm, we adapted several aspects of it, in order to tackle some specific challenges of the project. With respect to the version of Chapter 5, the extended version presented in this chapter can not only localize, but also to successfully detect and



(a) Results of detection and tracking algorithm.

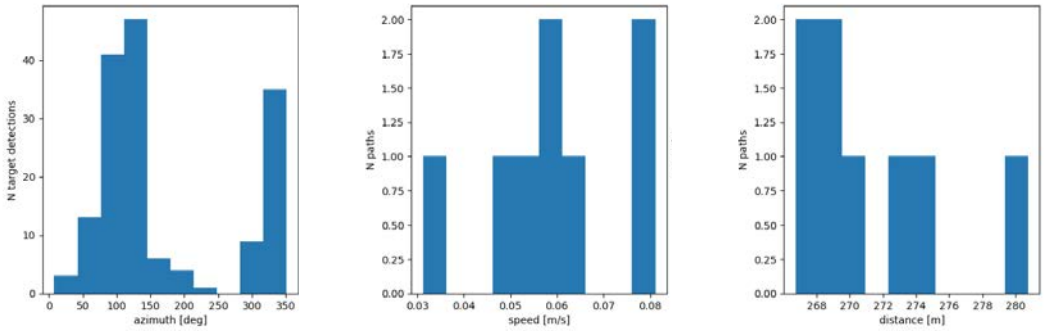


(b) Localization in depth plane



(c) Localization in azimuthal plane

Figure 6.7: Eilat dataset. a) NMF relevant peaks. Each dot represents a cluster of peaks with a possible target detection. Blue dots represent clusters that are defined as stationary. Lines represent clusters that are detected as possible target tracks with relevant speed. b) Depth plane localization of relevant target tracks limited by environment boundaries c) Azimuthal plane localization. Orange points represent all relevant target tracks. Other colors represent corresponding parts of relevant target tracks after filtering out detections with depth out of boundaries.



(a) Distribution of the azimuthal angles of arrival for each detected target

(b) Distribution of the speed for each targetpath tracked;

(c) Distribution of the average distance throughout eachtarget path tracked

Figure 6.8: Eilat dataset: Statistics of detections.

track underwater targets. We adapted the software to work in the presence of various challenges including lack of synchronization, changing number of acoustic channel recordings and various transmission parameters.

- We have tested the proposed algorithm in various natural water basins. Besides sweet water lake experiments in Chap. 5, we were able to test it in various salty water deployments. In this chapter, we consider two different tests environments: one in relatively deep waters in the Mediterranean sea, and another in very shallow waters in the Red sea. The results of both tests proved that the algorithm discriminates the environmental features of acoustic reflections from other target reflections of interest. In both cases the estimated information of target movements corresponds to the expected ones.

Summarizing the results of these tests, the outcomes show that the algorithm can localize not just highly reflective or active targets as in the tests of Chapter 5, but also smaller and weaker targets such as the fish species of interest for the SYMBIOSIS project. The algorithm works as expected in different environments. Further work in this area may consider further improvements of real-time signal processing performance through embedded graphical processors.

"How inappropriate to call this planet "Earth," when it is clearly "Ocean"."

Arthur C. Clarke (1917 - 2008)

7 Conclusions

Accurate, energy-efficient and minimally invasive underwater localization is an important step for many ocean sciences. In this thesis, we have tackled various aspects of underwater ranging and localization. We have identified several research challenges posed by the underwater acoustic propagation environment, and proposed algorithms to deal with these challenges. The key outcomes of this work are the following:

1. We studied the accuracy of range estimation based on time difference of arrival (TDoA) of multipath replicas under different conditions, and proposed a cost- and energy-efficient solution to improve range estimates in the presence of non-linear propagation without using additional equipment. Future extension of this problem may include an in-depth analysis of arrival history, in order to improve the matching between measured data and the geometry of the environment.

2. We proposed and evaluated a non-invasive and computationally effective algorithm for the localization of acoustically active underwater objects including vocalizing marine animals or autonomous underwater vehicles (AUVs). The proposed solution relies on a single hydrophone receiver. Further improvements in the computational effort may be achieved by optimizing the grid of locations used to characterize the environment (e.g., making the mesh finer at increasing distance and more coarse closer to the single receiver.

3. We designed an algorithm for the 3D localization and tracking of underwater targets that works with wideband acoustic arrays of arbitrary shapes, including those that do not obey proper array element spacing constraints. We demonstrated its effectiveness in several experiments and sea trials. Further extensions of this algorithm can tackle cases with imperfect array assembly and synchronization issues.

Overall, this work has shown the effectiveness of using additional information about acoustic signal propagation and its benefits in further application of this approach in similar problems.

Appendices

A ASUNA: A topology data set for underwater network emulation

Underwater communication devices have been steadily improving over time in terms of both reliability and bit rate [137], and can be arranged into underwater acoustic communication networks (UWANs) to support a broad variety of applications [138]. A multitude of protocols have been designed for UWANs to date, providing different functionalities at different layers of the ISO/OSI protocol stack [139–141].

Sea trials are considered a good option to test the performance of UWAN protocols. However, organizing and performing a sea experiment is usually time-consuming, effort-intensive, and implies high costs in terms of materials, rental of ship time, purchase, transport and deployment of underwater transceivers, etc. Moreover, sea experiment results are related to a local set of environment and channel conditions, which makes it difficult to extrapolate the results to different environments. Finally, a single experiment does not allow a fully fair comparison between algorithms.

As a result, simulations are often preferred when evaluating the performance of newly designed protocols against competing approaches in the literature. Simulations make it possible to approximately evaluate communication protocols and schemes by abstracting from specific hardware issues. However, on the one hand there are no statistical models of the underwater acoustic channel that are broadly agreed upon, so that a realistic simulation often has to rely on complex numerical propagation modeling (e.g., [30]); on the other hand a full-fledged evaluation must take into account the many practical issues that occur in actual underwater scenarios. This includes the time-dependency of the acoustic channel, the conditions of actual underwater environments, and possibly the behavior of hardware devices.

Hardware-in-the-loop systems are one of the means to improve the agreement between simulated protocol performance and actual performance at sea, at least in terms of the peculiarities of underwater transceivers. Examples of frameworks offering hardware-in-the-loop capabilities include DESERT Underwater [142], SUNSET [143], UNetStack [144], Aqua-Net-Mate [145], NETSIM [146], as well as the software-defined

cognitive communications architecture presented in [147]. These capabilities are made possible by interchanging the procedures that simulate underwater propagation and compute link budgets with software drivers for specific underwater modems. However, even the hardware-in-the-loop concept can reproduce actual underwater propagation and the variation thereof over time only to a limited extent.

The above discussion leads to the conclusion that, in the absence of both a fully detailed simulation model of an underwater acoustic communication system and of the resources to organize a sea experiment, a reliable performance evaluation method should preferably involve recordings from a real sea environment.

For the design and test of point-to-point underwater communication systems in realistic conditions, the community often resorts to publicly shared communication datasets in order to reproduce the broadest possible span of underwater channels (e.g., long or short delay spread, heavier or milder Doppler spread, single or multiple receivers, etc.). Examples include the measurements presented in [148], the SPACE08 and the KAM11 datasets, employed among others in [149–152]. More recently, the release of the Watermark benchmark [153] makes it possible to reproduce the distortion of acoustic waveforms transmitted through underwater channels that are either measured or stochastically replayed.

In the following, we propose a similar solution for the testing of underwater network protocols, named *ASUNA*, for "A shared underwater network emulation dataset." ASUNA is a collection of measurements from multiple sea experiments, and aims to be the first freely shared database that enables the replay of underwater acoustic networking trials, often referred to as *emulation*. To the best of our knowledge, this is the first attempt to assemble a dataset for the direct evaluation of the performance of network protocols. ASUNA provides a collection of time series of link quality indicators, collected over time during several sea experiments at different locations around Europe, Israel and West Africa. These experiments are representative of a broad set of conditions: different numbers of nodes, different deployments resulting in multiple network topologies, different transceivers, and multimodal setups (where communications are realized through a set of orthogonal technologies). Once this data has been loaded into a network emulator, the link quality time series can be used to reproduce the same realistic performance that could be experienced at the same location and time each experiment had been carried out. As a result, the user can evaluate networking solutions with a degree of accuracy that stands in between a simulation and a full-fledged sea experiment, in a fully reproducible setting, without having to actually go to sea.

In total, ASUNA includes 22 network topologies from 7 different sea experiments, for a total of more than 10 hours of underwater data packet transmissions. We make the data available in an Octave/Matlab format, so that it can be easily manipulated, converted to other formats, as well as integrated into existing Octave/Matlab code. For each dataset,

we document the experiment it is extracted from, so that the user knows the experiment's location and time; the location of the nodes; the conditions of the water body at that time; and the types of link quality measurements available for that experiment. The metrics provided by the dataset so far include received signal strength, bit error ratios, and "1/0" indicators conveying whether a given packet would be received correctly or not if transmitted at a given time. Different metrics may be embedded in the future as additional datasets are added to the collection.

Along with the database, we provide a network emulation code as an example of how to use ASUNA. The emulator runs a simple time-division-multiple-access (TDMA) protocol over the recorded topologies. Yet, by no means is the usage of ASUNA confined to such a solution.

While there is some novelty in our approach and it has been recently endorsed that reproducible and interactive research results bear significant value for the underwater community [154], the focus of this appendix is on the tool per se, rather than on novel results obtained through it. In the following, Section A.1 provides an account of related work; Section A.2 describes the ASUNA dataset; Section A.3 discusses the emulator provided with the dataset and some results obtained with it; Section A.4 concludes the work.

This work is a collaboration among several researchers and institutions. Our contribution involved the processing of sea trial data, the harmonization of the datasets throughout different experiments, and the programming of a demonstration code that exploits the dataset for emulation purposes. The code has been used to provide the results in this appendix.

A.1. Related Methods

In terrestrial radio networks, it is customary to evaluate the performance of wireless networking protocols by means of simulations, supported by different types of channel models [155–157]. Initial studies on channel modeling for underwater networks followed the same approach. For example, [158] modeled packet errors from the SubNet09 campaign using Markov and hidden Markov models. Typical statistical distributions of large-scale underwater channel gain [87, 159, 160] have been observed to be valid across a number of channel measurements.

Besides simulation, network performance can be evaluated through emulation or trace-based simulation. Emulation refers to the use of realistic networking hardware, or to the execution of actual applications on top of hardware components that reproduce the behavior of wireless networking equipment. For example, this implies running complex channel models in real-time in some dedicated hardware. Trace-based simulation [161–163], also described as channel replay-based, relies on the recording of the time series

(or "traces") of link quality metrics [164]. This makes it possible to exactly reproduce the same wireless channel conditions repeatedly, and to test different protocols in fully comparable scenarios. For tests that do not require to learn the channel evolution over time, the evaluation can be extended by suitably scrambling the measurements so that channel properties remain statistically coherent [165, 166].

In the underwater community, several works have tackled the reliable and validated reproduction of the communications performance measured during experimental campaigns. These studies mainly focused on the physical layer. For example, [148] proposed to collect underwater channel recordings in order to reproduce the impact of the acoustic channel on underwater modulation schemes. The collected dataset includes channel estimates from several sea experiments. More recently, Watermark [153] has been released as a benchmark for underwater modulation schemes. Watermark is based on the validated MIME tool, which enables both direct and stochastic underwater channel replay [151, 167]. In some cases, channel estimates can be directly obtained through deployed infrastructure that is shared with the community at large, typically for limited periods of time and under some form of collaboration agreement. This includes the NATO CMRE LOON [168], the equipment of Ocean Networks Canada [169], the SUNRISE testbed federation [170], as well as permanently online infrastructure such as the THEMO observatory [171].

Besides direct and stochastic channel replay, other methods have been considered to enable model-based channel reproduction. For example, in [172] the authors propose to evaluate the reliability of underwater communications through the multipath structure of previously measured underwater channels, which can be evaluated using numerical models rather than sea experiments. Realistic channel simulations obtained through the Bellhop ray tracing software [30] have been incorporated in the World Ocean Simulation System (WOSS) [27], a framework that automatically retrieves the environmental information required by Bellhop in order to compute attenuation figures and channel impulse responses. A similar integration of models based on parabolic equations in network simulations is discussed in [173]. Like many other channel simulators, both Bellhop and a parabolic equation solver present the issue that their output is deterministic for fixed boundary conditions. This was addressed, e.g., in [87], which provides time-varying channel realizations as would result from the movement of the transmitter and receiver around their nominal locations. When numerical models or stochastic replay are not sufficient, hardware-in-the-loop systems offer one additional degree of realism by allowing network protocol code (typically written for simulations) to run on actual underwater transceivers. Examples of this approach include DESERT Underwater [174], SUNSET [143], UNetStack [144], Aqua-Net-Mate [145] and NETSIM [146].

Replicating a real underwater communication experiment in network simulations is often challenging and necessarily leads to approximations. Typical approaches include:

placing nodes at random in an area and using acoustic models to predict the success of packet transmissions [10, 175–178]; simulating node motion, especially in the presence of autonomous underwater vehicles (AUVs) or other types of mobile nodes [179–182]; letting nodes drift, e.g., by using water current models [183–185]; and injecting the acoustic noise generated by ships and AUVs navigating near the network deployment [186].

While the above methods approximate realistic scenarios to some degree, only in sea experiments can all the details of actual underwater communications be taken into account. Experiments with a large number of nodes were demonstrated by large organizations or collaborations. Relevant examples include the joint TNO/FFI tests on the NILUS node [137] (7 nodes); the collaborative experiments promoted by the NATO STO CMRE, such as CommsNet13 [187] (up to 9 nodes); the MISSION 2013 campaign [144] (10 nodes); the final sea trial of the RACUN project [188] (15 nodes); as well as the Jaffe lab sub-mesoscale ocean sampling experiment, featuring 5 static pingers and 13 passive drifters [189].

Besides their complex logistics and cost, underwater networking experiments still capture only the local conditions of the underwater channel at a single location and time: such conditions are not easily extrapolated to different times and scenarios. Through ASUNA, we provide a number of experiment traces, each conveying recorded time series of link quality metrics for all links of several networking experiments. Our objective is to grow ASUNA into a rich and significant benchmark tool through contributions from the community: however, the experiments initially provided already represent a number of different conditions. ASUNA enables "network replay" in a form similar to [10] and [190], which employed previously recorded time series of the signal-to-noise ratio (SNR) or of successful packet receptions in order to test the performance of underwater scheduling and routing protocols, respectively. There are also similarities with the physical layer replay capabilities of the architecture in [147]. However, while the focus of the above approaches is on the performance evaluation of specific protocols or communication architectures, our objective here is to provide a growing collection of network communication traces. In doing so, we aim at making available a tool that remains positioned between pure simulation and pure experimentation, and that joins the repeatability of trace-based simulation with the rich representation of environments and contexts provided by a sea trial database.

A.2. Description of the Dataset

A.2.1. Overview and link reliability measurements

The ASUNA database is available for download at https://sites.google.com/marsci.haifa.ac.il/asuna/. ASUNA's databases are basically constructed as time series of link reliability metrics opportunistically collected from UWAN experiments at sea. In each

experiment, one or more network topologies were tested.

Link reliability signifies the integrity of the communications between adjacent nodes. It enables hard decisions about the existence of a link (e.g., by setting a threshold on the metric) or, alternatively, soft decisions (e.g., tying the bit error ratio to the probability of packet error). The link reliability is typically a time-varying property. This is especially true for underwater acoustic communications, where the channel impulse response and the ambient noise tend to change rapidly. While emulating physical layer reliability requires a fine time resolution (at least matching the symbol rate), the resolution constraint can be relaxed for the evaluation of underwater networks, where the most important aspect is typically the average (rather than instantaneous) link performance throughout the duration of a packet. In our experiments, we either (*i*) collected data on a per-packet rather than per-symbol basis, or (*ii*) relied on link metrics returned by the modems. The latter are derived either from a packet's preamble, or by observing whether packets are successfully received. We remark that such phenomena as flickering (a condition by which a link appears and disappears at a fast rate in the network's topology) are still present in our topologies at packet transmission time scales, and still enable the evaluation of adaptive protocols that specifically react to such phenomena.

We employ both physical layer and network layer metrics to characterize the link's reliability. Depending on the experiment, we provide: bit error rate (BER) values computed as the ratio of correctly received bits over the total number of bits in a received packet; Received Signal Strength Indicator (RSSI) values related to voltage readings at the receiver upon packet reception, or 1/0 flags that convey whether a link is available or not at a given time epoch. While these metrics can serve for experiment replay, future contributors of ASUNA are welcome to also record quality indices that are more specific to the setup of their experiment including, e.g., the packet error rate (PER) packet error ratio (PER) or the link throughput.[1] We remark that the datasets of ASUNA are opportunistically extracted from experiments originally designed to test specific communication protocol and schemes. As a consequence, the availability of link metrics depends on the logs collected from the experiment, and may vary across different sea trials. Moreover, the experiments were not necessarily focused on collision modeling. We leave the collection of collision-specific datasets to future extensions of ASUNA. In the meantime, it is still possible for ASUNA users to model collisions approximately by assuming that concurrently transmitted packets are always lost or that they are recovered with a given probability (e.g., as in the case of frequency-hopping schemes, where the recovery probability can be determined based on the hopping pattern).

[1] Providing fine-grained information about the packet transmission and reception times as well as about multipath propagation would be very convenient and would convey additional details about acoustic propagation at the time of the experiment. Unfortunately such accurate information is not available for the current version of ASUNA. We still plan to include it for any future datasets we will integrate, provided that these datasets can demonstrate sufficiently accurate time reckoning and multipath measurements.

A.2.2. Topology matrix information (TMI) structure

For each experiment, our database includes a description of the experiment's setup, an Octave/Matlab `.mat` file grouping link quality time series into a matrix for (called topology matrix information in the following, or TMI for short), and a reference to the publication(s) that convey the context of each experiment. The basic building block of each TMI is an instantaneous snapshot of the quality of all links. This can be seen as an $N \times N$ matrix, whose entry (i, j) reflects the link quality between nodes i and j as measured from the experiment, and where N is the number of nodes in the network.[2]

The time variation of the TMI is captured by adding a time dimension to each topology matrix. The sampling time depends on the context of the experiment and on the configuration of the communication protocols. For example, for an experiment based on a time-division-multiple-access (TDMA) schedule, the topology information is obtained for each time frame. Conversely, in experiments focusing on the physical layer, we update the topology information once for every transmitted packet. Still, the sampling time is sufficiently frequent to enable the interpretation of the topology information as a continuous process.[3] During replay processes it is then possible to, e.g., check the quality of a link at the time of each transmission in order to determine which data packets are correctly received, and how many useful application bits they carried, so as to compute the goodput (defined as the rate of reception of useful information bits over time); alternatively, it is possible to provide the communicating nodes with a noisy version of the TMI to emulate some form of topology instability.

