Summer 7-14-2014

Advances in Aquatic Target Localization with Passive Sonar

John Thomas Gebbie
Portland State University

Advances in Aquatic Target Localization with Passive Sonar

by

John Thomas Gebbie

A dissertation submitted in partial fulfillment of the
requirements for the degree of

Doctor of Philosophy
in
Electrical and Computer Engineering

Dissertation Committee:
T. Martin Siderius, Chair
Lisa M. Zurk
Richard L. Campbell
John S. Allen, III
Mark D. Sytsma

Portland State University
2014

Abstract

New underwater passive sonar techniques are developed for enhancing target localization capabilities in shallow ocean environments. The ocean surface and the seabed act as acoustic mirrors that reflect sound created by boats or subsurface vehicles, which gives rise to echoes that can be heard by hydrophone receivers (underwater microphones). The goal of this work is to leverage this "multipath" phenomenon in new ways to determine the origin of the sound, and thus the location of the target. However, this is difficult for propeller driven vehicles because the noise they produce is both random and continuous in time, which complicates its measurement and analysis. Further, autonomous underwater vehicles (AUVs) pose additional challenges because very little is known about the sound they generate, and its similarity to that of boats. Existing methods for localizing propeller noise using multiple hydrophones have approached the problem either purely theoretically, or empirically such as by analyzing the interference patterns between multipath arrivals at different frequencies, however little has been published on building localization techniques that directly measure and utilize the time delays between multipath arrivals while simultaneously accounting for relevant environmental parameters. This research develops such techniques through a combination of array beamforming and advanced ray-based modeling that account for variations in bathymetry (seabed topography) as well as variations of the sound speed of the water. The basis for these advances come from several at-sea experiments in which different configurations of passive sonar systems recorded sounds emitted by different types of targets, including small boats and an autonomous underwater vehicle. Ultimately, these contributions may reduce the complexity and cost of passive systems that need to be deployed close to shore, such as for harbor security applications. Further, they also create new possibilities

for applying passive sonar in remote ocean regions for tasks such as detecting illegal fishing activity.

This dissertation makes three key contributions:

1. Analysis of the aspect-dependent acoustic radiation patterns of an underway autonomous underwater vehicle (AUV) through full-field wave modeling.

2. A two-hydrophone cross-correlation technique that leverages multipath as well as bathymetric variations to estimate the range and bearing of a small boat, supported by a mathematically rigorous performance analysis.

3. A multi-target localization technique based on directly measuring multipath from multiple small surface vessels using a small hydrophone array mounted to the nose of an AUV, which operates by cross-correlating two elevation beams on a single bearing.

Dedication

To Donald and Jo Anne, my parents. You always believed in me.

To Tom, my uncle. You ignited my passion for science.

To Tom, my brother. You have always been there for me.

To Vanessa, the love of my life. You are my rock.

Acknowledgements

I would like to express my deep appreciation and gratitude to my advisor, Martin Siderius, for his constant encouragement, patience, and mentorship. Throughout the PhD program, not only have I been the beneficiary of his intellect and fantastic attitude, but his friendship as well. I consider myself truly fortunate to have had the opportunity to work with him. I would like to thank Lisa Zurk for initially recognizing that I belonged in this field, and for teaching me the how to think and speak like a scientist. I would like to thank John Allen for his advice, for making the Hawai‘i experiments possible, and for his valuable feedback on my research.

To my dissertation committee, Mark Sytsma, Richard Campbell, John Allen, and Lisa Zurk, thank you for your time and advice on how to strengthen my dissertation. To the many current and former colleagues I have had at the NEAR-Lab, thank you for your feedback, helpful conversations, and friendship. Finally, I would like to acknowledge and thank the Office of Naval Research for my fellowship, which has helped make this work possible.

Table of Contents

List of Tables

List of Figures

Chapter 1

Introduction

This dissertation is about using underwater passive sonar to localize targets in the ocean.[1] Simply put, this involves listening for sound waves created by surface or underwater vehicles, and then analyzing this information to estimate where the waves originate from. A passive system is one that operates only by listening. This is in contrast to an active system that listens for the echoes of sound waves that it emits. For example, human hearing is largely a passive process (Sutter, 2011). However, some mammals have evolved to use active sonar. Two notable examples are bats and whales that use echolocation to forage for food by making sounds and listening for echoes from prey.

Why use sonar in the ocean? For one, electromagnetic waves, such as radio waves or microwaves, can travel only very short distances in water before being absorbed by the water itself. This makes them essentially useless for detecting or communicating with submerged objects. A notable exception are extremely low frequency (ELF) band radio signals that have been used to send messages to submarines, but this technology lacks other applications due to the extreme cost and complexity (National

[1] In this work, the term *target* is used as a general-purpose descriptor for an object of interest, and does not necessarily imply hostility.

Research Council (U.S.), 1997). Much like radio waves in the atmosphere, sound waves can travel great distances in water. The deep sound channel in the ocean can carry sounds from ships and marine mammals thousands of kilometers across the oceans (Whitman, 2005). The use of sonar underwater can be dated back to 1490, when Leonardo da Vinci wrote "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" (National Defense Research Committee, 1946). Naval warfare over the past century has been a driving force behind research and development of sonar (Urick, 1967), and while this is still true today, interest has extended into other sectors such as civilian law enforcement, research, and industry.

This growth has not been without controversy. The Navy's use of mid-frequency active (MFA) sonar has been suspected as the cause of various marine mammal beachings (Carrington, 2013), and was even the subject of a recent Supreme Court case (Supreme Court, 2008). The challenge with MFA is it overlaps with the same frequency band that many mammals use potentially damaging their acoustic senses, thus "blinding" them. Another source of noise comes from the oil and gas industry, who regularly use seismic air guns for oil exploration (Bernstein, 2014). These systems work by making loud explosive-like sounds in the water and recording the echoes from sub-seabed geological layers. Beyond deliberate creation of sound, the increase in international trade has led to an increase in the number and frequency of cargo ships traversing the oceans (Li et al., 2010). Large ships can be very loud due to diesel engine noise coupling into the water through the hull, or the propeller blades cutting through the water (Arveson and Vendittis, 2000). The ocean is surely a noisy place.

Passive sonar does not add to these problems because no sound is produced during operation, thus greatly reducing its environmental impact. However, there is another advantage to not creating sound, in that a passive system is much less prone to counter detection. This makes it a vital tool for situations in which stealth is needed (Towle et al., 2007). It is no coincidence that tactical sonar systems on modern submarines make extensive use of passive sonar (Moreavek and Brudner, 1999).

The focus of this work is the application of passive sonar to shallow ocean environments. Why shallow water? Continental shelves can extend hundreds of kilometers from shore, meaning that coastal waters are generally quite shallow relative to the rest of the ocean. Underwater mountain ranges and sea mounts can also create shallow regions farther from shore, and can be host to unique ecosystems and significant biodiversity. Both coastal and remote regions are under constant assault from illegal fishing. In response, the U.S. Government, along with other governments around the world, have created Marine Protected Areas (MPAs) that define areas within coastal waters in which commercial fishing vessels and certain recreational vessels are not allowed (Wenzel et al., 2011). According to a recent study (Toropova et al., 2010), 1.17% of the marine area in the world is designated as protected, which equates to 4.2 million km^2. Unfortunately, enforcement of these laws is hindered by technical challenges and limited resources, which are needed to monitor and respond to violations (Bacher, 2010). These illegal fishing activities often take place in remote areas that are difficult to monitor with conventional methods (e.g., radar). In addition to governments, non-government agencies such as The Nature Conservancy are currently investigating how to monitor and patrol marine protected areas for illegal fishing boats (Lotz et al., 2007, Zurk et al., 2006, Cos, 2014).

Harbor security has become another concern for shallow water environments, especially in light of recent terrorist activity. The events on the morning of October 12, 2000 are a harrowing example. The USS Cole, a United States Navy guided-missile destroyer had tied up to a refueling dock earlier that morning (Burns and Myers, 2000). As the crew was lining up for lunch in the ship's galley, a small boat laden with explosives maneuvered up along the port side (National September 11 Memorial & Museum, 2014). The two occupants of the small boat made friendly gestures to the crew, then detonated their weapon, killing 17 US sailors and injuring at least 40 (Kean et al., 2004). Another example of harbor security being breached took place in 2001 at the Faslane naval base in Scotland. An anti-nuclear activist swam into the harbor, climbed aboard a nuclear submarine, and spray painted graffiti on the side. This was a major breach of security since it was later confirmed that the submarine was carrying nuclear-tipped Trident SLBMs (Staples, 2001). The concern over harbor security is not limited to terrorism. Narcotics smugglers are known to use so-called "go-fast boats" in an attempt to outrun Coast Guard ships. They have also been known to use small submarines, called "narco-subs" (Schmidt and Shanker, 2012) to avoid detection.

Current methods of monitoring harbors for suspicious activity typically involve the use of equipment mounted on the shore. RADAR, infrared imaging, and closed-circuit video are examples. The disadvantage of these devices is that they have a fixed range, so their coverage area is generally limited to a relatively short distance from the shore. They are also affected by inclement weather, such as precipitation, choppy surface conditions, or fog. However, underwater targets pose an entirely different challenge that these methods do not address. One method of detecting underwater targets is with LIDAR. This technology is currently used by the military for submerged mine

detection. LIDAR is a laser-frequency radar, and in this application is aimed directly at the water from an aircraft (Steele et al., 2005). It can penetrate on the order of 10's of meters in certain conditions (Churnside, 2012). However, the cost and limited coverage area of such systems make them impractical for continuous monitoring of large areas. In contrast to electromagnetic-based methods, underwater acoustics is less sensitive to weather and can be deployed in fixed locations far from shore.

In recent years, autonomous underwater vehicles (AUVs) have quickly gained prevalence, and have become essential tools in commerical (Wilkinson and Wadhams, 2008, Fernandes et al., 2003), military, (Hagen et al., 2003), and research sectors (Iwakami et al., 2002). The confluence of a number of technologies concurrently reaching a sufficient level of maturity has facilitated and fueled their evolution, increasing their capabilities and subsequent popularity. This includes, but is not limited to, battery life, navigational sensors, low power computational capabilities, data fusion algorithms, and underwater communications. These vehicles have a variety of practical applications, from oceanography to monitoring fisheries to inspecting hulls of ships. However, there is also growing awareness that AUVs also add a new kind of threat to harbor security (Kessel and Hollett, 2006). Similar types of threats in the past have come from manned underwater vehicles, semi-submersible vehicles, and divers. These threats have been identified by defense and security organizations such as the Coast Guard, Department of Homeland Security, the Navy, and other Department of Defense agencies (CBS News, 2013). Among the dangers are deploying chemicals in sensitive areas, delivery of explosives, deploying mines, smuggling contraband, and adversarial intelligence collection.

Using active sonar for detection of surface or subsurface vessels can be possible, but

the challenge with doing this in shallow water is clutter. When active systems emit a sound, it may reflect not only off the target, but also the seabed, ocean surface, and objects in the water (fish, coral, bubbles, etc.). All these reflections come back to the hydrophone receiver and can drown out the echo from the true target. Passive acoustics does not have this complication, and in fact reflections of the noise emitted by a target from the surface and seabed is positively utilized in this work to help locate the target. The flip side of not having to worry about clutter is that the sound emitted by the target is not something that is under the control of the passive system. Rather, the system must adapt to the radiation characteristics of the target, or even be designed with specific types of targets in mind. In any case, a key aspect of designing a passive sonar system is knowledge of an acoustic signature of potential targets.

While acoustic characterizations have been done for various kinds of boats (Ogden et al., 2011), AUVs are a new class of vehicle that are relatively unstudied from the standpoint of their acoustic characteristics. As such, the first part of this dissertation provides an aspect-dependent characterization of an underway autonomous underwater vehicle. Aspect-dependence refers to the direction relative to the vehicle itself in which sound is radiated. "Underway" is a term that describes a vehicle moving through the water under its own power. This is in contrast to a bollard condition, in which a vehicle is tethered to a fixed object and the propulsion system engaged. An example would be tethering a tugboat to the shore through a force meter in order to measure its tow capacity. However, there are differences in the hydrodynamics of the two conditions that likely affect the radiated acoustics. Ultimately, this analysis lays important groundwork for designing future passive acoustic detection methods for AUVs.

Beyond needing to understand and characterize the different types of threats that are found in shallow ocean environments, there is also a need to determine the most appropriate measurements and data processing methods of passively detecting these targets. If passive sonar is to achieve widespread adoption as a means of monitoring shallow waters, it needs to be practical to use and be able to operate in varied environments. The question that needs to be answered is how should a passive system be designed so that it meets the basic technical objective of effectively localizing targets of interest under different conditions while simultaneously minimizing cost and deployment complexity? Unfortunately, the characteristics of the sound radiated by different types of targets plays a significant role in its design, so a one-size-fits-all approach is impractical. For example, determining the direction of a narrow band signal requires the use of a hydrophone array with a spacing of one half the wavelength. In contrast, broadband signals can be processed using a technique called pulse compression, and can operate with fewer hydrophones that are spaced farther apart. In second part of this work, the latter method is used to locate small boats using only two hydrophones, thus minimizing cost and deployment complexity. It makes use of multipath and the natural contours of the seabed to improve localization accuracy. This is an improvement over existing methods that require additional hydrophones and treat multipath as a nuisance factor (Sutin et al., 2010).

For detecting small boat targets in remote locations for which power resources are limited, a passive sonar system must be designed to operate on a tight power budget. For example, recent designs consist of a hydrophone (an underwater microphone) attached to a battery-powered recording device containing an embedded computer and data storage (Sorensen et al., 2010). Advances in low-power computing capabilities have made new design possibilities such as these possible. Low power sonars are also

now being mounted to AUVs, which themselves have a limited power budget. For example, in one design, an array is towed behind a REMUS-100 AUV (Holmes and Kukulya, 2006). However, it is an open question as to what designs are most advantageous for different applications. The third part of this dissertation focuses on localizing small boats using a small metal-framed hydrophone array mounted to the nose of an AUV. The unique aspect about this array is it is volumetric, meaning that its hydrophones are not confined to a line or a plane, but are distributed in three spatial dimensions.

This volumetric array work essentially builds on earlier work in the related field of geoacoustic inversion (Siderius et al., 2006). It was discovered that beamforming and cross-correlation could be combined so that a drifting vertical line array could essentially "see" the seabed using only noise generated at the surface. As wind or breaking waves at the surface create noise, these waves travel down towards the seabed, then bounce off the seabed and travel back up towards the surface. The vertical line array is able to isolate the signals contained in each of these waves, and through cross-correlation, measure the travel time to the seabed. The formulation for this turned out to be fairly straightforward, and in this work a similar idea is applied to boat noise with a volumetric array. The major difference is that the boat could be far away from the array, so the algorithm is substantially altered to focus the beamformer at the boat. This overall method provides a way for an AUV to localize targets in remote areas, without needing to stay near shore. Many AUVs use satellite communications, such as Iridium, when at the surface to send data back to an operator. Thus, a combination of these capabilities implies a new ability to patrol shallow waters anywhere in the world while providing actionable information in near real time.

It should be noted that the basic idea of using multipath for localization in shallow water has been previously investigated. The novel contributions of this study are centered on methodology and signal processing. Until fairly recently estimating the range of a target using its received multipath (often referred to as "multipath ranging"), focused on the interference pattern created by the multipath arrivals (Weston and Stevens, 1972). This made sense because the broadband noise of a boat can be approximated as a random process, but between multipath arrivals there is coherence, or predictability. This shares some similarity to the way that laser light (which is highly self-coherent) is used to record an interference pattern between two beams on a film to create a hologram. Since the frequency dependent interference pattern changes with target range, this implies that by measuring the interference pattern, an estimate of the range can be obtained (Kapolka, 2008). The key point is that these interference patterns are created by multipath. The methods presented in this work take a different approach by directly measuring the physical directions and time delays between multipath arrivals, as opposed to the interference pattern. A discussion is included that relates the two by showing that the same information is being utilized, but the problem is being approached from a different perspective.

This work has a significant experimental component. Data collected from several deployments in Hawai'i and Italy form the basis for validating the presented methods. It is an important point that this work not only puts forward new methods and techniques, but provides a proof of concept for them. In other words, they are demonstrated to work. Simulations are used as tools to aid in developing an understanding of the expected behavior, but methodologically, a stand-alone simulation is insufficient evidence to support a claim that an algorithm will work in a real situation. This is mainly because simulators make assumptions about the phenomenon being

simulated, and it is non-trivial to accurately predict which assumptions will have an appreciable affect on the performance of a given algorithm. As such, much of what is presented here are based on measurements from experiments that were conducted during the course of this research. Simulations are included throughout this analysis, but mainly for the purpose of improving the understanding the underlying principals of operation.

The remainder of this dissertation is organized into chapters as follows. Chapter 2 reviews the current state of the literature. The first topic is on radiated noise from AUVs, including the methods for making underway measurements along with the published results. Target localization using fixed hydrophone systems and associated data processing methods is the next topic. The passive fathometer constitutes a method for time delay estimation of noise, and it forms the basis for the volumetric array processing algorithm described later in this dissertation, so it is described next. Since this has much in common with manifestations of multipath coherence commonly observed in spectrograms, literature pertaining to the principles of waveguide invariance are then discussed. Chapter 3 develops the mathematical framework around which the new techniques in this dissertation are based. Chapters 4, 5, and 6 describes each of these, analyzes the performance, and provides experimental results. Lastly, Ch. 7 provides a summary of these contributions and outlines possible future directions for this work.

Chapter 2

Background

This chapter reviews the state-of-the-art of passive sonar. Moreover, it identifies outstanding research questions that this dissertation aims to fill. This dissertation does not address all types of targets, all types of equipment, and all environments. Rather, it is focused on targets that are of particular concern for modern law enforcement agencies. Algorithmically, it brings together several processing techniques including the use of multipath for localization, cross-correlation and cross-beam processing.

An application area for the work in this dissertation is that of harbor security. Since this is an important potential use of the techniques developed in this dissertation, specific harbor security threats are identified and discussed in Sec. 2.1. While many types of threats exist, autonomous underwater vehicles (AUVs) are of particular interest because they represent a new type of threat. The open literature does not contain much information on the acoustic signatures of AUVs. This is surprising given that above-water systems are limited for these targets. Further, this information is necessary for developing detection and localization schemes using passive sonar. The first contribution of this dissertation is an analysis of radiated noise from AUVs, and Sec. 2.2 gives background on previous studies.

The use of passive sonar in shallow water is well established, yet many existing methods do not take advantage of all the information carried by the acoustic signal and contributed by the environment. Further, the explosion in computing power over the past decades has opened the door to new processing modalities that were once considered impractical. Traditional methods of passive target localization are discussed in Sec. 2.3. This begins with the category of methods that rely on the principle of time differences of arrival (TDOA), discussed in Sec. 2.3.1. New systems currently being developed for harbors are based on this principle (Sutin et al., 2010). Later algorithms expanded on this by incorporating multipath, which is discussed in Sec. 2.3.2 focusing mainly on recent work done with marine mammals.

The last contribution of this dissertation is on using small, low-power passive systems that can be attached to mobile platforms. The basis for this new technique comes from another technique known as the passive fathometer. In that formulation, an array is used to analyze a noise field instead of a single target. This foundational work is discussed in Sec. 2.4. The passive fathometer itself is discussed in Sec. 2.4.2.

Using multipath to determine target range using arrays is an idea that was studied extensively using a concept called matched field processing (MFP), discussed in Sec. 2.4.1. A closely related concept that utilizes multipath, and can be leveraged for localization, is waveguide invariance, which is discussed in Sec. 2.5.

2.1 Harbor Security Threats

Harbor security has been a technology driver for systems and methods aimed at detecting and tracking shallow water vehicles. Since the October 2000 bombing of the USS Cole (Burns and Myers, 2000), the concern over untracked small boats has increased. However, overt hostile action is generally less likely than other more mundane illegal activities. Small boats and small submarines have also been used by drug traffickers and smugglers of elicit or illegal goods in an effort to evade law enforcement. Divers also present threats in both scenarios. In 2001, an anti-nuclear activist swam into Faslane naval base in Scotland and spray painted graffiti on the side of a nuclear submarine carrying trident SLBMs (Staples, 2001).

Harbors face a variety of threats, and correspondingly a variety of remote sensing modalities have been applied to confront these threats, such as shore-based RADAR, LIDAR, optical, and thermal imaging. Major defense contractors spend considerable resources building integrated surveillance and monitoring systems based on these tools (Nor, 2011). Bruno et al. (2010) has investigated fusing above-ground with below-surface systems in an attempt to create an integrated detection system. The drawback of above-water methods is their limited range, and degraded performance in inclement weather For example, Steele et al. (2005) put forward the idea of using a RADAR-LIDAR as a detection methodology for divers, but the range of these systems is limited (Waldron and Mullen, 2009). They also have high overall cost per distance of coverage and the continuous need for an operator (i.e. personnel watching a video feed).

Active sonar has the disadvantage that high intensity acoustic signals may make it

relatively easy for intruders to detect the presence and location of any active acoustic security systems, and find ways to subvert them. Multipath propagation, sound absorption at high frequencies and interference from benign objects with similar acoustic scattering characteristics further limit their range and robustness. There is also the problem of a high false alarm rate, in which any object with similar target scattering properties, such as a large fish or seal mistaken as a diver, will trigger the system. Furthermore, effects on marine life from anthropogenic noise are not fully understood, and is an active area of research. These limitations have driven an interest in passive acoustic solutions.

The challenge with using passive acoustics is that each class of target can have significantly different acoustic signatures. As such, it is unlikely that a single one-size-fits-all processing approach will suffice. For example, boats and divers radiate broadband noise, and while both have high source levels, this is especially true of boats due to noise from the propeller (Vagle and Burch, 2005) and wake (Magliozzi et al., 1992). Engine noise couples into the water through the hull and propeller producing harmonically-spaced narrowband features (Ogden et al., 2011), that are not radiated from a diver. A detailed reference on the mechanics of boat-generated noise can be found in Ross (1976). Remotely operated vehicles (ROVs) and autonomous underwater vehicles (AUVs) mostly use motor driven propulsion systems, and move at much slower speeds. As such, the source levels are lower, with some vehicles having mainly tonal elements with few harmonics, and others producing broadband noise (Griffiths et al., 2001).

Several researchers at the Maritime Security Laboratory at the Stevens Institute of Technology (Hoboken, NJ) have studied technical solutions to harbor security,

including passive sonar. Borowski et al. (2008) investigated the detection distance of divers in terms of source level, transmission loss of acoustic signals, and ambient noise. In that study, they apply the passive sonar equation to estimate range to target. Stolkin et al. (2006) used used a similar approach with a single hydrophone to detect and classify divers and small boats.

Stevens has also deployed multiple hydrophone systems. Bunin et al. (2008) used cross-correlation of boat noise to obtain bearing estimates that were then fed to a data fusion system consisting of several different types of sensors. They also studied diver detection in an experiment involving a diver with a hand-held source. More recent studies have investigated small boat localization in the context of more complex tracking algorithms (Ou et al., 2011, Sherpa, 2012).

Only a few years ago, Sutin et al. (2010) described a system they designed called the Stevens Passive Acoustic Detection System (SPADES). This is a four-hydrophone system in a hub-and-spoke configuration that is deployed on the seabed with the hydrophones on the end of each spoke, and an electronics pod at the hub. A similar three-hydrophone system was described in Sutin et al. (2013). They used cross-correlation with pre-whitening to obtain bearings to a small boat target. By using two pairs of phones, the hyperbolic ambiguity from a single pair was resolved into a single absolute bearing. The spacing between the hydrophones was on the order of 10's of meters, and each hydrophone was attached using a flexible cable. This presented the problem of the hydrophones not being well localized as they most likely move during lowering to the seabed due to current or swell moving the boat. To compensate for this, they used an acoustic positioning system which involved broadcasting a waveforms from a boat-deployed source while capturing GPS coordinate data. They

then used a optimizer based on a least-mean-square-root metric to minimize the deviation of measured time delays with coordinate-derived ones. This sensitivity to potential errors in the array geometry and subsequent deployment complexity reduces the overall utility of such an approach. Specifically, the planar geometry of the array requires a rigid frame to maintain the relative spacing between hydrophones. However, a rigid frame of this size can be difficult to deploy. Further, any deformities in the frame that may occur during deployment will affect the measured time delays and will therefore have an impact on the localization performance.

Fillinger et al. (2010) outlined a joint effort betweenn the Netherlands Organization for Applied Scientific Research (TNO), and Stevens Institute of Technology to compare technologies for harbor security. Their aim was to compare SPADES (Stevens' system) to Delphinus (TNO's system), which is a passive towed line array. The interesting take-away from that paper was a technique they designed in which they drove a beamformer beyond the design frequency resulting in aliasing. Using cross-correlation of two phones on the array, they were able to disambiguate between aliases.

Stevens has also used the SPADES system to look at the detection of envelope modulation on noise (DEMON) spectrum for small boat classification (Chung et al., 2011). The DEMON spectrum has advantages for analyzing ship signatures, and is computed by taking the Fourier transform of the envelope of a band-pass filtered signal. In that paper, they computed the DEMON spectrum from cross-correlations of hydrophones at an offset corresponding to the vessel of interest.

All these studies involving cross-correlation of boat or diver noise only resolve target bearing. The studies that do consider target range only measure received level, using educated guesses for the various terms of the passive sonar equation. Also, none

of them discuss multipath propagation, or consider its geometric implications. In fact, as will be discussed later in this dissertation, multipath propagation can add additional features to the correlogram which complicate the TDOA measurements.

2.2 Radiated Noise from AUVs

Detection of AUVs poses a different detection challenge than boats or divers since no part of their hull extends above water and no surface wake is generated. Passive acoustic localization methods may be more practical than active methods in terms of cost and impact on the environment, but they require an accurate characterization of radiated noise from the target in terms of spectral content as well as aspect-dependent source levels.

The literature survey in this section outlines the work that has been done by others in the field in terms of reporting the radiated noise of AUVs. Most AUVs are driven by electromagnetic propulsion systems, which can be the largest contributor of radiated noise. Several studies have been performed involving the collection of acoustic data generated from AUVs and analyzed the spectral content of their acoustic signatures. Part of the contributions made in this dissertation are comprised of work done on characterizing the emissions from a REMUS-100 AUV (Gebbie et al., 2012b, 2011a,b).

AUVs come in all shapes and sizes. Fernandes et al. (2003) gave a list of them (circa 2003) showing 15 models in a certain size class and analyzed their applications to fisheries, for tasks such as fish stock assessment, as well as broader ecological studies. That number has undoubtedly grown since then.

Since so may AUVs exist, their acoustic signatures are also varied. Most AUVs are electrically powered, so as a class of vessel, they will differ in the overall structure of their signature from small boats. Although they travel underwater, their methods of locomotion will cause them to have a significantly different signature than divers.

However, AUVs can be further broken into several categories, based on the nature of their propulsion systems. Gliders do not require a propeller. Their method of propulsion is based a ballast pump which causes the vehicle to rise and fall in the water. Hydrodynamic elements of its design give it horizontal motion as it rises and falls.

Propeller-driven platforms also have several subcategories. Open-blade propellers are the most common, which are found on most recreational small boats. Pump-jets are another design that operate by placing the propeller inside of vehicle (in this configuration it is called an impeller). Examples are personal water craft (PWCs), which have gained popularity in recent years. In this design, water is drawn through an inlet, passes through the impeller, and is expelled through a jet creating thrust. Pump-jets have been observed to be quieter underwater despite being louder above water (Erbe, 2013). They are also more efficient, but only at higher engine speeds. It is perhaps for this reason that many entry-level battery-driven AUVs use open-bladed propellers in lieu of pump jets.