In some experiments, the time variation of the TMI was achieved through the dynamic relocation of one or more nodes in the same area. In this case, we provide link data for each topology separately in the same `.mat` file, with the understanding that the duration of the experiments may be different for each topology. The ASUNA dataset is generally obtained from static deployments. Some of these deployments include drifting nodes (e.g., the REP and Haifa Harbor datasets), which leads to limited mobility. To improve the possibilities for the user to simulate some form of mobility, as well as to emulate underwater networking scenarios where abrupt link quality changes occur, we also provide a global time series that covers a whole experiment across all tested topologies. This is obtained by concatenating the link measurements of each TMI. In fact, between subsequent topologies in a given dataset, some links typically disappear, some new links appear, and those that persist experience significant quality changes. Additionally, we remark that mobility can be approximately emulated by rotating the position of the nodes throughout the locations

[2]Note that the TMI may be asymmetric. This is the case when the SNR is location- or depth-dependent, and in scenarios involving near-far conditions, where interference blocks one end of the communication link.

[3]We remark that the link sampling time is a feature of the data provided in the dataset, and depends on the structure of the experiment from which we derived the link quality measurements. For this reason, it is not possible to configure this parameter.

indicated in each dataset.

In case several communication technologies are involved in an experiment, as is the case for multimodal network setups, a further dimension is added to the TMI. In this case, the time-varying TMI is provided per-technology. This makes it possible to have simultaneous or very close samples of the link quality perceived by different communication technologies. We remark that different technologies often have different transmission capabilities. For example, this is the case for the SC2R high-frequency (80-120 kHz) EvoLogics modem, which has a much higher nominal bit rate than the EvoLogics modem working in the 7-17 kHz band. Such different bit rates cause asynchronous channel sampling at unequal rates. Details about the sampling time are provided in the companion document of each dataset in ASUNA.

A.2.3. Analysis of TMIs

The resulting TMIs that create the heart of the database can be analyzed in different ways. For example, by setting a threshold over the link measurements, one may create an emulation system that avoids a physical layer and only uses realistic binary topologies to form time-varying communication links. This may become relevant when testing scheduling and routing protocols. The user can also treat the soft link quality measures to form a time-varying statistical model that generates links based on measured link reliability information. While some of our reported TMIs are small in terms of the number of nodes or short in terms of the testing time, the network size can be virtually increased by duplicating parts of it, and the time duration can be extended cyclically. In this manner, larger networks and longer deployment scenarios can be tested more reliably than using models, although such an extension to the network cannot be considered as a replay.

An illustration of the emulation process is given in Fig. A.1. The process begins with link quality data collection during a single sea experiment to form a matrix of time-varying TMIs. The experiment may include several arrangements of the network nodes into different topologies. The link quality data is used for *network replay*, where the time-varying link quality information determines the success of each data transmission. Similar to channel realizations used for channel replay [167, 191], the result is a reliable representation of the network performance in the sea conditions that occurred during the recorded network topology.

A.2.4. Structure and variety of the shared datasets

In this section, we describe the structure of the network TMIs currently available in ASUNA. When downloading ASUNA from the web site, TMIs come organized in separate folders. For a given TMI, call N the number of nodes, P the number of (physical layer)

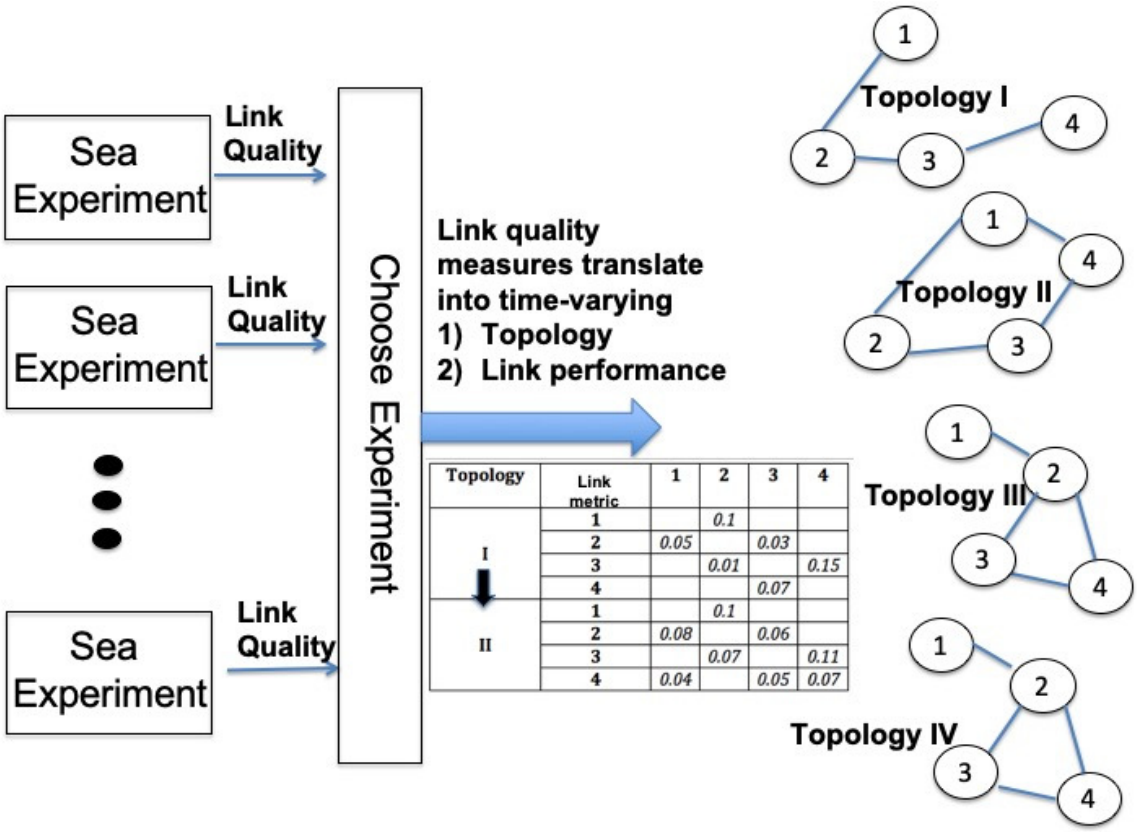


Figure A.1: Illustration of the network emulation process.

transmission technologies available to each node, and T the total number of link quality sampling epochs. Normally, these epochs are separated by an interval $\Delta t = 1$ second, unless otherwise stated in the experiment description. The `.mat` files of the TMIs have the same structure, and contain the following data:

- a `TopMat` matrix of size $T \times N \times N \times P$, where each entry `TopMat(t,i,j,p)` (using Octave/Matlab notation) conveys the link quality for the link between nodes `i` and `j` through physical layer technology `p` at time `t`;
- a `LocMat` matrix of size $T \times N \times 3$, where the three entries `LocMat(t,i,1:3)` represent the two UTM coordinates and the depth of node `i`, respectively;
- a `TechMat` matrix of size $T \times N \times P$, where each of the $k = 1, \ldots, P$ entries `TechMat(t,i,1:P)` is 1 if node `i` has technology k at epoch `t`, and 0 otherwise;
- an `AdjMat` matrix of size $T \times N \times N$, where each entry `AdjMat(t,i,j)` is 1 if nodes `i` and `j` are linked by any technology, at time `t`, and 0 otherwise.

A single experiment may contain measurements either for a single or for multiple TMIs. In the latter case, we provide the above matrices for each TMI separately, and

name them, e.g., `TopMat1`, `TopMat2`, etc. We also provide four matrices resulting from the concatenation of all matrices over the time dimension. The latter are called `FullTopMat`, `FullLocMat`, `FullTechMat` and `FullAdjMat`, respectively. This enables the emulation of abrupt link connectivity changes, as is often the case in UWANs. In particular, such changes may serve to emulate the performance of adaptive protocols.

The experiments from which the dataset has been retrieved were performed for a number of different purposes and applications, including the design of scheduling protocols, physical layer tests, and underwater communications security. As a result, each experiment has peculiarities which make it different from others in our database, and contributes to increasing the coverage of a variety of scenarios. In fact, there exist broad differences among the tests: from relatively large networks of 10 modems, to small link tests with 3 modems; from experiments of long duration (up to a few hours) to short experiments of a few tens of minutes; from tests including one type of modems to multimodal tests including multiple acoustic communication transceivers operating in orthogonal bands; and from tests involving commercial modems to tests that include custom modems and offline processing.

A.3. Example of results

We now present the results of a network emulator built upon the ASUNA database. We remark that these are just meant to serve as an example, and that the applications are by no means limited by the scope of our results. In Section A.3.1 we describe the structure of the emulator, whereas in Section A.3.2 we provide its results.

A.3.1. Structure of a network emulator

Our example of emulator is a discrete-event system written in an Octave/Matlab-compatible code, and comes with all datasets currently shared. These datasets are already placed in the right subdirectory structure to make it possible to load them correctly in the simulator. In this way, the user can open the main file, `TDMAsim.m`, and run it upfront to obtain some first results. The emulation code is freely provided along with the dataset on the ASUNA web site, and the users may employ, extend or modify it to suit their purposes. The code has also been uploaded to the Code Ocean platform [192], from where the results provided below can be reproduced.

The baseline emulator implements an interference-free TDMA scheduling protocol, where each node is assigned an exclusive time slot to transmit a unicast packet to any of its neighbors. The parameters of the protocol can be tuned via a configuration script named `setGlobals.m`. In the main file `TDMAsim.m`, a marked section instructs the user how to choose their desired dataset by commenting/uncommenting specific lines. After importing the data from the corresponding `.mat` files into the structures of the simulator,

the emulation sets up the TDMA schedule and arranges a periodic computation of network metrics.

The TDMA schedule is computed based on the distances among the nodes as derived from the `LocMat` matrix. For a given sound propagation speed (system parameter), the emulator computes the time slot length as the sum of the packet duration (also a system parameter) and of a guard interval as long as the maximum propagation delay in the network. For each TDMA transmission, the emulator uses the instantaneous TMI in order to infer the one-hop neighbors of the transmitter (through the `AdjMat` matrix). The unique destination is then chosen at random out of this list. In case a multi-modal communication dataset is chosen, the emulator also checks which communication technologies are in use both by the transmitter and by its receiver (through the `TechMat` matrix) and chooses one of them at random. The transmission outcome is finally determined by comparing the link quality from matrix `TopMat` to a threshold (system parameter). In the provided code, such threshold is pre-set in order to make it easier for the user to immediately operate with the data, but can be changed in order to obtain different results.

At tunable intervals, the emulator collects relevant metrics for post-processing. This includes a count of the transmitted and correctly received packets, as well as the network throughput. The metrics are plotted at the end of the emulation, and the resulting figures are saved as images. Next, we show results obtained from our TDMA emulation.

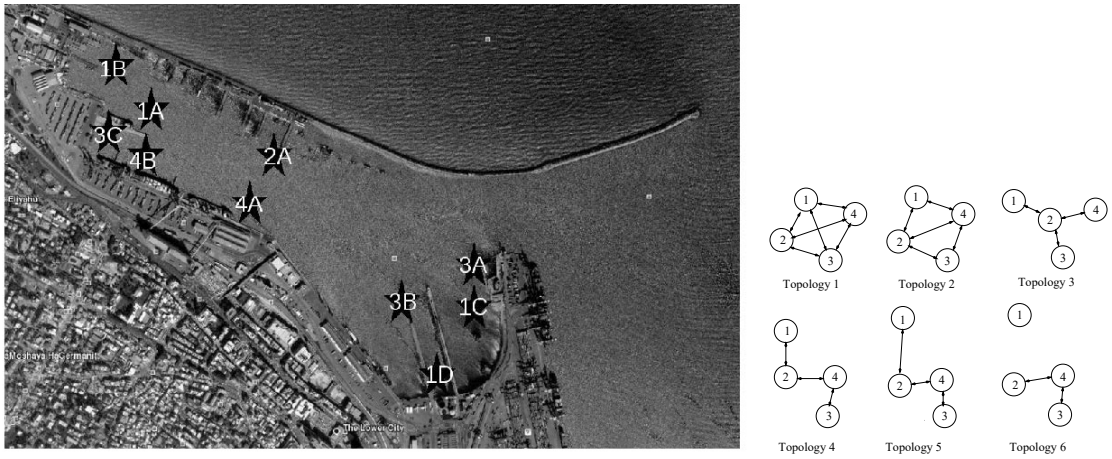
A.3.2. Results

A.3.2.1. Haifa Harbor

We first discuss results obtained for the "Haifa Harbor" dataset. The experiment was carried out in Israel, and included four boats carrying custom modems. The boats moved to different locations in the harbor at designated times. Due to the structure of the harbor, no communication between docks was possible in the absence of line of sight. Hence, the change in the boats' locations created a time-varying network topology. A map of the experiment location is shown in Fig. A.2a, and the formed topologies are illustrated in Fig. A.2b. The recorded dataset includes the per-link time-varying BER measurements arranged in a single TMI and in per-topology TMIs. The experiment included roughly six hours of data collection.

In order to obtain the longest possible emulation, we resort to the `FullTopMat` matrix, which contains the concatenation of the datasets corresponding to each TMI. In our emulation, we consider a successful packet delivery only if the instantaneous BER value is less than 10^{-2}. Considering this threshold,[4] the packet delivery ratio (PDR) and the

[4]This value has been chosen for demonstration purposes. However, we note that this BER regime may be easily related to PER regimes depending on the employed modulation and coding scheme. For example, a BER of 10^{-2} yields a PER of about 0.5 for 64-bit, uncoded packets transmitted using BPSK. In the same conditions, applying a convolutional code of rate 1/3 and soft Viterbi decoding would yield a



(a) Experiment map. (b) Tested network topologies.

Figure A.2: Information about the "Haifa Harbor" experiment. (From [10].) The letters indicate subsequent locations at which the nodes were moved to form the six deployments in Fig. A.2b. For example, node 1 was moved from location 1A to 1B, 1C, and finally 1D.

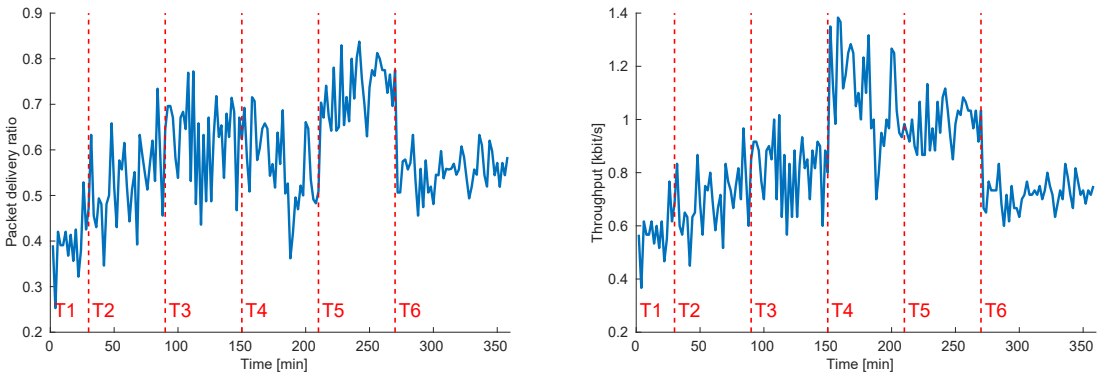
per-link throughput are shown in Fig. A.3. Metrics are collected every 120 seconds and plotted against the collection epoch. Vertical dashed lines mark the instant where the switch between different subsequent topologies occurs, and the TMI enumeration fits the number of topologies in Fig. A.2b.

We observe that the PDR changes over time due to both the topology configuration and the link quality measurements. The former is mostly observed when there is a transition between TMIs, while the consequences of the latter are observed when the TMI remains the same. We also remark that the node deployment affects the throughput, as the maximum propagation delay in the network determines the TDMA slot length, and therefore the packet transmission rate. In all topologies, the maximum propagation delay is about 1 s (corresponding to a maximum distance of about 1500 m), except in topologies 1 and 4, where the maximum propagation delay is 0.88 s and 0.55 s, respectively. For example, in topology T4, this means that the TDMA frame has a significantly shorter duration, which accommodates about 45% more transmissions than in topologies 2, 3, 5, and 6. For this reason, the throughput is larger for topology T4, despite a similar or lower PDR than in topology T3.

A.3.2.2. Berlin Multimodal

We now discuss network emulation results based on the "Berlin Multimodal" dataset, which provides a set of simultaneous measurements from three different acoustic communication technologies. The communication technologies used in the experiment are the EvoLogics SC2R 18-34 kHz (5×), 48-78 kHz (3×) and the 80-120 kHz (2×)

BER of 10^{-5}, which enables the transmission of 1024-bit packets with a PER of 0.01 [193, Section 8.2.8].



(a) Packet delivery ratio.

(b) Throughput.

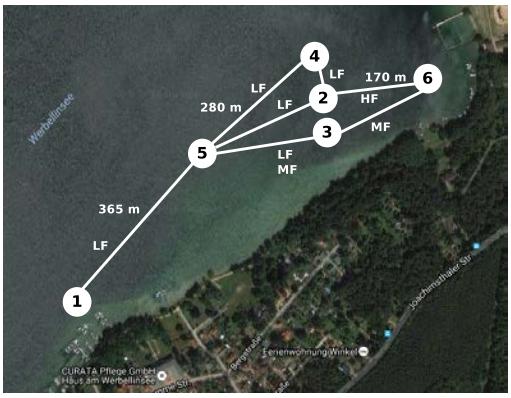
Figure A.3: Results for the full "Haifa Harbor" dataset. All topology data has been concatenated: dashed red lines indicate the transition between subsequent topologies.

modems, respectively named LF, MF, and HF in the following, as a shorthand for low-frequency, medium-frequency, and high-frequency. The TDMA emulator assumes that the transmission rates of each modem are 4 kbit/s, 16 kbit/s, and 32 kbit/s, respectively. The setup of the experiment and the tested topologies are shown in Fig. A.4.

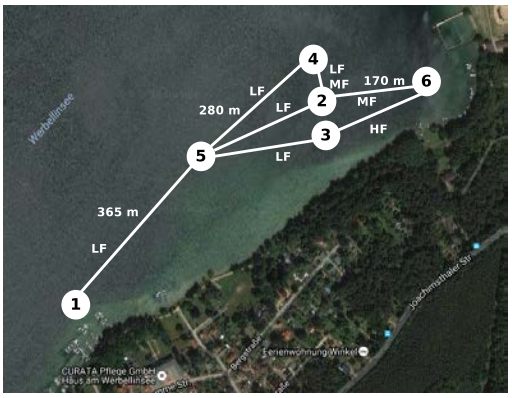
The results are given in Fig. A.5. Each point along the curves corresponds to average values taken over windows of 30 s. The most significant difference between the TMIs is the performance of the HF modem, which requires a low-noise, short-distance link, in order to operate at its maximum efficiency. Since the distance between the only two nodes with an HF modem was smaller in topologies T3, T4, and T5 than in topologies T1 and T2, the HF throughput is much higher and stable for T3, T4, and T5. We also observe that the success ratio for the LF and MF links is similar, and slightly lower for MF in topologies T4 and T5. Since the deployment includes a total of 3 MF and 5 LF modems, this explains the similar throughput achieved by LF and MF in Fig. A.5b.

The area plot in Fig. A.5c shows that the number of packets sent is about the same in each measurement window. The absolute values tend to remain stable over each window and depend on the connectivity of the sub-networks formed by each technology. For example, in topology T3, the nodes transmit fewer MF packets than in all other topologies. The reason is that, in topology T3, all nodes with MF also have LF. More specifically, we recall that in our tested TDMA scheduling protocol, a neighbor is chosen at random, and only then the transmission technology is determined. Since there are more LF modems, a node with both LF and MF is likely to have additional neighbors, and thus it is less likely to transmit using MF in T3 than in any other topology.

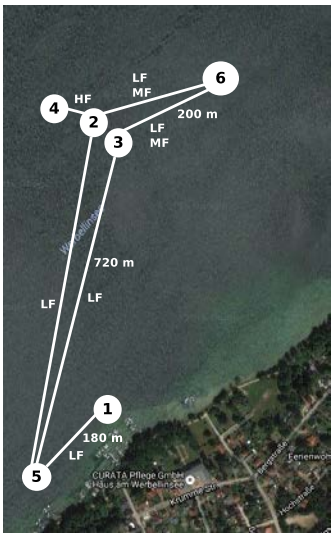
Finally, we demonstrate the flexibility of ASUNA by testing the optimal multimodal scheduling (OMS) scheme in [194]. OMS is an adaptive TDMA-based algorithm that exploits multimodal links in order to schedule transmissions that obey a number of



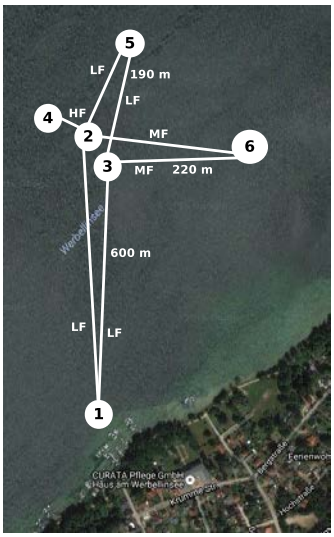
(a) Topology 1



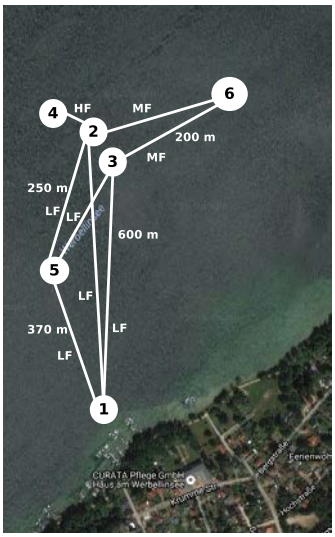
(b) Topology 2



(c) Topology 3



(d) Topology 4

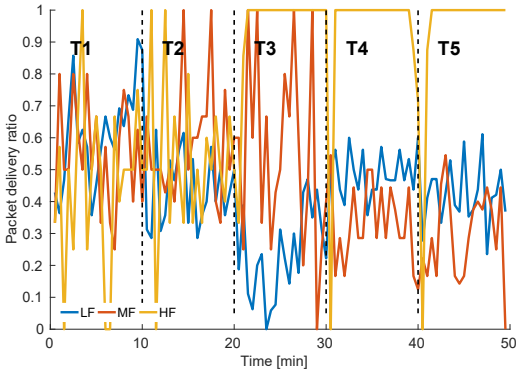


(e) Topology 5

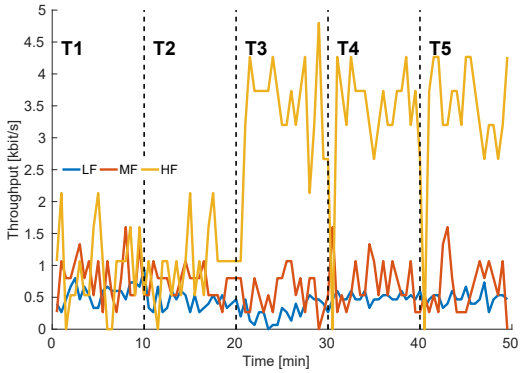
Figure A.4: Setup and tested topologies for the "Berlin Multimodal" dataset. (From [11].)

constraints. These include network topology structure, bounds to interference, and measures to favor multihop routing. OMS was already tested at sea via a dedicated experiment [194] (also part of ASUNA), hence here we rather test OMS using the "Berlin Multimodal" dataset. This also enables a direct comparison against the baseline TDMA protocol considered above.

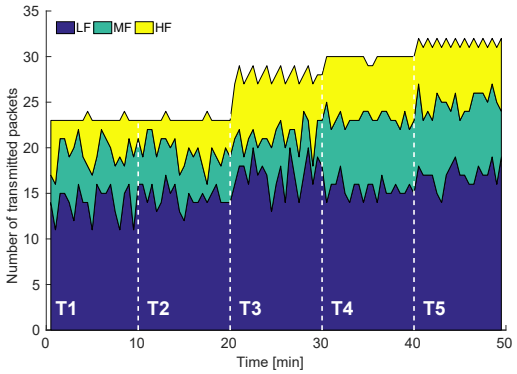
As before, we concatenate all topologies of the dataset, in order to obtain longer link quality time series, exhibiting significant connectivity changes across subsequent topologies. Figure A.6 shows the average throughput per technology. We observe that the OMS protocol adapts well to the characteristics of the topology by allowing simultaneous transmissions over different technologies and by balancing channel access throughout the network. By setting a slot length of 2.5 s, it adapts the packet length to fill this slot length minus the maximum propagation delay. This results in a slightly smaller number



(a) Packet delivery ratio.



(b) Throughput.



(c) Number of transmitted packets.

Figure A.5: Results for the full "Berlin Multimodal" dataset. All topology data has been concatenated: dashed lines indicate the transition between subsequent topologiess.

of transmissions being made, constantly equal to 12 packets per measurement interval of 30 seconds. However, OMS enables transmissions through multiple technologies at the same time, and additionally the above settings yield longer packets than for the baseline TDMA case of Fig. A.5. As a result, the throughput achieved by all technologies is higher (see also Fig. A.5c).

A.4. Conclusions

We presented ASUNA, a shared database containing recorded time-varying link quality measurements from various sea experiments. ASUNA serves as a tool to test underwater acoustic communication network algorithms through emulations or experiment replay. The ASUNA database includes an ensemble of time-varying link quality measures arranged as topology information matrices. The datasets cover different network configurations measured through a variety of acoustic communication devices, and using different network protocols. To demonstrate the use of ASUNA, we described

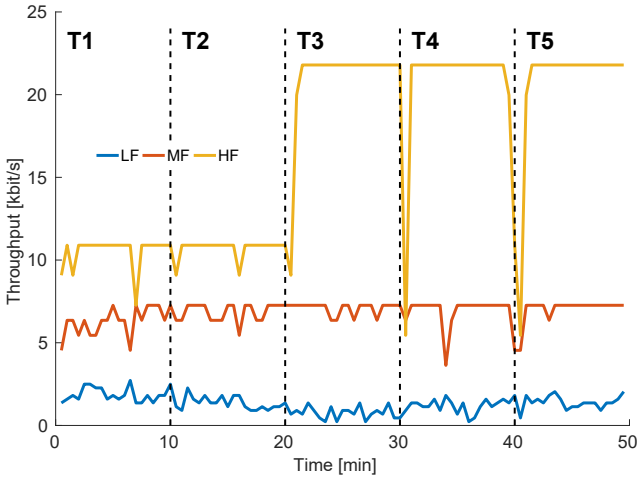


Figure A.6: Throughput for the OMS protocol tested over the "Berlin Multimodal" dataset. All topology data has been concatenated: dashed lines indicate the transition between subsequent topologies.

the details and results of an emulation system built to test a time-division multiple-access scheduling protocol over all collected topology matrices. For a multimodal communications dataset, we also test the optimal multimodal scheduling approach in [194]. We freely share ASUNA as well as the emulation code with the underwater communications community, with the hope that ASUNA will constitute a benchmark to test underwater acoustic networking solutions including, but not limited to, scheduling, routing, and automatic repeat query schemes. ASUNA is open to future contributions. With the expansion of the database that would result, we believe that this benchmark has the potential to greatly contribute to establishing and standardizing UWAN research.

B Underwater LiDAR signal processing for enhanced detection and localization of marine life

Detecting and tracking marine wildlife in its natural habitat is of interest for many fields of science, including behavioral ecology and sociobiology, and can help explain the social dynamics of different marine fauna species, as well as measure the impact of human interference [195]. A specific, important application is the prevention of harmful interactions with various types of human-made underwater equipment, such as marine hydro-kinetic devices [196]. However due to the challenging characteristics of the underwater environment, such as insufficient illumination, attenuation and scattering, it may be arduous to observe marine wildlife using conventional surface approaches, such as optical surveillance in the visible light spectrum. Extra lighting does not solve the issue, as it creates flares, and may also induce behavioral changes in the animals (e.g., many species are attracted by visible light sources [197]).

Available active acoustics technologies for fisheries monitoring can be categorized as either acoustic cameras or sounders. Active acoustic cameras, such as Dual Frequency Identification Sonar (DIDSON) use an acoustic lens technology which forms images with greater detail than found in conventional sonars. While reasonably good estimates of feature dimensions can be extracted from DIDSON images, that level of detail is only available at relatively short ranges and with an angular field of view where image resolution is highest. Consequently, data bandwidth and storage become problematic. 2D imaging sonars (e.g., Teledyne BlueView products), likewise provide a video-rate output, but with a limited angular field of view and range. Acoustic profilers such as echosounders, on the other hand, are a standard instrument of fisheries hydro-acoustics. These instruments can reach a long distance (hundreds of meters) but lack the resolution and evaluation intuitiveness of traditional optical imagery [198].

Therefore, the design of non-invasive systems for marine life observation remains an interesting research problem. A possible solution is provided by marine Light Detection And Ranging (LiDAR): using a portion of the light spectrum which is invisible to most

underwater species, LiDARs can image via serial laser scanning based on the received intensity of reflected light, with an additional layer of time of flight (ToF) data, which reflects the distance to the object (depth map). Therefore such systems may work without additional light sources, and due to the low average power of the emitted light, they do not affect underwater fauna [199]. In contrast to conventional optical systems, LiDARs can count on the depth map to improve the image quality, to perform accurate ranging and scaling, as well as to separate objects from one another and from the background. Beyond the single pulse per pixel type of LiDAR system that is described in this appendix, modulated-pulse versions of the serial-scanning underwater LiDAR imager have been tank-tested to study the improved ranging and intensity imaging performance in turbid environments by using radar waveforms and decoding techniques [200, 201].

The low average power system that is being developed at the Florida Atlantic University (FAU) Harbour Branch Oceanographic Institute (HBOI) is called Unobtrusive Multistatic Serial LiDAR Imager (UMSLI). The configuration of UMSLI described in this appendix consists of one laser transmitter, and one to three receivers. It has two working modes: the detection mode, when the laser does a sparse spatial scan, and the dense mode, when the scan area is reduced to the point of detection, so that the LiDAR can produce a detailed image of the scanned underwater object. The data collected by this system is further used for fish detection, localization and identification that has to be performed on the go. For the complete description of this project, we refer the interested reader to [202]. This particular LiDAR system uses relatively inexpensive components, and its overall price is foreseen to be about one order of magnitude lower than other systems available on the market. It uses low average power red laser diodes with a wavelength of 638 nm, which suffer from higher attenuation in water, but are unobtrusive for fish vision. Together with a relatively low sampling frequency, these and other aspects of UMSLI require a targeted approach to the processing of the obtained data.

The UMSLI prototype was developed and validated at the optical test facility at HBOI. The serial scanning technique on which the LiDAR design hinges has been employed in near-monostatic [203, 204] and bi-static [205] configurations. Serial scanning can be instrumental to help detect individuals within a cluster of animals better than by using a regular camera system. Furthermore, it can enable the formation of images even when the line-of sight between the target and the detector is occluded [206].

In the following sections we provide further details about the UMSLI design and experimental setup (Section B.1), we describe our data processing pipeline (Section B.2), and provide concluding remarks in Section B.3.

B.1. System Description and Experiment Setup

B.1.1. Description of the LiDAR system

The integrated system adopts a multi-tiered design that consists of the sensing hardware, of image enhancement, detection, and classification, as well as of data archiving.

The UMSLI sensing front end consists of six receivers (Rx), six transmitters (Tx), and a digital signal processor [202]. The transmitters scan a water volume in a bi-directional raster pattern. The scan field can be configured to be either sparse (with fewer pulses over a wider angle, used mainly for detection purposes), or dense (concentrating a higher pulse density through a narrower range of angles, typically after the detection of an object has been confirmed). On the one hand, the depth of field for each channel depends on the depth of the overlapping region between the laser beam and the receiver's field of view. On the other hand, the image resolution depends on the diameter of the pulsed laser beam as it intersects a hard target. The receivers consist of a high-speed photomultiplier relying on focusing optics and on a spectral bandpass filter, designed to efficiently collect time-resolved back-irradiated light from the emitted laser pulses. Depending on the conditions of the water, such light may consist of either reflections from objects or from backscattering. The signals received by the photomultipliers are digitized at a rate of one sample every 0.5 ns, before they are conveyed to digital processing elements that reconstruct the areas of interest in the water volume.

Given its unobtrusiveness with respect to marine fauna vision and behavior, the UMSLI system can be used in the proximity of marine installations to ensure the monitoring of endangered/threatened species, and to assess the impact of the installation on marine wildlife. The image formation approach has been demonstrated to be especially suitable for turbid or low-light applications. This makes the UMSLI more convenient than traditional camera solutions, potentially achieving persistent surveillance through day and night, while still providing the operational advantages of optical solutions.

B.1.2. Test scenario and data sets

In this appendix, we consider two data sets from previous experiments taken in a controlled test tank environment at HBOI. One data set contains experimental results from a system with a single light emitter and a single receiver, which captured images of a static fish model in different poses, with 200×200 pulse resolution, at 128 samples per 128-ns pulse record. The frames are taken in clear water, as well as by controlling the turbidity such that the total attenuation coefficient at 638nm ranges up to $c = 0.75$ m^{-1} (turbid water) using sparse and dense mode scans. Altogether, the data set contains 79 frames.

The second data set has a similar structure, but is obtained from another version

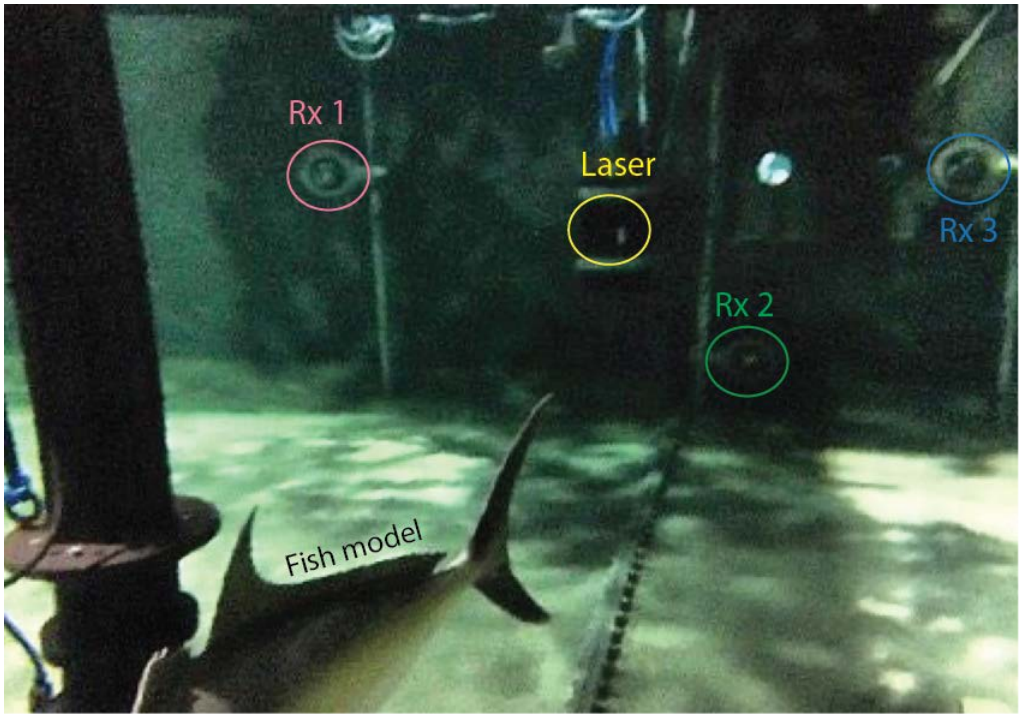


Figure B.1: Photo of the test tank environment, showing the location of the laser transmitter and of the three receivers, as well as the model fish target (in this case, an amberjack model).

of UMSLI with three receivers having a higher sampling resolution of 256 samples per 128-ns pulse record. A picture of the test tank with this configuration of the LiDAR is shown in Fig. B.1. The laser occupies a central position, whereas the three receivers are deployed to the side of the laser and below it. Fig. B.1 also shows part of the model fish target (an amberjack in this case). Seven different scenes were taken for each turbidity setting, with moving and static fish models, as well as with a calibration board. Each scene was taken 15 times, hence the data set contains altogether over 500 frames per receiver. These challenging data sets provide an interesting test platform to seek an optimal solution for the enhancement of LiDAR data under power and computational complexity constraints. The following section explains our proposed processing pipeline to produce classifier-friendly images and depth maps for further post-processing stages.