In the literature, a few studies have been conducted on AUV signatures. In 2003 Cuschieri and Frandsen (2003) gave a presentation on acoustic signature of an Ocean Explorer AUV using data collected in an open-water environment. They compared this to measurements in a test tank. In 2009, Carey et al. (2009), presented on the applicability of a small autonomous vehicle towed array system to ocean acoustic measurements and signal processing. Some acoustic measurements of self noise of the AUV were recorded, but this was limited by flow noise. A similar experiment using a towed array a was also outlined in Holmes and Kukulya (2006). Zimmerman et al. (2005) performed an analysis of different methods of reducing self noise of an

Odyssey IIb AUV. This vehicle used a ducted fan design for its propulsion. Changing the motor gear ratios, and trapezoidal drive with a sine drive ended up reducing noise levels by 20-50 dB. Griffiths et al. (2001) analyzed the acoustic emissions of an Autosub AUV and found that it produced such low source levels that it was only detectable to within a few meters of their passive sonar equipment. In one instance, the AUV drove straight into the hydrophone, but the hydrophone didn't pick anything up until the moment of impact. It was also hypothesized that propeller cavitation was the primary noise producer of this AUV since broadband emissions were detected up to roughly 4 kHz, with most acoustic energy concentrated in the lower band up to roughly 400 Hz. However, the Autosub may have been an anomalously quiet vehicle since other studies have since reported that higher source levels due to motor-borne vibrations. Also, the lack of ground truth measurements of the actual location of the AUV precluded calculation of source levels and beam patterns.

In contrast with broadband cavitation noise, these motor vibrations have narrow-band characteristics. This was reported by Holmes et al. (2010). This study gave an extensive noise audit of REMUS-100 AUV. In it, several AUVs were compared based on data available in the literature, including the potentially anomalous data reported by Griffiths et al. (2001). Most levels concentrated in the 100-120 dB range above 250 Hz. In Holmes's dissertation, (Holmes, 2007), a detailed analysis was given for sources of radiated noise. Radiation from wetted portions of an AUV's motor stator, and Helmholtz resonance of the fluid connection gap between the rotor and stator can lead to monopole radiation patterns, whereas alternating thrust (caused by non-sinusoidal motor currents, for example) can produce hull vibrations that radiate in a dipole manner.

Several studies have taken a more rigorous approach to measuring source levels and beam patterns by placing a REMUS-100 AUV in a bollard condition inside a tank (Holmes and Kukulya, 2006, Holmes et al., 2005). A bollard condition is an arrangement in which the vehicle under test is tethered in place and its propulsion system activated. This differs from an "underway" condition in which the vehicle is moving through the water. The hydrodynamics and engine torque of each are different, and this likely has an effect on the radiated emissions. However, aside from the references mentioned earlier, limited published information exists on radiated noise characteristics of AUVs in this condition.

2.3 Time Delay Localization

This section discusses work that has been done on localization of various targets. Early localization methods, both in electromagnetics as well as acoustics, employed a principal called time-difference of arrivals (TDOA). Essentially, if two receivers detect the same signal but at different times, that time delay limits the source to a set of physical locations that has the shape of a hyperbola. This class of "hyperbolic fixing" algorithms is described in Sec. 2.3.1.

In its simplest form, multipath effects are largely ignored by hyperbolic fixing, and in fact are often treated as a nuisance factor. In shallow water, multipath is often a stable phenomenon that depends on the target location. That is, multipath is primarily a function of the relative positions of the source and receiver within the ocean waveguide. Section 2.3.2 discusses the work that has been done in marine mammal localization, which is heavily based on the use of multipath. Certain species of marine mammals use signals that are impulsive in nature (Nosal and Frazer, 2006). Nevertheless, the principles behind how the multipath is processed for this kind of target shares some similarity to this work, but a different approach is taken to processing the data.

In contrast to impulsive whale noises, boat noise is is continuous in time. Section 2.3.3 describes the work that has been done in localizing broadband noise sources.

2.3.1 Hyperbolic Fixing / Multilateration

Hyperbolic fixing is a localization technique that uses signal timing to estimate position. The idea requires two or more known points and one unknown point. The goal is to estimate the location of the unknown point relative to the known points. Waves are passed between known points and the unknown point, but due to physical constraints, the quantity that is measured are the time delays between pairs of known points and the unknown point. If only two known points exist, a single time delay will be measured. Considering 2-D space to begin with, we can conclude that the unknown point lies somewhere on a hyperbola for which the two known points are the foci. In 3-D space, this locus of points is a hyperboloid. Note this rests on the assumption that waves travel at the same speed everywhere in space. Including additional known points creates additional pairs of known points, each of which will trace out its own hyperboloid. The unknown point lies on the surface of all the hyperboloids, so the "solution" corresponds to the point or points in which all the hyperboloids intersect. Hyperbolic fixing is a general approach to a general problem, and is often referred to as time difference of arrival, or TDOA. In the passive sonar problems discussed here, the known points are the receiver elements, and the unknown point is the target.

When hyperbolic fixing is used for navigation, it is often referred to as multilateration. One of the earliest examples of multilateration was the GEE and LORAN systems (Launer, 2009). The latter was developed during World War II by the United States so that naval ships could determine their location (Staff, 1946). These systems consisted of fixed broadcast stations distributed along the coast (see Fig. 2.1 for an historical LORAN-C navigation chart). Each station emits a unique radio pulse that is picked up by a receiver on a vessel. The receiver then measures the time-differences

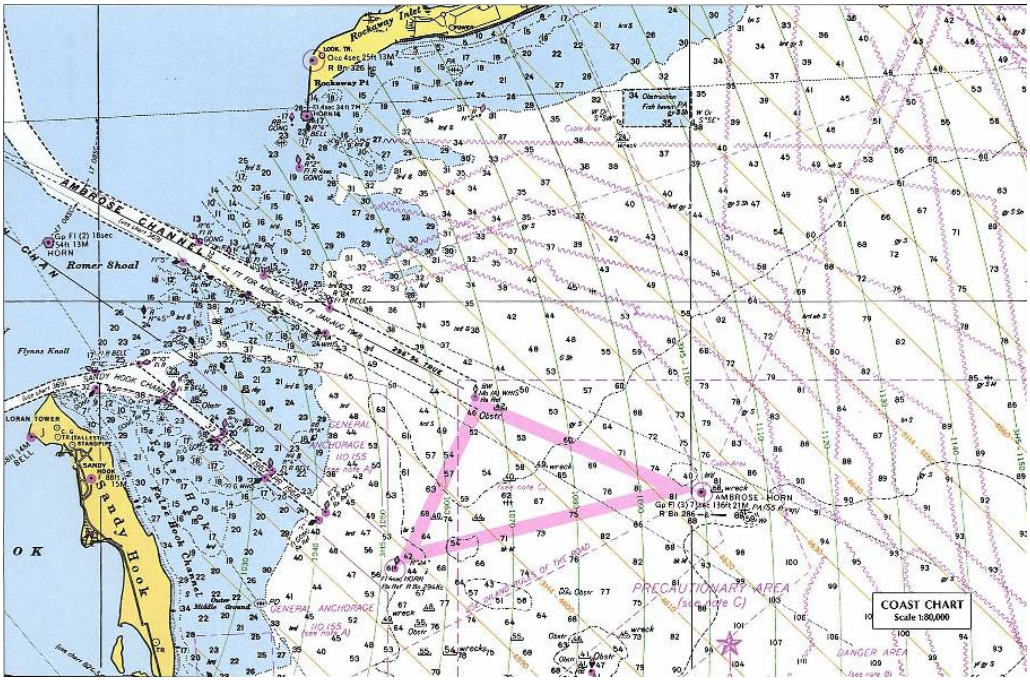


Figure 2.1: A LORAN nautical chart of New York Harbor (source: Wikimedia Commons, freely licensed media)

between pulses originating from different stations. These time-differences correspond to hyperbolic location lines, on which the receiver must lie.

Recently, in the context of localizing marine mammals acoustically, Spiesberger and Wahlberg (2002) introduced the term "isodiachron" using the Greek words "iso" for same, "dia" for difference, and "chron" for time. This defines any surface that would conceptually be formed in two or three dimensions that satisfy a given TDOA measurement between two fixed points. Inhomogeneity in the sound speed, or boundary reflections cause isodiachrons to have non-regular shape (Spiesberger, 2004).

The efficiency of hyperbolic localization has been studied for the better part of a century. For example, Marchand (1964) developed a mathematical model for the error distribution of the best estimate position of a vehicle determined form a set of TDOA

measurements. Schmidt (1972) expanded the hyperbolic ambiguity to generalized lines, when measurements are made on the surface of a sphere. This naturally had significance for navigation purposes since a non-negligible amount of earth curvature may exist between a ship and ground stations. Lee (1975) developed a procedure for determining the accuracy of hyperbolic multilateration systems. This expanded on an earlier methodology called the geometric dilution of precision (GDOP). This was later reframed using a Cramér-Rao lower bounds (CRLB) methodology by Chan and Ho (1994) which placed mathematical limits on the lowest possible error for a given system. Efforts have been made to improve accuracy, as well as quantify errors. Location errors due to insufficient numbers of fixed points was studied by Spiesberger (2001) in the context of marine mammal localization.

In the context of passive sonar, multilateration techniques are also based on measuring TDOAs. This often involves estimating the time-lag between two hydrophones, or possibly delayed replicas of a signal from a single hydrophone (i.e. multipath "echoes"). The topic of estimating time-delays from signals has been extensively covered by Carter (1993), who has compiled an anthology of works on the topic.

2.3.2 Marine Mammal Localization

In 2008, the US Supreme Court heard a case in which the Natural Resources Defense Council sued the US Navy over the latter's use of mid-frequency active (MFA) sonar during anti-submarine warfare exercises (Supreme Court, 2008). At issue was the assertion that high source levels harm marine mammals. There have been several instances over the past few decades in which use the of the MFA has correlated with beachings of marine mammals, however no concrete linkage has been made. The

Supreme Court ruled in favor of the US Navy partly based on this fact, and partly based on the judgement that the potential harm to marine mammals is outweighed by national security interests.

Research into behaviors of marine mammal is an active area of research, partly from a basic science point-of-view on the part of the marine biology community, but also from a public policy point-of-view in an effort to determine the extent to which anthropogenic noise actually affects these animals. Nevertheless, understanding behavioral patterns relates to tracking their movements. As such, passive sonar has been studied extensively in this area as a means to ascertain animal movements. The work that has been done to support these capabilities shares many principals with the techniques presented in this dissertation, which is why marine mammal localization is discussed here.

Marine mammals in the biological order cetacea (specifically odontocetes) use active sonar. They emit broadband clicks for echolocation while performing different kinds of behaviors such as foraging, hunting, and communication (Wood and Evans, 1980). Whale clicks are short duration acoustic events that pack a significant amount of acoustic energy into a short time span. In signal processing, these kinds of events are described as "impulsive." This is important because much of signal processing is built on the concept of a delta function, which can be thought of as an idealized impulse. The waveform of a delta function spans an infinitely short interval, is infinitely tall, and is zero everywhere else. It has important mathematical properties in that it can be used to characterize the propagation of waves between a source and receiver through a specific environment. It is for this reason that a whale's delta-function-like waveform has significant implications for method used to localize it, particularly

methods that employ multipath. The remainder of this section discusses different methods that have been utilized to localize marine mammals that make broadband clicks.

Humans inherently know that the closer a sound source is, the louder it will be perceived. This intuition is formalized in the passive sonar equation, which relates incoherent received levels. For a single receiver, a simplified sonar equation might look like

$$\mathrm{RL} = \mathrm{SL} - \mathrm{TL} \tag{2.1}$$

in which SL is the source level, TL is the transmission loss, and RL is the received level. All levels are in decibels (dB). The received level is measured, the source level is estimated, and modeling tools predict the transmission loss at different ranges. Humans can actually perform all these functions cognitively to estimate the range. A computer approaches the problem by adjusting the range in order to satisfy the equation, thus yielding the answer. The challenge with using the passive sonar equation for whale ranging is it requires guesswork for the source level; and errors in estimating this parameter translate to errors in range (Cummings and Holliday, 1987, Watkins et al., 1987).

In parallel with source-level ranging, using time-differences between arrivals (TDOA) on multiple hydrophones was employed to estimate bearing. That means determining the direction in which a whale is located, without necessarily knowing anything about its range. Conceptually, this is same principle as multilateration described in the previous section. An early work that employed a four-element volumetric hydrophone array was done by Watkins and Schevill (1972). Magyar et al. (1978) applied this to

entire grids of hydrophones.

In recent years, the marine mammal community began drawing more heavily on the body of work developed for using multipath to localize other kinds of underwater targets. An early comparison of the source-level method with multipath ranging was given by Cato (1998). Later, in a study on Hawai'ian spinner dolphins, Aubauer et al. (2000) used multipath ranging from both surface and bottom bounces on a single hydrophone for range and depth localization. They point out that one advantage of this is that only a simple environmental model is needed, consisting of the depth of the bottom and hydrophone. Thode (2004) analyzed multipath data from two hydrophones towed behind a boat at different ranges and depths. By plotting spectrograms of the two phones side-by-side, time differences between multipath arrivals for each hydrophone, and across hydrophones, could be manually read off. Skarsoulis et al. (2004) considered the effects of a stratified seabed and refractive sound speed profile (SSP) using a simulation. It was determined that adding separation between hydrophones increases accuracy. Laplanche et al. (2005) showed that as a sperm whale changed its geometric pitch (i.e. whether the animal was facing the surface or seabed), so would the different multipath arrivals from the surface and seabed change intensity. From this, they used a single hydrophone to localize the whale and estimate its pitch by looking at the multipath energy distribution. It is known that many whales have focused beam patterns, which is the principal behind this approach. Laplanche (2007) later expanded on that work to only require multipath arrivals from the sea surface.

For multiple hydrophone depoloyments, Nosal and Frazer (2006) used multipath ranging from five widely distributed hydrophones to obtain the full 3-D location of the

animal. This worked by finding intersections of rings of ambiguity from each hydrophone. Nosal and Frazer (2007) then applied the same method, except they used a model beam pattern of the whale to predict its orientation based on source level differences between phones. Miller and Dawson (2009) used several time-synced and coordinate-tracked drifting buoys to record whale sounds. By fusing the TDOA between each pair, as well as the multipath-with-direct TDOAs on each hydrophone, they were able to obtain a 3-D position estimate of the animal.

Perhaps the closest related work to the one presented in this dissertation for small boat localization was put forward by Tiemann et al. (2006). They demonstrated that only a single hydrophone was needed to produce a full 3-D localization of a sperm whale by comparing measured and modeled time differences between multipath arrivals using a raytracer. In that work, multipath arrivals were determined directly from the received time-series (after some filtering) by time-gating each whale click and stacking the resulting snapshots with the direct arrivals vertically aligned. Viewed in this manner, all arrivals appeared as a time-evolving striation lines with multipath striations curving over time as the animal maneuvered relative to the hydrophone. Variations in bathymetry were opportunistically utilized to provide range, depth, and azimuthal discrimination by leveraging the fact that the relative arrival time of bottom-interacting multipath rays are functionally dependent on the bathymetry along each azimuthal radial. Similar work was done by Mouy et al. (2012) for localizing walruses.

There are a few key differences between Tiemann et al. (2006) and the work presented in this dissertation. Multipath arrivals were determined directly from the received time-series (after some filtering) by time-gating each whale click and stack-

ing the resulting snapshots with the direct arrivals vertically aligned. Viewed in this manner, all arrivals appeared as a time-evolving striation lines with multipath striations curving over time as the animal maneuvered relative to the hydrophone. Variations in bathymetry were opportunistically utilized to provide range, depth, and azimuthal discrimination by leveraging the fact that the relative arrival time of bottom-interacting multipath rays are functionally dependent on the bathymetry along each azimuthal radial. Whale clicks are impulsive acoustic events which allows for clean separation of mulitpath arrivals in the received time-series. As such, this technique does not hold for continuous broadband noise sources, such as that produced by a small boat. Further, the surface reflection of a whale click can be exploited (typically the strongest reflected multipath arrival) which is not the case for surface ships. Also, the inversion approach adopted by Tiemann et al. (2006) was based on a scoring mechanism whereby each click event was assigned to a range-depth-azimuth coordinate bin. Each assignment was done in a a winner-takes-all scheme, and patterns in the data were realized by aggregating over several click events. The approach taken in this dissertation is to apply a functional projection into the entire spatial domain producing an ambiguity surface for *each* snapshot.

2.3.3 Noise Source Localization

Unlike the whale species that produce clicks, propeller-driven vehicles produce continuous noise. In a shallow water waveguide that supports strong multipath characteristics, a whale click will show up in the received waveform as a series of impulses having offsets corresponding to the travel time of each multipath ray. For a noise-producing source, such as a boat, looking at the waveform is useless for discerning anything

about the multipath structure. Noise, by definition, is a sequence of random numbers, or samples. In other words, the value at any given time is a random variable. If the source produces a random sequence of samples, then the received waveform will be the sum of delayed copies of this sequence. To be precise, the delayed sequence may not exactly match the original sequence due to propagation effects, but the delayed sequence will be "coherent" with respect to the source. In other words, there is a nonzero degree of predictability between them. Even though the received waveform is the sum of coherent versions of the same sequence, it still appears as a random sequence of samples.

The correlation operation is a very useful tool for analyzing this kind of sequence. Auto-correlation is applied to a single channel, whereas cross-correlation is applied to two different channels. The auto-correlation can thus be thought of as a specialization of the cross-correlation in which the same channel is used both times. Cross-correlation is applied not to absolute points in time, but rather to relative spans of time. In the case of boat noise for which a direct arrival and a sequence of multipath arrivals are being recorded, the cross-correlation operation can expose the coherence between each pair of arrivals. This operates by "compressing" the broadband noise into a broadband impulse, hence the name "pulse compression." This will be described in greater detail in the next chapter. However, it is important to point out that this is the underlying mechanism that is used extensively for localization of noise-type sources. Pulse compression, as it is used here, is formally based on the generalized correlation algorithm from Carter and Knapp (1976). Knapp and Carter (1976) showed that the Cramér-Rao lower bound on time delay estimation is reached by the generalized cross correlation algorithm.

The advantages of using the correlations of multipath versus only the correlation of direct arrivals have been studied extensively over the past several decades. These investigations have mostly been formulated by deriving the optimal theoretical accuracy of source location using the Cramér-Rao lower bound (CRLB) as a metric. The CRLB is the lowest achievable variance of an unbiased estimator, and works by considering how much information about the unknown parameters (i.e. source location) is contained in the measured data (i.e. recorded source noise including multipath). These bounds hold under certain conditions such as no relative target motion, stochastic Gaussian signals, and additive uncorrelated Gaussian noise (Rendas and Moura, 1991). It also assumes that all parameters (i.e. depth of seabed reflections, sound speeds of the medium) are fixed and known *a priori*.

There has been a multitude of work in this area. Target localization using multipath obtained from an autocorrelator has been studied by Ianniello (1986). For multiple hydrophones, the CRLB of the target's range and depth based on pulse compression of radiated noise from the target was given by Friedlander (1988). Two paths (one direct and one reflected) from each of two vertically aligned receivers to a submerged source were considered. Bandwidth and signal-to-noise ratio (SNR) determined the CRLB for target location. A similar study was conducted by Rendas and Moura (1991). Abel and Lashkari (1987) investigated using multipath information from multiple hydrophones in different orientations. Yuan and Salt (1993) and Yuan et al. (2000) used simulations to try to characterize the maximum possible efficiency from pre-extracted time delay measurements between vertical line array elements. Lee et al. (2002) used bearing angles and time-difference of arrival between two multipaths on a bottom-mounted array to perform underwater source localization. Cramér-Rao analysis has also been applied to matched field processing by Hursky et al. (2004).

In 1968, Van Trees (2001a) expanded the CRLB to include non-deterministic parameters, in a formulation that is often referred to as the Bayesian CRLB. The term "Bayesian" is used to indicate that prior information about the parameters are being accounted for. This was the basis for the study by Hamilton and Schultheiss (1993) that examined the performance impact of having imperfect knowledge of the bathymetry by treating the depth of the seabed reflections as normally-distributed random variables. Later studies have expanded on this by applying the CRLB or Bayesian CRLB to gauge the performance limits of trackers that reduce location estimation variance by modeling the Markov chain nature of measurements observed from moving targets (Dauwels, 2005, Lehmann and Williamson, 2007).

Refraction is a difficult phenomenon to model with the CRLB analysis due to the need to obtain partial derivatives of the cross spectrum with respect to parameters that define the refraction (i.e. sound speed profile slope). Franchi and Jacobson (1972) and Rendas and Moura (1990)are two studies that have examined this issue.

2.4 Beamforming and Multipath

The previous section discussed multipath arrivals at individual hydrophones. This section shifts focus onto work that has been done using array beamforming to analyze multipath. Why use arrays? Most single hydrophone elements are reasonably approximated as being omnidirectional, meaning they are equally sensitive to noise arriving from any direction. Hydrophone arrays comprised of omnidirectional sensors can increase sensitivity in some directions while decreasing it in others. This is called spatial filtering. Vector sensors are single elements that measure particle velocity on three independent axes. However, the focus of this work is on omnidirectional hydrophones since they are more common, simpler in design, and less costly in general. For localization built on time delays, the basic instrument of measurement is a single omnidirectional hydrophone which provides no directionality. In the array beamforming context, the basic instrument is a set of synchronized omnidirectional hydrophones that work together to provide directional information.

Array beamforming operates on the idea that closely spaced receivers will observe the acoustic field very similarly, but from slightly different vantage points. These slight differences are utilized to determine the direction in which waves are arriving. The concept is built on the assumption that a source located somewhere in space will produce coherent responses on the individual elements of the array. Coherency is an abstract idea but has a concrete definition [see Carter (1993)]. Two signals are coherent with respect to one another if some information from one signal is shared with the other signal. In other words, if by observing one signal it is possible to perfectly predict the waveform of other signal through a deterministic function, then the signals are perfectly coherent. Note that the two signals need not be identical, they just need

to be deterministically related. Coherency is a measure of the predictability between two signals, and varies from zero (totally unpredictable) to one (totally predictable). Two signals that share no information, and thus do not have any shared predictability, are called incoherent. Pairs of multipath arrivals at a single hydrophone are mostly coherent. Also considering that two closely spaced hydrophones in a field share a high degree of spatial coherence, what methods exist for leveraging this property for processing element-level data from an array? Two such methods are discussed here: matched field processing and the passive fathometer.

Matched field processing (MFP) uses array beamforming to exploit the spatial coherence between hydrophones while simultaneously leveraging the coherence that exists between multipath arrivals. It is typically used for localizing underwater targets, and is described in Sec. 2.4.1. One advantage of it is it can operate on narrow-band data. A disadvantage, however, is that it requires modeling the full field, which can be sensitive to environmental parameters. This is because by modeling the full field, it is effectively modeling the coherent sum of all multipath arrivals.

The passive fathometer was borne out of the field of geoacoustic inversion, and also uses array beamforming to analyze coherence between multipath arrivals. Geoacoustic inversion is the study of the seabed using sound. The passive fathometer is a recently developed method that uses ocean ambient noise as a sound source. The way in which it works represents a new method of exploiting coherence between multipath. Whereas MFP focuses the beamformer on all multipath arrivals simultaneously, the passive fathometer used beamforming to isolate and cross-correlate beams containing individual arrivals. This is a fundamentally different processing modality than MFP, and is described in Sec. 2.4.2. The third contribution of this dissertation is based

on applying the same basic processing modality to noise data that originates from a distant source using a three dimensional array, thus applying principals from both modalities.

2.4.1 Matched Field Processing

As described in Sec. 2.4, Matched field processing (MFP) attempts to utilize the coherence between multipath for estimating the location of a source. It is inherently a narrowband technique, meaning that it can be used to localize sources that only emit distinct tones. The field measured at a receiver depends on several factors: the water, the seabed, and the geometry. MFP takes the approach of directly modeling the field, and so it can be used not only to estimate the source location, but the environmental parameters as well. In the source localization context, usually some model of the environment is assumed, so the only "unknown" is the source location. The challenge comes when the model chosen for the environment does not sufficiently represent the actual acoustic propagation through the waveguide.

The hypothesized source location is called a replica, and the model of the acoustic field from that point is called a forward model. For multiple receivers, the modeled pressure values for a given replica are called a replica vector. Note that the modeled pressure values consist of the coherent sum of all multipath arrivals. The replica vector is then matched against the measured pressure field values, much in same way that a manifold vector from a beamformer is used to match the measured pressure field values. This involves conjugating the replica vector, multiplying it to the spectral pressure values, and then summing the results. As the replica position approaches the true source position, and provided the waveguide is accurately modeled, the output

of the matched field beamformer will reach a maxima. This was first noted by Clay (1966) for a general waveguide. It was later applied to an underwater acoustics context with real ocean propagation considerations by Bucker (1976) and later by Heitmeyer et al. (1985) and is generally referred to today as "conventional MFP."

Following the success of adaptive beamforming, the same principles were later applied to matched field replica vectors by Fizell and Wales (1985) and Lee et al. (1999). In a related work, source motion compensation was studied by Zurk et al. (1999) and Zurk et al. (2003). One of the challenges to applying MFP is its sensitivity to environmental mismatch (Huang et al., 2006). Recently, Bayesian methdologies have been applied in an attempt to shore up the robustness of MFP by Huang et al. (2006) and Dosso and Dettmer (2011). A good summary of matched field processing is found in works Tolstoy (1993), Baggeroer et al. (1993), Porter and Tolstoy (1994), Mecklenbrauker (2001). A comparison of using time-delays of multipath arrivals with narrowband matched field techniques was done by Blanc-Benon (1995) and Shaoyin and Yudong (1998). MFP has also been adapted for use with broadband signals (Soares and Jesus, 2003, Michalopoulou and Porter, 1996). Another variation was developed in two separate papers (Yang and Yates, 1998, Yang et al., 1998) in which MFP was applied in beam space.

MFP has documented shortcomings in the underwater acoustics community in recent years due to sensitivity to environmental modeling errors. More recent works have instead focused on methods that have less sensitivity. An example from the geoacoustic community is the passive fathometer, which is described in the next section.

2.4.2 Passive Fathometer

The ambient noise field in the ocean has increasingly become an object of study in recent years. This field is created mainly by wind and breaking waves at the surface. A hydrophone observes this spatially-distributed noise in addition to reflections of this noise from the seabed. It is this interaction with the seabed that has fueled much recent interest, such as Siderius and Harrison (2004), Siderius et al. (2013), Muzi and Siderius (2013), Quijano et al. (2013), Walker (2012), Quijano et al. (2012). When noise is generated at the surface, it radiates a random signal in all directions. The waves that arrive at the seabed are distributed over a wide range of incidence angles and a wide band of frequencies. The reflections of these waves thus contain information about the reflection coefficient of the seabed. The reflection coefficient is a useful quantity to know since it helps describe acoustic propagation through the waveguide.

An interesting feature of the ambient noise field is that for most common ocean environments for which the bathymetry is roughly flat, it is azimuthally symmetric in power. If it were possible to measure the power of the noise field as a function of all possible 3-D directions, the measured power would be the same along all azimuthal angles for a single elevation angle. This is because the waveguide is essentially the same along each radial direction (starting at the hydrophone and going out in range).

This azimuthal symmetry is efficiently measured with a vertical line array, which has a conical beam pattern. By measuring the power arriving at different elevation angles, it is possible to estimate the amount of energy lost in reflections from the seabed. This, in turn, gives an estimate of the magnitude of the reflection coefficient as a

function of frequency and grazing angle, as described by Harrison and Simons (2002). It is important to note that by comparing the energy arriving at different elevation angles does not consider coherence between multipath. In fact, by only considering beam power, much of the coherence information is purposely lost.

In 2006, an algorithm dubbed the "passive fathometer" was introduced by Siderius et al. (2006) for extracting the multipath coherence from the ambient noise field using a vertical line array. This technique was later refined by Gerstoft et al. (2008), Harrison and Siderius (2008), and Means and Siderius (2009). In this formulation, the two-way travel time from the array to the seabed was measured by beamforming at the end fire directions of the vertical line array to isolate the downward traveling noise and the multipath reflection from the seabed. Later, Siderius et al. (2010) showed that switching to adaptive beam forming (Capon, 1969) improved time delay estimated by reducing interference from non-vertically-traveling waves. Additional modeling work was done by Traer et al. (2011), and in Traer and Gerstoft (2011) described a method of reducing the averaging times. Some details surrounding an anomalous sign was examined in Harrison (2009) and Traer et al. (2009). Other related work showed improvement similar to using adaptive beamforming through the use of "super gain" (Siderius, 2012). Super gain refers to driving an array just past its endfire look directions.

The passive fathometer is a methodology of finding coherent echoes from the seabed. However, coherent echoes of downward traveling noise should also come from any objects directly under the array. Harrison (2008) demonstrated this was indeed the case. Leroy et al. (2012) has also examined extraction of coherent wavefronts using array processing techniques.