B.2. LiDAR Data Processing Pipeline and Results

In this section, we present the steps of the LiDAR data processing pipeline. We start with water turbidity estimation in Section B.2.1. The following steps include image enhancement via backscattering removal (Section B.2.2), and contrast enhancement for 2D visualization (Section B.2.3). Finally, the LiDAR scan is transformed into a 3D cloud of points. Those that are not part of the target are identified and removed via a clustering

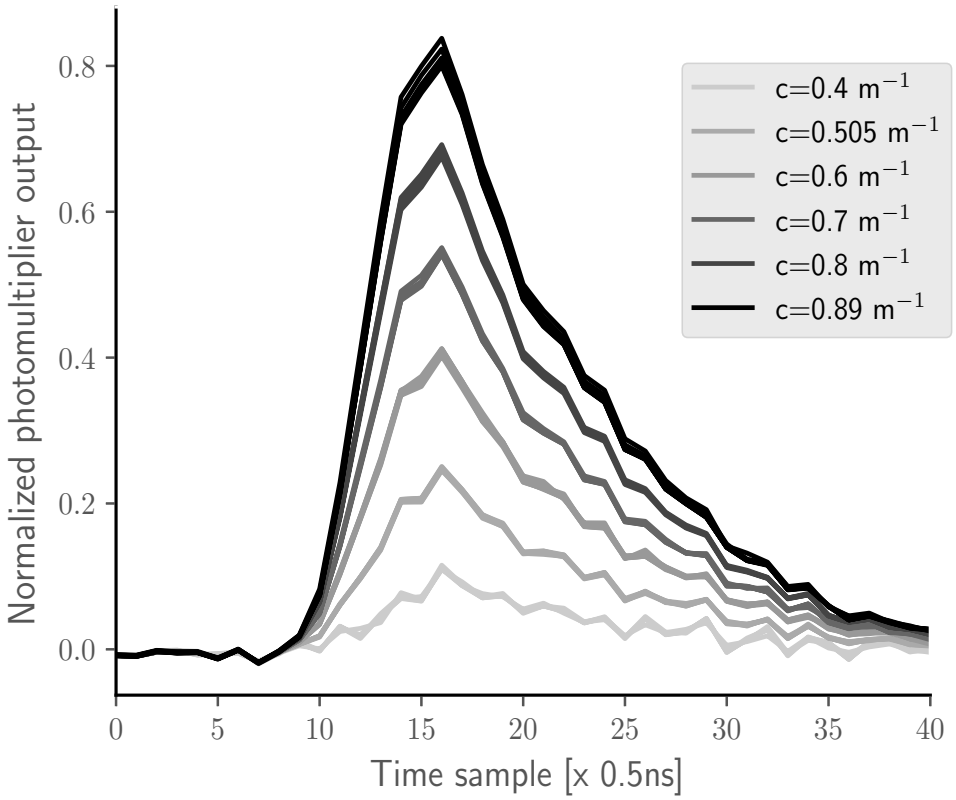


Figure B.2: Backscattered light intensity over time for different turbidity values.

algorithm; the remaining points are then employed to construct a 3D model of the target, which can be employed for biomass evaluation (Section B.2.4).

B.2.1. Turbidity estimation

An important preliminary step to achieve good-quality LiDAR imaging is to estimate environmental parameters, and specifically turbidity, which is expressed in terms of the total light intensity attenuation coefficient of the water c [m^{-1}]. In the current setup for test tank experiments, additional equipment is used for precise turbidity measurements, which is advantageous for further image processing. However, once the LiDAR is deployed at sea, it might be challenging to measure the ocean water's turbidity without such additional equipment, especially in the presence of mixing phenomena that may induce turbidity variations over time. In these cases, the backscattering intensity can be a good measure for turbidity [207]. Several parameters of the LiDAR affect backscattering. These include the wavelength, power and gain of the incident laser beam, the scan area, and the relative position of the transmitter and receiver [208]. By including these parameters in a backscattering model, the LiDAR can be calibrated to infer turbidity from backscattering measurements.

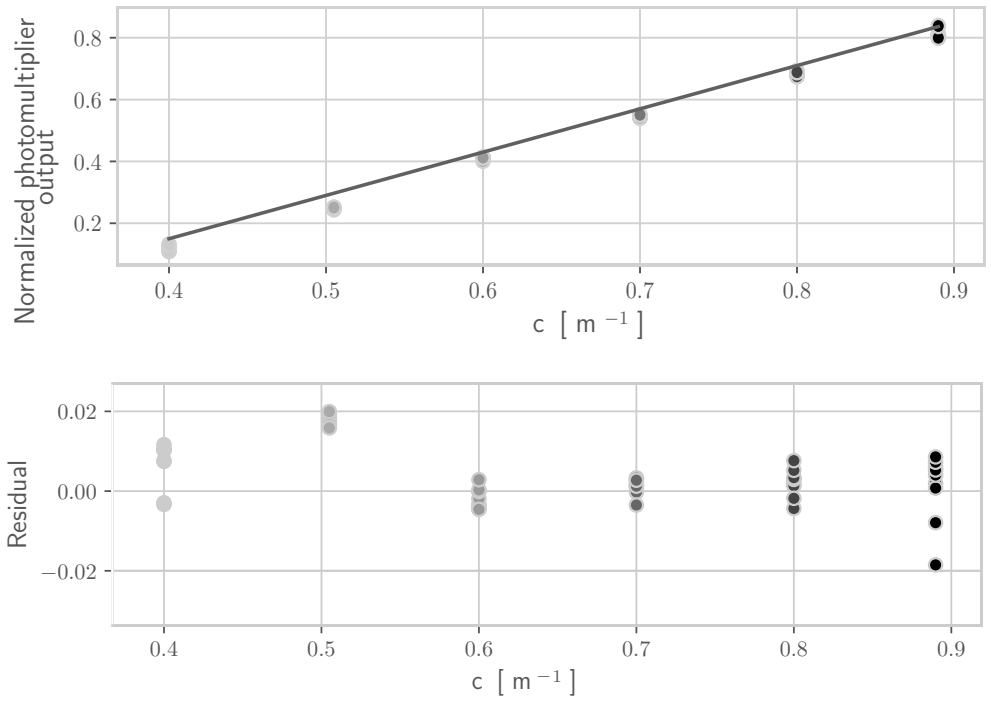


Figure B.3: (Top) Estimated linear fit of the average backscattering peak value against the total attenuation coefficient of the tank water. (Bottom) Residual errors in the estimation of the turbidity value.

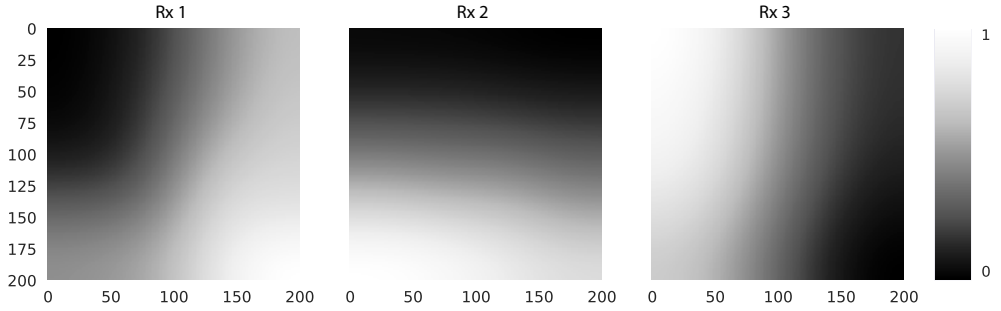


Figure B.4: Backscattering peak amplitude heatmaps for the three receivers.

We employed the FAU test tank to carry out several backscattering calibration measurements in different water turbidity conditions. The results can be observed in Fig. B.2, which presents a superposition of several backscattered light intensity signals sampled at intervals of 0.5 ns. For the same value of c, the signals show approximately the same backscattering peak value. However, this peak varies significantly by varying c. As observed in Fig. B.3, the relationship between the backscattering intensity peak value and the turbidity level expressed through c is roughly linear. A least-squares fit (top panel in Fig. B.3) is in fact sufficient to achieve very low residual error (bottom panel), on the order of ± 0.02 m^{-1} or less, which is sufficiently accurate to inform the subsequent

processing steps.

B.2.2. Backscattering removal

As backscattered light from particles suspended in water can strongly corrupt the quality of the LiDAR imaging process, it becomes important to compensate for, at least, the backscattering intensity peak. To achieve this, we consider the same calibration data set employed to estimate the relationship between the total attenuation coefficient c and the backscattering peak intensity. We recall that a LiDAR scan comprises a matrix of 200×200 received pulses, and that 256 samples are extracted for every channel of each pulse at a sampling interval of 0.5 ns. We employ this data to generate an average backscattering pulse to be employed for calibration. This pulse is obtained by summing the detected light intensity samples both over all 200×200 pulses in a frame and over several frames, and by finally normalizing the pulse so that it has a maximum value of 1. We then consider a Gamma probability density function (PDF) of the form

$$f(x;k,\theta) = \frac{x^{k-1}\,\mathrm{e}^{-x/\theta}}{\theta^k\,\Gamma(k)}\,, \quad x>0,\ k>0,\ \theta>0\,, \tag{B.1}$$

where $\Gamma(k)$ is Euler's Gamma function, and we employ least-squares fitting to estimate the parameters of the function that best approximate the shape of the pulse.

Before we can actually compensate backscattering, we need to create another average map that conveys the peak backscattering intensity, and that will be used to scale the Gamma function approximation. We obtain these maps from the calibration frames, by recording the average backscattering peak intensity for each of the 200×200 signals of a frame. In Fig. B.4, we present three heatmaps that show, for each signal, the amplitude of the backscattering peak. The values are normalized with respect to the digital counts corresponding to the maximum expected backscattering value. Such digital counts are the output of the digitizer of the photomultiplier output, and are proportional to the instantaneous detector current, which in turn is proportional to the time-resolved scene irradiance (in $\mathrm{W/m^2}$) at 638 nm. After scaling each signal in accordance with the heatmap, the backscattering component of the signal can be subtracted in order to enhance the contrast of any other peaks reflected by a target, if present.

In Fig. B.5, we provide an example of backscattering removal from light intensity signals. The blue line is the detected signal before backscattering removal, whereas the grey line is its approximation through a Gamma pdf. The red curve is obtained after the subtraction of modeled backscattering. We observe that the red curve already contains a distinguishable component related to the target, located between samples 145 and 180. In order to remove residual noise, as a last step we apply a Savitzky-Golay (SG) filter. The resulting smoothed signal is shown using a black line.

The main advantage of the above processing method is that the residual backscattering

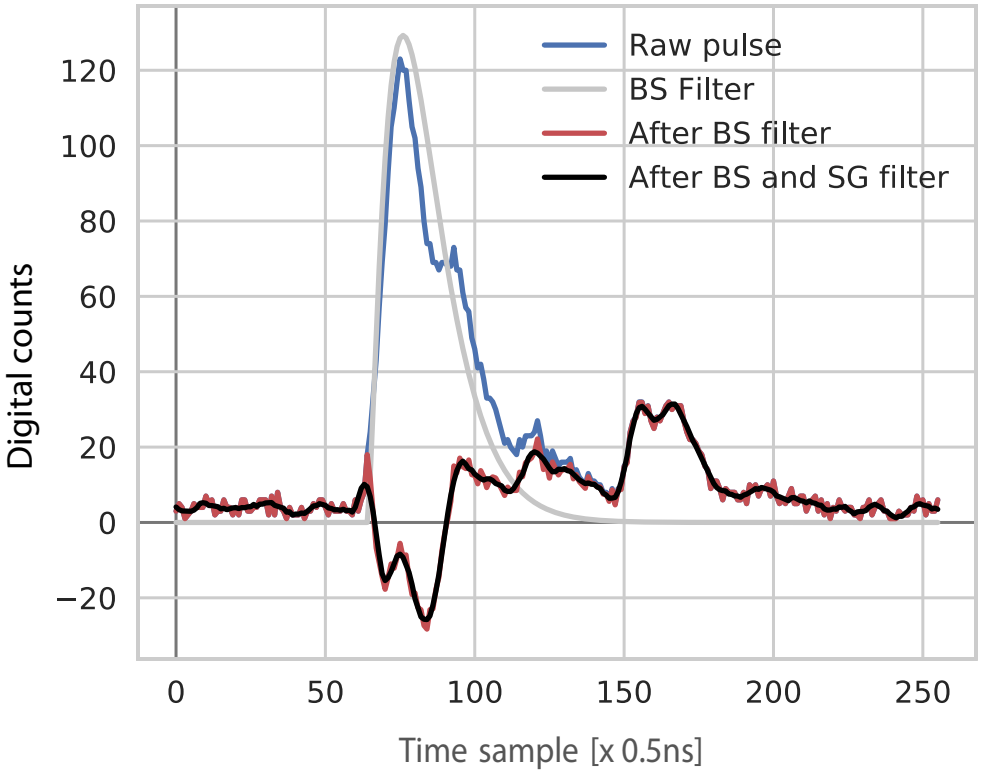


Figure B.5: Example of pulse backscattering removal based on the estimated turbidity and the backscattering model of Figs. B.2 and B.4.

intensity is now significantly smaller that the target reflection peak, so that a simple maximum search operation suffices to identify the signal coming from the target. At the cost of a simple preliminary calibration step, this method does not require to average multiple LiDAR frames, and requires fewer LiDAR scans, besides having very low complexity.

B.2.3. 2D image enhancement and visualization

In order to visualize the effect of the backscattering removal, the intensity of the target peak for each receiver is coded using a red/green/blue (RGB) color vector. Each color channel conveys the intensity of one receiver: with reference to the circles in Fig. B.1, we use red for receiver 1 (to the top right of the laser when facing towards the water in the tank setup), green for receiver 2 (bottom-central position in the tank), and blue for receiver 3 (top left position). The three panels on the top of the figure refer to low turbidity conditions ($c = 0.36$ m^{-1}), whereas the three bottom panels refer to relatively high turbidity ($c = 0.75$ m^{-1}). The left panel in Fig. B.6 shows the RGB coding for unprocessed signals which have not undergone any backscattering compensation. Due to the location of the receivers, the top-left corner is dominated by blue hues, the top-right corner by red hues, and the bottom side by greener hues. A pixel color turning towards

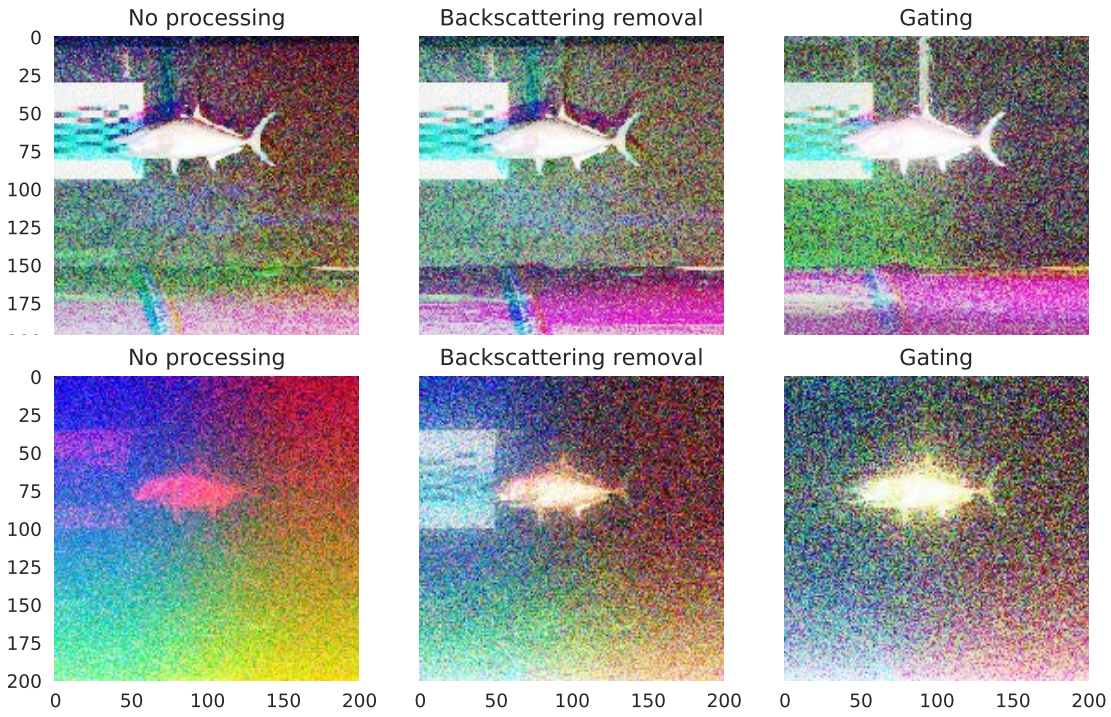


Figure B.6: Example of the RGB representation of a LiDAR image after pulse backscattering removal, compared to a simpler gating solution. (Top) $c = 0.36$ m^{-1} (Bottom) $c = 0.75$ m^{-1}. The distance between the transmitter and the fish is 5.7 m (4.272 beam attenuation lengths at $c = 0.75$ m^{-1})

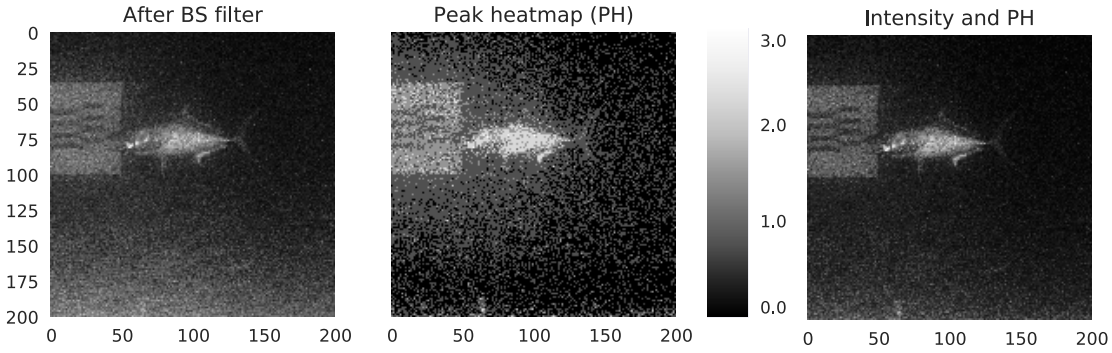


Figure B.7: Contrast enhancement for 2D image representation using peak heatmaps. $c = 0.75$ m^{-1}.

white represents a high intensity detected by all receivers.

For the clear water case, the image produced by the system without any processing (Fig. B.6 top-left) is clear enough for further processing. However even here backscattering removal (top-middle) improves the visibility of features towards the far side of the tank. The results are better than a manual gating procedure focusing on the target-reflected light portion (top-right). In the bottom-left panel of Fig. B.6, in relatively high turbidity

conditions, the shape of the fish target can be observed from the purple hues in the center of the image, whereas the near-field calibration target is barely seen along the left side of the image. After backscattering removal, we obtain the image in the bottom-middle panel. Here we observe that the fish target is the brightest element, and its characteristic biological features (e.g., the shape and the fins) are much easier to distinguish. For comparison, zeroing out the backscattering component of the received signal through manual gating would not be as effective in enhancing the image. An example of a gated image is shown in the bottom-right panel: here, the overall image is more noisy, and the sharpness of the fish contours and biological features is greatly reduced.

We finally derive a black and white version of the target image by fusing the data from the three receivers of the LiDAR as follows. Starting from the image with backscattering removed (bottom middle panel of Fig. B.6), we sum the output intensities over all channels, and normalize the intensity values to occupy the whole interval $[0, 1]$. The resulting image for the $c = 0.75$ m^{-1} case is plotted in gray scale in the left panel of Fig. B.7. While both the fish target and the calibration target on the left already stem clearly from the darker bottom, it is often useful to generate an image with better contrast. We do so by producing a peak heatmap where, for each pixel in the image, we count the number of peaks in the corresponding LiDAR signal. The result is shown in the middle panel, where we can observe that the target and its surroundings indeed present a higher number of intensity peaks per pixel. By performing a pixel-wise multiplication of each pixel's intensity value by the corresponding number of peaks from the heatmap, and by rescaling the resulting grey intensity values back to the interval $[0, 1]$, we obtain the image in the right panel of Fig. B.7. Here light intensity differences due to the shape of the fish model are amplified, and the fish contours stem more clearly out of a darker background. The same conclusion applies to the calibration panel on the left.

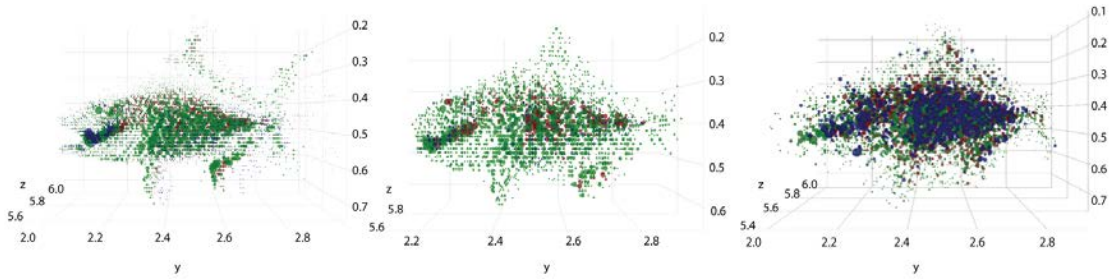


Figure B.9: Sets of points corresponding to the fish target after clustering. The color of each point corresponds to the receiver (see also Fig. B.1 for reference), whereas the size of marker is proportional to intensity. The coordinates are expressed in meters. (Left) low turbidity, $c = 0.36$ m^{-1}); (Center) medium turbidity, $c = 0.57$ m^{-1}); (Right) high turbidity, $c = 0.75$ m^{-1}).

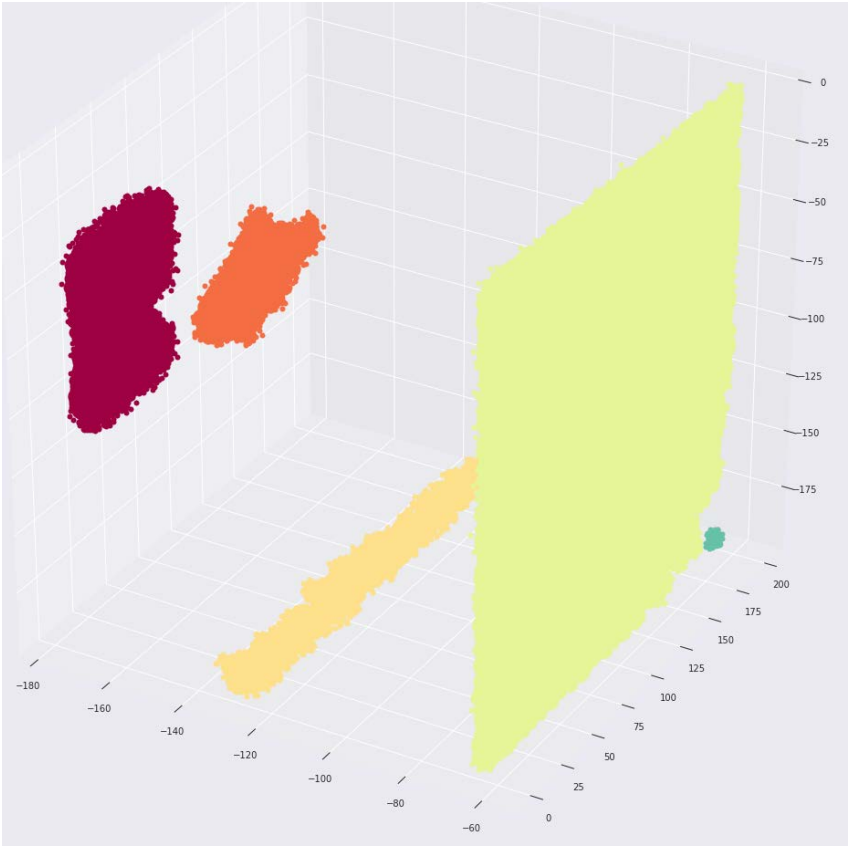


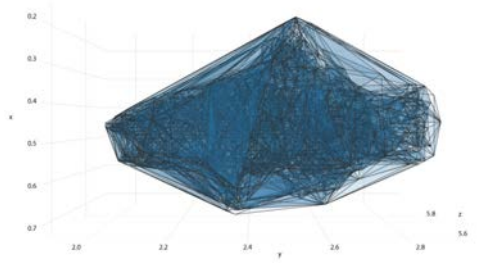
Figure B.8: Example of result obtained by applying the DBSCAN clustering algorithm to a LiDAR frame. The shape of the fish target (orange) can be clearly distinguished in the image, along with the calibration target (dark red) and with the backscattering residual (light yellow on the right).

B.2.4. Derivation of a 3D target model

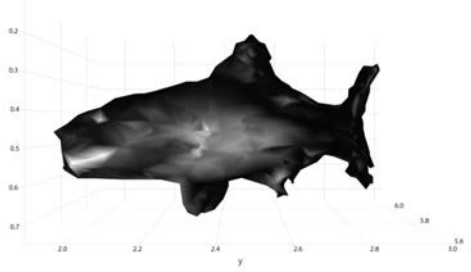
Besides image processing for the enhancement of 2D images, it is also of interest to create a 3D model of the target. To do so, further processing is required to separate the light reflected by the target from backscattered or background light. To achieve this separation, we employ the Density-Based Spatial Clustering of Applications with Noise (DBSCAN) algorithm [131].

The DBSCAN algorithm is commonly used for clustering spatial data with density constraints: its run time is almost linear with the number of points to be clustered, which makes it amenable to real-time applications. Moreover, DBSCAN enables clustering without any prior knowledge about the number of clusters to be formed. As an input, the algorithm requires a definition of the meaning of "dense" clusters, conveyed by the parameters ε and $N_{\min}$. Specifically, ε is the maximum distance between two samples that still allows both of them to belong in the same neighborhood, and $N_{\min}$ is the minimum number of samples in any cluster.

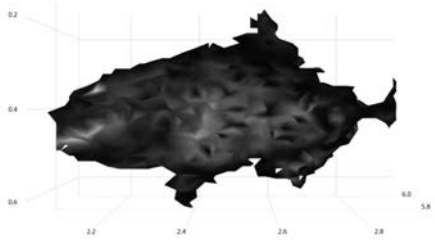
The data passed to DBSCAN is formatted as a set of samples with four features: three



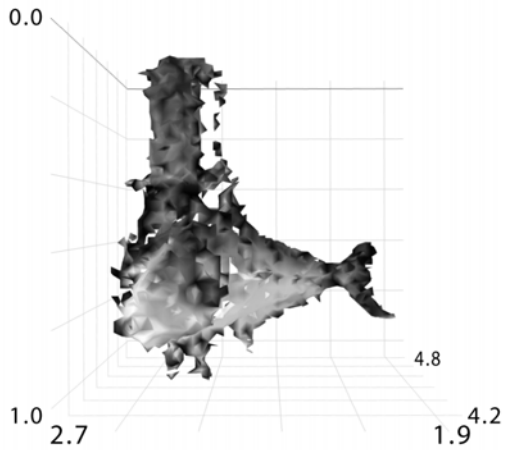
(a) Delaunay triangulation applied to the points of the identified fish cluster



(b) α-shape of the identified fish cluster, $c = 0.36$ m^{-1}.



(c) α-shape of the identified fish cluster, $c = 0.57$ m^{-1}.



(d) α-shape of the identified fish cluster in a more complex pose facing towards the LiDAR's receivers. In this case, the pole mount of the fish model becomes visible. $c = 0.36$ m^{-1}.

Figure B.9

coordinates (to identify the location of the point in space) and an intensity value. The coordinates are expressed as sample indices, or "units:" the x- and y-coordinates can take values in $0, 1, \ldots, 199$ (since the scanning grid contains is 200×200 points), whereas the z-coordinate can take values in $0, 1, \ldots, 255$ (as there are 256 samples in each received signal). Since the LiDAR transmitter and receivers are synchronized, the depth of each point can be derived from the coordinates towards which the transmitter is pointed, and the round-trip time required for the light to irradiate back to each receiver. For our results, we set the parameters of DBSCAN to be $\varepsilon = 4$ units and $N_{\min} = 30$.

An example of the output of the clustering algorithm applied to a LiDAR frame is shown in Fig. B.8. We observe that the cluster of points identifying the fish shape clearly stands out (orange cluster), along with the near-field calibration target (dark red cluster), whose border on the right side is interrupted due to the shadowing of the fish target. Spurious clusters are also present (e.g., a reflective element on the bottom of the tank (dark yellow) and a large planar cluster near the transmitter (light yellow, on the

right of the picture) which corresponds to residual backscattering intensity that could not be completely canceled via the method described in Section B.2.2. Still such spurious clusters are typically well separated, so that it is possible to at least distinguish the shape and size of the fish.

Fig. B.9 shows three different point clusters pertaining to the fish target, in three different turbidity conditions: low turbidity ($c = 0.36$ m^{-1}, left panel), intermediate turbidity ($c = 0.57$ m^{-1}, middle panel), and high turbidity ($c = 0.75$ m^{-1}, right panel). In each figure, the size of each point is proportional to the intensity of the signal received from that location, whereas the red, green and blue colors encode the receiver of the signal from which the point was estimated. As expected, in the presence of low water turbidity, several fish features can be already distinguished directly from the cloud of points. In the left panel of Fig. B.9, this includes all fins, and the shape of the open mouth. As turbidity increases, the weakest reflectors of the fish (e.g., the fins) mix up with backscattered light, resulting in an overall increase of the fish size, with less sharp transitions along the contour of the fish's body. In the highest turbidity conditions, the blurring of the borders becomes even worse, although it is still possible to distinguish the shape of the fish from the most reflective components of its body (in this case, the side facing the LiDAR's receivers and the caudal fin). Still, we remark that even in high turbidity conditions the backscattering compensation and the subsequent processing of the LiDAR signals make it possible to create sufficiently separated sets of points, from which the target-related points can be successfully singled out by DBSCAN.

As a last processing step, we are interested in forming a 3D model of the imaged target, which is typically useful for classification and biomass evaluation purposes. In order to achieve this, we start from the clustered samples derived from the previous step and obtain a first 3D model via Delaunay triangulation. For $c = 0.36$ m^{-1}, the result is shown in Fig. B.9a. The plain application of the triangulation algorithm creates triangular simplices that altogether enclose the fish figure within a convex hull. However the outer simplices tend to connect much farther points, and thus have larger sides, compared to the simplices that connect the denser points of the cloud corresponding to the body and fins of the fish. This makes it possible to create an alpha-shape [209] by ruling out the simplices of the Delaunay triangulation that have a circumscribing sphere of radius greater than α. The value of α can be chosen to strike a balance between the accuracy of the 3D fish model and the need to achieve a connected complex of simplices. In this work, we set $\alpha = 17$, which corresponds to eliminating simplices which have a circumscribed sphere of radius larger than about 6 cm.

Fig. B.9b shows the alpha-shape computed for the low turbidity point set in the left panel of Fig. B.9. The shade of grey on each face encodes the intensity of the samples at the corner of the respective simplices. We observe that besides having a good model of the features of the fish, its length can be estimated quite reliably from the figure to a total

head-to-tail extension of 1.14 m. While the actual length of the fish model is 1.3 m the difference can be explained by the relatively low sampling frequency, which translates to a coarse depth sensing where a displacement of one sample corresponds to a length of about 11 cm in the tank. Part of the error is also due to the slight smoothing of protruding features, such as the fish nose and tail. In the presence of higher turbidity (e.g., as in the middle panel of Fig. B.9), the alpha-shape is more noisy as can be observed in Fig. B.9c, and some fish details are lost after the execution of the triangulation algorithm. In any event, the shape and size of the fish are still recognizable, and the length can be measured with a sufficient degree of accuracy for biomass assessment. To conclude, we consider the imaging of the fish target rotated counter-clockwise by an angle of about 60° with respect to the horizontal pose of Figs. B.9a and B.9c. The result of the LiDAR scan in low turbidity conditions is shown in Fig. B.9d. Interestingly, in this case the anchoring pole to which the fish target is attached is included in the fish point cluster singled out by DBSCAN, due to its greater proximity to the rest of the fish samples. Even in this case, the fish is rendered correctly, with several visible biological details, and the grey scale intensity coding reveals that the tail of the fish is farther from the LiDAR receiver than the head and pole.

B.3. Conclusions

In this work, we presented the design of the FAU HBOIs Unobtrusive Multistatic Serial LiDAR Imager. This LiDAR system can scan its surrounding using a low average power red laser at a wavelength of 638 nm, which makes it non-invasive for fish vision and behavior. The LiDAR is based on inexpensive components. We presented several in-tank measurements carried out both in the absence and in the presence of a model fish target. Our results show that a simple calibration step is sufficient to estimate near-field water turbidity and thus inform the following image processing steps. These steps include both 2D image contrast enhancement, and further processing to extract the LiDAR samples related to the target from the background as well as from backscattering residuals, and construct a 3D sensed target model. The relatively low complexity of such processing makes the system amenable to be implemented in real time.

C Dissemination and communication activities

During the time I have been working on this thesis, I had several chances to share the progress with other specialists from academy and industry.

- *Annual Harbor Branch Summer Intern Symposium August 5, 2016*

The 43rd Annual Harbor Branch Summer Intern Symposium Johnson Education Center is an annual symposium I attended at the end of my summer internship at Florida Atlantic University. The symposium unites about 20 students and their mentors from various fields related to marine science. I presented the preliminary results of my work on "Underwater Lidar Signal Processing for Enhanced Detection and Localization of Marine Life."

- *Symposium on Hydroacoustics May 23-26, 2017*

XXXIV Symposium on Hydroacoustics May 23-26, 2017 (SHA'17) in Jastarnia, Poland. An annual international event organized in Poland that brings together people from different scientific fields related to hydroacoustics. I gave a talk reporting preliminary results of multipath-aided range and depth estimation and received feedback on possible applications and improvements.

- *Workshop on Positioning, Navigation and Communications 25-26 October 2017*

This workshop was held at Jacobs University in Bremen, Germany, and proved to be a great opportunity to unite industry and academic researchers that work on the very specific field of positioning and navigation. In this workshop, I was chairing a session and presenting our work "Anchorless underwater acoustic localization." The work was recognized with the "Best paper award."

Figure C.1: WPNC'2017 Best paper award ceremony

- *OCEANS conference and exposition*

This bi-annual conference organized by the IEEE Ocean Engineering Society and the Marine Technology Society (MTS) is one of the top known conferences in the field of ocean sciences. It gathers over 750 registrants, engineers and scientists from all over the world, and over 7k visitors. Serving as a session chair and presenting our work "Underwater LiDAR signal processing for enhanced detection and localization of marine life" (Kobe, Japan, 2018) and "Underwater Direction of Arrival Estimation using Wideband Arrays of Opportunity" (Marseille, France, 2019), I was able to reach many contacts working in the same field and discuss further work.

- *Breaking the surface (BTS) October 2019*

This event is an international multidisciplinary workshop and field training program that enables participants to share experience and knowledge in maritime robotics and related fields of research. During this workshop, I had a chance to present our ongoing work on the EU H2020 SYMBIOSIS project during the demo program. I also co-organized and held a tutorial together with the team of Evologics GmbH on the topic "Software Defined USBL-Modem."

Figure C.2: BTS'19. Preparing a tutorial together with Evologics GmbH team.

- *Publications in news media*

Our work on "Anchorless underwater acoustic localization" was covered by multiple publications in several scientific news portals including [210], [211], [212], [213]. We would like to thank IMDEA Networks's operations team, and especially to Rebeca de Miguel, for the help with the dissemination of these scientific results to the press. Numerous articles also covered the overall SYMBIOSIS project, e.g., [214], [215]. The full list of publications about the project can be found on project website [216].

D Experimental activities

Experimental activities are one of the key challenges in research related to underwater localization. Sea trials are complex to organize, and usually expensive to implement. Various issues including harsh weather or failure of key equipment during deployments may prevent researchers from completing the planned activities. Taking part in a sea trial therefore provides the participants with a lot of experience about what can fall outside the researchers' control, how to mitigate risks, and how to foresee and/or compensate for events that slow down or impede experimental activities.

While working on this thesis, I had the chance to take part in the organization and/or execution of several experimental campaigns and field trial, often in collaboration with other universities and research centers. A brief list of these opportunities follows.

- *FAU premises*

  Lab experiments at FAU. During my internship at FAU HBOI, I had the chance to take part in experiments executed in a controlled environment provided by a test tank (changing turbidity conditions), as well as design an experiment for calibration of LiDAR working in different turbidity conditions. This work led to the publication from Append. B

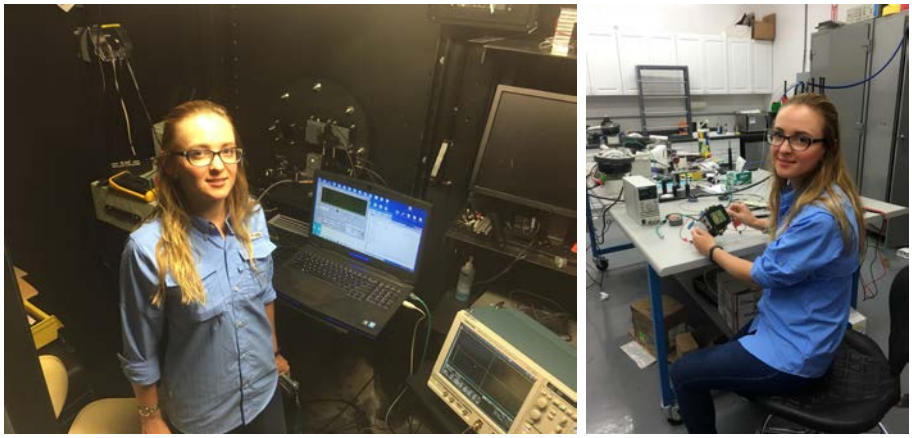


Figure D.1: Lab experiment at FAU Harbor branch.