2.5 Multipath Coherence and Waveguide Invariance

The passive fathometer described in the previous section is a method for exposing the inherent coherence between multipath arrivals. However, this technique was introduced relatively recently, and was designed to operate on surface-generated noise. An older class of techniques that operate on broadband noise from distant, more point-like, sources have centered around a phenomenon known as waveguide invariance. Waveguide invariance is relevant to this work since waveguide invariance is essentially a manifestation of multipath coherence.

The following example describes the concept of waveguide invariance. Take a noise-producing source, such as a boat, or underwater vehicle, that is moving past a hydrophone. Both are in a waveguide that supports multipath, and the noise is detected by the hydrophone through multiple arrivals. In this situation, the spectrogram from the hydrophone will show striations curves in the noise. These appear as oscillations of intensity along the frequency axis. The centers of each intensity peak changes over time as the source range decreases, then increases. The principle of waveguide invariance says (loosely) that the collective slope of these striations is "invariant" to the specific parameters of the waveguide. This is a powerful idea, since the environmental parameters of the waveguide are known to significantly impact acoustic propagation. Measuring something that depends on the target range but not the environmental parameters implies it can be used for target ranging while simultaneously being robust to environmental parameters.

The waveguide invariant concept was originally introduced and formalized by Chuprov (1982), in which he showed that the slope of striations in frequency-range space can

be described by the formula

$$\frac{\partial \omega}{\partial r} = \beta \frac{\omega}{r}. \tag{2.2}$$

The parameter β describes the slope of the striations, and its value does not vary much to environmental parameters. This parameter thus provides a link between the source range and the measured acoustic data, while having minimal dependence on environmental parameters. Recent work by Oesterlein et al. (2010) showed that striation slopes are easily extracted from spectrograms thus providing an estimate of the left hand side of Eq. 2.2.

The modal regime has traditionally been used to describe the waveguide invariant phenomenon D'Spain and Kuperman (1999). Modal decomposition of the pressure field results in a weighted orthogonal basis set of conical waves. An individual multipath arrival can thus be distributed across several modes. While analysis has traditionally been formulated in a modal regime by describing interfering interactions between modes, Harrison (2011) showed that the relative eigenray travel times give the striations their shape, and computed β for several analytic sound speed profiles.

The use of the ray regime has some advantages, in that it provides a somewhat more intuitive picture of what gives rise to the striations. Each of the multipath eigenrays between a source and receiver will have a different travel time. Each will also have a unique set of phase and amplitude changes as a result of boundary reflections. Since noise has a short correlation time, then at a given moment in time, the hydrophone receiver will simultaneously observe all the arrivals as multiple incoherent sources. However, spectrograms are computed with a snapshot that is typically longer than the majority of the time delays between multipath arrivals. Within the time span

of a snapshot the hydrophone will observe the arrivals as multiple delayed copies of the same noise signal. So, in other words, the hydrophone observes the multipath coherence if the snapshot length is long enough. An auto-correlation of the received waveform produces several delayed pulses on the time delay axis. Some pulses are stronger than others depending on the amplitude of the arrivals, and the amount of signal common to the arrivals within the snapshot. The power spectrum is the Fourier transform of the auto-correlation. Hence, the time-domain pulses in the auto-correlation correspond to sinusoidal patterns in the frequency domain, which give rise to the oscillating pattern in the power spectrum along the frequency axis. As the target range changes, so do the position of the pulses on the time axis, which affect the frequency of the sinusoids in the power spectrum thus manifesting as continuously varying striations in the spectrogram.

The principals of waveguide invariance have been shown to be useful for passive determination of range and radial velocity from sources (Rakotonarivo and Kuperman, 2012, Kapolka, 2008). These are single hydrophone applications, which mean they can be utilized with relatively little deployment complexity. However, recently waveguide invariance has been applied to more complex receiver configurations. For example, it has been utilized by a modified conventional beamformer to increase array gain by (Rouseff and Zurk, 2011). It has also been shown to improve active sonar beamforming by (Zurk and Rouseff, 2012).

Chapter 3

Theory

This chapter gives the necessary technical background forming the basis for the techniques presented in this dissertation. The discussion begins with the foundation of classical sonar processing: the sonar equation (Sec. 3.1). Sec. 3.2 then delves into some fundamental signal processing tools needed to process raw time-series data. Sec. 3.3 presents the acoustic wave equation that links the signal processing to fundamental physical laws. Sec. 3.4 then extends this to describe the Helmholtz equation and Green's functions. It goes into several theoretical regimes for modeling propagation of acoustic waves through the ocean, all of which are based on the wave equation. Returning to signal processing, the concept of beamforming, also known as spatial filtering, is then discussed in Sec. 3.5. This begins with the simplest algorithm, the conventional beamformer, and then proceeds to more sophisticated algorithms that have the ability to adapt to the data being measured. Matched field processing (MFP) is another technique that uses formulations related to beamforming but is built around physical waveguide models. The techniques of this dissertation are not based on MFP, but it is enlightening to examine how the a similar problem is solved in a different way.

3.1 Signal Intensity Detection and Ranging

Traditionally, passive sonar detection has been framed in the context of measurement and analysis of incoherent signal intensity (Urick, 1962). Much of this has been outlined in the seminal text Urick (1967), and is captured in a set of equations, commonly referred as "the sonar equations." This material has recently been reformulated using a generalized source model in Ainslie (2010).

The sonar equations define the relationships between the *equipment*, the *medium*, and the *target*. In this detection modality, only the intensity of each quantity is considered. The passive sonar equation, as it is defined by Urick (1967), is

$$\mathrm{SL} - \mathrm{TL} = \mathrm{NL} - \mathrm{DI} + \mathrm{DT} \tag{3.1}$$

in which SL is the source level, TL is the transmission loss, NL is the noise level, DI is the receiving directivity index, and DT is the detection threshold, and each is specified in decibels. For a single hydrophone, array gain (AG) is not needed. For an array, it can be combined with DI by treating the beamformed receiver array as a unified receiver system with a given directivity pattern. Reverberation level (RL) is also not a factor in the passive sonar equation, so is omitted. An abundant amount of complexity underlies each variable in Eq. 3.1. Considering TL, for example, propagation of an acoustic wave through a waveguide can be affected by numerous environmental factors, some of which are time-varying, such as the sea surface. Much effort has been spent modeling propagation using a variety of mathematical approaches, as outlined in the seminal work by Jensen et al. (2011a). The sonar equation also does not consider many temporal features of a waveform. An impulse, such as might be produced by an

explosion, will have distinct arrivals in certain waveguides, each of which undergoes a different amount of transmission loss. The sonar equation is nevertheless a useful conceptual framework since it places clear bounds on overall system performance.

This intensity framework can also be used for target ranging. To illustrate, the passive sonar equation is rewritten as

$$\mathrm{SL} - \mathrm{TL}(R) + \mathrm{DI} = \mathrm{X} \ominus \mathrm{NL}. \tag{3.2}$$

TL is parameterized with range, R, and represents the modeled propagation loss at range R. DI is the amount of array gain, which amplifies SL. The last term, X, is the intensity recorded by the receivers. The $\ominus$ operator is subtraction in linear space (both arguments are converted to linear space from log space, subtracted there, and the result transferred back to log space). The source level is reduced in amplitude by some amount (TL) when it gets to the receiver, at which point it is amplified by the receiver (DI). The received intensity (X) is the sum of the signal plus noise, so if the intensity of the noise is subtracted off, what remains is the signal level. By varying the parameter, R to satisfy this equation, an estimate of range is obtained.

3.2 Signal Processing

This section provides some context for signal processing in underwater acoustics, and defines the mathematical constructs used in this study. One of the main themes of this dissertation is measuring time-lags between noise recorded on two hydrophones. Finding this lag is generally accomplished by cross-correlation, and is usually implemented in the frequency domain. These sections define what the frequency domain is, what cross-correlation is in the signal processing context, and how it is implemented in the frequency domain.

In general, a signal is an ordered sequence of measurements of a time-varying quantity. In underwater acoustics, the most common type of signal is the type that is recorded by a hydrophone, which measures slight pressure fluctuations and converts them to voltages. These voltages are then digitally sampled using an A2D (analog-to-digital) converter. The rate at which samples are recorded is called the "sample rate," often denoted f_s. The amount of voltage change for every unit of pressure change is captured in a metric called sensitivity that is different for each hydrophone. The units convention for pressure in underwater acoustics is micro-Pascals, or μPa. Sensitivities are therefore in units of μPa/V, but are usually expressed as decibels, which is a log scale. The result of A2D conversion is a digital signal of amplitudes of the form

$$x[n], \tag{3.3}$$

in which n is the sample number that monotonically increases. Underlying the sam-

pled process is a continuous process

$$x(t). \tag{3.4}$$

3.2.1 Fourier Analysis

The Fourier transform is based on the idea that any real-valued signal can be decomposed into an infinite sum of sin's and cos's. These sums are called the Fourier series expansion of a signal. The Fourier series can also be created using complex exponentials, which are a generalization of the trigonometric functions. This relationship is captured in Euler's famous formula[1]

$$e^{i\theta} = \cos(\theta) + i\sin(\theta). \tag{3.5}$$

The forward and inverse Fourier transform pair are defined as

$$X(\omega) = \mathcal{F}\left[x(t)\right] \triangleq \int_{-\infty}^{\infty} x(t)e^{-i\omega t}\,\mathrm{d}t \tag{3.6}$$

$$x(t) = \mathcal{F}^{-1}\left[X(\omega)\right] \triangleq \frac{1}{2\pi}\int_{-\infty}^{\infty} X(\omega)e^{+i\omega t}\,\mathrm{d}\omega. \tag{3.7}$$

The Fourier transform also exists in discrete time, and is known as the discrete Fourier

[1]The importance of this formula was eloquently noted by the late physicist Richard Feynman. "In our study of oscillating systems, we shall have occasion to use one of the most remarkable, almost astounding, formulas in all of mathematics. From the physicists' point of view, we could bring forth this formula in two minutes or so and be done with it. But science is as much for intellectual enjoyment as for practical utility, so instead of just spending a few minutes, we shall surround the jewel by its proper setting in the grand design of that branch of mathematics called elementary algebra. " –Richard Feynman, in "The Feynman lectures on physics" (Feynman et al., 2011)

transform (DFT). These pair of transform equations are defined as

$$X[k] = \sum_{n=0}^{N-1} x[n] e^{-2\pi i \frac{kn}{N}} \tag{3.8}$$

$$x[n] = \frac{1}{N} \sum_{k=0}^{N-1} X[k] e^{2\pi i \frac{kn}{N}} \tag{3.9}$$

where $x[n]$ is the input time-domain signal of length N and $X[k]$ is the corresponding frequency domain signal also of length N.

In discrete space, the DFT takes the place of the continuous Fourier transform. The DFT is almost always computed using the Fast Fourier Transform (FFT) algorithm. The reason for this is evident from their respective runtimes. The DFT is an $O(n^2)$ algorithm, and the FFT is $O(n \log n)$. The FFT makes use of $O(n \log n)$ memory in order to store intermediate calculations, whereas the DFT only requires $O(n)$ memory. This trade-off may be relevant only in the most tightly constrained embedded environments, but on modern hardware the additional memory utilization is negligible.

3.2.2 Cross-Correlation

Correlation processing has been studied for the better part of a century (Faran and Hills, 1952). The concept of correlation is that of a statistical measure of the amount of shared information between two data streams. In classical statistics, correlation is defined as a normalized covariance, in which the correlation coefficient is

$$\rho_{X,Y} = \frac{E\left[(X - \mu_X)(Y - \mu_Y)\right]}{\sigma_X \sigma_Y}. \tag{3.10}$$

The variables X, and Y are two random variables, μ is the mean, and σ is the standard deviation.

In signal processing, the correlation between signals is often defined as the conjugate operation of convolution. Convolution of two signals $x_1(t)$ and $x_2(t)$ is

$$x_1(t) * x_2(t) \triangleq \int_{-\infty}^{\infty} x_1(\tau) \cdot x_2(t-\tau)\ \mathrm{d}\tau \tag{3.11}$$

whereas cross-correlation is defined as

$$\Gamma_{x_1,x_2}(t) = x_1(t) \otimes x_2(t) \triangleq \int_{-\infty}^{\infty} x_1(\tau) \cdot x_2(t+\tau)\ \mathrm{d}\tau. \tag{3.12}$$

Note that both of Eq. 3.11 and Eq. 3.12 do not use the normalization found in classical statistics (i.e. the denominator of Eq. 3.10). The variable τ shifts x_2 on the time axis. At each value of τ, the integral is computed over the entire time axis. If $x_1(t)$ and $x_2(t+\tau)$ are statistically similar, the result will be a value with a large magnitude. In terms of notation, the cross-correlation function Γ is sometimes specified in terms of τ, such as $\Gamma_{x_1,x_2}(\tau)$. However, this does not change its meaning, but is rather a useful way to clearly distinguish the cross-correlation time axis from absolute time t, which are fundamentally different.

Cross-correlation in the time-domain is equivalent to conjugate-multiplication in the frequency domain. This fact can be leveraged to compute cross-correlations in the

frequency domain as

$$\Gamma_{x_1,x_2}(t) = x_1(t) \otimes x_2(t) \tag{3.13}$$

$$= \mathcal{F}^{-1}\left\{S_{x_1,x_2}(\omega)\right\} \tag{3.14}$$

$$= \mathcal{F}^{-1}\left\{\mathcal{F}\left[x_1(t)\right]\mathcal{F}\left[x_2(t)\right]^*\right\} \tag{3.15}$$

in which $*$ is the complex conjugation operation, ω is the angular frequency, and S_{x_1,x_2} is the cross-spectrum of x_1 and x_2.

One important consideration when computing the cross-correlation in the frequency domain is the circular convolution behavior of this method. The DFT is computed over a finite-length snapshot of data, $\mathcal{F}\left\{x_1[n]\right\}$ for which $n \in 1 \ldots N$. Samples outside of $[1, N]$ on the time axis are not used in the computation. However, according to Eq. 3.9, if $x[n]$ is computed outside of $[1, N]$, the exponential term will "wrap" causing $x[n]$ to repeat. In other words,

$$x[n] = x[n - qN] \tag{3.16}$$

for integer q. This implies that when $x_1[n]$ is cross-correlated with $x_2[n]$ using Eq. 3.15, that in actuality it is a correlation of the repeated pattern of each process. This effectively implies that as one process slides on the time axis, it is conceptually "wrapping" into the other side of the snapshot window. For many time-delay problems, this is not desirable. The solution is to zero-pad each snapshot prior to computing the DFT. The zero-pad should be exactly the length of the snapshot. This way, the part of the snapshot that wraps is all zeros.

3.2.3 Random Processes and Stationarity

Noise is typically thought of as the part of any measurement that is not related to the signal, or the waveform of interest. In this way, it serves to corrupt the signal making its measurement more difficult. However, signals of interest are not necessarily deterministic, and can be characterized as a random process.

What is a random process? A random process, a.k.a. a stochastic process, extends the concept of a random variable to a time-series. A single random variable is defined by its probability distribution function (PDF). A random process can be thought of as a set of random variables defined on the time axis. In discrete time, each sample would be a separate random variable. Note that each sample may depend on the other samples in the time series. This means that the process is fully characterized by the joint PDF across all samples.

A single realization of a random variable would be to draw a single value from that PDF. An infinite number of realizations together would converge to the PDF. Likewise, a single realization of a random process is a draw from the joint PDF. Note that the joint PDF is a multivariate function with dimension equal to the number of samples, which can possibly be infinite. A Gaussian process is one in which the joint PDF is a multivariate normal distribution. This is defined in Kay (1998) as

$$p(x) = \frac{1}{(2\pi)^{n/2}\det^{1/2}(C)} \exp\left[-\frac{1}{2}(x-\mu)^T C^{-1}(x-\mu)\right] \tag{3.17}$$

in which C is the covariance matrix, n is the dimensionality of the PDF, μ is a vector of means, det is the determinant operator, and T is the transpose operator.

A stationary process is a specialization of a random process in which the joint PDF does not vary with time. It is important to note this is a theoretical construct. In practice, the random processes encountered in passive acoustic localization are never truly stationary, but within a very short snapshot of time, it is assumed to be.

A "white" process is a process for which each sample is independent and identically distributed (IID). A common representation of noise emitted from a propeller is zero-mean white Gaussian noise, such that each sample is drawn from a normal distribution centered at the origin with a given variance. The variance determines the source level, and is the only parameter needed to fully characterize the process, making it easy to work with. Note that in a waveguide, noise emitted by a source may reflect from several boundaries before arriving at a receiver. This received waveform now contains several delayed copies of the original white Gaussian process. Hence, the received waveform is no longer white since there exist time delays for which the samples are correlated (i.e. not independent).

The whiteness of a process is defined as each sample being drawn from IID random variables. This implies that the spectral levels of the process are flat across frequency. Purely white noise is a theoretical construct, mainly because a process that is defined as having the same level over all frequencies out to infinity, and will have infinite energy. As with stationarity, a real sonar noise process are sometimes approximated as being white, but more commonly are approximated as having a flat spectrum within a specific frequency band. Hence, analytically, a white Gaussian process is often used, but algorithmically, a flat-spectrum band-limited Gaussian process is more appropriate.

3.2.4 Pre-Whitening

The last section discussed band-limited white Gaussian processes, which are useful for modeling noise emitted from a propeller-driven vehicle. However, vehicles that use combustion engines to drive their propellers often emit other sounds from the different components of the drive train. These components are "narrowband" in that the acoustic energy is concentrated at specific frequencies instead of smeared out over a wider band. An example would be the crankshaft of an engine, which turns at a certain RPM. This RPM rate is proportional to the rate of firing of the pistons, also the rate of the camshafts, as well as the turn rate propeller itself. Most propellers have multiple blades, so the blade-rate is a multiple of the propeller RPM. The base crankshaft RPM and the various components that turn at constant proportions to it will typically impart strong vibrational energy into the water that show up as "tonal" components at a receiver. These superimpose with the noise produced by cavitation from the propeller. In that case, both signatures show up in the received waveform. The problem is that these tonal components show up in the cross-correlation (Eq. 3.15), which hinder time delay estimation.

This can be described with an example. Assume $x(t)$ is a white Gaussian process but contains an unwanted strong signal component of the form $B\cos(\omega' t)$, in which ω' is any single frequency. This is detected at receiver 1 after a time delay of t_1, and receiver 2 after a time delay of t_2. The superimposed signals are

$$y_1(\omega) = x(t - t_1) + B\cos[\omega'(t - t_1)] \tag{3.18}$$

$$y_2(\omega) = x(t - t_2) + B\cos[\omega'(t - t_2)]. \tag{3.19}$$

In the frequency domain,

$$Y_1(\omega) = X(\omega)e^{-i\omega t_1} + B\delta(\pm\omega')e^{-i\omega t_1} \tag{3.20}$$

$$Y_2(\omega) = X(\omega)e^{-i\omega t_2} + B\delta(\pm\omega')e^{-i\omega t_2}. \tag{3.21}$$

The cross-correlation of receivers 1 and 2 is

$$S_{X_1,X_2}(\omega) = Y_1(\omega)Y_2^*(\omega) \tag{3.22}$$

$$= |X|^2 e^{-i\omega(t_1-t_2)} \tag{3.23}$$

$$+ XB^* e^{-i\omega(t_1-t_2)}\delta(\pm\omega') \tag{3.24}$$

$$+ X^* B e^{+i\omega(t_1-t_2)}\delta(\pm\omega') \tag{3.25}$$

$$+ |B|^2 e^{-i\omega(t_1-t_2)}\delta(\pm\omega'). \tag{3.26}$$

Let $\Delta = t_1 - t_2$, then in the time domain,

$$\Gamma_{x_1,x_2}(\tau) = |X|^2\delta(\tau - \Delta) \tag{3.27}$$

$$+ XB^* \cos[\omega'(\tau - \Delta)] \tag{3.28}$$

$$+ X^* B \cos[\omega'(\tau + \Delta)] \tag{3.29}$$

$$+ |B|^2 \cos[\omega'(\tau - \Delta)]. \tag{3.30}$$

If B is large enough, the last term dominates and adds a significant amount of sinusoidal corruption to $\Gamma(\tau)$ that has the same frequency as the unwanted tonal signal component. Time delay estimation is tantamount to finding the offset peak associated with the first term, so sinusoidal corruption will make this difficult.

One solution is called the "phase transform", or PHAT, and is often referred to the

"pre-whitening" (Knapp and Carter, 1976). It is one method of compensating for noise that is not initially spectrally flat and involves forcibly whitening the received signal in a finite band as

$$\hat{Y}(\omega) = \frac{Y(\omega)}{|Y(\omega)|}. \tag{3.31}$$

Pre-whitening reduces strong tonal components, while amplifying weak spectral values. Essentially, after the operation is complete, all the information is contained in the phase, since all amplitudes are one. This is usually done using discrete time-series with finite bandwidth. For signals that are constrained to a finite band, pre-whitening is applied only to that band while suppressing (setting to zero) spectral values outside of the band. This is critical because if parts of the frequency axis in which no signal exists are pre-whitened, this has the effect of amplifying those spectral components, which shows up as additional noise in the output time-series. In the example above, if the band consists of N discrete bins and the tonal corruption occupies one bin, then B is reduced to $|X|/N$, regardless of its true level, which may be much greater than X.

3.2.5 Time Delay Estimation

As was demonstrated in the last section in the context of pre-whitening, noise recorded at two receivers will manifest as a single pulse in the cross-correlation output. This process is called pulse-compression. Noise is broadband, but so are short-duration time-domain pulses. Pulse compression is thus a method to compress the broadband energy of noise into a short-duration pulse.

Similar to the definition in Eqs. 3.18 and 3.19, two receivers record a white Gaussian noise process $x(t)$ in the presence of additive white Gaussian noise $n(t)$,

$$y_1(\omega) = x(t - t_1) + n_1(t) \tag{3.32}$$

$$y_2(\omega) = x(t - t_2) + n_2(t). \tag{3.33}$$

The additive noise processes are separate IID processes unique to each receiver.

The cross-correlation of receivers 1 and 2 is

$$S_{X_1,X_2}(\omega) = Y_1(\omega)Y_2^*(\omega) \tag{3.34}$$

$$= |X|^2 e^{-i\omega(t_1-t_2)} \tag{3.35}$$

$$+ XN_2^* e^{-i\omega t_1} \tag{3.36}$$

$$+ X^* N_1 e^{+i\omega t_2} \tag{3.37}$$

$$+ N_1 N_2^*. \tag{3.38}$$

This is often computed by averaging the cross-correlation of multiple snapshots together. The rationale for doing this is based on the fact that $n(t)$ is uncorrelated with $x(t)$. Averaging thus reduces the relative magnitude of the cross-terms, Eqs. 3.36 and 3.37. It is useful to consider the ensemble average,

$$S_{X_1,X_2}(\omega) \approx \mathrm{E}\left[Y_1(\omega)Y_2^*(\omega)\right] = |X|^2 e^{-i\omega(t_1-t_2)}, \tag{3.39}$$

in which E is the expected value operator. This form eliminates any terms that are comprised of the product of two uncorrelated spectra. Transforming to the time

domain,

$$\Gamma_{x_1,x_2}(\tau) = |X|^2 \delta[\tau - (t_1 - t_2)] \tag{3.40}$$

for which $t_1 - t_2$ is the time delay. This method can thus be used to estimate the time-delay of broadband noise between two hydrophones by measuring the offset of the delta function in Eq. 3.40. Carter (1993) has compiled an anthology of papers on this subject of using cross-correlation for time-lag estimation.

3.3 The Wave Equation

The acoustic wave equation is derived in this section. The material presented here borrows heavily from Richard Feynman's lectures on physics [Volume 1, Chapter 47 of Feynman et al. (2011)].

The wave equation follows from three physical phenomena.

I As the fluid[2] moves, it changes density.

II The change in fluid density causes a change in pressure.

III The pressure gradient causes acceleration.

The following sections derive the wave equation for a plane wave traveling along the x-axis. Initially, however, a definition is needed. At equilibrium, a fluid has a constant pressure p_0 and density ρ_0. Waves are carried by small fluctuations in pressure p_e and density ρ_e whose magnitude is significantly less than equilibrium. The total pressure and density are then

$$p = p_0 + p_e \qquad\qquad \rho = \rho_0 + \rho_e. \tag{3.41}$$

3.3.1 Continuity Equation (I)

The continuity equation is essentially a restatement of the law of the conservation of mass. It basically says that as the volume expands, the density decreases. The

[2]Here, fluids are considered to be compressive or weakly compressive, which gives rise to linear acoustics.

displacement, or movement, of the fluid located at position x and at time t is $\chi(x,t)$. The new position is then

$$\chi(x,t) + x. \tag{3.42}$$

Now considering a nearby point a small distance away $x + \Delta x$, the displacement of that point is $\chi(x + \Delta x, t)$ making its new position

$$\chi(x + \Delta x, t) + x + \Delta x. \tag{3.43}$$

The mass of the air, per unit area, that is moved is between the first and second point initially is

$$\rho_0(x + \Delta x - x). \tag{3.44}$$

After displacement, the volume expands so the mass is

$$\rho[\chi(x + \Delta x, t) + x + \Delta x - \chi(x,t) - x] \tag{3.45}$$

in which ρ is the new density. The masses must be equal,

$$\rho_0 \Delta x = \rho[\chi(x + \Delta x, t) - \chi(x,t) + \Delta x]. \tag{3.46}$$

Since Δx is small,

$$\chi(x + \Delta x, t) - \chi(x,t) = \frac{\partial \chi}{\partial x} \Delta x. \tag{3.47}$$

Therefore,

$$\rho_0 \Delta x = \rho \left(\frac{\partial \chi}{\partial x} \Delta x + \Delta x \right) \tag{3.48}$$

$$= (\rho_0 + \rho_e) \left(\frac{\partial \chi}{\partial x} \Delta x + \Delta x \right). \tag{3.49}$$

Since

$$\rho_e \ll \rho_0, \tag{3.50}$$

this now becomes

$$\rho_0 \Delta x = \rho_0 \frac{\partial \chi}{\partial x} \Delta x + \rho_0 \Delta x + \rho_e \Delta x. \tag{3.51}$$

Which yields

$$\rho_e = -\rho_0 \frac{\partial \chi}{\partial x}. \tag{3.52}$$

3.3.2 Equation of State (II)

The equation of state describes the behavior of the fluid, or its "state" in a chemistry sense, in response to changes in pressure, volume, density, and temperature. In general, the pressure is some function of the density, such that

$$p = f(\rho) \tag{3.53}$$

so

$$p_0 + p_e = f(\rho_0 + \rho_e). \tag{3.54}$$

Since

$$\rho_e \ll \rho_0, \tag{3.55}$$

this allows for the first-order approximation

$$p_0 + p_e = f(\rho_0) + \rho_e f'(\rho_0) \tag{3.56}$$

where $f'(\rho_0)$ is the derivative of f evaluated at ρ_0. Since $p_0 = f(\rho_0)$,

$$p_e = \rho_e f'(\rho_0) \tag{3.57}$$

$$f'(\rho_0) = \frac{p_e}{\rho_e} = \left(\frac{\partial p}{\partial \rho}\right)_0 = B \tag{3.58}$$

which is the derivative of pressure with respect to density evaluated at equilibrium. The equation of state is thus defined as,

$$p_e = B\rho_e. \tag{3.59}$$

The product $B\rho_0$ is known as the bulk modulus, and is based on the properties of the fluid.

An adiabatic process is defined in Kinsler et al. (1999) as

$$\frac{p}{p_0} = \left(\frac{\rho}{\rho_0}\right)^{\gamma}. \tag{3.60}$$

For an ideal gas, $\gamma = 1$. Compression of a gas causes it to rise in temperature. Through diffusion, excess heat in one volume will migrate to an adjacent volume of lower temperature. However, if the distance between crests of the wave (i.e. the wavelength) are much greater than the separation between the elementary particles of the fluid, this heat flow will be much slower than the movement of the wave. This allows for the approximation to be made that no energy is lost to the movement of heat, hence the term adiabatic. In ocean water, some acoustic energy is lost due to heat, which gives rise to an attenuation constant that increases with frequency. The attenuation constant is fairly small, and is commonly approximated as zero if waves are propagating over short distances, or if losses due to other factors outweigh it.

3.3.3 Equation of Motion (III)

The equation of motion is essentially a restatement of Newton's second law. For a free body this is

$$F = ma \tag{3.61}$$

in which F is the force, m is the mass, and a is the acceleration. For a small volume of fluid, the net pressure in the positive and negative x direction corresponds to a net force per unit area. A pressure gradient will therefore lead to an acceleration of the

small volume. The mass per unit area of the fluid having thickness Δx is

$$\rho_0 \Delta x. \tag{3.62}$$

The acceleration is simply the second derivative of its position,

$$\frac{\partial^2 \chi}{\partial t^2}. \tag{3.63}$$

The pressure at x in the positive x direction is $p(x,t)$ and at a nearby point the pressure in the negative x direction is $p(x+\Delta x,t)$ so the net force per unit area in the positive x direction is is

$$p(x,t) - p(x+\Delta x,t). \tag{3.64}$$

This can be specified as

$$p(x,t) - p(x+\Delta x,t) = -\frac{\partial p}{\partial x}\Delta x = -\frac{\partial p_e}{\partial x}\Delta x \tag{3.65}$$

since the derivative of p_0 with respect to x is zero. The negative sign makes sense since if the pressure decreases between x and $x+\Delta x$, the net pressure will be positive. Lastly, since force equals mass times acceleration,

$$\rho_0 \Delta x \frac{\partial^2 \chi}{\partial t^2} = -\frac{\partial p_e}{\partial x}\Delta x \tag{3.66}$$

$$\rho_0 \frac{\partial^2 \chi}{\partial t^2} = -\frac{\partial p_e}{\partial x} \tag{3.67}$$

which is the equation of motion.

3.3.4 The Wave Equation

To summarize so far,

$$\rho_e = -\rho_0 \frac{\partial \chi}{\partial x} \qquad \text{continuity} \tag{3.68}$$

$$p_e = B\rho_e \qquad \text{state} \tag{3.69}$$

$$\rho_0 \frac{\partial^2 \chi}{\partial t^2} = -\frac{\partial p_e}{\partial x} \qquad \text{motion.} \tag{3.70}$$

Substituting Eq. 3.69 into Eq. 3.68,

$$\frac{1}{B} p_e = -\rho_0 \frac{\partial \chi}{\partial x}. \tag{3.71}$$

Taking the derivative with respect to x,

$$\frac{\partial p_e}{\partial x} = -\rho_0 B \frac{\partial^2 \chi}{\partial x^2}. \tag{3.72}$$

Substituting in Eq. 3.70,

$$-\rho_0 \frac{\partial^2 \chi}{\partial t^2} = -\rho_0 B \frac{\partial^2 \chi}{\partial x^2}. \tag{3.73}$$

Rearranging and canceling,

$$\frac{\partial^2 \chi}{\partial x^2} - \frac{1}{B} \frac{\partial^2 \chi}{\partial t^2} = 0. \tag{3.74}$$

The multiplier is

$$\frac{1}{B} = \frac{1}{c^2} \tag{3.75}$$

in which c is the speed of the wave. So,

$$\frac{\partial^2 \chi}{\partial x^2} - \frac{1}{c^2}\frac{\partial^2 \chi}{\partial t^2} = 0 \tag{3.76}$$

which is the wave equation in terms of displacement, χ. Taking the derivative again with respect to x,

$$\frac{\partial^2 \frac{\partial \chi}{\partial x}}{\partial x^2} - \frac{1}{c^2}\frac{\partial^2 \frac{\partial \chi}{\partial x}}{\partial t^2} = 0. \tag{3.77}$$

From Eq. 3.71,

$$-\frac{\rho_0}{B} p_e = \frac{\partial \chi}{\partial x}. \tag{3.78}$$

Substituting this into Eq. 3.77, canceling the extra constants, and redefining the excess pressure

$$p \triangleq p_e, \tag{3.79}$$

yields

$$\frac{\partial^2 p}{\partial x^2} - \frac{1}{c^2}\frac{\partial^2 p}{\partial t^2} = 0 \tag{3.80}$$

which is the wave equation in terms of (excess) pressure.

In three dimensions, the pressure wave equation in a homogeneous medium is defined by taking spatial derivatives along each cartesian axis as

$$\nabla^2 p - \frac{1}{c^2}\frac{\partial^2 p}{\partial t^2} = 0. \tag{3.81}$$

The symbol ∇^2 is the Laplacian operator defined as

$$\nabla^2 = \frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} + \frac{\partial^2}{\partial z^2} \tag{3.82}$$

in cartesian coordinates. Equation 3.81 is a second-order homogeneous differential equation and describes how the pressure within an infinitely small volume of space behaves. Equation 3.81 relates the Laplacian of the pressure field with the second derivative of the same pressure field. The Laplacian can be conceptualized as the curvature of the field in space, whereas the second derivative is the acceleration of the pressure. The wave equation states that these quantities are equal. As the curvature of the pressure field increases, so does the acceleration of the pressure.

3.4 Propagation Modeling

The wave equation describes the rules that govern a pressure field. It is based on fundamental laws of physics and holds at every point for which the assumptions are valid. By itself, the wave equation is of limited use since it describes the behavior of a fluid within an infinitesimally small volume. More often, the object of study is how waves, whose behavior is governed by this equation, propagate between points. In the ocean, the surface and seabed act as boundaries that reflect water-borne acoustic waves, thus creating a "waveguide." Typically, something in the water or on a boundary is creating waves, which is called a "source." These waves "propagate" through the waveguide to some other point in which they are observed with a "receiver," such as a hydrophone. The goal is to solve the wave equation in such a way as to capture the essential characteristics of this propagation, while remaining agnostic with respect to the specific signals being generated by the source. The formal way to describe the field at a receiver point with respect to a source point within a specific environment is called a Green's function. A propagation model is a method of estimating a Green's function.

The following sections describe the Helmholtz equation (Sec. 3.4.1) that describe how the wave equation is modified to accommodate a source point. An overview of several popular propagation models are given in Sec. 3.4.2. Lastly, Sec. 3.4.3 and Sec. 3.4.4 describe wavenumber integration and ray tracing models that are used later.

3.4.1 The Helmholtz Equation

Much of this discussion closely follows the derivation described in [Chapter 2 of Jensen et al. (2011a)].

Taking the Fourier transform of Eq. 3.81 allows the time derivative to be replaced with a constant. This reveals the a homogeneous Helmholtz equation,

$$\left[\nabla^2 + k^2\right] p(\mathbf{r}, k) = 0 \tag{3.83}$$

in which k is the wavenumber defined as

$$k \triangleq \frac{\omega}{c} \tag{3.84}$$

and $\mathbf{r}$ is the position in space defined as

$$\mathbf{r} \triangleq (x, y, z). \tag{3.85}$$

The plane-wave solution to Eq. 3.83 in a homogeneous medium is

$$p(\mathbf{r}) = \begin{cases} Ae^{i\mathbf{k}\cdot\mathbf{r}} \\ Be^{-i\mathbf{k}\cdot\mathbf{r}} \end{cases} \tag{3.86}$$

where $\mathbf{k}$ is the wavenumber vector and A and B are arbitrary wave amplitudes. These represent outward and inward propagating waves, respectively. Only outward propagating waves are of interest, so we discard the $Be^{-i\mathbf{k}\cdot\mathbf{r}}$ solution.

If a point source is placed in a medium that is infinite and homogeneous, only the

distance from the source matters. In spherical coordinates, Eq. 3.83 can be written as

$$\left[\frac{1}{r^2}\frac{\partial}{\partial r}r^2\frac{\partial}{\partial r}+k^2\right]p(r,k)=0 \tag{3.87}$$

in which r is the distance from the source. This has solutions

$$p(r)=\begin{cases}\frac{A}{r}e^{ikr}\\ \frac{B}{r}e^{-ikr}\end{cases}. \tag{3.88}$$

As with Eq. 3.86, these correspond to outgoing and incoming spherical waves, so the latter is omitted. The r in the denominator indicates that the amplitude of the wave decays with $\frac{1}{r}$. This also implies that the intensity decays with $\frac{1}{r^2}$.

A point source in an infinite homogeneous medium is modeled as a small sphere with radius a whose volume oscillates. It is assumed the change in radius is much less than the wavelength, such that

$$ka \ll 1. \tag{3.89}$$

The surface displacement is

$$u(t,r)=U(t) \tag{3.90}$$

which is the boundary condition for the sphere. The term u can be defined in terms of the displacement potential,

$$u(t,a)=\frac{\partial\psi(r,t)}{\partial r}. \tag{3.91}$$

In the frequency domain, the boundary condition is

$$u(\omega, r) = U(\omega). \tag{3.92}$$

The displacement potential, like pressure, is another variable that obeys the wave equation, so by Eq. 3.87 and Eq. 3.88,

$$\psi(r) = A\frac{e^{irk}}{r} \tag{3.93}$$

for outgoing spherical waves. The displacement is found through Eq. 3.91 as

$$u(r) = Ae^{ikr}\left(\frac{ik}{r} - \frac{1}{r^2}\right). \tag{3.94}$$

By the assumption in Eq. 3.89,

$$u(\omega, a) \approx -\frac{A}{a^2} \tag{3.95}$$

for which

$$A = -a^2 U(\omega). \tag{3.96}$$

The source strength is defined as

$$S_\omega \triangleq 4\pi a^2 U(\omega), \tag{3.97}$$

so

$$\psi(r) = -S_\omega \frac{e^{ikr}}{4\pi r}. \tag{3.98}$$

This basically says that the field at range r is the product of the source spectrum and a linearly decaying complex exponential. This latter part is called the Green's function,

$$g_\omega(r, 0) = \frac{e^{ikr}}{4\pi r}. \tag{3.99}$$

For a source located at an arbitrary point $\mathbf{r}_0$ and observed at point $\mathbf{r}$,

$$g_\omega(\mathbf{r}, \mathbf{r}_0) = \frac{e^{ik|\mathbf{r}-\mathbf{r}_0|}}{4\pi|\mathbf{r} - \mathbf{r}_0|}. \tag{3.100}$$

It is a simple matter to see that the Green's function satisfies the inhomogeneous Helmholtz equation

$$\left[\nabla^2 + k^2\right] g_\omega(\mathbf{r}, \mathbf{r}_0) = -\delta(\mathbf{r} - \mathbf{r}_0). \tag{3.101}$$

The principle of reciprocity follows from Eq. 3.101 which is defined as

$$g_\omega(\mathbf{r}, \mathbf{r}_0) = g_\omega(\mathbf{r}_0, \mathbf{r}). \tag{3.102}$$

The full inhomogeneous Helmholtz equation follows from Eq. 3.101 and Eq. 3.98,

$$\left[\nabla^2 + k^2\right] \psi(\omega, \mathbf{r}) = S_\omega \delta(\mathbf{r} - \mathbf{r}_0). \tag{3.103}$$

For a source that is distributed in space according to $f(\mathbf{r})$, the Helmholtz equation

becomes

$$\left[\nabla^2 + k^2\right] \psi(\omega, \mathbf{r}) = f(\mathbf{r}). \tag{3.104}$$

In the presence of boundaries, the Green's function takes on a new form, called the general Green's function,

$$G_\omega(\mathbf{r}, \mathbf{r}_0) = g_\omega(\mathbf{r}, \mathbf{r}_0) + H_\omega(\mathbf{r}) \tag{3.105}$$

in which $H_\omega(\mathbf{r})$ is any solution to the homogeneous Helmholtz equation,

$$\left[\nabla^2 + k^2\right] H_\omega(\mathbf{r}) = 0, \tag{3.106}$$

and $G_\omega(\mathbf{r}, \mathbf{r}_0)$ solves the inhomogeneous Helmholtz equation,

$$\left[\nabla^2 + k^2\right] G_\omega(\mathbf{r}, \mathbf{r}_0) = -\delta(\mathbf{r} - \mathbf{r}_0). \tag{3.107}$$

There is no general closed-form expression for $G_\omega(\mathbf{r}, \mathbf{r}_0)$ as it requires integrating around the boundaries. However, under certain conditions, this can be solved using different methods, some of which are listed in the next section.

While the derivation for the general Green's function and inhomogeneous Helmholtz equation is performed with respect to the displacement potential, the same equations hold for pressure. In signal processing terms, the Green's function can be considered the transfer function between a source and a receiver. This is a useful abstraction, since it allows propagation modeling tools to solve for the Green's function which can be then be utilized to predict how a particular source signal will manifest at the

receiver.

3.4.2 Overview of Common Models

Propagation models are techniques of predicting the Green's function, either as a whole, or in constituent pieces. Several methods exist to solve Eq. 3.107. What does it mean to "solve" this equation? Solving means determining a value for $G_\omega(\mathbf{r}, \mathbf{r}_0)$ at $\mathbf{r}$ and $\mathbf{r}_0$ in a waveguide having a particular set of parameters, such as water depth, sound speed gradient, attenuation constants, etc.

The popular methods of solving Eq. 3.107 are

- finite-difference
- normal mode
- parabolic equation
- wavenumber integration
- ray tracing

Only the latter two are used in this dissertation, and they will be described in the following sections.

3.4.3 Wavenumber Integration

The derivation in this section draws heavily from Section 2.4.1 of Jensen et al. (2011a). A common approximation to the ocean is that of an axisymmetric environment. That

means that with respect to a single point, the 2-D range-depth slice of the environment is the same regardless of the chosen azimuth angle. This symmetry can be used to simplify the Helmholtz equation, Eq. 3.104. This is done with the Hankel transform pair,

$$f(r,z) = \int_0^\infty f(k_r,z) J_0(k_r r) k_r \, \mathrm{d}k_r \tag{3.108}$$

$$f(k_r,z) = \int_0^\infty f(r,z) J_0(k_r r) r \, \mathrm{d}r \tag{3.109}$$

which produces the depth-separated wave equation

$$\left[\frac{\mathrm{d}^2}{\mathrm{d}z^2} + (k^2 - k_r^2)\right] \psi(k_r,z) = S_\omega \frac{\delta(z-z_s)}{2\pi} \tag{3.110}$$

such that

$$k = \sqrt{k_r^2 + k_z^2}. \tag{3.111}$$

Equation 3.110 is a second order partial differential equation of one variable, and is a much simpler form than Eq. 3.104. However, it comes at the cost of having to evaluate the infinite integral of the Hankel transform to obtain the frequency-domain solution. Hence the name *wavenumber integration*.

The subject of applying wavenumber integration to environments of different types is somewhat tangential to this work. The salient aspects that pertain here are that wavenumber integration is an exact solution to the wave equation, under the stated assumptions (axisymmetric and range independent). OASES (Schmidt, 2004) is an implementation of this method that requires a range-independent environment.

3.4.4 Ray Tracing

Acoustical ray tracing is similar to optical ray tracing. In optics, rays travel in straight lines through space until they intersect a boundary that causes them to reflect or refract. In computer graphics, ray tracing involves following the path of ray from each screen pixel through the environment to determine the surfaces it interacts with, which ultimately determines the color of the pixel. Acoustic ray tracing in the ocean is similar in principle, but there are some important differences. The first thing to note is that optical wavelengths are on the order of the size of atoms, whereas acoustic wavelengths commonly found in the ocean can be on the order of meters to even hundreds of meters long. The other big difference is that the ocean water itself can cause refraction due to sound speed changes over depth. An example of this in optics would be observing an object through rising hot gasses. For example, this causes the shimmer the horizon above pavement on a hot day, or the flutter of objects located on the other side of the gasses rising from a flame. Slight variations in the temperature and density of the air cause the speed of light to vary as well. This causes bending of optical rays. In computer graphics, this is generally a minor effect, but it can be much more significant in ocean acoustics.

The mathematical derivation in this section draws heavily from Section 3.3 of Jensen et al. (2011a). Like wavenumber integration, acoustic ray methods start with the Helmholtz equation, but the sound speed is now a function of depth. This is,

$$\nabla^2 p + \frac{\omega^2}{c^2(\mathbf{x})} p = -\delta(\mathbf{x} - \mathbf{x}_0), \tag{3.112}$$

in which $c(\mathbf{x})$ is the sound speed at point $\mathbf{x}$. The ray solution is assumed to take the

form

$$p(\mathbf{x}) = e^{i\omega\tau(\mathbf{x})} \sum_{j=0}^{\infty} \frac{A_j(\mathbf{x})}{(i\omega)^j}. \tag{3.113}$$

Substituting Eq. 3.113 into Eq. 3.112, the following set of equations emerge:

$$\begin{aligned} O(\omega^2) \quad &: |\nabla\tau|^2 = \tfrac{1}{c^2(\mathbf{x})} \\ O(\omega) \quad &: 2\nabla\tau\cdot\nabla A_0 + (\nabla^2\tau)\, A_0 = 0 \\ O(\omega^{1-j}) \quad &: 2\nabla\tau\cdot\nabla A_j + (\nabla^2\tau)\, A_0 j = -\nabla^2 A_{j-1}, \qquad j = 1, 2, \ldots \end{aligned} \tag{3.114}$$

The first equation, Eq. 3.114, is called the *eikonal* equation. Solving only this equation is equivalent to making a high-frequency approximation, in which the wavelength is essentially zero. The other equations are called the *transport* equations.

The eikonal equation in two dimension (range and depth) can be decomposed into

$$\frac{\partial r}{\partial s} = c\xi(s), \qquad \frac{\partial \xi}{\partial s} = -\frac{1}{c^2}\frac{\partial c}{\partial r} \tag{3.115}$$

$$\frac{\partial z}{\partial s} = c\zeta(s), \qquad \frac{\partial \zeta}{\partial s} = -\frac{1}{c^2}\frac{\partial c}{\partial z} \tag{3.116}$$

in which $[r(s), z(s)]$ is the ray trajectory in range and depth. Inclusion of the first transport equation results conceptually transforms the ray from a line to a tube of rays such that the acoustic energy carried in the tube obeys the conservation of energy law. Interested readers are referred to Section 3.3.2 of Jensen et al. (2011a).

Bellhop (Porter and Bucker, 1987) is a popular ray tracing tool for analyzing acoustic propagation between two points in an ocean waveguide. It is based on Gaussian ray tracing, and can operate on depth-varying and range-varying environments. Besides estimating the Green's function between two points, it can also provide information

for each individual arrival. An important concept here is the eigenray. Bellhop works by casting a ray "fan" into the water column. Pairs of adjacent rays that pass near the receiver are used to define a unique path, or eigenray, between the source and receiver. Depending on the environment, multiple eigenrays may exist. The advantage of this approach is it is relatively fast compared to wavenumber integration. However, it is an approximate solution to the wave equation which is valid only at higher frequencies.

3.5 Spatial Filtering

In the field of signal processing, the concept of filtering plays a central role. A filter can be thought of as a processing unit that takes an input signal, applies some function to it, and produces a "filtered" output signal. A commonly used filter is a finite impulse-response (FIR) filter that is used to amplify or suppress certain frequencies through linear combinations of sequential samples. However, in multichannel hydrophone systems, filtering can also be applied in the spatial domain. This is based on the idea that a set of two or more receivers may observe a source differently if it were located in different positions relative to the receivers. In a similar manner to a FIR filter, spatial filters work to isolate waves traveling in a particular direction while suppressing waves traveling in other directions. Spatial filtering is commonly referred to as beamforming.

As a matter of terminology, a beam is the primary direction that is being amplified, a side-lobe is a different direction that is being amplified unintentionally, and a null is a direction that is being suppressed. There exist methods to control the position side-lobes and nulls directly, but the first-order method ignores that and only considers the primary beam. This is called conventional beamforming, and is described in Sec. 3.5.1. The next class of algorithms that are discussed in Sec. 3.5.2 automatically adapt the side-lobes and nulls to the data being processed, hence the name adaptive beamforming. Lastly, the relationship of this to Matched Field Processing is given in Sec. 3.5.3, in which complex propagation models that account for multipath are used directly with beamforming. This is presented since it is an alternative to the new multipath processing methodology described in Sec. 6.

3.5.1 Conventional Beamforming

The plane-wave response on an array of arbitrary geometry is

$$\mathbf{v}(\mathbf{u}) = e^{-ik\mathbf{a}^T\mathbf{u}}, \tag{3.117}$$

in which k is the wavenumber, $\mathbf{a}$ is the array manifold that specifies the position of each element, and $\mathbf{u}$ is a unit vector indicating the direction of wave propagation (Van Trees, 2002a). Note that this bears striking similarity to the plane-wave Green's function in free space Eq. 3.86, but uses a different sign convention. The array response, $\mathbf{v}$, is also called the array manifold, and that the terminology adopted here[3]. This is an important quantity as describes what the expected response on the array should be. The parameters of the model is often specified in the parenthesis (i.e. $\mathbf{v}(\mathbf{u})$ for a plane-wave). It can take on more complex forms, such as incorporating wavefront curvature, or even pressure field simulations akin to those used in Matched Field Processing (MFP). In those examples, the parameter space can be much larger.

A conventional beamformer conceptually operates by delaying the signals from individual channels differently such that waves coming from the primary beam direction are aligned perfectly in time. Then, summing the delayed signals produces a coherent gain in the primary beam direction. In the frequency domain, time delays correspond to phase changes based on frequency. So, delay-and-sum can be implemented as a single inner vector product, as will be shown next.

[3] "Steering vector" is another commonly used term, but there is some ambiguity whether $\mathbf{u}$ is the steering vector or $\mathbf{v}$, so to avoid confusion that nomenclature is not used

The conventional frequency-domain phase-and-sum beamformer is defined as

$$b = \mathbf{w}^H \mathbf{p}, \tag{3.118}$$

in which $\mathbf{w}$ is the weight vector, $\mathbf{p}$ is the received spectral value, and H denotes the Hermitian transpose operation. In the conventional beamforming context, the weight vector $\mathbf{w}$ is proportional to the manifold

$$\mathbf{w} = \frac{1}{N}\mathbf{v}(\mathbf{u}), \tag{3.119}$$

such that it satisfies the "distortionless" constraint

$$\mathbf{w}^H \mathbf{v} = 1. \tag{3.120}$$

The vector $\mathbf{u}$ is the direction of wave travel,

$$\mathbf{u} = \begin{bmatrix} u_x \\ u_y \\ u_z \end{bmatrix}. \tag{3.121}$$

The power output of an arbitrary beamformer with weights $\mathbf{w}$ can be derived as

$$B = \left|\mathbf{w}^H \mathbf{p}\right|^2 \tag{3.122}$$

$$= \left(\mathbf{w}^H \mathbf{p}\right)\left(\mathbf{w}^H \mathbf{p}\right)^H \tag{3.123}$$

$$= \mathbf{w}^H \left(\mathbf{p}\mathbf{p}^H\right)\mathbf{w}. \tag{3.124}$$

The term in parenthesis is the outer product of the received pressure spectral values.

To improve estimates, of the beamformer output, several snapshots are averaged as

$$\mathbf{R} \triangleq \frac{1}{M} \sum_{m=1}^{M} \mathbf{p}_m \mathbf{p}_m^H, \tag{3.125}$$

in which $\mathbf{p}_m$ is the mth snapshot, often containing some amount of overlap from the previous snapshot. Thus, the improved estimate of B is

$$B = \mathbf{w}^H \mathbf{R} \mathbf{w}. \tag{3.126}$$

In the limit as the number of snapshots goes to infinity,

$$\mathbf{K} \triangleq \mathrm{E}\left[\mathbf{p}\mathbf{p}^H\right] \tag{3.127}$$

in which E[ ] is the expected value operator. The matrix $\mathbf{R}$ is referred to as the sample-averaged cross-spectral density matrix, and $\mathbf{K}$ is the ensemble cross-spectral density matrix. Cross-spectral density matrix is sometimes abbreviated CSDM, or CSD.

3.5.2 Adaptive Beamforming

While conventional beamforming places side-lobes in set places, this can have adverse affects in situations in which loud, unwanted signals overlap with the side-lobes. The distortionless constraint, Eq. 3.120 uses one degree of freedom. For an array of N elements, that leaves $N-1$ degrees of freedom that can be used to place the side-lobes and nulls. Ideally, if there are $N-1$ interferers, then each would get assigned its own null. In a sense, a perfect beamformer would automatically "adapt" to the

data and put these nulls where the interferers are. This is the subject of adaptive beamforming. The challenge with beamforming constraints is intuitively summed up as the "bubble under the carpet."[4] The basic idea is to automatically place the nulls to remove the interference while steering the side-lobes in quiet directions. What follows is mathematical description of the algorithm.

The derivation in this section is based on the method outlined in Antenna-Theory.com for an arbitrary array of N elements. The signal of interest signal spectra $s(\omega)$. It is carried by a plane wave traveling in direction $\mathbf{u}_s$, a unit vector. This is called the wave propagation vector and has cartesian components

$$\mathbf{u}_s \triangleq \begin{bmatrix} u_{s,x} \\ u_{s,y} \\ u_{s,z} \end{bmatrix}. \tag{3.128}$$

The geometry of the array specifies the coordinates of each element as

$$\mathbf{a}_n \triangleq \begin{bmatrix} a_{n,x} \\ a_{n,y} \\ a_{n,z} \end{bmatrix}. \tag{3.129}$$

The medium is assumed to have a constant wave speed, c so the wave number is

$$k = \frac{\omega}{c}. \tag{3.130}$$

[4] If you step on it, it just pops up somewhere else.

The array manifold for a particular geometry is therefore

$$\mathbf{v}(\mathbf{u}) = \begin{bmatrix} e^{-ik\mathbf{u}\cdot\mathbf{a}_1} \\ e^{-ik\mathbf{u}\cdot\mathbf{a}_2} \\ \vdots \\ e^{-ik\mathbf{u}\cdot\mathbf{a}_N} \end{bmatrix}. \tag{3.131}$$

The noise response on the array consists of uncorrelated noise signals on each channel,

$$\mathbf{N}(\omega) = \begin{bmatrix} n_1(\omega) \\ n_2(\omega) \\ \vdots \\ n_N(\omega) \end{bmatrix}. \tag{3.132}$$

In addition to noise, there are G interfering sources having signals $I_a(\omega)$ with wave propagation vectors $\mathbf{u}_a$. The total received spectra is then

$$\mathbf{X}(\omega) = s(\omega)\mathbf{v}(\mathbf{u}_s) + \mathbf{N}(\omega) + \sum_{a=1}^{G} I_a(\omega)\mathbf{v}(\mathbf{u}_a). \tag{3.133}$$

The desired output of the beamformer is simply the signal of interest

$$Y_d(\omega) = s(\omega). \tag{3.134}$$

The beamformer is implemented as a weighted sum of the received signals, so the actual output of the beamformer is

$$Y(\omega) = \mathbf{W}^H(\omega)\mathbf{X}(\omega). \tag{3.135}$$

Henceforth, the explicit dependence on ω is omitted in the notation.

The error between the actual and desired output is

$$\epsilon = Y - Y_d. \tag{3.136}$$

The mean-squared error is then

$$\mathrm{E}\left[\epsilon\epsilon^*\right] \tag{3.137}$$

in which E[ ] is the expected value operator. Substituting,

$$\begin{aligned}
\mathrm{E}\left[\epsilon\epsilon^*\right] &= \mathrm{E}\left[\left(Y - Y_d\right)\left(Y - Y_d\right)^*\right] && (3.138)\\
&= \mathrm{E}\left[\left(\mathbf{W}^H\mathbf{X} - s\right)\left(\mathbf{W}^H\mathbf{X} - s\right)^*\right] && (3.139)\\
&= \mathrm{E}\left[\mathbf{W}^H\mathbf{X}\mathbf{X}^H\mathbf{W} - \mathbf{W}^H\mathbf{X}s^* - \mathbf{X}^H\mathbf{W}s + ss^*\right] && (3.140)\\
&= \mathrm{E}\left[\mathbf{W}^H\mathbf{X}\mathbf{X}^H\mathbf{W}\right] - \mathrm{E}\left[\mathbf{W}^H\mathbf{X}s^*\right] - \mathrm{E}\left[\mathbf{X}^H\mathbf{W}s\right] + \mathrm{E}\left[ss^*\right]. && (3.141)
\end{aligned}$$

The weights are independent of $\mathbf{X}$, so they can be pulled outside of the expected value operator,

$$\mathrm{E}\left[\mathbf{W}^H\mathbf{X}\mathbf{X}^H\mathbf{W}\right] = \mathbf{W}^H\mathrm{E}\left[\mathbf{X}\mathbf{X}^H\right]\mathbf{W}. \tag{3.142}$$

The quantity $\mathrm{E}\left[\mathbf{X}\mathbf{X}^H\right]$ is the ensemble cross-spectral density matrix,

$$\mathbf{K} \triangleq \mathrm{E}\left[\mathbf{X}\mathbf{X}^H\right], \tag{3.143}$$

and

$$\mathrm{E}\left[ss^*\right] \tag{3.144}$$

is the mean signal power, σ_s^2.