- *Haifa University*

During the 3 month of my internship at Haifa University and later during several visits to Israel, thanks to Dr. Roee Diamant, I had a chance to take part in multiple sea trials of different scale, as well as some laboratory experiments. Sea trials included 1-2 days boat trips to THEMO observatories [171].

Figure D.2: Experiment near North coast of Israel

- *Sea trial in Eilat, Israel, 1-4 July 2018*

A first experiment related to Symbiosis project. Despite the temperature reaching up to 50°C and the damage to hydrophones due to overheating, we were able to successfully perform a first localization experiment.

Figure D.3: Sea trial in Eilat, Israel

- *Large pool experiment (Berlin, Germany) 25-27 May 2019*

A preliminary work on software control of ultra short baseline (USBL) Software defined modem (SDM) from SYMBIOSIS platform and calibration of acoustic equipment installation.

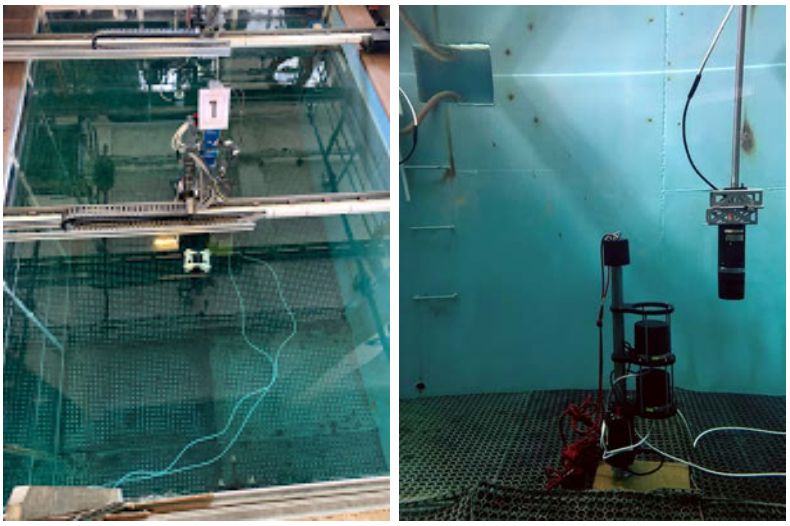


Figure D.4: Lab experiment at Evologics GmbH premises

- *Werbellin lake, Germany May-June 2019*

Several experiments at Werbellin lake near Evologics GmbH's premises to test the SYMBIOSIS acoustic platform. The results of these experiments are presented in Chap. 5

Figure D.5: SYMBIOSIS platform acoustic tests at Werbellin lake in May and June 2019

- *BTS'19, Biograd Na Moru, Croatia, October 2019*

This workshop was a unique opportunity to work on the SYMBIOSIS platform and test it in real conditions. We performed multiple experiments in a test pool, as well as in a sheltered shallow sea environment, accessible from a pier and a boat.

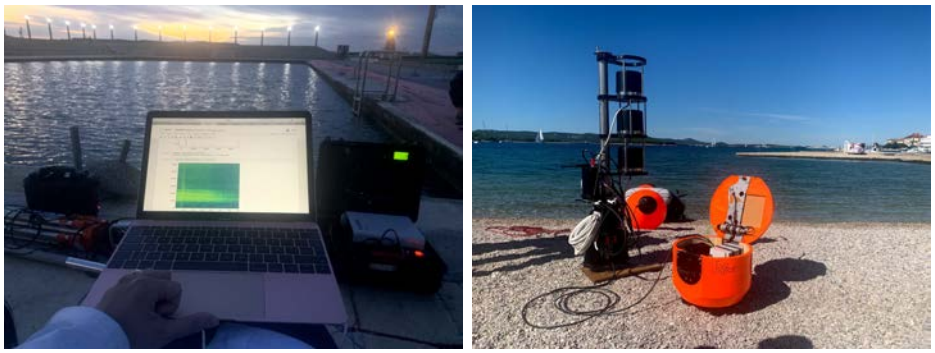


Figure D.6: SYMBIOSIS platform acoustic tests preliminary tests at Biograd na Moru

- *SYMBIOSIS platform test near the Deep THEMO site, Israel, November 2019*

This experiment is one of the final sea trials performed in the context of the SYMBIOSIS project. Besides having the opportunity to test localization algorithms for a full-day cycle of about 24 hours, I could perform further tests on the localization of a rehabilitated turtle while it was set free at sea.

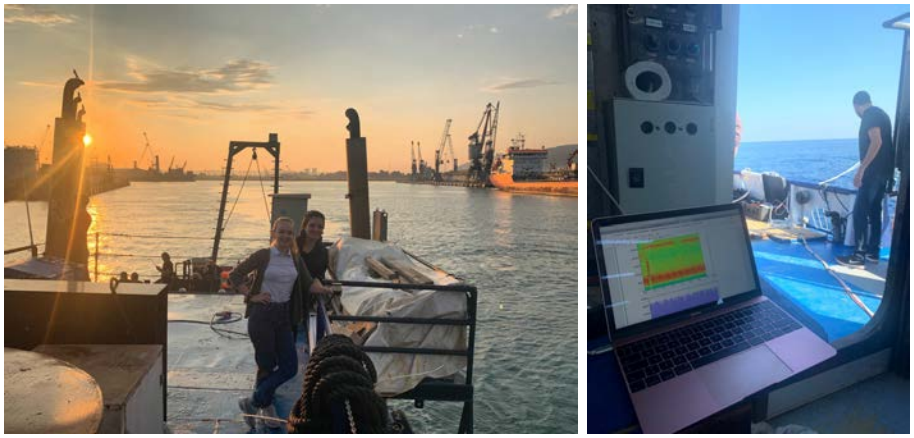


Figure D.7: SYMBIOSIS platform first test near deep THEMO deployment

References

[1] E. Dubrovinskaya, I. Nissen, and P. Casari, "On the accuracy of passive multipath-aided underwater range estimation," in *2016 IEEE Third Underwater Communications and Networking Conference (UComms)*. IEEE, 2016, pp. 1–5.

[2] E. Dubrovinskaya, R. Diamant, and P. Casari, "Anchorless underwater acoustic localization," in *2017 14th Workshop on Positioning, Navigation and Communications (WPNC)*. IEEE, 2017, pp. 1–6.

[3] E. Dubrovinskaya, F. Dalgleish, B. Ouyang, and P. Casari, "Underwater lidar signal processing for enhanced detection and localization of marine life," in *2018 OCEANS-MTS/IEEE Kobe Techno-Oceans (OTO)*. IEEE, 2018, pp. 1–8.

[4] E. Dubrovinskaya and P. Casari, "Underwater direction of arrival estimation using wideband arrays of opportunity," in *OCEANS 2019-Marseille*. IEEE, 2019, pp. 1–7.

[5] E. Dubrovinskaya, P. Casari, and R. Diamant, "Bathymetry-aided underwater acoustic localization using a single passive receiver," *The Journal of the Acoustical Society of America, S.I. on Acoustic Localization*, vol. 146, no. 6, pp. 4774–4789, 2019.

[6] E. Dubrovinskaya, V. Kebkal, O. Kebkal, K. Kebkal, and P. Casari, "Underwater localization via wideband direction-of-arrival estimation using acoustic arrays of arbitrary shape," *Sensors*, vol. 20, no. 14, p. 3862, 2020.

[7] P. Casari, F. Campagnaro, E. Dubrovinskaya, R. Francescon, A. Dagan, S. Dahan, M. Zorzi, and R. Diamant, "Asuna: A topology dataset for underwater network emulation," *IEEE Journal of Oceanic Engineering*, 2019.

[8] J. D. Colladon, *Souvenirs et mémoires*. Library of the Musée d'histoire des sciences de Genéve, 1893.

[9] "World ocean atlas," last visited: Jan. 2020. [Online]. Available: www.nodc.noaa.gov/OC5/WOA05/pr_woa05.html

[10] R. Diamant, G. N. Shirazi, and L. Lampe, "Robust spatial reuse scheduling in underwater acoustic communication networks," *IEEE J. Ocean. Eng.*, vol. 39, no. 1, pp. 32–46, Jan. 2014.

[11] R. Diamant, P. Casari, F. Campagnaro, O. Kebkal, V. Kebkal, and M. Zorzi, "Fair and throughput-optimal routing in multimodal underwater networks," *IEEE Trans. Wireless Commun.*, vol. 17, no. 3, pp. 1738–1754, Mar. 2018.

[12] L. Brekhovskikh and Y. P. Lysanov, *Fundamentals of ocean acoustics*. Acoustical Society of America, 2004.

[13] J. E. Peck, "Development of hearing. part i: Phylogeny," *Journal-American Academy of Audiology*, vol. 5, pp. 291–291, 1994.

[14] D. R. Lindberg and N. D. Pyenson, "Things that go bump in the night: evolutionary interactions between cephalopods and cetaceans in the tertiary," *Lethaia*, vol. 40, no. 4, pp. 335–343, 2007.

[15] L. Bjørnø, *Applied underwater acoustics*. Elsevier, 2017.

[16] P. C. Etter, *Underwater acoustic modeling and simulation*. CRC press, 2018.

[17] A. Wirgin, "On the velocity of sound in water: theoretical aspects of colladon's nineteenth century experiments," *arXiv preprint arXiv:1710.03777*, 2017.

[18] V. Del Grosso and C. Mader, "Speed of sound in pure water," *the Journal of the Acoustical Society of America*, vol. 52, no. 5B, pp. 1442–1446, 1972.

[19] C.-T. Chen and F. J. Millero, "Speed of sound in seawater at high pressures," *The Journal of the Acoustical Society of America*, vol. 62, no. 5, pp. 1129–1135, 1977.

[20] N. P. Fofonoff and R. Millard Jr, *Algorithms for the computation of fundamental properties of seawater*. Unesco, 1983.

[21] S. Domenico, "Acoustic wave propagation in air-bubble curtains in water. part I: History and theory," *Geophysics*, vol. 47, no. 3, pp. 345–353, 1982.

[22] T. Wen, G. Garrison, R. Francois, R. Stein, and W. Felton, "Sound speed, reflectivity, and absorption measurements in arctic ice in 1988," WASHINGTON UNIV SEATTLE APPLIED PHYSICS LAB, Tech. Rep., 1991.

[23] M. Stojanovic, "On the relationship between capacity and distance in an underwater acoustic communication channel," *ACM SIGMOBILE Mobile Computing and Communications Review*, vol. 11, no. 4, pp. 34–43, Oct. 2007.

[24] M. A. McDonald, J. A. Hildebrand, and S. M. Wiggins, "Increases in deep ocean ambient noise in the northeast pacific west of san nicolas island, california," *The Journal of the Acoustical Society of America*, vol. 120, no. 2, pp. 711–718, 2006.

[25] N. Jones, "Ocean uproar: saving marine life from a barrage of noise," *Nature*, vol. 568, no. 7751, pp. 158–161, 2019.

[26] K. Kebkal, O. Kebkal, E. Glushko, V. Kebkal, L. Sebastiao, A. Pascoal, J. Gomes, J. Ribeiro, H. Silva, M. Ribeiro *et al.*, "Underwater acoustic modems with integrated atomic clocks for one-way travel-time underwater vehicle positioning," in *Proceefings of the Underwater Acoustics Conference and Exhibition (UACE)*, 2017.

[27] P. Casari, C. Tapparello, I. Calabrese, F. Favaro, G. Toso, S. Azad, R. Masiero, and M. Zorzi, "Open-source suites for the underwater networking community: WOSS and DESERT Underwater," *IEEE Network, S.I. on Open Source for Networking: Development and Experimentation*, vol. 28, no. 5, pp. 38–46, Sept. 2014.

[28] I. Nissen, *Adaptive Systems for Mobile Underwater Communications with a p(oste)riori Channel Knowledge.* FWG-Report, Wehrtechnische Dienststelle 71, Eckernförde.

[29] M. B. Porter, "The kraken normal mode program," Naval Research Lab Washington DC, Tech. Rep., 1992.

[30] M. Porter *et al.*, "Bellhop code," last visited: Jan. 2020. [Online]. Available: http://oalib.hlsresearch.com/Rays/index.html

[31] J. Liu, Z. Zhou, Z. Peng, J.-H. Cui, M. Zuba, and L. Fiondella, "Mobi-sync: efficient time synchronization for mobile underwater sensor networks," *IEEE Transactions on Parallel and Distributed Systems*, vol. 24, no. 2, pp. 406–416, 2012.

[32] F. Lu, D. Mirza, and C. Schurgers, "D-sync: Doppler-based time synchronization for mobile underwater sensor networks," in *Proceedings of the Fifth ACM International Workshop on UnderWater Networks*, 2010, pp. 1–8.

[33] R. Diamant and L. Lampe, "Underwater localization with time-synchronization and propagation speed uncertainties," *IEEE Transactions on Mobile Computing*, vol. 12, no. 7, pp. 1257–1269, 2012.

[34] A. Turetta, G. Casalino, E. Simetti, A. Sperindè, and S. Torelli, "Analysis of the accuracy of a lbl-based underwater localization procedure," in *2014 Oceans-St. John's.* IEEE, 2014, pp. 1–7.

[35] H.-E. De Bree *et al.*, "The microflown: An acoustic particle velocity sensor," *Acoustics Australia*, vol. 31, no. 3, pp. 91–94, 2003.

[36] A. Nehorai and E. Paldi, "Acoustic vector-sensor array processing," *IEEE Transactions on signal processing*, vol. 42, no. 9, pp. 2481–2491, 1994.

[37] L. Paull, M. Seto, and J. J. Leonard, "Decentralized cooperative trajectory estimation for autonomous underwater vehicles," in *2014 IEEE/RSJ International Conference on Intelligent Robots and Systems*. IEEE, 2014, pp. 184–191.

[38] M. Erol, L. F. Vieira, A. Caruso, F. Paparella, M. Gerla, and S. Oktug, "Multi stage underwater sensor localization using mobile beacons," in *2008 Second International Conference on Sensor Technologies and Applications (sensorcomm 2008)*. IEEE, 2008, pp. 710–714.

[39] T. LN Nguyen and Y. Shin, "An efficient RSS localization for underwater wireless sensor networks," *Sensors*, vol. 19, no. 14, p. 3105, 2019.

[40] R. Diamant, H.-P. Tan, and L. Lampe, "LOS and NLOS classification for underwater acoustic localization," *IEEE Trans. Mobile Comput.*, vol. 13, no. 2, pp. 311–323, Feb. 2014.

[41] J. Palacios, G. Bielsa, P. Casari, and J. Widmer, "Single- and multiple-access point indoor localization for millimeter-wave networks," *IEEE Trans. Wireless Commun.*, vol. 18, no. 3, pp. 1927–1942, Mar. 2019.

[42] X. Sheng and Y.-H. Hu, "Maximum likelihood multiple-source localization using acoustic energy measurements with wireless sensor networks," *IEEE Transactions on Signal Processing*, vol. 53, no. 1, pp. 44–53, 2004.

[43] Z. Yuyi, G. Zhenbang, W. Lei, Z. Ruiyong, and L. Huanxin, "Study of underwater positioning based on short baseline sonar system," in *2009 International Conference on Artificial Intelligence and Computational Intelligence*, vol. 2. IEEE, 2009, pp. 343–346.

[44] M. B. Larsen, "Synthetic long baseline navigation of underwater vehicles," in *OCEANS 2000 MTS/IEEE Conference and Exhibition. Conference Proceedings (Cat. No. 00CH37158)*, vol. 3. IEEE, 2000, pp. 2043–2050.

[45] D. Marx, M. Nelson, E. Chang, W. Gillespie, A. Putney, and K. Warman, "An introduction to synthetic aperture sonar," in *Proceedings of the Tenth IEEE Workshop on Statistical Signal and Array Processing (Cat. No. 00TH8496)*. IEEE, 2000, pp. 717–721.

[46] G. Tuna and V. C. Gungor, "A survey on deployment techniques, localization algorithms, and research challenges for underwater acoustic sensor networks," *International Journal of Communication Systems*, vol. 30, no. 17, p. e3350, 2017.

[47] L. Paull, S. Saeedi, M. Seto, and H. Li, "AUV navigation and localization: A review," *IEEE Journal of Oceanic Engineering*, vol. 39, no. 1, pp. 131–149, 2014.

[48] L. Karim, N. Nasser, Q. H. Mahmoud, A. Anpalagan, and T. Salti, "Range-free localization approach for m2m communication system using mobile anchor nodes," *Journal of Network and Computer Applications*, vol. 47, pp. 137–146, 2015.

[49] T. Zhang, L. Chen, and Y. Li, "AUV underwater positioning algorithm based on interactive assistance of SINS and LBL," *Sensors*, vol. 16, no. 1, p. 42, 2015.

[50] R. Diamant and L. Lampe, "Underwater localization with time-synchronization and propagation speed uncertainties," *IEEE Transactions on Mobile Computing*, vol. 12, no. 7, pp. 1257–1269, 2013.

[51] H. G. Thomas, "GIB buoys: an interface between space and depths of the oceans," in *Autonomous Underwater Vehicles, 1998. AUV'98. Proceedings of the 1998 Workshop on*, Aug 1998, pp. 181–184.

[52] C. O. Tiemann, M. B. Porter, and L. N. Frazer, "Localization of marine mammals near Hawaii using an acoustic propagation model," *J. Acoust. Soc. Am.*, vol. 115, no. 6, pp. 2834–2843, 2004.

[53] B. Friedlander, "Accuracy of source localization using multipath delays," *IEEE Transactions on Aerospace and Electronic Systems*, vol. 24, no. 4, pp. 346–359, 1988.

[54] P. Blanc-Benon and C. Jauffret, "TMA from bearings and multipath time delays," *IEEE Transactions on Aerospace and Electronic Systems*, vol. 33, no. 3, pp. 813–824, 1997.

[55] R. Morrissey, J. Ward, N. DiMarzio, S. Jarvis, and D. Moretti, "Passive acoustic detection and localization of sperm whales (physeter macrocephalus) in the tongue of the ocean," *Applied Acoustics*, vol. 67, no. 11, pp. 1091–1105, 2006.

[56] J. J. Leonard and A. Bahr, "Autonomous underwater vehicle navigation," in *Handbook of Ocean Engineering*. Springer, 2016, pp. 341–358.

[57] Z. Zhou, Z. Peng, J.-H. Cui, Z. Shi, and A. Bagtzoglou, "Scalable localization with mobility prediction for underwater sensor networks," *IEEE Trans. Mobile Comput.*, vol. 10, no. 3, pp. 335–348, 2011.

[58] K.-C. Lee, J.-S. Ou, M.-C. Huang, and M.-C. Fang, "A novel location estimation based on pattern matching algorithm in underwater environments," *Elsevier Applied Acoustics*, vol. 70, no. 3, pp. 479–483, 2009.