Define the correlation of the signal with the received spectral values from each element as

$$\mathbf{\Lambda} \triangleq \mathrm{E}\left[\mathbf{X}s^*\right] \tag{3.145}$$

so

$$\mathrm{E}\left[\mathbf{W}^H\mathbf{X}s^*\right] = \mathbf{W}^H\mathbf{\Lambda}. \tag{3.146}$$

Similarly,

$$\mathrm{E}\left[\mathbf{X}^H\mathbf{W}s\right] = \mathbf{\Lambda}^H\mathbf{W}. \tag{3.147}$$

So, the mean square error is

$$\mathrm{E}\left[\epsilon\epsilon^*\right] = \mathbf{W}^H\mathbf{K}\mathbf{W} - \mathbf{W}^H\mathbf{\Lambda} - \mathbf{\Lambda}^H\mathbf{W} + \sigma_s^2. \tag{3.148}$$

To minimize the mean square error, first the derivative with respect to the conjugate weight vector is taken. Since this is a multi-variate function, the gradient operator is used

$$\nabla\mathrm{E}\left[\epsilon\epsilon^*\right] = \mathbf{K}\mathbf{W} - \mathbf{\Lambda}. \tag{3.149}$$

Setting to zero and solving,

$$\mathbf{W} = \mathbf{K}^{-1}\mathbf{\Lambda}. \tag{3.150}$$

Substituting Eq. 3.146 into Eq. 3.133,

$$\mathbf{\Lambda} = \mathrm{E}\left[\left(\mathbf{X}(\omega) = s(\omega)\mathbf{v}(\mathbf{u}_s) + \mathbf{N}(\omega) + \sum_{a=1}^{G} I_a(\omega)\mathbf{v}(\mathbf{u}_a)\right) s^*\right] \tag{3.151}$$

$$= \sigma_a^2 \mathbf{v}(\mathbf{u}_s) \tag{3.152}$$

so

$$\mathbf{W}_{\text{MSE}} = \sigma_a^2 \mathbf{K}^{-1}\mathbf{v}(\mathbf{u}_s). \tag{3.153}$$

The distortionless constraint is then applied such that

$$\mathbf{W}_{\text{MSE}}^H \mathbf{v}(\mathbf{u}_s) = 1, \tag{3.154}$$

which produces the minimum variance distortionless response (MVDR) weight vector

$$\mathbf{W}_{\text{MVDR}} = \frac{\mathbf{K}^{-1}\mathbf{v}(\mathbf{u}_s)}{\mathbf{v}^H(\mathbf{u}_s)\mathbf{K}^{-1}\mathbf{v}(\mathbf{u}_s)}. \tag{3.155}$$

Capon (1969) is credited for initially coming up with this methodology.

Using almost the same formula, minimum power distortionless response (MPDR) adaptive beamformer is defined as

$$\mathbf{w}_{\text{MPDR}} = \frac{\mathbf{R}^{-1}\mathbf{v}(\mathbf{u}_s)}{\mathbf{v}^H(\mathbf{u}_s)\mathbf{R}^{-1}\mathbf{v}(\mathbf{u}_s)} \tag{3.156}$$

which uses $\mathbf{R}$ instead of $\mathbf{K}$. This is a least-squares solution to

$$\min\left(\mathbf{w}_{\text{MPDR}}^H \mathbf{p}\right) \quad \text{constrained by} \quad \mathbf{w}_{\text{MPDR}}^H \mathbf{v}(\mathbf{u}_s) = 1. \tag{3.157}$$

Mathematically, the matrix inversion can be applied to ensemble CSD matrices for just the noise, or noise plus interference. In those cases, the beamformer is called the minimum *variance* distortionless response (MVDR) beamformer. However, if the signal itself is present in all relevant snapshot data, then the CSD matrix will contain noise plus interference plus signal, and if that is the matrix being inverted, the quantity being minimized is power. Hence the name difference. This is made explicit in Van Trees (2002a).

The degree of adaptivity can be a controlled using a technique called diagonal loading. This involves adding a constant value to the diagonal elements of the CSD matrix. This is also the same thing as artificially increasing the white noise gain (WNG). Diagonal elements of the CSD correspond to auto-correlations of each channel, so larger values along the diagonal in the noise covariance matrix correspond to greater amounts of uncorrelated white noise. The weight vector can be parameterized with WNG, ϵ, as

$$\mathbf{w}_{\text{MPDR+WNG}} = \frac{\left(\mathbf{R} + \epsilon\frac{\text{tr}(\mathbf{R})}{N}\mathbf{I}\right)^{-1}\mathbf{v}(\mathbf{u}_s)}{\mathbf{v}^H(\mathbf{u}_s)\left(\mathbf{R} + \epsilon\frac{\text{tr}(\mathbf{R})}{N}\mathbf{I}\right)^{-1}\mathbf{v}(\mathbf{u}_s)}. \tag{3.158}$$

A value of $\epsilon = 0$ implies pure MPDR, and $\epsilon = \infty$ implies pure conventional beamforming. This technique can be used to stabilize a rank deficient $\mathbf{R}$ for inversion, or to compensate for mismatch errors in element location. Cox et al. (1987) defined a framework for determining ϵ in an optimal way. That technique is often referred to

as the white noise gain constraint (WNGC) beamformer.

3.5.3 Matched Field Processing

Conventional MFP is formulated in the same beamforming context as Eq. 3.118 and Eq. 3.120. The primary exceptions are that $\mathbf{v}$ is computed using a propagation model, and instead of using beam-angle replicas, the replicas span a range-depth grid. The n^{th} element of $\mathbf{v}(x)$ is $v_n(x)$, which corresponds to a single element in the array at position, a_n, and is the Green's function between a replica at position x and a_n.

Computing the Green's function in an ocean waveguide can be done using a variety of tools, such as those outlined in Jensen et al. (2011a). The Green's function depends on environmental characteristics, such as ocean depth, bathymetry, ocean sediment type and layering, surface roughness, etc. To be explicit, it may actually be appropriate in situations in which the true environment is poorly described to use a fully parameterized model, $v_n(x, \Xi)$, in which Ξ contains all the parameters for the environment.

One reason why MFP is attractive is because it is not limited by the "design frequency" of an array in the same way that a conventional plane wave beamformer is. Rather, it is limited by modeling error, which can take several forms in the MFP modality, and is often called "mismatch." There is array mismatch, environmental mismatch error, and statistical errors (Baggeroer et al., 1993). An example of array mismatch is any perturbation of an array element from it's modeled position. It can also be improperly calibrated, leading to corruption of recorded spectral values. Environmental mismatch errors is a bit more open ended given the complexity that can

exist in a given waveguide. The seabed may consist of a different number of layers than the model, and each layer may have different acoustic impedances. The sound speed profile may also deviate from the model, particularly in areas of strong current in which the profile may be varying on a time-scale that is on the same order as the data collection period. Undetected internal waves would also cause unmodeled refraction. Rough surface scattering from the seabed may also modify the multipath characteristics. The surface is a time-varying rough surface, which needs to be accurately captured in a model. The seabed and its layers are also not flat in many areas of the ocean, so bathymetric variations must be taken into account. An example of statistical errors is a rank-deficient CSD matrix.

Chapter 4

Line Array Processing

This chapter discusses the use of line arrays for target localization. Line arrays have proved to be useful tools mainly for narrowband underwater target detection and tracking. These types of arrays are typically used for the submarine detection problem, and often the arrays are towed from a moving platform such as a ship. One of the unknowns regarding detecting autonomous underwater vehicles is the target signature. That is, the spectral levels and aspect dependency. Therefore, the first step in developing any detection scheme for a given underwater target using passive sonar is to characterize its radiated noise. Are the signatures narrowband or is there broadband content, and what are the source levels and aspect dependencies? These are critical questions and define the parameters needed to design a detection and tracking system.

AUVs are relatively new devices, and little work has been done on what sounds they make in an "underway" condition. In this chapter, the source signature is determined for a REMUS-100 AUV using measurements from a horizontal line array. Developing this underway characterization of the AUV, and providing an analysis of possible methods that could be used to detect and track the AUV has implications for law

enforcement. Agencies tasked with harbor and littoral security can leverage this work in guiding technical and policy-based decisions. This work has been published in Gebbie et al. (2012b). It was also the subject of an invited talk at the Acoustical Society of America Meeting in May 2014 in Providence, RI (Gebbie et al., 2014a). Initial work on processing the acoustic data for different types of targets was published in the Marine Technology Society Journal (Gebbie et al., 2011a).

The organization of this chapter is as follows. Sec. 4.1 gives an overview of the Hawai'i 2010 experiment in which the AUV was deployed. Sec. 4.2 discusses the research that was conducted into characterizing the aspect-dependent source levels emitted by an AUV using the line array.

4.1 Hawai'i 2010 Experiment

The AUV experiments were conducted in July 2010 at the Kilo Nalu Nearshore Reef Observatory (University of Hawai'i-Manoa) near Honolulu Harbor, Honolulu, HI. This work was done in collaboration with the University of Hawai'i. A map of this region is shown in Fig. 4.1. The observatory had several underwater electronics nodes, which supplied power and ethernet connectivity to scientific equipment. One node was at a fixed site about 0.4 km from shore in roughly 12 m of water. Honolulu Harbor was located 1 km to the northwest of this node, which served as a port for large commercial vessels. In addition, the nearby areas were popular sites for recreational fishing, boating, and SCUBA diving. There was also a significant amount of biological noise primarily from snapping shrimp, which created loud, broadband clicking noise (Lammers et al., 2008).

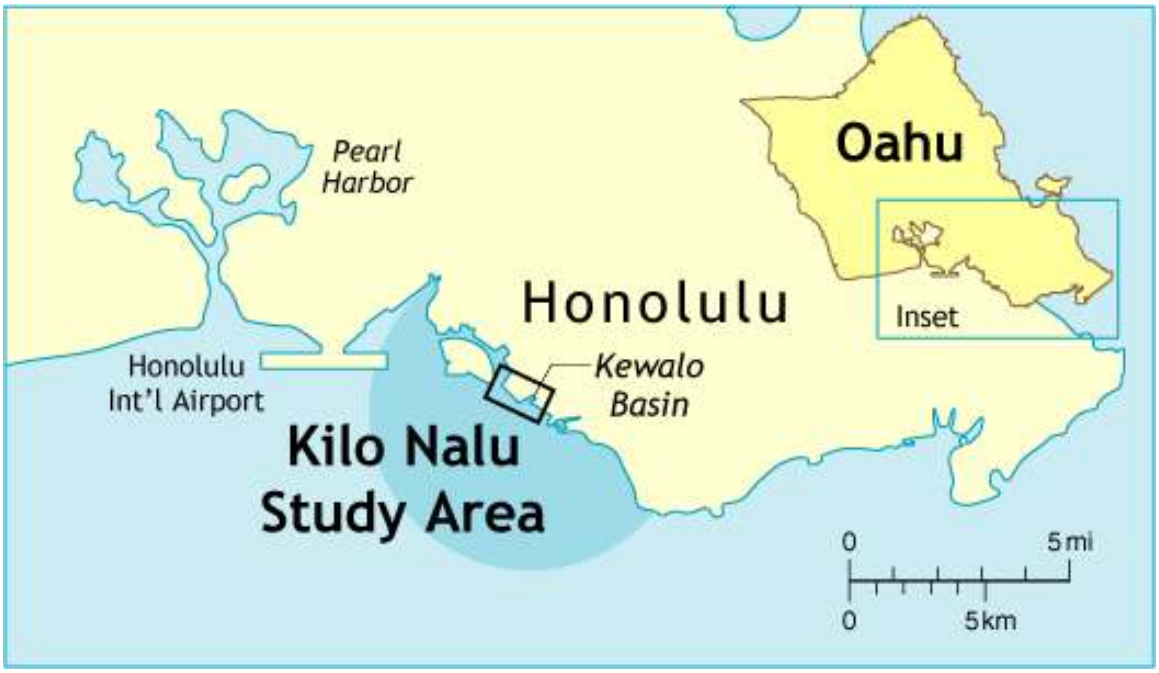


Figure 4.1: Map of the Kilo Nalu deployment site. Reproduced from Kilo Nalu Nearshore Reef Observatory (2011) website (Kilo Nalu Study Area) with permission.

4.1.1 POEMS-A Hydrophone Arrays

Two hydrophone arrays (HTI-92-WB, with hydrophone sensitivities of -160 dB re 1V/μPa) were connected to the node for a two-week period. The arrays were manufactured by High Tech, Inc. based in Gulfport, Mississippi, and are shown in Fig. 4.2. The model name of each array is POEMS-A, so the array is sometimes referred to by that name. One of the array systems was owned by the University of Hawai‘i, and the other by Portland State University.

The arrays were configured in a vertical and horizontal configuration, with the latter stretched out along the seafloor. This is shown in a sketch in Fig. 4.3, and in a photo taken with an underwater camera in Fig. 4.4. The horizontal orientation provided

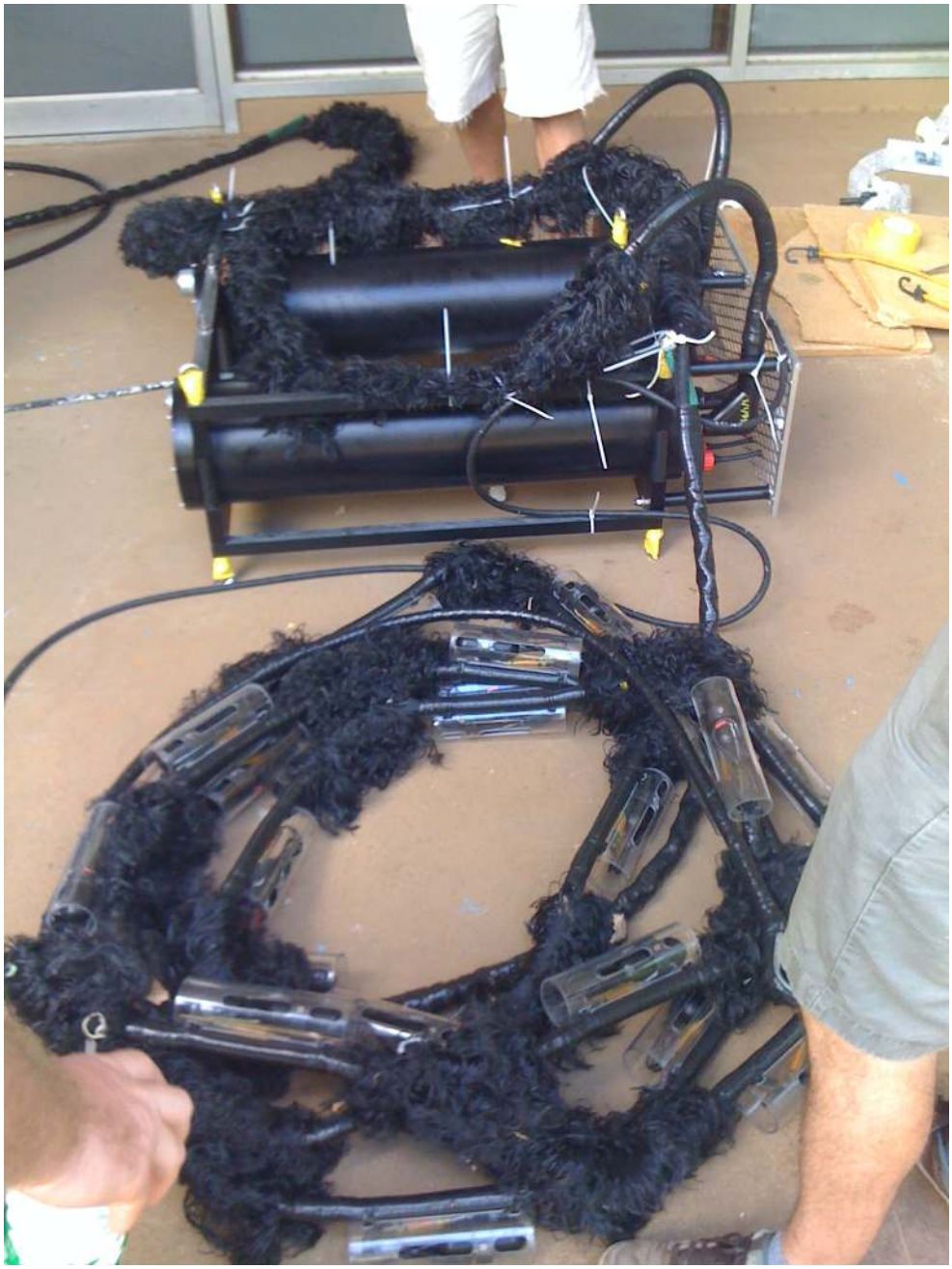


Figure 4.2: Photo of the pressure vessels containing the acquisition systems for both POEMS-A arrays in the Hawai'i 2010 experiment. Also visible is the coiled horizontal line array. The black material between array elements is "hair fairing," which adds drag when the cable is stretched taught while in a current. This reduces strumming, which can corrupt acquisition of the desired field-borne acoustic data.

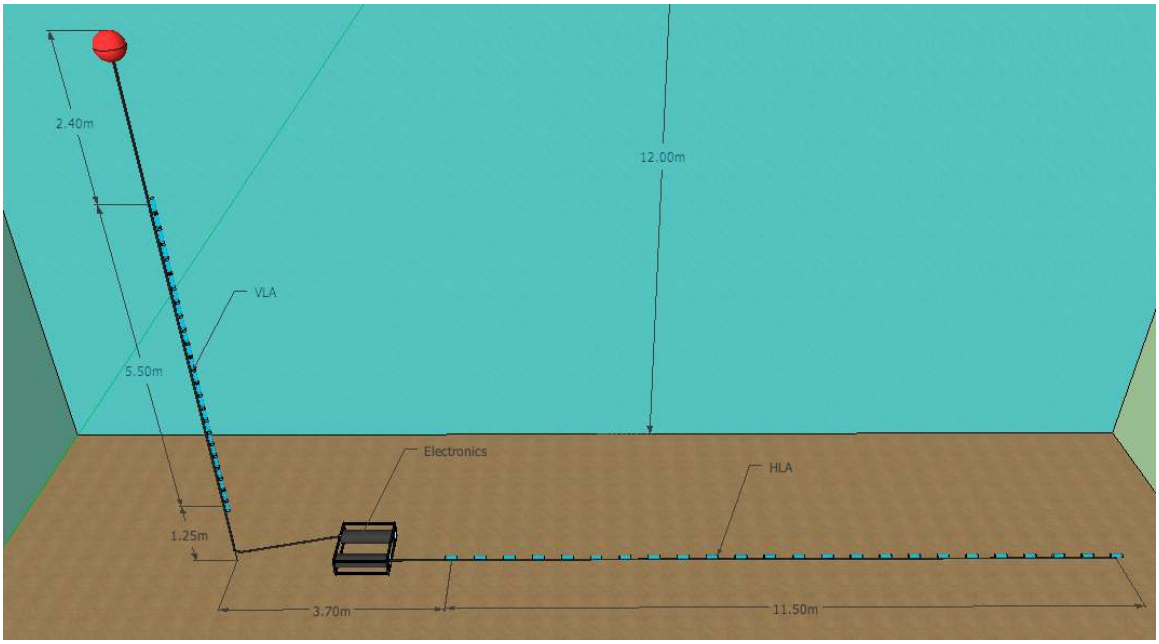


Figure 4.3: Sketch of the POEMS-A arrays with one in a horizontal configuration and the other in a vertical configuration (Gebbie et al., 2011a).

azimuthal beam resolution facilitating direction-of-arrival (DOA) estimation of radiated noise from targets. The vertical array enabled discrimination between multipath arrivals arriving at different elevation angles. Each had a total of 24 elements, with the vertical array evenly spaced at 0.25 m, and the horizontal evenly spaced at 0.5 m. The sample rate was 102.4 kHz with a dynamic range of 24 bits. The arrays used a high-pass filter to suppress frequencies below 300 Hz and a 110 dB anti-aliasing filter giving a usable acoustic band from 300 Hz to 46.4 kHz.

The first channel of each array was connected together with a synchronization device that generated a known electrical signal. The waveform is shown in Fig. 4.5. The synchronization signal was developed specifically as part of this research to time-align the CPUs so that both arrays could be used together as one large array. A photo of the synchronizer device is shown in Fig. 4.6. The reason for using the synchronizer was

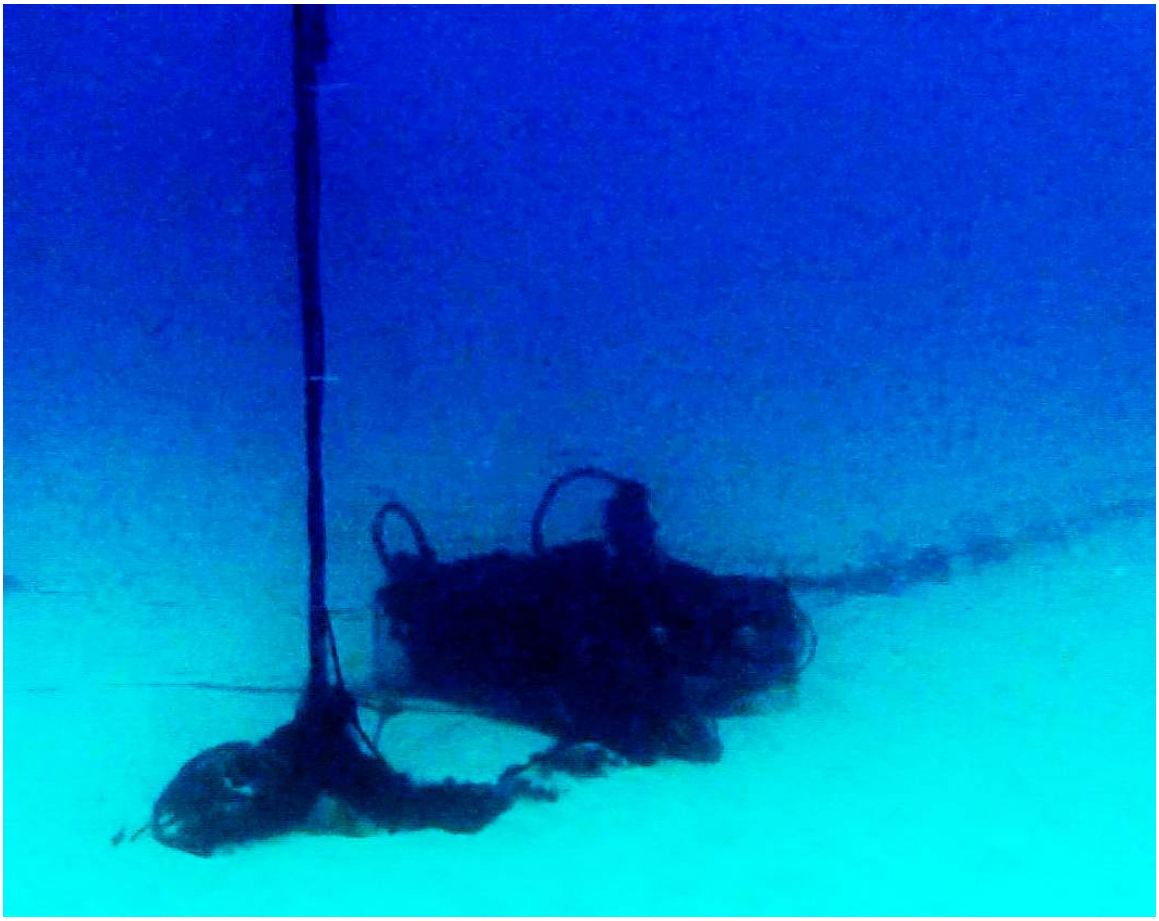


Figure 4.4: Underwater photo of the POEMS-A arrays with one in a horizontal configuration and the other in a vertical configuration.

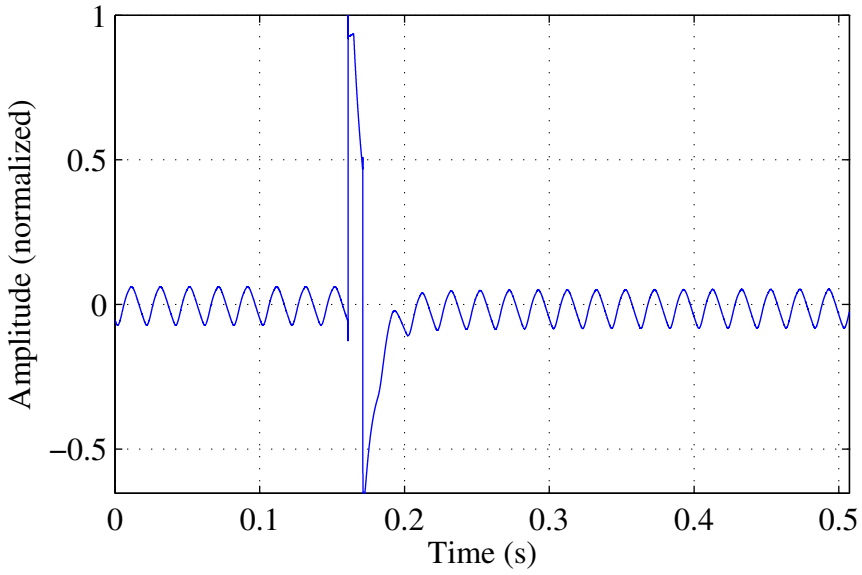


Figure 4.5: Waveform of POMES-A synchronizer.

that each array was connected to its own recording unit. The recording units were each closed systems consisting of a separate CPU and clock, so there was no way to time-synchronize their clocks directly without making major modifications to the vendor's design. The waveform being sent to the first channel by the synchronizer consisted of a broadband pulse and a tonal, which facilitated accurate time-synchronization during post-processing.

4.1.2 REMUS-100 AUV

During the deployment, a REMUS-100 AUV was operated using programmed tracks at the deployment site of the arrays. Photos of this AUV are shown in Fig. 4.7 and Fig. 4.8. The REMUS-100 was powered by a battery that can last up to 22 hours, and driven by a direct-drive DC brushless motor, which turned an open three-blade propeller. It was 160 cm in length, 19 cm in diameter, and weighed 37 kg. It had a

Figure 4.6: Photo of the synchronizer that fed a known electrical waveform to the first channel of each array, allowing the arrays to be time synchronized in post-processing.

Figure 4.7: Photo of the REMUS-100 AUV as it is deployed from the small boat. Bubbles are likely due to air entrainment at the surface.

Figure 4.8: Photo of the REMUS-100 AUV as it moves several meters below the surface. Despite the lack of clarity in this photo, note the lack of bubbles behind the propeller. This suggests that the propeller is not cavitating in these conditions, which affects the amount of broadband radiated noise.

maximum water speed of 2.6 m/s and a maximum depth of 100 m. Onboard navigational sensors estimated its location via several methods. At the surface, satellite GPS was used to obtain a coordinate fix. An acoustic doppler current profiler (ADCP) was used to measure its velocity over the seabed after submerging. Long baseline navigation was also used, which involved acoustic interrogation of separately deployed transponders. When both of these data streams were unavailable, the AUV defaulted to dead reckoning. Depth measurements were obtained from a pressure depth sensor. The AUV logged these navigational records in the form of downloadable, time stamped 3-D coordinates. An onboard acoustic modem, which operated in the 20-30 kHz band, was used to send data back to the operator in the boat. The modem served as a broadband source that proved useful for cross-correlation based DOA estimation.

The track for the pre-programmed mission is shown in Fig. 4.9. This plot is annotated with the location of the acoustic array. Only a small portion of the overall AUV track is actually shown, as the full track extends to a maximum distance of almost 500 m. This is shown in Fig. 4.10.