[59] H.-P. Tan, R. Diamant, W. K. Seah, and M. Waldmeyer, "A survey of techniques and challenges in underwater localization," *Ocean Engineering*, vol. 38, no. 14-15, pp. 1663–1676, 2011.

[60] J. Reis, M. Morgado, P. Batista, P. Oliveira, and C. Silvestre, "Design and experimental validation of a usbl underwater acoustic positioning system," *Sensors*, vol. 16, no. 9, 2016. [Online]. Available: https://www.mdpi.com/1424-8220/16/9/1491

[61] M. Erol, H. Mouftah, and S. Oktug, "Localization techniques for underwater acoustic sensor networks," *IEEE Commun. Mag.*, vol. 48, no. 12, pp. 152–158, Jun. 2010.

[62] S. Gezci, "A survey on wireless position estimation," *Wireless Personal Communications*, vol. 44, no. 3, pp. 263–282, Feb. 2008.

[63] Z. Zhang, C. L. Law, and Y. L. Guan, "Modified phase-only correlator with kurtosis-based amplified-noise suppression," *IEEE Trans. Wireless Commun.*, vol. 9, no. 11, pp. 3341–3345, Nov. 2010.

[64] L. Mu, G. Kuo, and N. Tao, "A novel ToA location algorithm using LOS range estimation for NLOS environments," in *Proc. of the IEEE Vehicular Technology Conference (VTC)*, Melbourne, Australia, May 2006, pp. 594–598.

[65] J. Ash and R. Moses, "Acoustic time delay estimation and sensor network self-localization: Experimental results," *Journal of Acoustic Society of America*, vol. 118, no. 2, pp. 841–850, August 2005.

[66] D. McCrady, L. Doyle, H. Forstrom, T. Dempsey, and M. Martorana, "Mobile ranging using low-accuracy clocks," *IEEE Trans. Microwave Theory and Techniques*, vol. 48, no. 6, pp. 951–957, June 2000.

[67] S. Fischer, H. Grubeck, A. Kangas, H. Koorapaty, E. Larsson, and P. Lundqvist, "Time of arrival estimation of narrowband TDMA signals for mobile positioning," *Proc. of the IEEE International Symposium on Personal, Indoor and Mobile Radio Communications (PIMRC)*, pp. 451–455, September 1998.

[68] J. Gebbie, M. Siderius, and J. S. Allen III, "A two-hydrophone range and bearing localization algorithm with performance analysis," *J. Acoust. Soc. of America*, vol. 137, no. 3, pp. 1586–1597, 2015.

[69] A. B. Baggeroer, W. A. Kuperman, and P. N. Mikhalevsky, "An overview of matched field methods in ocean acoustics," *IEEE Journal of Oceanic Engineering*, vol. 18, no. 4, pp. 401–424, 1993.

[70] J. S. Perkins and W. Kuperman, "Environmental signal processing: Three-dimensional matched-field processing with a vertical array," *J. Acoust. Soc. of America*, vol. 87, no. 4, pp. 1553–1556, 1990.

[71] K. L. Gemba, W. S. Hodgkiss, and P. Gerstoft, "Adaptive and compressive matched field processing," *The Journal of the Acoustical Society of America*, vol. 141, no. 1, pp. 92–103, 2017.

[72] S. Shakeri and G. Leus, "Underwater ultra-wideband fingerprinting-based sparse localization," in *Signal Processing Advances in Wireless Communications (SPAWC), 2014 IEEE 15th International Workshop on.* IEEE, 2014, pp. 140–144.

[73] S. Rakotonarivo and W. Kuperman, "Model-independent range localization of a moving source in shallow water," *The Journal of the Acoustical Society of America*, vol. 132, no. 4, pp. 2218–2223, 2012.

[74] C. M. Verlinden, J. Sarkar, B. D. Cornuelle, and W. A. Kuperman, "Determination of acoustic waveguide invariant using ships as sources of opportunity in a shallow water marine environment," *The Journal of the Acoustical Society of America*, vol. 141, no. 2, pp. EL102–EL107, 2017.

[75] L. Liao, Y. V. Zakharov, and P. D. Mitchell, "Underwater localization based on grid computation and its application to transmit beamforming in multiuser UWA communications," *IEEE Access*, vol. 6, pp. 4297–4307, Jan. 2018.

[76] M. S. Ibn Seddik, L. Jaulin, and J. Grimsdale, "Phase based localization for underwater vehicles using interval analysis," *Mathematics in Computer Science*, vol. 8, no. 3, pp. 495–502, Sep. 2014.

[77] H. Liu, H. Darabi, P. Banerjee, and J. Liu, "Survey of wireless indoor positioning techniques and systems," *IEEE Trans. Syst., Man, Cybern. C*, vol. 37, no. 6, pp. 1067–1080, Nov. 2007.

[78] K. Kaemarungsi and P. Krishnamurthy, "Modeling of indoor positioning systems based on location fingerprinting," in *Proc. IEEE INFOCOM*, vol. 2, Hong Kong, China, Mar. 2004, pp. 1012–1022.

[79] Y. Jin, W. Soh, and W. Wong, "Indoor localization with channel impulse response based fingerprint and nonparametric regression," *IEEE Trans. Wireless Commun.*, vol. 9, no. 3, pp. 1120–1127, Mar. 2010.

[80] X. Guo and N. Ansari, "Localization by fusing a group of fingerprints via multiple antennas in indoor environment," *IEEE Trans. Veh. Technol.*, vol. 66, no. 11, pp. 9904–9915, Nov. 2017.

[81] P. Tseng, Y. Chan, Y. Lin, D. Lin, N. Wu, and T. Wang, "Ray-tracing-assisted fingerprinting based on channel impulse response measurement for indoor positioning," *IEEE Trans. Instrum. Meas.*, vol. 66, no. 5, pp. 1032–1045, May 2017.

[82] N. Etemadyrad, "A sequential detection approach to indoor positioning using RSS-based fingerprinting," Master's thesis, George Mason University, 2017.

[83] P. K. Yadav and W. Meng, "Mobile targets localization in a field area using moving gaussian peaks and probability map," in *Proc. IEEE ICCA*, Taichung, Taiwan, Jun. 2014.

[84] Z. Xiao, H. Wen, A. Markham, and N. Trigoni, "Lightweight map matching for indoor localisation using conditional random fields," in *Proc. IPSN*, Berlin, Germany, Apr. 2014.

[85] Z. Wei, Y. Zhao, X. Liu, and Z. Feng, "DoA-LF: A location fingerprint positioning algorithm with millimeter-wave," *IEEE Access*, vol. 5, pp. 22 678–22 688, Sep. 2017.

[86] F. Jensen, W. Kuperman, M. Porter, and H. Schmidt, *Computational Ocean Acoustics*, 2nd ed. New York: Springer-Verlag, 1984, 2nd printing 2000.

[87] P. Qarabaqi and M. Stojanovic, "Statistical characterization and computationally efficient modeling of a class of underwater acoustic communication channels," *Oceanic Engineering, IEEE Journal of*, vol. 38, no. 4, pp. 701–717, 2013.

[88] L. M. Wolff, E. Szczepanski, and S. Badri-Hoeher, "Acoustic underwater channel and network simulator," in *Proc. MTS/IEEE OCEANS*, Yeosu, Korea, May 2012, pp. 1–6.

[89] R. Diamant, "Closed form analysis of the normalized matched filter with a test case for detection of underwater acoustic signals," *IEEE Access*, vol. 4, pp. 8225–8235, 2016.

[90] Q. Wang, L. Wei, and R. Kennedy, "Iterative Viterbi decoding, trellis shaping, and multilevel structure for high-rate parity-concatenated TCM," *IEEE Trans. Commun.*, vol. 50, no. 1, pp. 48–55, Jan 2002.

[91] L. Marchetti and R. Reggiannini, "An efficient receiver structure for sweep-spread-carrier underwater acoustic links," *IEEE J. Ocean. Eng.*, vol. 41, no. 2, pp. 440–449, Apr. 2016.

[92] C. Boya, M. Ruiz-Llata, J. Posada, A. Garcia-Souto, and A. José, "Identification of multiple partial discharge sources using acoustic emission technique and blind source separation," *IEEE Trans. Dielectr. Electr. Insul.*, vol. 22, no. 3, pp. 1663–1673, Jun. 2015.

[93] S. Byun, S. Kim, C. Park, K. Kim, and C. Lee, "Cyclostationary analysis of underwater noise for vehicle propeller monitoring," in *Proc. MTS/IEEE OCEANS*, Monterey, CA, Sep. 2016.

[94] D. Divins and D. Metzger. (2016) US coastal relief model. [Online]. Available: http://www.ngdc.noaa.gov/mgg/coastal/coastal.html

[95] R. Diamant, A. Knap, S. Dahan, I. Mardix, J. Walpert, and S. DiMarco, "THEMO: the Texas A&M - University of Haifa - eastern mediterranean observatory," in *IEEE International Conference on Ultra-Wideband (ICUWB)*, Nanjing, China, Sep. 2010.

[96] T. Alexandri and R. Diamant, "A reverse bearings only target motion analysis for autonomous underwater vehicle navigation," *IEEE Trans. Mobile Comput.*, vol. 18, no. 3, pp. 494–506, March 2019.

[97] "Low-frequency pingers a quick and easy way to relocate an underwater site," http://www.jwfishers.com/products/pingers-low.html, 2019.

[98] D. Levy, Y. Belfer, E. Osherov, E. Bigal, A. P. Scheinin, H. Nativ, D. Tchernov, and T. Treibitz, "Automated analysis of marine video with limited data," in *Proc. IEEE/CVF CVPRW*, 2018.

[99] B. Zerr, G. Mailfert, A. Bertholom, and H. Ayreault, "Sidescan sonar image processing for auv navigation," in *Proc. MTS/IEEE Oceans*, vol. 1, Brest, France, Jun. 2005, pp. 124–130.

[100] E. Gallimore, E. Terrill, R. Hess, A. Nager, H. Bachelor, and A. Pietruszka, "Integration and evaluation of a next-generation chirp-style sidescan sonar on the remus 100," in *Proc. IEEE/OES AUV Workshop*, Porto, Portugal, Nov. 2018, pp. 1–6.

[101] Z. Tang, G. Blacquiere, and G. Leus, "Aliasing-free wideband beamforming using sparse signal representation," *IEEE Trans. Signal Process.*, vol. 59, no. 7, pp. 3464–3469, Jul. 2011.

[102] M. Arkhipov, "An approach to using basic three-element arrays in tetrahedral-based USBL systems," in *Proc MTS/IEEE OCEANS*, San Diego, CA, Sep. 2013, pp. 1–8.

[103] A. Tesei, S. Fioravanti, V. Grandi, P. Guerrini, and A. Maguer, "Localization of small surface vessels through acoustic data fusion of two tetrahedralarrays of hydrophones," in *Proc. ECUA*, Jul. 2012.

[104] M. R. Devi, N. S. Kumar, and J. Joseph, "Sparse reconstruction based direction of arrival estimation of underwater targets," in *Proc. SYMPOL*, Nov. 2015, pp. 1–6.

[105] W. Liu and S. Weiss, *Wideband beamforming: concepts and techniques.* John Wiley & Sons, Mar. 2010.

[106] S. Stergiopoulos, *Advanced signal processing handbook: theory and implementation for radar, sonar, and medical imaging real time systems.* CRC press, 2018.

[107] F. Andersson, M. Carlsson, J.-Y. Tourneret, and H. Wendt, "A method for 3D direction of arrival estimation for general arrays using multiple frequencies," in *Proc. CAMSAP*, 2015.

[108] K. T. Wong and M. D. Zoltowski, "Closed-form underwater acoustic direction-finding with arbitrarily spaced vector hydrophones at unknown locations," *IEEE J. Ocean. Eng.*, vol. 22, no. 3, pp. 566–575, 1997.

[109] U. K. Chandrika and V. N. Hari, "A vector sensing scheme for underwater acoustics based on particle velocity measurements," in *Proc. MTS/IEEE OCEANS 2015*, Washington, DC, 2015, pp. 1–7.

[110] A. Bereketli, M. B. Guldogan, T. Kolcak, T. Gudu, and A. L. Avsar, "Experimental results for direction of arrival estimation with a single acoustic vector sensor in shallow water," *Hindawi Journal of Sensors*, vol. 2015, Apr. 2015.

[111] EvoLogics GmbH, "S2C R 7/17 USBL communication and positioning device," last visited: July 1, 2020. [Online]. Available: https://evologics.de/acoustic-modem/7-17/usbl-serie

[112] R. J. Vaccaro, "The past, present, and the future of underwater acoustic signal processing," *IEEE Signal Process. Mag.*, vol. 15, no. 4, pp. 21–51, Jul. 1998.

[113] X. Zhong, A. B. Premkumar, and W. Wang, "Direction of arrival tracking of an underwater acoustic source using particle filtering: Real data experiments," in *Proc. IEEE Tencon–Spring*, Apr. 2013, pp. 420–424.

[114] W. Chen, T. Luo, F. Chen, F. Ji, and H. Yu, "Directions of arrival and channel parameters estimation in multipath underwater environment," in *Proc. MTS/IEEE OCEANS*, Apr. 2016, pp. 1–4.

[115] H. Amindavar and A. M. Reza, "A new simultaneous estimation of directions of arrival and channel parameters in a multipath environment," *IEEE Trans. Signal Process.*, vol. 53, no. 2, pp. 471–483, Feb. 2005.

[116] W. A. P. van Kleunen, K. C. H. Blom, N. Meratnia, A. B. J. Kokkeler, P. J. M. Havinga, and G. J. M. Smit, "Underwater localization by combining time-of-flight and direction-of-arrival," in *Proc. MTS/IEEE OCEANS*, Apr. 2014, pp. 1–6.

[117] X. Guo, K. Yang, X. Yan, and Y. Ma, "Theory of passive localization for underwater sources based on multipath arrival structures," in *Proc. MTS/IEEE OCEANS*, Sep. 2017.

[118] H. Song, J. Qin, C. Yang, and M. Diao, "Compressive beamforming for underwater acoustic source direction-of-arrival estimation," in *Proc. IEEE ICCE-TW*, May 2018.

[119] J. Lin, X. Ma, C. Hao, and L. Jiang, "Direction of arrival estimation of sparse rectangular array via two-dimensional continuous compressive sensing," in *Proc. ICIST*, May 2016, pp. 539–543.

[120] J. Li, Q.-h. Lin, C.-y. Kang, K. Wang, and X.-T. Yang, "DoA estimation for underwater wideband weak targets based on coherent signal subspace and compressed sensing," *MDPI Sensors*, vol. 18, no. 3, Mar. 2018.

[121] A. Abdi, H. Guo, and P. Sutthiwan, "A new vector sensor receiver for underwater acoustic communication," in *Proc. MTS/IEEE OCEANS*, Sep. 2007, pp. 1–10.

[122] T. Zhao, H. Chen, H. Zhang, Z. Li, and L. Tong, "Design of underwater particle velocity pickup sensor," in *Proc. IEEE/OES COA*, Jan. 2016, pp. 1–6.

[123] J. Cao, J. Liu, J. Wang, and X. Lai, "Acoustic vector sensor: reviews and future perspectives," *IET Signal Processing*, vol. 11, no. 1, pp. 1–9, 2017.

[124] G. Sun, J. Hui, and Y. Chen, "Fast direction of arrival algorithm based on vector-sensor arrays using wideband sources," *Springer J. Marine. Sci. Appl.*, vol. 7, pp. 195–199, Sep. 2008.

[125] P. K. Gentner, W. Gartner, G. Hilton, M. E. Beach, and C. F. MecklenbrÃ¤uker, "Towards a hardware implementation of ultra-wideband beamforming," in *Proc. ITG WSA*, Bremen, Germany, 2010, pp. 408–413.

[126] L. Wang, C. Heng, and Y. Lian, "CMOS UWB beamforming radar system," in *Proc. IEEE ICECS*, Marseille, France, 2014, pp. 810–813.

[127] B. Wang, H. Gao, M. K. Matters-Kammerer, and P. G. M. Baltus, "Interpolation based wideband beamforming architecture," in *Proc. IEEE ISCAS*, Baltimore, MD, 2017, pp. 1–4.

[128] M. Stojanovic, J. A. Catipovic, and J. G. Proakis, "Reduced-complexity spatial and temporal processing of underwater acoustic communication signals," *J. Acoust. Soc. Am.*, vol. 98, no. 2, pp. 961–972, 1995.

[129] K. F. Nieman, K. A. Perrine, T. L. Henderson, K. H. Lent, T. J. Brudner, and B. L. Evans, "Wideband monopulse spatial filtering for large receiver arrays for reverberant underwater communication channels," in *Proc. MTS/IEEE OCEANS*, Sep. 2010, pp. 1–8.

[130] X. Li, H. Jia, and M. Yang, "Underwater target detection based on fourth-order cumulant beamforming," *ASA Proc. of Meetings on Acoustics*, vol. 31, no. 1, pp. 1–9, 2017.

[131] M. Ester, H.-P. Kriegel, J. Sander, and X. Xu, "A density-based algorithm for discovering clusters in large spatial databases with noise," *Proc. AAAI KDD*, vol. 96, no. 34, pp. 226–231, 1996.

[132] Pedregosa, F.; *et al.*, "Scikit-learn: Machine learning in Python," *Journal of Machine Learning Research*, vol. 12, pp. 2825–2830, 2011.

[133] "Scikit-learn – clustering," last visited: July 1, 2020. [Online]. Available: https://scikit-learn.org/stable/modules/clustering.html#clustering

[134] R. Diamant, V. Voronin, and K. G. Kebkal, "Design structure of SYMBIOSIS: An opto-acoustic system for monitoring pelagic fish," in *Proc. MTS/IEEE OCEANS*, Marseille, France, 2019, pp. 1–6.

[135] J. Bezanson, A. Edelman, S. Karpinski, and V. B. Shah, "Julia: A fresh approach to numerical computing," *SIAM Review*, vol. 59, no. 1, pp. 65–98, 2017.

[136] "Themo - deep sea observatory," http://themo.haifa.ac.il/buoys/mmp, 2020, [Online; accessed 30-July-2020].

[137] H. S. Dol, P. Casari, T. van der Zwan, and R. Otnes, "Software-defined underwater acoustic modems: Historical review and the NILUS approach," *IEEE J. Ocean. Eng.*, vol. 42, no. 3, pp. 722–737, Jul. 2017.

[138] E. Felemban, F. K. Shaikh, U. M. Qureshi, A. A. Sheikh, and S. B. Qaisar, "Underwater sensor network applications: A comprehensive survey," *SAGE International Journal of Distributed Sensor Networks*, vol. 11, no. 11, Nov. 2015.

[139] P. Casari and M. Zorzi, "Protocol design issues in underwater acoustic networks," *Elsevier Computer Communications*, vol. 34, no. 17, pp. 2013–2025, Nov. 2011.

[140] K. Chen, M. Ma, E. Cheng, F. Yuan, and W. Su, "A survey on MAC protocols for underwater wireless sensor networks," *IEEE Commun. Surveys Tuts.*, vol. 16, no. 3, pp. 1433–1447, Third quarter 2014.

[141] M. Ahmed, M. Salleh, and M. I. Channa, "Routing protocols based on node mobility for underwater wireless sensor network (UWSN): A survey," *Elsevier Journal of Network and Computer Applications*, vol. 78, pp. 242–252, Jan. 2017.

[142] P. Casari, C. Tapparello, F. Guerra, F. Favaro, I. Calabrese, G. Toso, S. Azad, R. Masiero, and M. Zorzi, "Open-source suites for underwater networking: WOSS and DESERT Underwater," *IEEE Network, special issue on "Open Source for Networking: Development and Experimentation"*, vol. 28, no. 5, pp. 38–46, Sep. 2014.

[143] C. Petrioli, R. Petroccia, J. R. Potter, and D. Spaccini, "The SUNSET framework for simulation, emulation and at-sea testing of underwater wireless sensor networks," *Elsevier Ad Hoc Networks, S.I. Advances in Underwater Commun. and Networks*, vol. 34, pp. 224–238, Nov. 2015.

[144] M. Chitre, R. Bhatnagar, and W. S. Soh, "UnetStack: an agent-based software stack and simulator for underwater networks," in *Proc. MTS/IEEE OCEANS*, St. John's, NL, Canada, Sep. 2014.

[145] Y. Zhu, S. Le, L. Pu, X. Lu, Z. Peng, and J. Cui, "Towards experimental evaluation of software-defined underwater networked systems," in *Proc. MTS/IEEE OCEANS*, San Diego, CA, Sep. 2013.

[146] T. Schneider and H. Schmidt, "NETSIM: A realtime virtual ocean hardware-in-the-loop acoustic modem network simulator," in *Proc. UCOMMS*, Lerici, Italy, Aug. 2018.

[147] R. Petroccia, G. Zappa, T. Furfaro, J. Alves, and L. D'Amaro, "Development of a software-defined and cognitive communications architecture at CMRE," in *Proc. MTS/IEEE OCEANS*, Charleston, SC, Oct. 2018, pp. 1–10.

[148] R. Diamant and L. Chorev, "Emulation system for underwater acoustic channel," in *Proc. UDT*, Amsterdam, The Netherlands, Jun. 2005.

[149] B. Tomasi, J. Preisig, and M. Zorzi, "On the predictability of underwater acoustic communications performance: The KAM11 data set as a case study," in *Proc. ACM WUWNet*, Seattle, Washington, Dec. 2011.

[150] K. Pelekanakis and M. Chitre, "Adaptive sparse channel estimation under symmetric alpha-stable noise," *IEEE Trans. Wireless Commun.*, vol. 13, no. 6, pp. 3183–3195, Jun. 2014.

[151] R. Otnes, P. A. van Walree, H. Buen, and H. Song, "Underwater acoustic network simulation with lookup tables from physical-layer replay," *IEEE J. Ocean. Eng.*, vol. 40, no. 4, pp. 822–840, Oct. 2015.

[152] R. Ahmed and M. Stojanovic, "Grouped packet coding: A method for reliable communication over fading channels with long delays," *IEEE J. Ocean. Eng.*, 2018, in press.

[153] P. A. van Walree, F. Socheleau, R. Otnes, and T. Jenserud, "The watermark benchmark for underwater acoustic modulation schemes," *IEEE J. Ocean. Eng.*, vol. 42, no. 4, pp. 1007–1018, Oct. 2017.

[154] M. Chitre, "Editorial on writing reproducible and interactive papers," *IEEE J. Ocean. Eng.*, vol. 43, no. 3, pp. 560–562, Jul. 2018.

[155] A. Konrad, B. Y.Zhao, and A. D. Joseph, "Determining model accuracy of network traces," *Elsevier Journal of Computer and System Sciences*, vol. 72, no. 7, pp. 1156–1171, Nov. 2006, special issue: Performance modelling and evaluation of computer systems.

[156] A. O. Kaya, L. Greenstein, and W. Trappe, "Characterizing indoor wireless channels via ray tracing, and validation via measurements," in *Proc. IEEE GLOBECOM*, New Orleans, LA, Nov. 2008.

[157] P. Ferrand, M. Amara, S. Valentin, and M. Guillaud, "Trends and challenges in wireless channel modeling for evolving radio access," *IEEE Commun. Mag.*, vol. 54, no. 7, pp. 93–99, Jul. 2016.

[158] B. Tomasi, P. Casari, M. Zorzi, G. Zappa, and K. McCoy, "Experimental study of the acoustic channel properties during subnet 2009," University of Padova, Tech. Rep., 2010. [Online]. Available: http://telecom.dei.unipd.it/pages/read/75/

[159] J. Llor, M. Stojanovic, and M. P. Malumbres, "A simulation analysis of large scale path loss in an underwater acoustic network," in *Proc. IEEE/OES OCEANS*, Santander, Spain, Jun. 2011.

[160] J. F. Paris, "Statistical characterization of $\kappa-\mu$ shadowed fading," *IEEE Trans. Veh. Technol.*, vol. 63, no. 2, pp. 518–526, Feb. 2014.

[161] P. Owezarski and N. Larrieu, "A trace based method for realistic simulation," in *Proc. IEEE ICC*, vol. 4, Paris, France, Jun. 2004.

[162] E. Nordström, P. Gunningberg, C. Rohner, and O. Wibling, "Evaluating wireless multi-hop networks using a combination of simulation, emulation, and real world experiments," in *Proc. ACM MOBIEVAL*, San Juan, Puerto Rico, Jun. 2007.

[163] M. Jacobsson and C. Rohner, "Comparing wireless flooding protocols using trace-based simulations," *EURASIP Journal on Wireless Communications and Networking*, vol. 2013, no. 1, Jun. 2013.

[164] A. Abedi, A. Heard, and T. Brecht, "Conducting repeatable experiments and fair comparisons using 802.11n MIMO networks," *ACM SIGOPS Oper. Syst. Rev.*, vol. 49, no. 1, pp. 41–50, Jan. 2015.

[165] M. C. Weigle, P. Adurthi, F. Hernández-Campos, K. Jeffay, and F. D. Smith, "Tmix: A tool for generating realistic TCP application workloads in ns-2," *ACM SIGCOMM Comput. Commun. Rev.*, vol. 36, no. 3, pp. 65–76, Jul. 2006.

[166] P. Agrawal and M. Vutukuru, "Trace based application layer modeling in ns-3," in *Proc. NCC*, Guwahati, India, Mar. 2016.

[167] R. Otnes, P. A. van Walree, and T. Jenserud, "Validation of replay-based underwater acoustic communication channel simulation," *IEEE J. Ocean. Eng.*, vol. 38, no. 4, pp. 689–700, Oct. 2013.

[168] J. Alves, J. Potter, P. Guerrini, G. Zappa, and K. LePage, "The LOON in 2014: Test bed description," in *Proc. UComms*, Sestri Levante, Italy, Sep. 2014.

[169] S. M. Taylor and B. Bornhold, "Connecting the dots: Ocean research and public policy," in *Proc. MTS/IEEE OCEANS*, Sydney, Australia, May 2010.

[170] C. Petrioli, R. Petroccia, D. Spaccini, A. Vitaletti, T. Arzilli, D. Lamanna, and A. Galizia, "The SUNRISE GATE: accessing the SUNRISE federation of facilities to test solutions for the Internet of underwater things," in *Proc. UCOMMS*, Sestri Levante, Italy, Sep. 2014.

[171] R. Diamant, S. Dahan, and I. Mardix, "Communication operations at THEMO: the Texas A&M - University of Haifa - eastern mediterranean observatory," in *Proc. UComms*, Lerici, Italy, Aug. 2018.

[172] R. Diamant, Y. Bucris, and A. Feuer, "An efficient method to measure reliability of underwater acoustic communication links," *Journal Of Ocean Engineering And Science*, vol. 1, no. 2, pp. 129–134, 2016.

[173] M. Zuba, A. Song, and J. Cui, "Exploring parabolic equation models for improved underwater network simulations," in *Proc. UComms*, Sestri Levante, Italy, Sep. 2014.

[174] F. Campagnaro, R. Francescon, F. Favaro, F. Guerra, R. Diamant, P. Casari, and M. Zorzi, "The DESERT Underwater framework v2: Improved capabilities and extension tools," in *Proc. UComms*, Lerici, Italy, Sep. 2016.

[175] B. Peleato and M. Stojanovic, "Distance aware collision avoidance protocol for ad hoc underwater acoustic sensor networks," *IEEE Commun. Lett.*, vol. 11, no. 12, pp. 1025–1027, Dec. 2007.

[176] M. Zorzi, P. Casari, N. Baldo, and A. F. Harris III, "Energy-efficient routing schemes for underwater acoustic networks," *IEEE J. Sel. Areas Commun.*, vol. 26, no. 9, pp. 1754–1766, Dec. 2008.

[177] A. Syed, W. Ye, and J. Heidemann, "Comparison and Evaluation of the T-Lohi MAC for Underwater Acoustic Sensor Networks," *IEEE J. Sel. Areas Commun.*, vol. 26, pp. 1731–1743, Dec. 2008.

[178] F. Guerra, P. Casari, and M. Zorzi, "World Ocean Simulation System (WOSS): a simulation tool for underwater networks with realistic propagation modeling," in *Proc. of ACM WUWNet 2009*, Berkeley, CA, Nov. 2009.

[179] G. E. Burrowes, J. Brown, and J. Y. Khan, "Adaptive space time - time division multiple access (AST-TDMA) protocol for an underwater swarm of AUV's," in *Proc. MTS/IEEE OCEANS*, Bergen, Norway, Jun. 2013.

[180] P. A. Forero, S. K. Lapic, C. Wakayama, and M. Zorzi, "Rollout algorithms for data storage- and energy-aware data retrieval using autonomous underwater vehicles," in *Proc. ACM WUWNET*, Rome, Italy, Nov. 2014.

[181] W. Zhuo, G. Hongmei, J. Longjie, and F. Xiaoning, "AUV-aided communication method for underwater mobile sensor network," in *Proc. MTS/IEEE OCEANS*, Shanghai, China, Apr. 2016.

[182] S. M. Ghoreyshi, A. Shahrabi, and T. Boutaleb, "An efficient AUV-aided data collection in underwater sensor networks," in *Proc. IEEE AINA*, Krakow, Poland, May 2018.

[183] M. Erol-Kantarci, L. F. M. Vieira, A. Caruso, F. Paparella, M. Gerla, and S. Oktug, "Multi stage underwater sensor localization using mobile beacons," in *Proc. SENSORCOMM*, Cap Esterel, France, Aug. 2008.

[184] D. Pompili, T. Melodia, and I. F. Akyildiz, "Three-dimensional and two-dimensional deployment analysis for underwater acoustic sensor networks," *Elsevier Ad Hoc Networks*, vol. 7, no. 4, pp. 778–790, Jun. 2009.

[185] Y. Ren, W. K. G. Seah, and P. D. Teal, "Performance of pressure routing in drifting 3D underwater sensor networks for deep water monitoring," in *Proc. ACM WUWNet*, Los Angeles, California, Nov. 2012.

[186] E. Coccolo, F. Campagnaro, A. Signori, F. Favaro, and M. Zorzi, "Implementation of AUV and ship noise for link quality evaluation in the DESERT Underwater framework," in *Proc. ACM WUWNet*, Shenzhen, China, Dec. 2018.

[187] G. Toso, I. Calabrese, F. Favaro, L. Brolo, P. Casari, and M. Zorzi, "Testing network protocols via the DESERT underwater framework: The CommsNet'13 experience," in *Proc. MTS/IEEE OCEANS*, St. John's, Canada, Sep. 2014.

[188] P. Casari, J. Kalwa, M. Zorzi, S. Nasta, S. Schreiber, R. Otnes, P. van Walree, M. Goetz, A. Komulainen, B. Nilsson, J. Nilsson, T. Öberg, I. Nissen, H. Strandberg, H. Dol, G. Leus, and F. Pacini, "Security via underwater acoustic networks: the concept and results of the RACUN project," in *Proc. JNIC*, Leon, Spain, Aug. 2015.

[189] J. S. Jaffe, P. J. S. Franks, P. L. D. Roberts, D. Mirza, C. Schurgers, R. Kastner, and A. Boch, "A swarm of autonomous miniature underwater robot drifters for exploring submesoscale ocean dynamics," *Nature Communications*, vol. 8, Jan. 2017, art. no. 14189.

[190] S. Basagni, C. Petrioli, R. Petroccia, and D. Spaccini, "Channel replay-based performance evaluation of protocols for underwater routing," in *Proc. MTS/IEEE OCEANS*, St. John's, Canada, Sep. 2014.

[191] G. Deane, J. Preisig, and A. Singer, "Making the most of field data to support underwater acoustic communications R&D," in *Proc. UCOMMS*, Lerici, Italy, Aug. 2018.

[192] "ASUNA Octave/Matlab emulator's Code Ocean capsule," last visited: Jan. 2020. [Online]. Available: https://codeocean.com/capsule/3164355

[193] J. G. Proakis, *Digital Communications*, 3rd ed. McGraw-Hill, 1995.

[194] F. Campagnaro, P. Casari, M. Zorzi, and R. Diamant, "Optimal transmission scheduling in small multimodal underwater networks," *IEEE Wireless Commun. Lett.*, vol. 8, no. 2, pp. 368–371, Apr. 2019.

[195] J. Karnowski, E. Hutchins, and C. Johnson, "Dolphin detection and tracking," in *Proc. IEEE WACVW*, Waikoloa, HI, Jan. 2015.

[196] L. Dubbs, A. G. Keeler, and T. O'Meara, "Permitting, risk and marine hydrokinetic energy development," *The Electricity Journal*, vol. 26, no. 10, pp. 64–74, 2013.

[197] A. W. Stoner, C. H. Ryer, S. J. Parker, P. J. Auster, and W. W. Wakefield, "Evaluating the role of fish behavior in surveys conducted with underwater vehicles," *Canadian Journal of Fisheries and Aquatic Sciences*, vol. 65, no. 6, pp. 1230–1243, 2008.

[198] A. E. Copping, B. Polagye, R. Suryan, S. Kramer, J. B. Saracino, and C. Smith, "Instrumentation for monitoring around marine renewable energy converters:

Workshop final report," Pacific Northwest National Lab. (PNNL), Richland, WA (United States), Tech. Rep., 2014.

[199] Z. Cheng, K. Yang, J. Han, Y. Zhou, L. Sun, W. Li, and M. Xia, "Improved time-of-flight range acquisition technique in underwater lidar experiments," *Applied optics*, vol. 54, no. 18, pp. 5715–5725, 2015.

[200] S. P. O'Connor, L. J. Mullen, and B. Cochenour, "Underwater modulated pulse laser imaging system," *Optical Engineering*, vol. 53, no. 5, 2014.

[201] R. W. Lee, J. K. Nash, B. Cochenour, and L. J. Mullen, "Pulse compression techniques to improve modulated pulsed laser line scan systems," *Proc. SPIE*, vol. 9459, 2015.

[202] F. Dalgleish, B. Ouyang, A. Vuorenkoski, B. Ramos, G. Alsenas, B. Metzger, Z. Cao, and J. Principe, "Undersea lidar imager for unobtrusive and eye safe marine wildlife detection and classification," in *Proc. MTS/IEEE OCEANS*, Aberdeen, Scotland, Jun. 2017.

[203] F. M. Caimi. and F. R. Dalgleish, "Performance considerations for continuous-wave and pulsed laser line scan (LLS) imaging systems," *Journal of the European Optical Society*, vol. Rapid Publications 5, 2010.

[204] F. R. Dalgleish, F. M. Caimi, W. B. Britton, and A. C. F., "Improved LLS imaging performance in scattering-dominant waters," in *Proc. of SPIE*, vol. 7317, 2009.

[205] F. Dalgleish, A. Vuorenkoski, G. Nootz, B. Ouyang, and F. Caimi, "Experimental study into the performance impact of environmental noise on undersea pulsed laser serial imagers," *Journal of Underwater Acoustics (USN)*, vol. 61, Oct. 2011.

[206] F. R. Dalgleish, A. K. Vuorenkoski, and B. Ouyang, "Extended-range undersea laser imaging: Current research status and a glimpse at future technologies," *Marine Technology Society Journal*, vol. 47, pp. 128–147, 2013.

[207] C. A. Brock, H. H. Jonsson, J. C. Wilson, J. E. Dye, D. Baumgardner, S. Borrmann, M. C. Pitts, M. T. Osborn, R. J. DeCoursey, and D. C. Woods, "Relationships between optical extinction, backscatter and aerosol surface and volume in the stratosphere following the eruption of mt. pinatubo," *Geophysical research letters*, vol. 20, no. 22, pp. 2555–2558, 1993.

[208] W. Schröer, J. Köser, and F. Kuhnen, "Light–scattering in turbid fluids: The single-scattering intensity," *Journal of Molecular Liquids*, vol. 134, no. 1, pp. 40–48, 2007.

[209] H. Edelsbrunner, D. Kirkpatrick, and R. Seidel, "On the shape of a set of points in the plane," *IEEE Trans. Inf. Theory*, vol. 29, no. 4, pp. 551–559, 1983.

[210] "Underwater acoustic localization of marine mammals and vehicles," https://phys.org/news/2017-11-underwater-acoustic-localization-marine-mammals.html, 2017, [Online; accessed 26-January-2021].

[211] "Underwater acoustic localization of marine mammals and vehicles," https://www.sciencedaily.com/releases/2017/11/171122112945.htm, 2017, [Online; accessed 26-January-2021].

[212] "Underwater acoustic localization of marine mammals and vehicles," https://new.eurekalert.org/pub_releases/2017-11/ini-ual112217.php, 2017, [Online; accessed 26-January-2021].

[213] "Underwater acoustic localization of marine mammals and vehicles," https://www.innovations-report.com/information-technology/underwater-acoustic-localization-of-marine-mammals-and-vehicles/, 2017, [Online; accessed 26-January-2021].

[214] "Innovative autonomous system for identifying schools of fish," https://phys.org/news/2018-06-autonomous-schools-fish.html, 2018, [Online; accessed 26-January-2021].

[215] "Can schools of fish be identified without human intervention?" https://phys.org/news/2018-10-schools-fish-human-intervention.html, 2018, [Online; accessed 26-January-2021].

[216] "SYMBIOSIS Project: A holistic opto-acoustic system for monitoring biodiversities," http://symbiosis.networks.imdea.org, 2018, [Online; accessed 26-January-2021].