The AUV was deployed from a boat that motored a short distance away from the programmed AUV track. The boat engine was on for a few minutes and subsequently turned off. The modem transponder continued to transmit from the stationary boat. The AUV moved slowly at the surface for about 30 seconds before diving and accelerating to about 2.25 m/s. It followed a programmed track of several turns and course reversals passing the end of the array three times. On one turn, it reached a maximum distance of 100 meters from the array. The AUV's acoustic modem broadcasted information to the boat every 30 seconds.

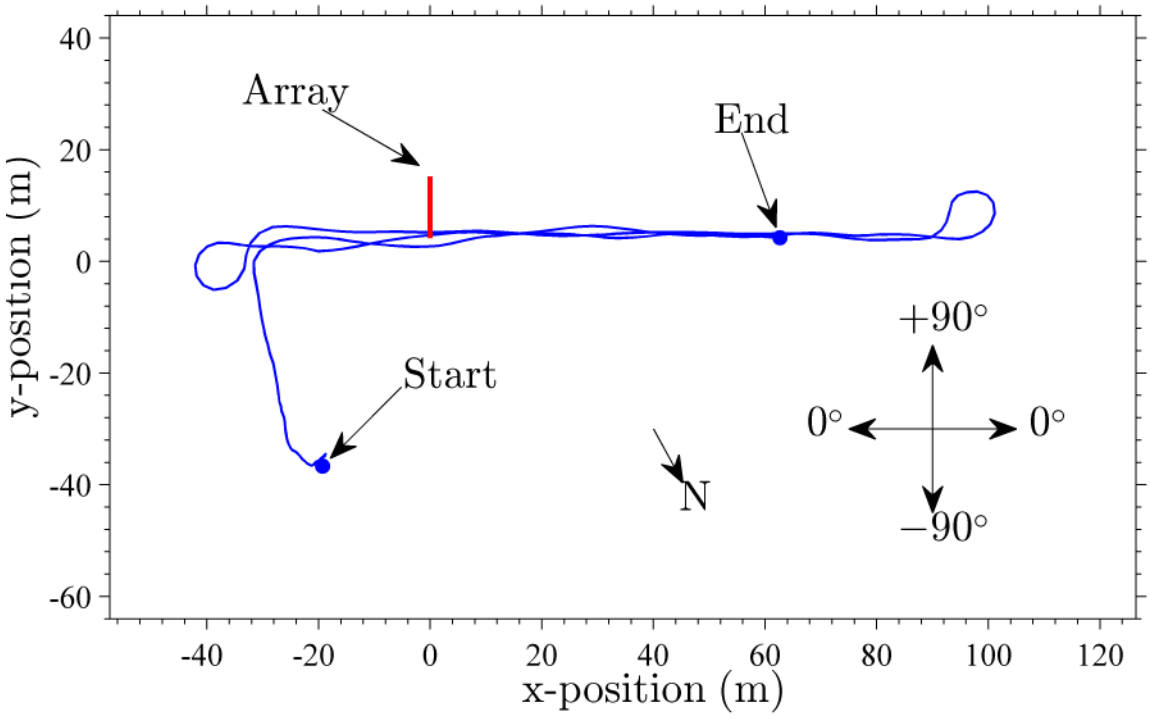


Figure 4.9: Track of AUV showing its position relative to the array. Navigation records taken from the AUV were used to generate the track line.

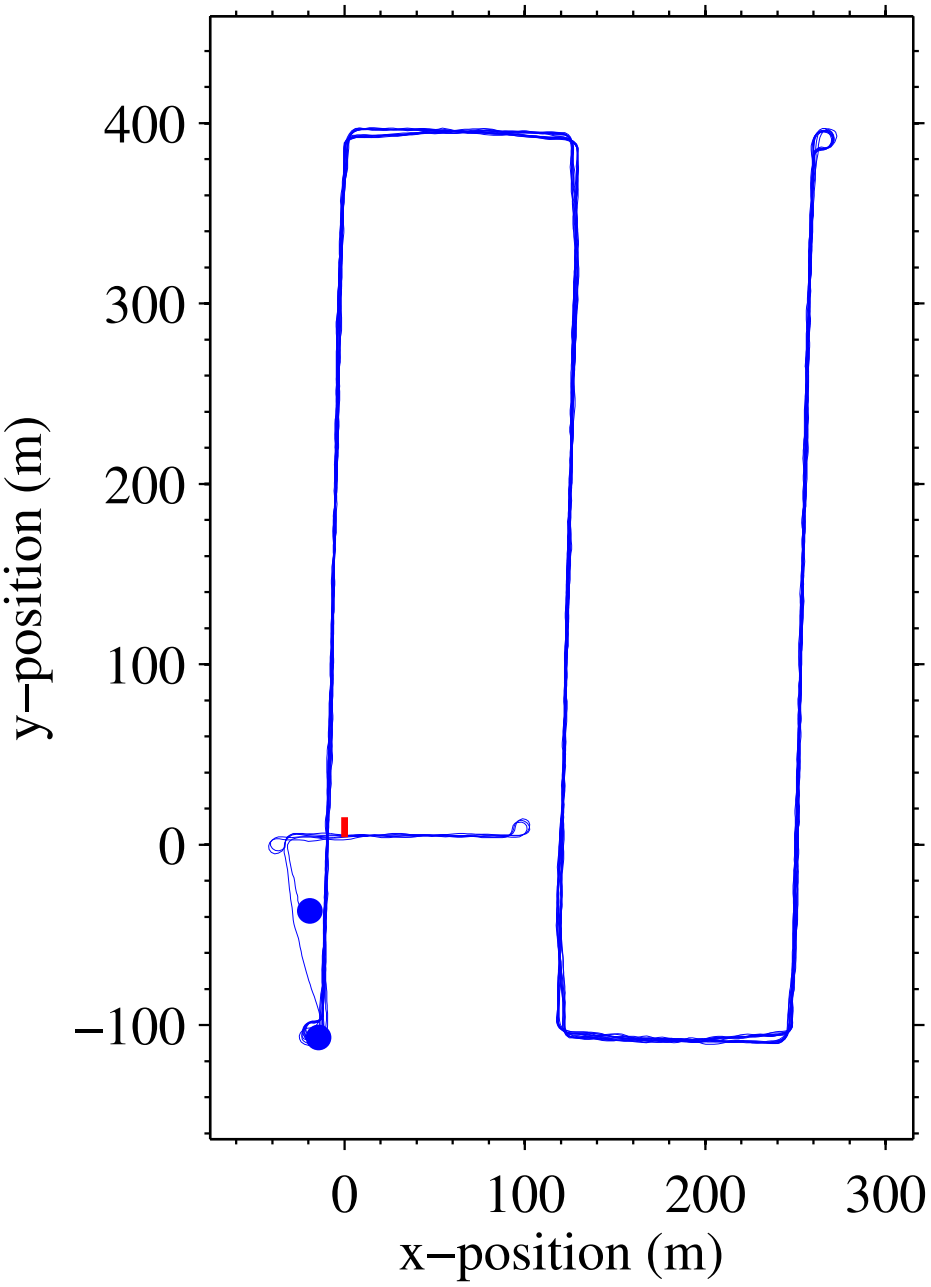


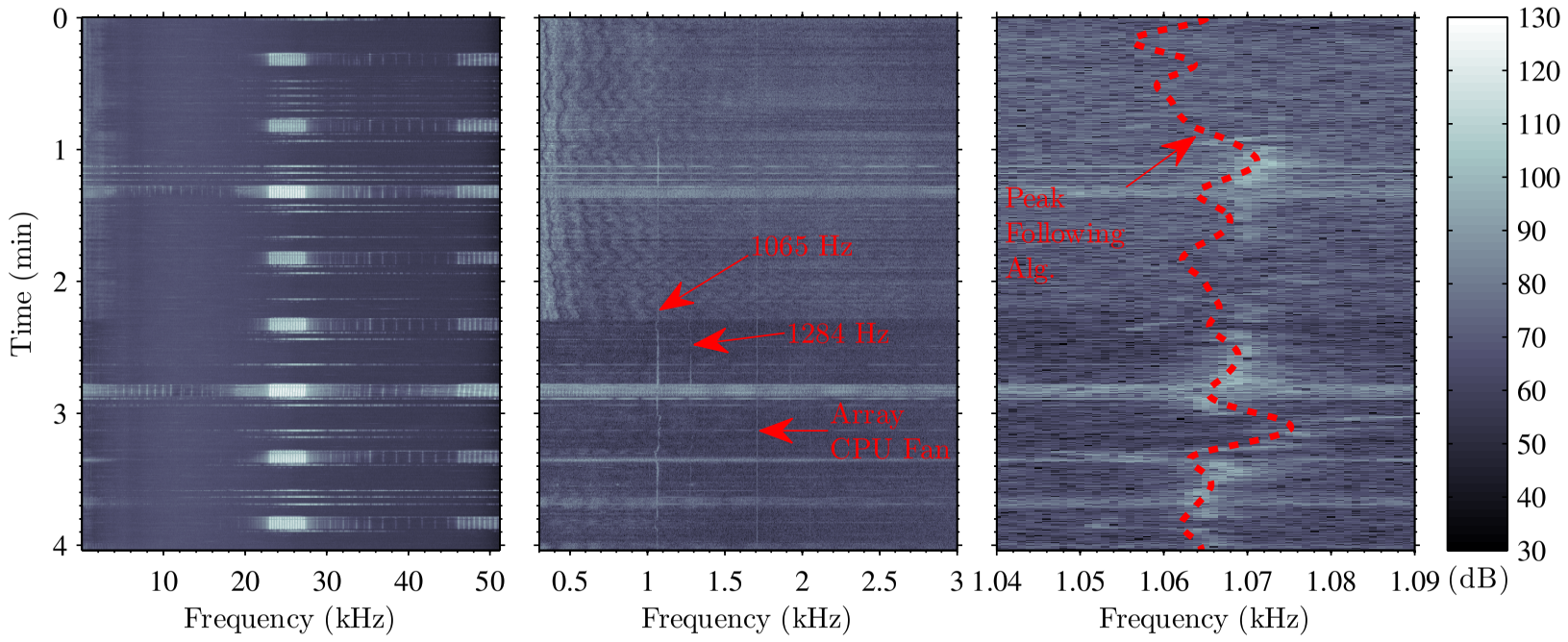
Figure 4.10: Full Track of AUV over four hours.

4.2 Source Level Estimation Method

This section outlines a technique for characterizing the source levels and beam patterns of a REMUS-100 AUV in an underway condition by fusing navigational records from the AUV with acoustic data collected on POEMS-A, a passive, bottom-mounted HLA. The AUV's navigational records are used to identify which signature components are originating from the its propulsion system using beamforming. Using a basic characterization of the environment, propagation modeling tools produce an estimate of transmission loss (TL) between each position of the AUV and each hydrophone. Directivity index (DI) or array gain (AG) are not considered since each hydrophone is processed separately. Source levels are then computed with the passive sonar equation using the received levels and TL. The AUV's aspect with respect to the HLA is then derived from the AUV's position and heading at each time step, facilitating the calculation of the beam pattern.

4.3 Spectrograms

Acoustic data from the northernmost element of the array during the first four minutes of the mission was used for the spectrograms shown in Fig. 4.11. In Fig. 4.11(a) the modem transmits data in the 20-30 kHz band in five second bursts every 30 seconds. Fig. 4.11(b) illustrates various tones from the AUV propulsion system, which are visible upon magnification of the frequency axis to 300 Hz - 3 kHz. The strongest tone was centered at about 1065 Hz, a weaker one at 1284 Hz, and a very faint one at 1369 Hz. The 1065 Hz tone was by far the strongest, and so was used for this study. Further magnification (Fig. 4.11(c)) reveals that this tone wandered within



(a) Full Spectrum (b) Lower Band (c) REMUS Tone

Figure 4.11: Spectrograms from the first phone of the array showing the signature of the REMUS-100 AUV and boat. Units are in dB re 1 μPa^2/Hz. (a) The full 300 Hz - 51.2 kHz band in which the modem noise is visible. (b) Magnification of the 300 Hz - 3 kHz band shows the various propulsion tones. (c) Magnification of the 1065 Hz propulsion tone showing the peak following algorithm (dashed line).

a 10 Hz band. The tone faded toward the end of the first minute and again at the beginning of the third minute. This corresponds to times during which the AUV turned 180°. The gap at the beginning of the dataset was most likely due to the AUV slowly moving at the surface. The 1700 Hz tone originated from a CPU fan inside the array electronics pressure vessel. The source level of the modem was sufficiently high that when within a few meters of the array it overcame the 110 dB anti-aliasing filter in the array and leaked energy into the lower frequency bands. This is most apparent at the time offsets of 1.3, 2.7, 3.2, and 3.5 minutes.

4.4 Cross-correlation of Broadband Modem Noise

The broadband modem noise was the highest SNR signal emanating from the AUV, as seen in Fig. 4.11(a). This was used, along with the propagation delay of sound in water, to estimate the direction of the signal relative to the orientation of the array. With two hydrophones, the time-of-arrival difference was easily measured by cross-correlating their signals. The phones at opposite ends of the array were chosen because they had the maximum separation distance. Cross-correlating time series data from each phone produced another time series in which the time-axis corresponded to time-of-arrival differences. For a sufficiently coherent broadband signal, such as that of the modem, a peak appeared at an offset corresponding to the delay of that signal between the two hydrophones. This was computed in the frequency domain with a pre-whitening filter. Pre-whitening preserves the phase information while forcing the power spectrum to be white. This technique improves results for source signals which are not initially spectrally-white (see Sec. 3.2.4 for details).

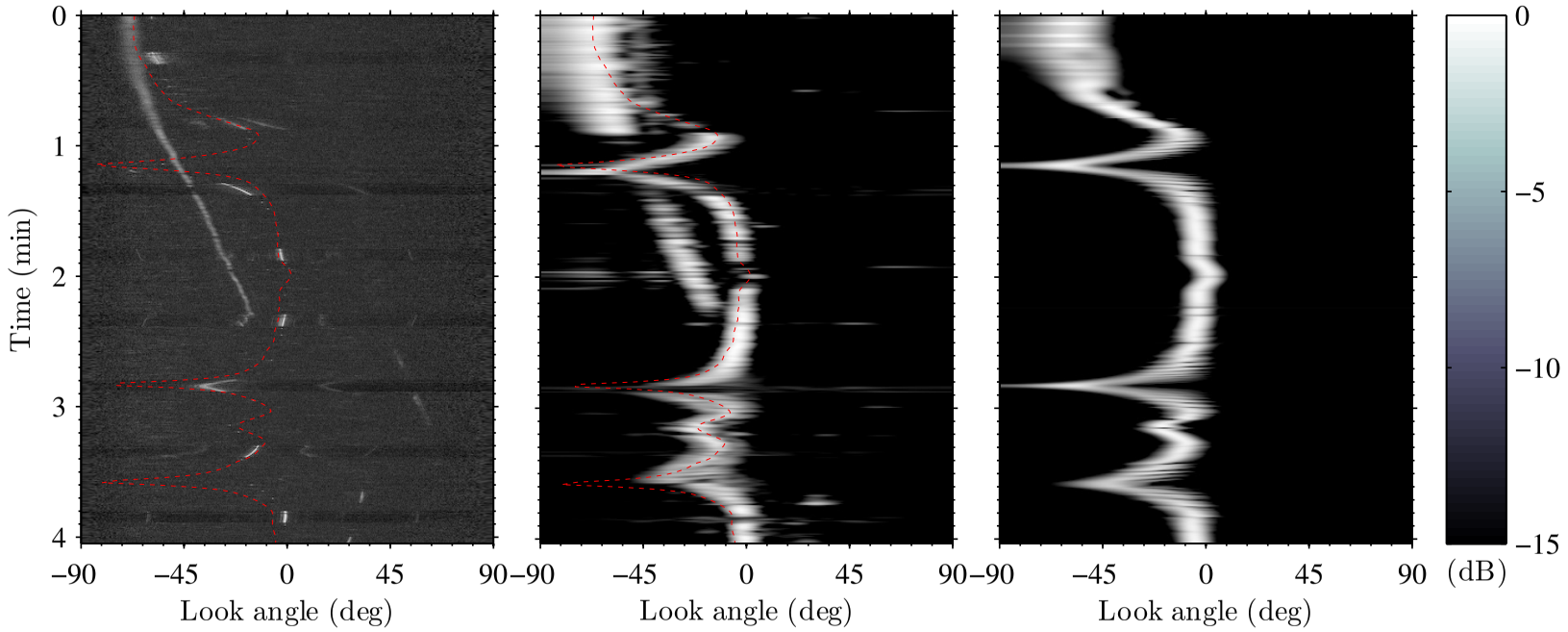
Each peak in the cross-correlation output produced an isodiachron, a hyperboloid

surface in 3-D space (assuming an isospeed), corresponding to all possible target locations having the same time-of-arrival differences. In the far field, this was approximated using the asymptotes of the hyperboloid, which trace out a cone. In 2-D, this reduced to a single direction-of-arrival (DOA) with the angle measured from the broadside and originating from the center of the array.

Fig. 4.12(a) shows a bearing-time-record (BTR) plot in which the cross-correlation time series was converted to DOA angles and plotted along the horizontal axis. The data was divided into short time windows, and DOA angles from each are plotted on the vertical axis. The dashed line shows the DOA angle based on navigational records downloaded from the AUV after the mission. The boat track is visible during the first two minutes but disappeared after its engine was shut off. Strong peaks in the BTR correspond to the modem, but vanish during the time intervals when the modem is not transmitting. A good agreement exists between the navigational records and the cross-correlation analysis during the "on" periods and for sufficient angular separation from the nearby boat.

4.5 Genetic Algorithm for Array Element Localization

A genetic algorithm was used to find the "best set" of coordinates and orientation for the array, where "best set" is defined by that which minimizes an objective function computed as the sum square difference between DOA angles determined from the cross-correlation of modem noise with DOA angles derived from the AUV coordinate data. The DOA angles from cross-correlation were found over an entire four minute dataset when the AUV was near the array. The position and orientation of the array were the adjustable parameters manipulated by the genetic algorithm. From



(a) Cross-correlation: Data (b) Beamforming: Data (c) Beamforming: Oases

Figure 4.12: Bearing-time-record (BTR) plots. Dashed lines are DOA angles computed from navigational records downloaded from the AUV. (a) Two-hydrophone cross-correlation algorithm. The boat track is visible during the first few minutes. The AUV's modem produces a very strong peak every 30 seconds. (b) Conventional narrowband beamforming applied to acoustic data recorded on the array at around 1065 Hz. (c) Same algorithm using simulated data at 1065 Hz created with navigational records from the AUV using Oases. The color scale represents intensity decibels.

each array position, the DOA to each fixed GPS coordinate could be computed. So, the optimization minimized the distances between DOA's while moving and rotating the array to find the best fit. The optimum array location was about 7.6 meters northwest from its initial estimated location, and its orientation rotated about 4 degrees counter clockwise. This was a reasonable amount of correction based on the amount of navigational precision used during deployment.

The Optimization Toolbox in MATLAB (MATLAB, 2010) was used as a framework for running the genetic algorithm. It was configured with a population size of 100. The objective function was defined as

$$f(\mathbf{x}) = \sum_{\forall n} \left[\theta_{n,\text{data}} - \theta_{n,\text{model}}(\mathbf{x})\right]^2 , \tag{4.1}$$

in which $\mathbf{x}$ is the parameter vector for the array location and orientation, θ is the DOA, n is the index of each DOA, of which there are a fixed number in the four-minute period.

4.6 Bearing-Time-Record (BTR) of Narrowband Emissions

In Fig. 4.11(b), several tones can be attributed to the propulsion noise with the strongest centered about 1065 Hz. Given the extremely narrowband nature of this source, a conventional (delay-and-sum) beamformer computed in the frequency domain at the bin corresponding to this specific frequency tracked the DOA angle accurately, as shown in Fig. 4.12(b). A Hann window was applied to individual, overlapping snapshots in the time domain during averaging, and a Taylor window was multiplied to the steering vector for additional side lobe suppression (i.e. array

shading). The maximum achievable gain (in terms of signal power) for an unshaded conventional beamformer on a N element array is

$$10 \log_{10} \left[\frac{N}{1 + (n-1)\rho} \right], \tag{4.2}$$

where ρ is the correlation coefficient of noise (Urick, 1967). For this 23-element array, the array gain was estimated at 11.7 dB, with a standard deviation of 2.6 dB, and a ρ of 0.04. Decibels for array gain are in reference to a unit-less quantity. The correlation coefficient, ρ, is also unit-less.

The cross-correlation BTR in Fig. 4.12(a) localizes targets having broadband signatures, whereas the beamforming BTR in Fig. 4.12(b) is specific to targets producing energy in a very narrow band. However, broadband sources that overlap the narrow band gap were also detected by the beamformer. Hence the boat track, which produced broadband energy from a few Hz to roughly 6 kHz, appears in both the cross-correlation BTR as well as the beamforming BTR (~1065 Hz). Notably, a continuous track for the AUV is visible using beamforming (Fig. 4.12(b)), but only appears in limited segments using cross-correlation (Fig. 4.12(a)).

4.7 Tonal Peak Following

One challenge of this AUV was the wandering nature of the tone seen in Fig. 4.11(c). A simple peak following algorithm was created in an attempt to keep the beamformer locked onto the frequency of the tone. In each row of the spectrogram, the tallest peak was chosen inside of a narrow band centered at 1065 Hz. These data were smoothed using a low pass filter to obtain the curve shown in Fig. 4.11(c). It was expected that

improvement from peak following could be more substantial for data collected over a longer period of time. At the end of the four-hour AUV mission the 1065 Hz tone had drifted to about 1000 Hz, possibly caused by a lower current draw from the battery.

4.8 Propagation Modeling of Narrowband Emissions

For further verification of the AUV track (Fig. 4.12(b)), the AUV's navigational records were passed into the Oases software for simulated snapshots of data at every position. Oases is a full-wave propagation model based on a numerical wavenumber integration solution to the acoustic wave equation (Schmidt, 2004). The simulation was configured with a pressure release surface, flat half-space seabed, and isovelocity water with parameters c = 1500 m/s. The seabed parameters, based on a composition of medium/coarse sand that were consistent with grab samples taken at the deployment site, were c = 1836 m/s, α = 0.88 dB/λ, and ρ = 2.151 g/cm^3 (Ainslie, 2010). The results of this simulated data using the same beamformer are shown in Fig. 4.12(c). There is excellent agreement between the experimental data and the simulated results indicating that the 1065 Hz tone was indeed emitted by the AUV.

4.9 Aspect-Dependent Source Level Measurements

Source level (SL) measurements of the AUV's propulsive emissions were computed by adding the received level (RL) to an estimate of the transmission loss (TL). Oases was used to estimate TL between the AUV source and different elements of the array. The AUV's heading was computed by taking the time-derivative of the AUV's position vector. Ocean currents measured by the observatory were on the order of

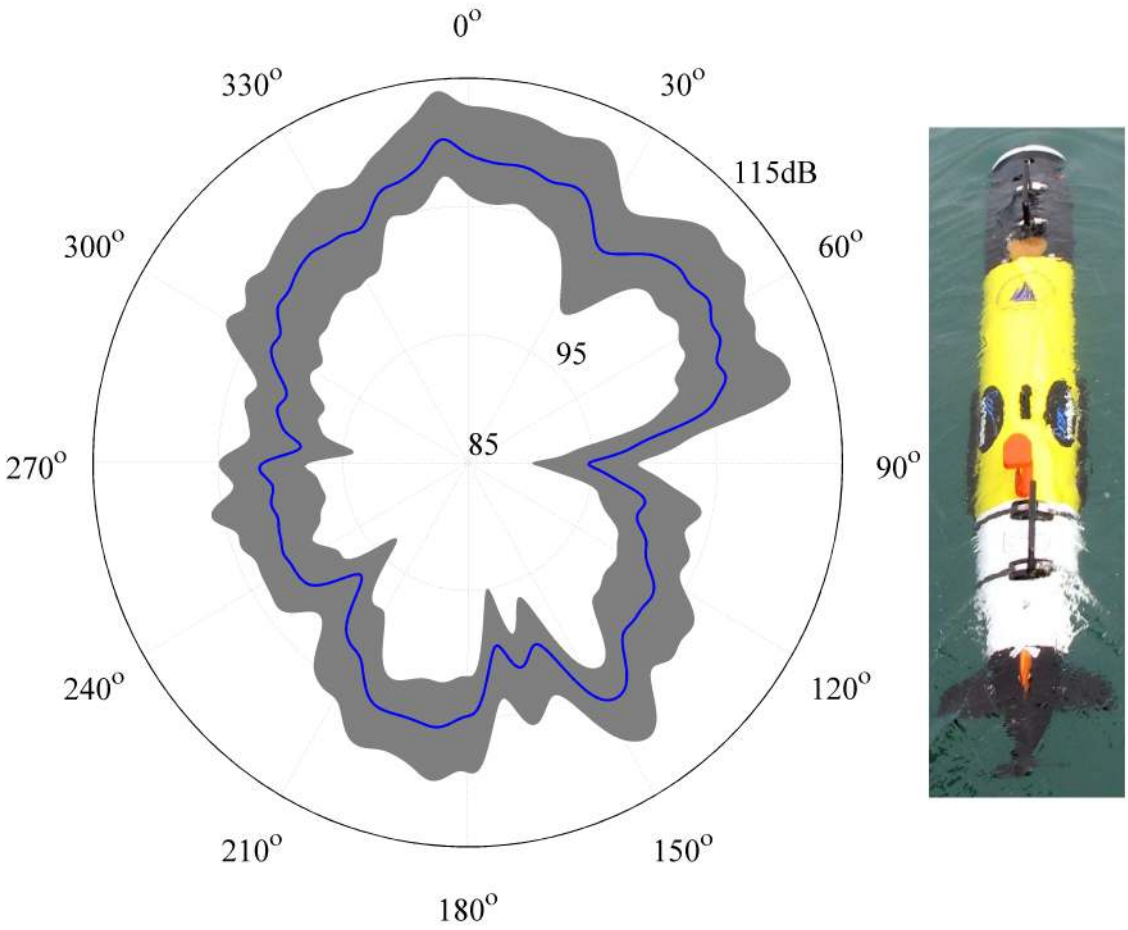


Figure 4.13: Estimated source level (SL) of the REMUS-100 primary propulsion tone as a function of aspect. The mean is represented by the dark line, and the shaded region denotes the interquartile range. The front of the vehicle points towards zero degrees and is viewed top-down, as shown in the image on the right. Units are in dB re 1 μPa^2/Hz at 1 m.

2-3 cm/s, so deviation between the AUV's heading and course were negligible. Combining the AUV position and heading with coordinate estimates of array elements allowed the AUV's aspect relative to each element to be computed. For this study, a subarray of 8 equally-spaced elements from the array were selected and the SL was computed as a function of aspect independently for each element. Since the 1065 Hz tone disappeared when the AUV maneuvered, this had the potential to bias the SL estimate. To avoid this, only time segments corresponding to constant speed, heading, and depth of the AUV were considered. The SL of the propulsive tone as a function of aspect is shown in Fig. 4.13. The dark line represents the mean, and the shaded region denotes the interquartile range. Using multiple elements spread out over the entire array allowed for greater aspect coverage than would have for a single element. An additional 37 minutes of acoustic data was added to this part of the analysis, which included additional AUV passes by the HLA. The mean SL normalized over the full 360° aspect range for the 1065 Hz propulsion tone was 104.8 dB re 1 μPa^2/Hz at 1 m, with a standard deviation of 8.4 dB re 1 μPa^2/Hz at 1 m. The levels reported here are slightly lower than, but consistent with, the values measured using tank experiments reported by Holmes et al. (2010), Holmes (2007). Ambient noise received level (RL) was measured at 60.9 dB re 1 μPa^2/Hz, with a standard deviation of 6.0 dB re 1 μPa^2/Hz. When the AUV was between 10 and 50 meters from the hydrophones, the mean signal to interference plus noise ratio was 19.4 dB with a standard deviation of 7.6 dB. All decibel calculations were consistent with the methods used by Ainslie (2010). Variations of SLs between snapshots were partly attributed to mismatch between actual and modeled environmental parameters. Additionally, localized obstructions, such as coral reefs extending a few meters above the seabed in some areas, might have impeded propagation. Unmodeled factors such as bathymetric variations and surface scattering might also have contributed to the

observed variability.

4.10 Scissorgrams

The SL analysis of the previous section treated each hydrophone as an independent receiver element. In this section, the time-synchronous nature of all the hydrophones in the array is utilized to further analyze the spectral signature of the REMUS-100. The method employed is a scissorgram. A scissorgram is essentially the output of a beamformer steered at a target as it maneuvers. This is shown in Fig. 4.14 in for the initial four minutes of the AUV's track. Figure 4.14(a) is a single-hydrophone spectrogram that shows the two tones being emitted by the AUV. Figure 4.14(c) is a BTR for the strongest tone, and the red overlay shows the bearing of the vehicle based on its navigational records. Figure 4.14(b) is the scissorgram for the red line using a conventional, plane-wave beamformer. It can be seen that the boat noise is greatly reduced as it maneuvers away into a separate beam. The secondary tonal at ∼1280 Hz is also enhanced, and it is evident that it is highly correlated with the primary tonal, which suggests that the respective sources are derived from the same mechanism within the propulsion system.

A similar analysis conducted for a later segment of data in which the AUV is making a maneuver close to the array reveals the directionality if the propulsion tones. Figure 4.15(d) shows the track, in which the AUV is initially traveling away from the array. It makes a left turn bringing the port side in view of the array. Next, it turns to the right. The aspects of the AUV thus are exposited to the array in the following order: stern, port, stern, starboard, and bow. In the BTR of Fig. 4.15(c), strong received levels can clearly be observed initially, which then drop, elevate, drop

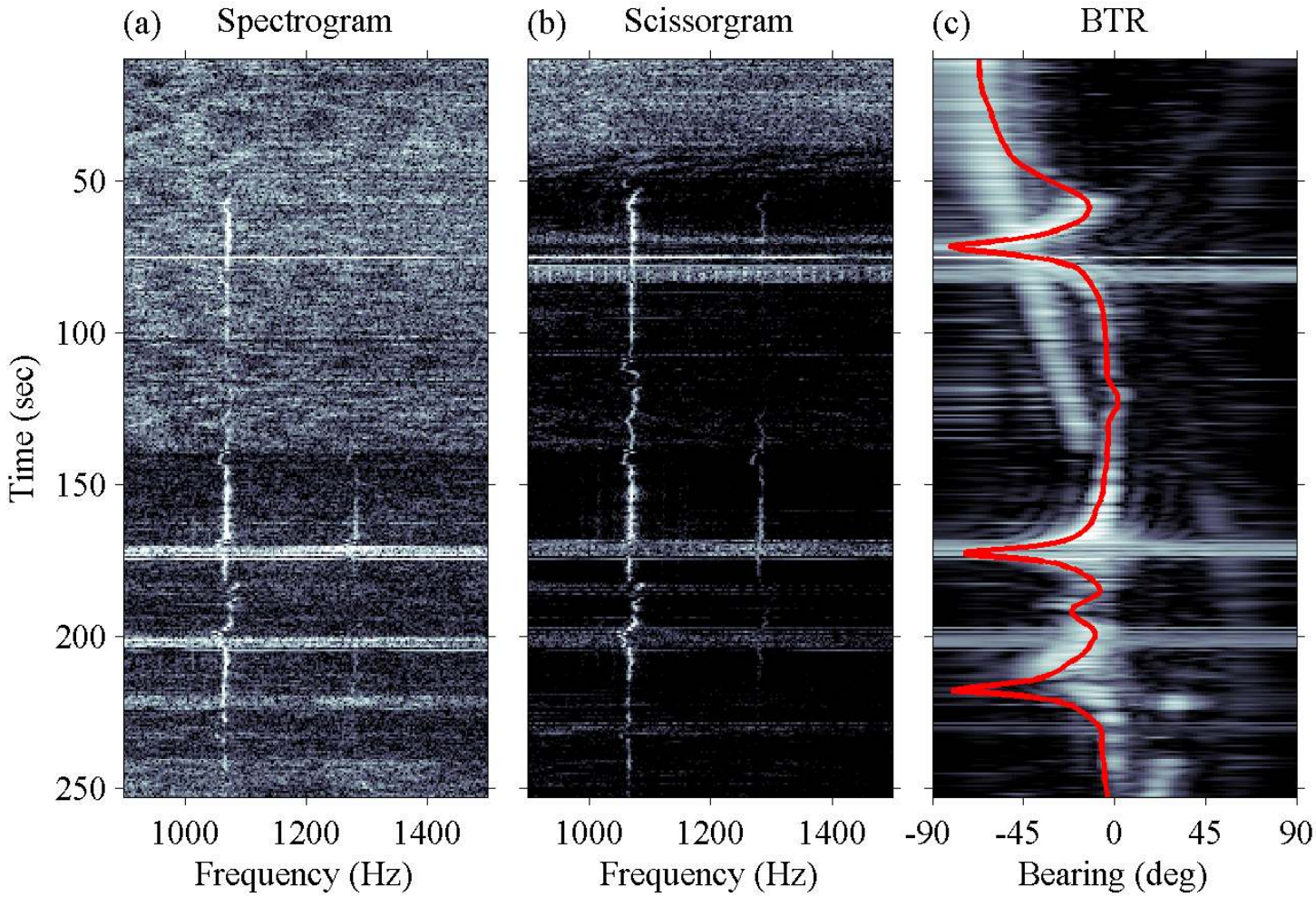


Figure 4.14: Scissorgram showing the REMUS-100 at the start of its track. (a) shows a single hydrophone spectrogram. (b) is computed by beamforming at the red line in (c). The red line in (c) is the expected bearing based on the AUV's navigation records. The boat track is visible in first few minutes of (c). The three features of the AUV track in (c) at times ∼75 s, ∼170 s, and ∼220 s correspond to the AUV passing near one end of the array. All plots use 30 dB of dynamic range.

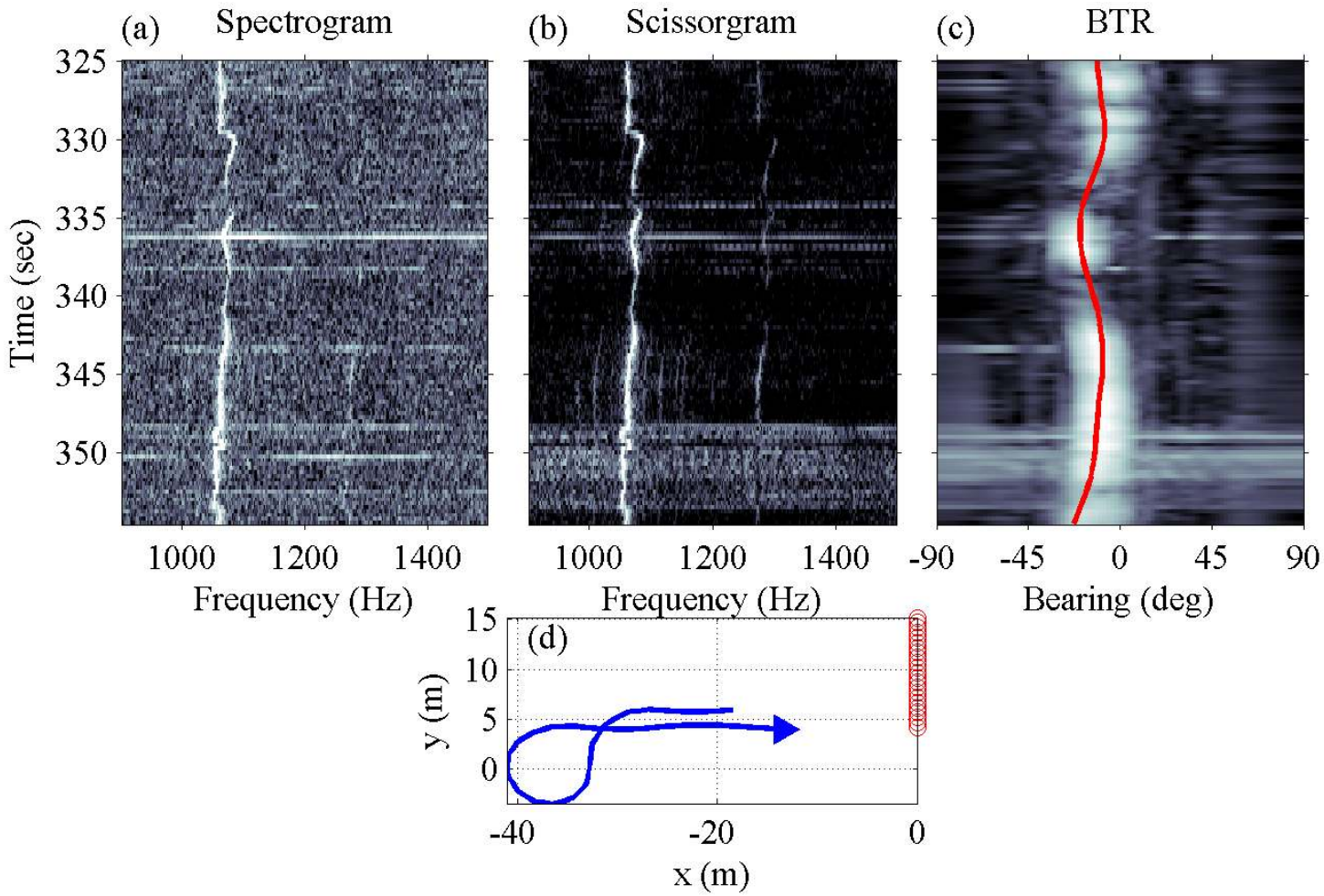


Figure 4.15: Scissorgram showing the REMUS-100 executing a complete turn near the array. (a) shows a single hydrophone spectrogram. (b) is computed by beamforming at the red line in (c). The red line in (c) is the expected bearing based on the AUV's navigation records. (d) shows the track of the REMUS (blue line) relative to the array (red circles). All plots use 30 dB of dynamic range.

again, and finally elevate again. The two drops correspond to the port and starboard aspects. In the scissorgram of Fig. 4.15(b), both tones are clearly evident. An interesting feature occurs around time 330 in which the tone jumps a few Hz before settling. What is likely happening here is this is the point at which the first maneuver is initiated. To turn, the rudder is actuated to one side, which causes additional drag and slows the vehicle slightly. The reduced torque on the propeller causes the motor to increase speed briefly. Each maneuver causes a similar "wobble" in the tones.

The tones emitted by the AUV are a promising sound source that can be used for detection. Experimentally, what are the ranges at which these tones can be heard? Figure 4.16 shows another segment of data in which the AUV is at the far end of its track, out to a distance of a little over 400 meters. First, the AUV is traveling away from the array. It makes a sharp turn to starboard, exposing that aspect to the array. Then, it turns again to starboard, exposing the bow to the array. It is evident in the scissorgram of Fig. 4.16(b) when the starboard aspect is exposed to the array, as the tone disappears into the background noise. The secondary tone is not detectable at this range. Another useful feature of the scissorgram is the fact that it is performing a spatial filter. In the spectrogram, there are several additional tones that are evident, but these do not show in the scissorgram indicating they are arriving on a separate bearing.

4.11 Chapter Summary

This work quantifies, for the first time, the acoustic source levels as a function of aspect of the strongest propulsion-system-borne tone emitted from an underway REMUS-100 AUV by fusing acoustic data from a fixed HLA with the AUV's navi-

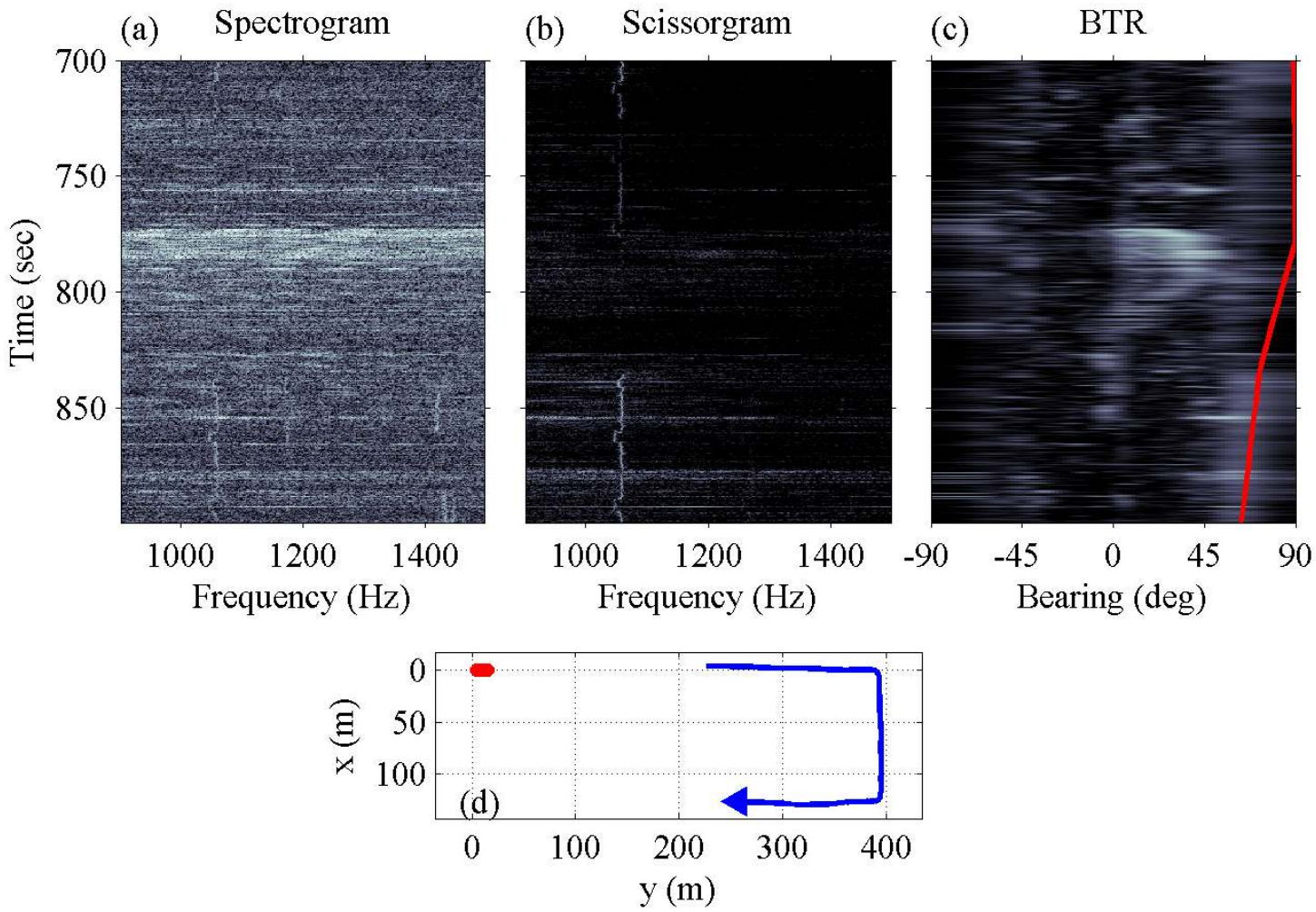


Figure 4.16: Scissorgram showing the REMUS-100 executing two turns far from the array. (a) shows a single hydrophone spectrogram. (b) is computed by beamforming at the red line in (c). The red line in (c) is the expected bearing based on the AUV's navigation records. (d) shows the track of the REMUS (blue line) relative to the array (red circles). All plots use 30 dB of dynamic range.

gation records. The directionality of the tone and the broadband modem signal with respect to the HLA are consistent with the navigational records, thus validating that these signals were indeed emitted by the AUV. Propagation analysis was performed using OASES in which the time-varying TL was calculated between the maneuvering AUV and the array. A complementary simulation of the radiated tone is also performed with OASES enabling discrimination between the AUV target and a boat interferer. These results are published in the Journal of the Acoustical Society of America Express Letters (Gebbie et al., 2012b). An additional contribution to this dissertation is the scissorgram analysis of the radiated noise from the AUV showing the spatial filtering capability of the array and showing detectability out to over 400 m. These results were included in an invited talk at the Acoustical Society of America Meeting in May 2014 in Providence, RI (Gebbie et al., 2014a).

Chapter 5

Two-Hydrophone Processing

Passive acoustic localization has been the subject of recent work in the field of marine mammal bioacoustics. It has been shown that certain species of whales, which use broadband clicks for echolocation, can be localized by measuring the time delay between the direct and bottom or surface-reflected multipaths to estimate range and depth (Aubauer et al., 2000, Nosal and Frazer, 2006). Multipath arrivals are the sequence of echoes of a target's radiated sound as the waves reflect from the surface and seabed in different sequences before reaching the receiver.

Since whale clicks are generally loud and have a short duration (Weirathmueller et al., 2013), the delay between the direct and multipath arrivals can be discerned directly from the received time series (Thode, 2004). Tiemann et al. (2006) showed that by using bathymetry to pre-compute expected arrival times, full 3-D localization could be performed from a single hydrophone. Each arrival corresponds to an eigenray, which is a unique acoustic path between a source and a receiver. For natural variations in bathymetry, if the animal is at a certain range and depth along one bearing, then the relative travel times of the eigenrays differ from those that would be observed if the animal were at the same range and depth along another bearing.

Unlike whale clicks, boat noise is distributed continuously in time, so the same measurement methodology does not directly apply. With boat noise, multipath arrivals overlap in time obscuring individual arrivals and their relative arrival times. Pulse compression is a method of gathering the energy in broadband noise and compressing it into a single broadband pulse, and based on the generalized correlation algorithm from Carter and Knapp (1976). The difficulty that arises with using pulse compression is that it produces peaks for all combinations of multipath arrivals. The problem becomes more tractable if the relative delays between the multipath arrivals are large, which can be realized by placing the hydrophones directly on the seabed.

In Gebbie et al. (2013a), it was shown that two bottom-mounted hydrophones constitute a simple yet favorable geometry for small boat localization, in that bathymetric variations can be leveraged to improve range localization and array side discrimination. Much of that work was presented at the European Conference of Underwater Acoustics in 2012 in Edinburgh, Scotland (Gebbie et al., 2012a). However, beyond giving a description of the algorithm along with experimental validation, it remained an open question as to what the performance limits of the algorithm were, and also what factors gave rise to the array side discrimination capability.

Over the past decades, many studies have been conducted into the performance limits of algorithms for localizing noise-producing targets. This has mainly been done by deriving and calculating the Cramér-Rao lower bound (CRLB). The CRLB represents the lowest achievable variance of an unbiased estimator, and conceptually operates by computing the amount of information passed from a set of hidden variables (i.e. the source location) to the set of observed variables (i.e. received waveforms). These bounds hold under the assumption that only the hidden variables are random, and

all other parameters are deterministic. Friedlander (1988) used this to compute the range and depth accuracy of a submerged source from two vertically aligned receivers using the direct arrival and a single multipath arrival.

Van Trees (2001b) expanded on the CRLB to also handle non-deterministic parameters, in a formulation often referred to as the Bayesian CRLB. The term "Bayesian" is used to indicate that prior information about these random parameters is being used. This was the basis for the study by Hamilton and Schultheiss (1993) that examined the performance impact of having imperfect knowledge of the bathymetry by treating the depth of reflections from the seabed as normally-distributed random variables.

This chapter describes the localization algorithm and demonstrates its operation experimentally. However, to ascertain the theoretical performance of this algorithm, the CRLB of target range and bearing for different hydrophone spacings is derived. A two-hydrophone geometry is adopted, similar to Friedlander (1988), but the Bayesian CRLB approach of Hamilton and Schultheiss (1993) is applied while treating several additional parameters as non-deterministic. This more accurately represents the imperfect knowledge typically available in a real deployment scenario. Bathymetric variations also support an ability to discriminate which side of the array the target is on. This capability is based on slight differences in the first two eigenrays along each leg of the hyperbolic ambiguity. One important question this chapter aims to answer is what kind of bathymetric variations are necessary to support this capability, and what is the probability of localizing the target on the correct side of the array?

A full-length manuscript was recently submitted to the Journal of the Acoustical Society of America (Gebbie et al., 2014b) that expands on Gebbie et al. (2013a) by providing a rigorous performance analysis of the capabilities based on an improved

ambiguity surface function. Additional experimental data was employed to test different hydrophone separation distances, and to validate performance analysis results. Also, the manual processing of correlogram striations was replaced with an automated particle filter design, without any loss of overall algorithm generality.

The rest of this chapter is organized in the following manner. Ray-based models of the multipath structure in shallow water are described (Sec. 5.1), and the processing algorithm is presented (Sec. 5.2). Theory and simulation pertaining to the Bayesian CRLB on range (Sec. 5.3) and factors affecting array side discrimination (Sec. 5.4) are then discussed. The experimental design is reported in Sec. 5.5. The design of a particle filter for extracting correlogram striations is in Sec. 5.6. Experimental results are given in Sec. 5.7.

5.1 Multipath Structure in Shallow Water

This section outlines how boat noise propagation through a shallow water waveguide produces the passive acoustic observations. The geometry of this problem is illustrated in Fig. 5.1(a) and (b). It shows two bottom-mounted hydrophones and a source (target) on the surface. The top two diagrammatic plots in Fig. 5.2 show the received time series, $r_1(t)$ and $r_2(t)$. The bottom plot shows their cross-correlation, $\Gamma_{1,2}(\tau)$. The source waveform is represented as a single pulse for illustration purposes, but for a small boat it would consist of continuous broadband noise. This would result in the $r_1(t)$ and $r_2(t)$ signals appearing as noise, obscuring the clean separation between multipath arrivals. However, this does not affect $\Gamma_{1,2}(\tau)$ because the noise is pulse compressed through the cross-correlation operation. Environmental factors such as bottom loss and rough-surface scattering serve to decorrelate high-

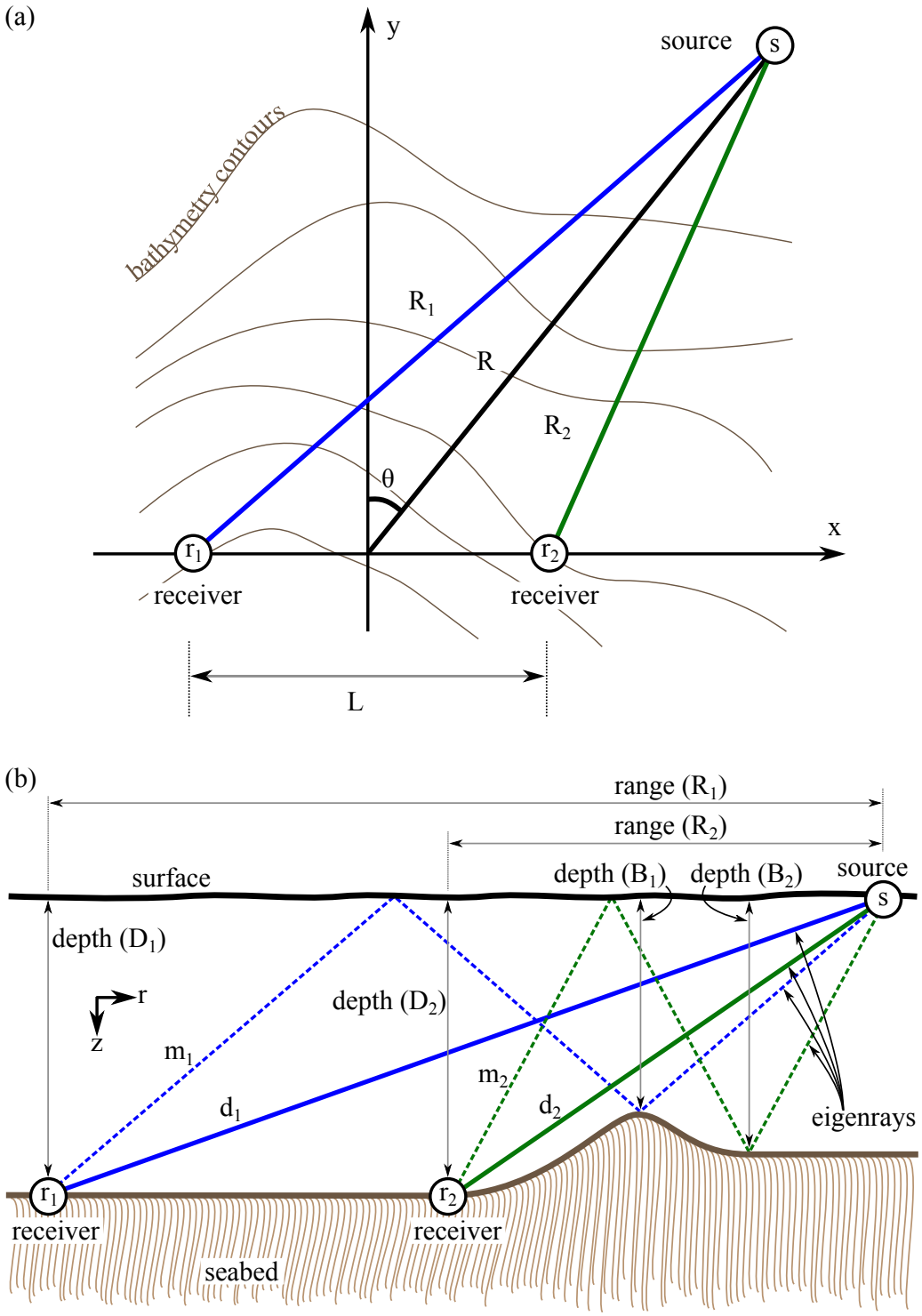


Figure 5.1: (a) A top-down view of the geometry. (b) A side view of the geometry.

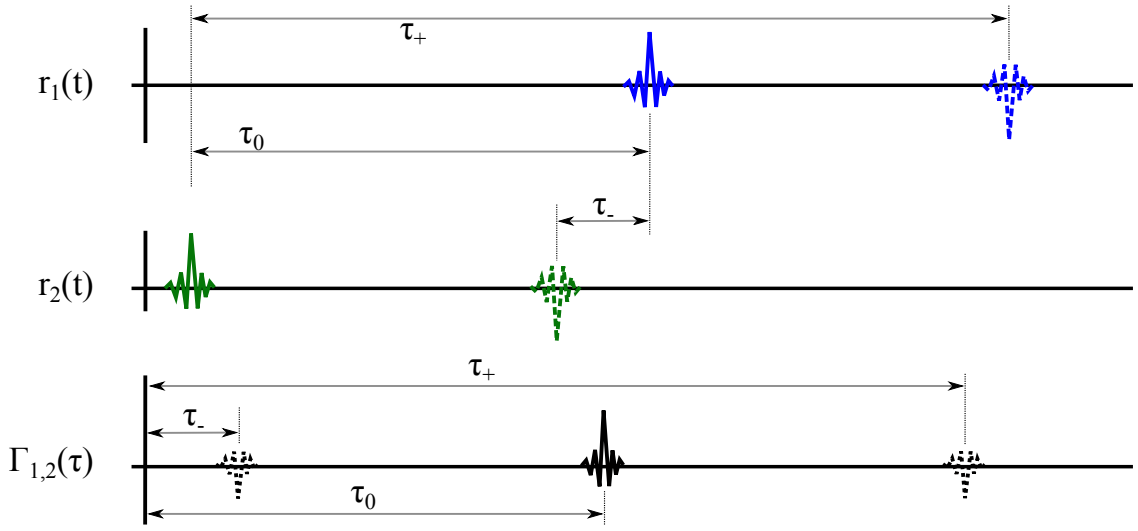


Figure 5.2: Diagrammatic plots of the relative multipath arrival times in Fig. 5.1. The top two plots are the time series at each receiver, and the bottom is their cross-correlation. If the source was continuous boat noise (instead of the pulse in this illustration) the $r_n(t)$ would appear to be solely noise, but the $\Gamma(\tau)$ would still have the peaks due to pulse compression.

order eigenrays, whereas low-order eigenrays often retain enough coherence to appear as stable features in $\Gamma(\tau)$; therefore, only the first-order multipath arrival is used.

Since receiver 1 is farther from the source than receiver 2, both peaks in $r_1(t)$ are shifted later in time to account for the additional travel time. The term τ_0 denotes the time delay between direct arrivals, and $\tau_\pm$ are the delays between the direct arrival at one receiver with the multipath arrival at the other receiver. In $\Gamma_{1,2}(\tau)$, the strongest peak is in the center, with an absolute offset at τ_0, and is used to constrain the ambiguity surface to a hyperbola on the surface. A correlation peak between just the multipath arrivals is not shown in $\Gamma_{1,2}(\tau)$ because it shows up on the time axis close to τ_0, and is often hidden by the direct correlation. This makes measuring that time delay difficult, but information contained in the multipath-only correlation is included in the Bayesian CRLB calculation, described later. As the target initially

moves into the far field of the hydrophone pair, τ_- and τ_+ start to converge but are sufficiently large that the flanking peaks are distinct from the center peak, but this separation eventually vanishes in the distant far field. Bathymetric variations affect only τ_+ and τ_- because the length of reflected eigenrays depend on the depth at the location of the bottom reflection.

Let $t_{n,d}$ and $t_{n,m}$ be the travel times along each eigenray to the n^{th} receiver for direct and multipath rays respectively. The absolute time delays are defined as

$$\tau_0 = t_{1,d} - t_{2,d} \tag{5.1}$$

$$\tau_+ = t_{1,m} - t_{2,d} \tag{5.2}$$

$$\tau_- = t_{1,d} - t_{2,m}. \tag{5.3}$$

Assuming a constant sound speed, the ray travel times can be written as

$$t_{n,d} = \frac{\sqrt{R_n^2 + D_n^2}}{c} \tag{5.4}$$

$$t_{n,m} = \frac{\sqrt{R_n^2 + (D_n + 2B_n)^2}}{c} \tag{5.5}$$

in which the law of cosines yields

$$R_1 = \sqrt{R^2 + (L/2)^2 + RL\sin\theta} \tag{5.6}$$

$$R_2 = \sqrt{R^2 + (L/2)^2 - RL\sin\theta}. \tag{5.7}$$

5.1.1 Cross-Correlation and Cross-Spectrum

The time delays between arrivals contain the information needed for target localization, but these quantities are only observable in the second order statistics of the received signals. While time delay information is contained in both the auto- and cross-spectra, only the cross-spectra is used as an observation since it contains the time delay of the direct arrivals, which is strongly informative of bearing. However, both auto- and cross- spectra are used in Sec. 5.3 to compute the Bayesian CRLB.

The received time series at each receiver is modeled with two arrivals and additive uncorrelated Gaussian noise, $n(t)$. The source waveform is also Gaussian noise represented by $s(t)$. The received waveforms are thus defined as

$$r_1(t) = s(t - t_{1,d}) + \gamma s(t - t_{1,m}) + n_1(t) \tag{5.8}$$

$$r_2(t) = s(t - t_{2,d}) + \gamma s(t - t_{2,m}) + n_2(t), \tag{5.9}$$

in which γ represents the additional amount of propagation loss relative to the first arrival, and is real-valued in the range $(-1, 1)$. The same γ is used in both Eqs. 5.8 and 5.9 based on the assumptions that the hydrophone spacing is small relative to the target range implying that the ray grazing angles on the seabed are similar, and that the composition of the seabed is locally homogeneous in the vicinity of the two reflection points. Phase changes due to boundary reflections are not modeled as this information is later discarded using the envelope operation. The Fourier

transform of the spectra at each receiver is

$$R_1(\omega) = S(\omega)e^{-i\omega t_{1,d}} + \gamma S(\omega)e^{-i\omega t_{1,m}} + N_1(\omega) \tag{5.10}$$

$$R_2(\omega) = S(\omega)e^{-i\omega t_{2,d}} + \gamma S(\omega)e^{-i\omega t_{2,m}} + N_2(\omega) \tag{5.11}$$

Henceforth, we drop the explicit dependence on ω.

Cross-correlation consists of holding one signal constant while sliding the other signal on the time axis, multiplying, and integrating. The fixed signal is referred to here as the "correlated" signal, and the sliding signal as the "reference" signal. Let r_1 be the correlated signal and r_2 be the reference signal. The cross-spectrum is computed by multiplying the correlated spectrum by the complex conjugate of the reference spectrum,

$$K_{1,2} = \mathrm{E}[R_1 R_2^*] \tag{5.12}$$

$$= |S|^2 e^{-i\omega(t_{1,d}-t_{2,d})} + |S|^2\gamma^2 e^{-i\omega(t_{1,m}-t_{2,m})} + |S|^2\gamma e^{-i\omega(t_{1,m}-t_{2,d})} + |S|^2\gamma e^{-i\omega(t_{1,d}-t_{2,m})} \tag{5.13}$$

in which E[ ] is the expected value operator, and $*$ denotes complex conjugation. The second term in Eq. 5.13 is the correlation of multipath arrivals on each receiver, and it is effectively hidden by the correlation of direct arrivals (first term). This is because both will roughly have the same time delay but the second term has a lower amplitude due to γ^2.

For a finite bandwidth signal with a flat spectrum, each exponential term will manifest in the time domain as a sinc function. For a baseband signal of bandwidth b, the sinc function is defined as $\sin(2\pi b)/(2\pi b)$. If γ were complex and had a constant phase

with respect to frequency, as might be the case with sub-critical bottom reflections from a single-layer seabed, this would cause skewing of the sinc function such that the maximal value would not align exactly with the time delay. The envelope operation is thus applied to produce positive valued peaks having a maximal value at the time delay, which simplifies time delay estimation. The envelope of an arbitrary signal $x(t)$ is computed by $|x(t) + \mathcal{H}[x(t)]|$, in which $\mathcal{H}$ is the Hilbert transform. This is equivalent to computing the inverse Fourier transform of only the positive frequencies, followed by taking an absolute value and multiplying by a factor of two.

A Gaussian is a suitable approximation to the envelope of a sinc function, and is used here to model the time domain pulses. It is defined a

$$g(x, a, \sigma) \triangleq e^{-\frac{1}{2}\left(\frac{x-a}{\sigma}\right)^2}. \tag{5.14}$$

Neglecting the γ^2 term in Eq. 5.13, the cross correlation time series is then

$$\Gamma_{1,2}(\tau) \approx g(\tau, \tau_0, \sigma) + \gamma g(\tau, \tau_-, \sigma) + \gamma g(\tau, \tau_+, \sigma) \tag{5.15}$$

in which the time axis is $x = \tau$, and the pulse offsets occur at $a \in \{\tau_0, \tau_-, \tau_+\}$. The width of the pulse depends on the signal bandwidth b, and since the variable σ specifies the half-width of the Gaussian, it is defined as

$$\sigma \triangleq \frac{1}{2b}. \tag{5.16}$$

The auto-spectrum for each receiver is defined as

$$K_{1,1} = \mathrm{E}[R_1 R_1^*] \tag{5.17}$$

$$= |S|^2 + |S|^2|\gamma|^2 + |S|^2\gamma e^{-i\omega(t_{1,m}-t_{1,d})} + |S|^2\gamma^* e^{-i\omega(t_{1,d}-t_{1,m})} + |N_n|^2 \tag{5.18}$$

$$= |S|^2\{1 + |\gamma|^2 + 2|\gamma|\cos[\omega(t_{1,m} - t_{1,d})]\} + |N_1|^2 \tag{5.19}$$

with a similar definition for $K_{2,2}$. This is required later in the calculation of the range CRLB in Sec. 5.3. Note that unlike the cross spectrum, the auto spectrum includes the noise term. The next section describes the methodology for processing the acoustic data into a localization.

5.2 Localization Algorithm

This section describes the localization algorithm. It consists of the following parts: the extraction of multipath time delays, the ray model, and the ambiguity function. The first step processes the acoustic data, the second step runs the model at discretized points on the surface, and the third step combines them into an ambiguity surface that yields the localization.

5.2.1 Acoustic Processing

The measured cross-correlation function, $\Gamma(\tau)$, is obtained from raw data by the following procedure. A snapshot of data is obtained from the channels; the time span of this data constitutes the total averaging time and should be short enough so that the effects of target motion are negligible. Overlapping segments are formed within the

snapshot using the weighted overlapping segment averaging method (Carter et al., 1980). Segments are windowed using a Hann function, then zero-padded to twice the original length to avoid wrapping effects of the discrete Fourier transform (DFT). Time windowing provides better spectral estimates at the expense of a small amount of frequency resolution. The cross-spectrum is computed for each segment by multiplying one channel with the complex-conjugate of the result, followed by a bandpass filter. The average for the snapshot is computed across all segments. The resulting cross-spectrum is pre-whitened. Pre-whitening preserves phase information while enforcing a flat power spectrum and is defined as $X(\omega)/|X(\omega)|$ for an input spectrum $X(\omega)$, and is based on the phase transform (PHAT) algorithm (Knapp and Carter, 1976). This prevents loud tonal components that may stand out in the target's acoustic signature from corrupting $\Gamma(\tau)$. Lastly, this is transformed back into the time domain with the inverse DFT while simultaneously taking the envelope.

A correlogram is then formed by stacking $\Gamma(\tau)$ from each snapshot vertically. Viewed in this manner, striation lines appeared that correspond to the correlations of individual arrivals as they evolve over time. The center striation, which has an offset of τ_0 at a given snapshot, is often the strongest as it corresponds to correlations of the direct arrivals on each hydrophone. The nearest flanking striations are the multipath arrival at one hydrophone correlating with the direct arrival at the other hydrophone, and have offsets $\tau_\pm$ for a given snapshot.

A sequential importance resampling (SIR) particle filter (Ristic et al., 2004) is then used to extract the striation offsets from the correlogram. The output of the tracker are the measured values τ_0, τ_-, and τ_+ at each snapshot. The details of the tracker construction are given in Sec. 5.6.

5.2.2 Ray Model Processing

In the second stage of processing, predictions of eigenray propagation-time differences, $\bar{\tau}_0(\mathbf{x})$, $\bar{\tau}_-(\mathbf{x})$, and $\bar{\tau}_+(\mathbf{x})$, are computed for each possible (Easting, Westing) target position, $\mathbf{x}$. A ray tracer, such as described in Jensen et al. (2011b), can be used to compute the travel time of the direct and bottom-surface eigenrays between each receiver and each point $\mathbf{x}$ on the surface. An eigenray between a point on the surface and a receiver on the seabed in a region of varying bathymetry could involve a path that is not confined to a vertical plane due to three dimensional seabed variations. While this could be done in three dimensions, a simpler N-by-2D approach is used here. A N-by-2D ray model invokes a 2D ray tracer separately along N radial lines for at each receiver, and the bathymetry is interpolated along each radial line. This flattens the problem to two dimensions, range and depth, and a single ray fan starting at the receiver can be used to determine the eigenrays for all points along the radial which greatly reduces computational complexity. The $\bar{\tau}_0(\mathbf{x})$, $\bar{\tau}_-(\mathbf{x})$, and $\bar{\tau}_+(\mathbf{x})$ at each radial point can then be gridded using a technique such as Delaunay triangulation (de Berg, 2008). A constant water sound speed profile (SSP) is adopted in this analysis, but the technique can be readily applied to environments having a varying SSP. The choice of constant SSP is based on the experiment described in Sec. 5.5, which was close to isovelocity.

5.2.3 Ambiguity Surface for Target Location

In the last stage of processing, the measured time delays are matched against modeled time delays computed at regular points on the surface. An ideal ambiguity function

would be a delta function at the true target location, but in practice the aim is to maximize the value at the true target location relative to all other locations. At the true target location, the error between the predicted time delays and the measured time delays will be minimized, provided the model is sufficiently accurate. It is important to point out that *all* three delay errors that are minimized together at this point: $|\tau_0 - \bar{\tau}_0(\mathbf{x})|$, $|\tau_+ - \bar{\tau}_+(\mathbf{x})|$, and $|\tau_- - \bar{\tau}_-(\mathbf{x})|$. When $|\tau_0 - \bar{\tau}_0(\mathbf{x})|$ is at a minimum, this corresponds to a hyperbola on the $\mathbf{x}$ plane. When the other two are also at a minimum, this corresponds to a range at some distance down one leg of the hyperbola. However, since the hyperbola has two legs, this also corresponds to a "false" target position on the other leg. The false target position may retain a greater amount of overall error due to bathymetric variations.

Thus, in order to constrain the ambiguity function to produce a large value at the true target position, it is constructed as the product of three Gaussians. Using Eq. 5.14,

$$\Phi(\mathbf{x}) \triangleq g[\tau_0, \bar{\tau}_0(\mathbf{x}), \sigma] \times g[\tau_+, \bar{\tau}_+(\mathbf{x}), \sigma] \times g[\tau_-, \bar{\tau}_-(\mathbf{x}), \sigma]. \tag{5.20}$$

The first term constrains the target location to the hyperbola determined by the direct arrivals. The latter two terms are maximized at the true target range, and possibly the false target range. Values of $\Phi(\mathbf{x})$ are in the range $(0, 1]$ and are interpreted as yielding information about the relative certainty of the target being at a particular location, $\mathbf{x}$, on the water surface.

5.3 Cramér-Rao Lower Bound on Range

This section examines the theoretical limits on how accurately the target range can be determined. Localization is an estimation problem that uses a set of observed quantities (i.e. acoustic data) to estimate the value of a set of hidden parameters that cannot be directly observed (i.e. the location of the target). Different algorithms may estimate the hidden parameters with different degrees of variance, but the question that naturally arises is what is the best (lowest) possible variance that can be achieved? The CRLB gives a theoretical lower bound on the variance of an unbiased estimator. This operates by considering the amount of information about the hidden parameters that is contained in the observed quantities. As the amount of information about a hidden parameter increases, the variance with which it can be estimated decreases, and vice-versa. In fact, the definition of the CRLB is that it is the inverse of the Fisher information matrix. The i, j element of the Fisher information matrix is defined as

$$J_{D_{i,j}} = \mathrm{E}\left[\frac{\partial \log p(y|\xi_i)}{\partial \xi_i}\frac{\partial \log p(y|\xi_j)}{\partial \xi_j}\right], \tag{5.21}$$

in which $p(y|\xi)$ is the conditional probability density function of an observed quantity y given a hidden parameter ξ.

The basis for this analysis starts with an article by Friedlander (1988), in which the target range and depth CRLB were derived. In that scenario, Gaussian noise is emitted by a submerged source and travels along two paths to a pair of vertically arranged receivers. One path is direct and the other reflected. Propagation delays are specified as a function of source range and depth, and are considered the hidden

parameters. Instead of applying Eq. 5.21 directly, Whittle's theorem [Whittle (1953), Theorem 9] is invoked to compute the Fisher information matrix by integrating a kernel based on the cross spectral density matrix and its derivatives over a band of frequencies. In that formulation, the i, j element of the Fisher information matrix is

$$J_{D_{i,j}} = \frac{N}{2\pi f_s} \int_{\omega_{\min}}^{\omega_{\max}} \mathrm{tr} \left[\frac{\partial K}{\partial \xi_i} K^{-1} \frac{\partial K}{\partial \xi_j} K^{-1} \right] \mathrm{d}\omega, \tag{5.22}$$

in which tr is the trace operator, $[\omega_{\min}, \omega_{\max}]$ is the signal frequency band, ξ_i and ξ_j are parameters (e.g. range, bearing), N is the number of samples, and f_s is the sample rate. The term K is the cross spectral density matrix for which $K_{1,1}$ and $K_{2,2}$ are described by Eq. 5.19, and $K_{1,2}$ is defined by Eq. 5.13. Note that $K_{2,1} = K_{1,2}^*$. All the parameters (the ξ's) are assumed to be hidden. Parameters that are deterministic are simply used in the calculation of K and the derivatives of K with respect to the hidden parameters. In Friedlander (1988), the environment is deterministic, and the only hidden parameters are the target range and depth.

The Fisher information divides parameters into two categories: those for which no prior information is available and those that are known exactly. Obviously, not all problems can be defined in this way. However, Van Trees (2001b) proposed a "Bayesian" version of the CRLB, often referred to simply as the Bayesian CRLB, that allowed a hidden parameter to have some (but not necessarily perfect) prior information. Hamilton and Schultheiss (1993) used this formulation to determine the CRLB for a target's range, but treated the depth of the multipath reflection from the seabed as a hidden parameter having a finite prior variance.

The Bayesian "prior" information about the variances of the parameters is specified

in the matrix

$$J_P = \text{diag}[\text{var}(\xi_1)^{-1}, \text{var}(\xi_2)^{-1}, \ldots, \text{var}(\xi_N)^{-1}], \tag{5.23}$$

in which diag puts the elements on the diagonal of a matrix with zeros off the diagonal, and var(ξ) is the variance of ξ. The total Fisher information matrix is then

$$J_T = J_D + J_P, \tag{5.24}$$

and the lower bounds on the variances of the individual parameters fall on the main diagonal of

$$\text{Bayesian CRLB} = J_T^{-1}. \tag{5.25}$$

At the extremes, an infinite prior variance implies no prior information about that parameter exists, which drives that element of J_P to zero thus increasing the CRLB. Conversely, a small variance for the prior implies accurate knowledge of a parameter, thus increasing the value in J_P and decreasing the CRLB. Intuitively, as more prior information about a parameter is included (corresponding to a larger value somewhere on the diagonal of J_P), that parameter, and potentially other parameters as well, can be estimated with a smaller variance corresponding to an overall smaller CRLB.

Whittle's theorem, Eq. 5.22, requires second-order derivatives of the cross spectral density matrix with respect to the parameters. For this reason, constant sound speeds are assumed throughout much of the literature since relationships between the water sound velocity profile and the observed relative ray travel times are difficult to determine analytically. For simplicity, a constant sound velocity profile is adopted here

Parameter (ξ)	Value (μ_ξ)	Bayesian prior (σ_ξ)
R	100-1500 m	∞
θ	45°	∞
L	11, 2 m	0.02 m
D_1	120 m	2 m
D_2	120 m	2 m
B_1	120 m	10 m
B_2	120 m	10 m
γ	-0.9	0.01
c	1530 m/s	5 m/s

Table 5.1: Hidden Parameters for CRLB Calculation

and used to show what affect treating other parameters as uncertain has on range localization accuracy. The derivatives of K are taken for all pairs of the parameters R, θ, L, D_1, D_2, B_1, B_2, γ, and c, such that J_D is a 9×9 matrix. These derivatives are computed analytically using a symbolic math engine (MathWorks, 2013) and used directly in the simulation described in the next section. The individual equations that make up J_D are large, and the details are not particularly germane so are omitted here. Many other works have presented explicit derivations [e.g. Friedlander (1988), Hamilton and Schultheiss (1993)], and the method presented here follows essentially the same procedure.

5.3.1 Simulation

In this simulation, the Bayesian CRLB for range (specified as σ_R) and bearing (specified as σ_θ) are computed as a function of the true target range. This is done for long and short hydrophone separation distances in order to examine what effect this has. The hydrophone separation is also modeled to have a small amount of error. The seabed is assumed to be roughly flat with some variations, and the depth of the receivers also has a slight amount of error. The water has a constant sound

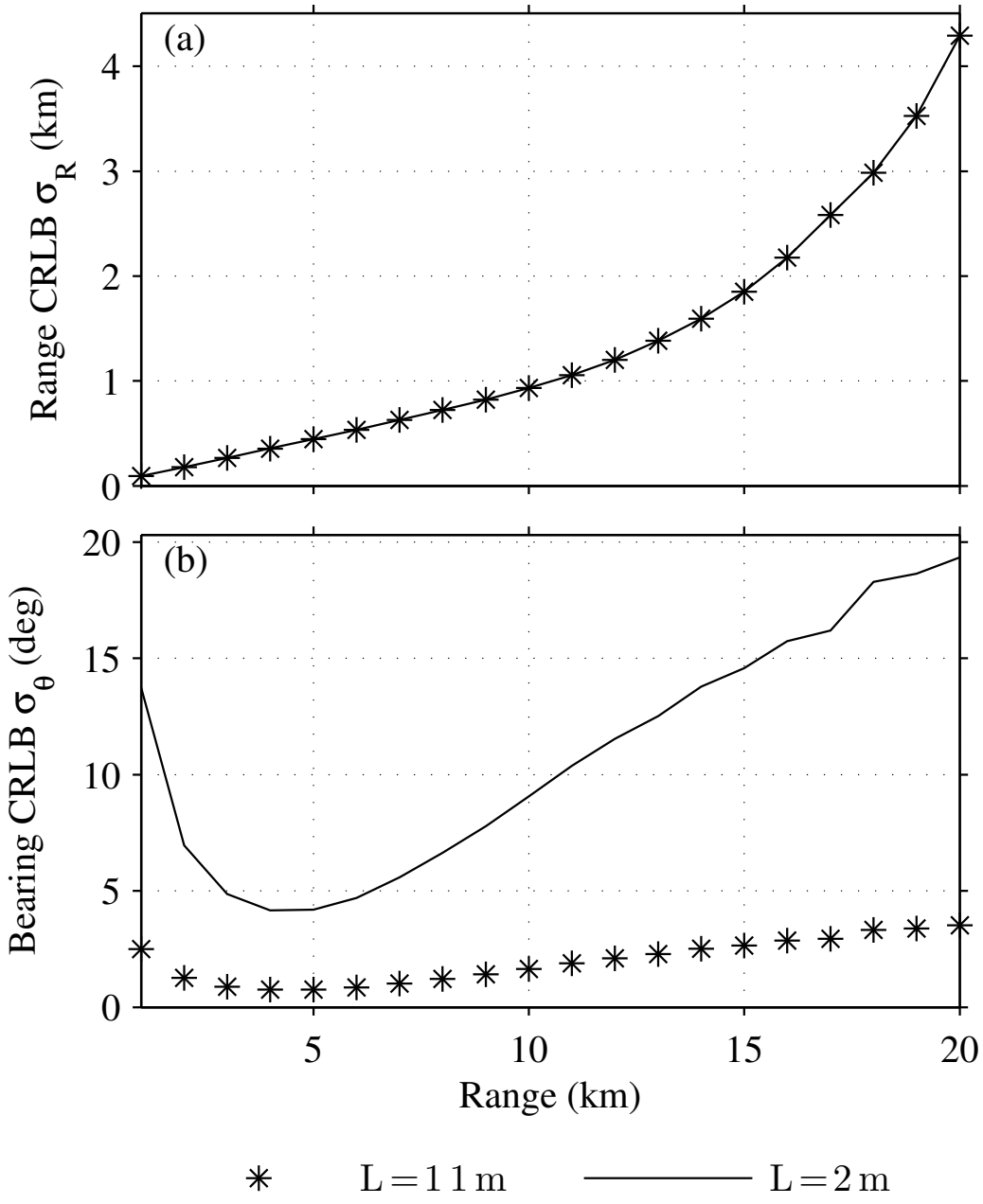


Figure 5.3: (a) Bayesian CRLB for range estimation as a function of true range. Starred and solid lines representing two phone separations essentially overlap indicating that this factor does not impact range resolution. (b) For comparison, the same metric is applied to bearing estimation, in which the starred and solid lines show significant deviations, which indicates bearing resolution decreases as the phone separation decreases.

speed, but there is some uncertainty in the exact speed. In Eq. 5.22, $N = 4096$, $f_s = 102.4$ kHz, $[\omega_{\min}, \omega_{\max}] = 2\pi[0, 3000]$. The source level emits 130 dB and the noise level is 60 dB. As the target range increases, the signal to noise ratio (SNR) at the receiver decreases. The parameter values (μ_ξ) for the simulation are given in Table 5.1, along with the Bayesian priors (σ_ξ). Range (R) and bearing (θ) are the parameters under investigation, and so their infinite priors cause their entries in J_P to be zero.

The Bayesian CRLB for range (R) and bearing (θ) are shown in Fig. 5.3(a) and (b), respectively, for both hydrophone separations (L). As the target moves farther from the array, the lower bound on range estimation variance increases, which corresponds to the direct and multipath eigenray travel times converging, and therefore containing less information about the target range. The lines for the 2 m and 11 m hydrophone separations essentially overlap, indicating that hydrophone separation does not affect the lower bound on range variance. However, it does have a significant impact on the bearing estimation, σ_θ, shown in Fig. 5.3(b), as expected. The increase in bearing CRLB at closer ranges are due to the uncertainty in the receiver depths.

5.4 Array Side Discrimination

An interesting feature of the ambiguity surface defined in Eq. 5.20 is that in addition to estimating the target range and bearing, it can also predict (with some uncertainty) which "side" of the array the target is on. This predictive capability is manifest as a taller peak in the ambiguity surface on the "true" side of the array (the side of the array where the target is actually located). This section examines what aspects of the ambiguity surface construction give rise to this capability, and explores the factors

that support it through simulation.

Consider a single snapshot of acoustic data that is short enough to safely neglect target motion. The bathymetry is varied such that depths on one side of the array differ from the corresponding "mirror" points on the other side of the array. This difference affects propagation which causes the multipath travel times to differ. Equation 5.20 will now produce an ambiguity surface with two peaks, one on "true" side and one on the "false" side. The height of these peaks depend on error between modeled and measured time delays for each pair of arrivals on each side. These errors are random variables based on the inherent uncertainty in measuring time delays with the generalized cross correlator (Carter and Knapp, 1976). Errors due to environmental uncertainty are ignored and the environment is assumed to be perfectly characterized.

On the true side, there exists a possible target position on the surface for which these random variables will be normally distributed with zero mean, i.e. the true target position. However, such a position does not necessarily exist on the false side, implying that some of the means will deviate from zero. The first term of Eq. 5.20 constrains the ambiguity surface to a hyperbola based on the delay between direct arrivals. This is oriented with a leg extending onto each side of the array. Conceptually, sliding along the hyperbola on the true side, at some point the multipath delays will exactly match the modeled predictions (no environmental uncertainty) at the true target range. However, on the false side, the multipath delays may become close, but do not perfectly match the modeled predictions at the false target range. This small amount of error, if it exists, is amplified by Eq. 5.20 and is often sufficient to produce a taller peak on the true side.

Let point $\mathbf{x}_t$ be the (x, y) position of the peak on the true side, and $\mathbf{x}_f$ be the tallest

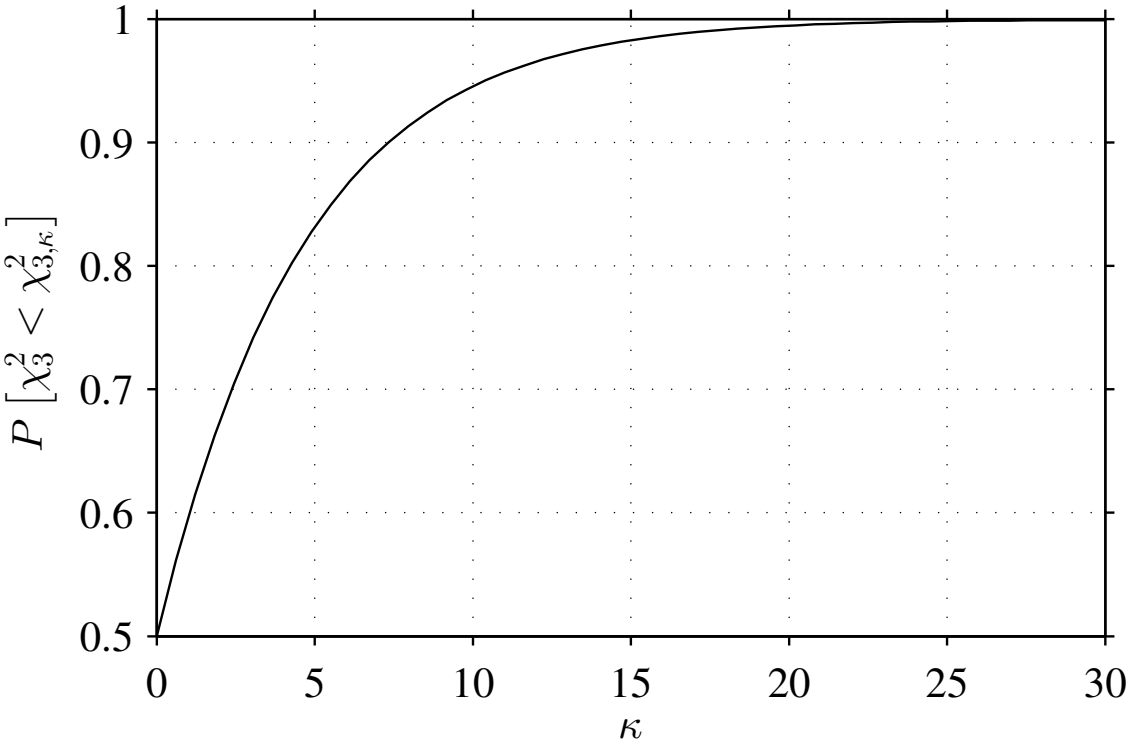


Figure 5.4: Probability that a chi-squared random variable with 3 degrees of freedom, χ_3^2, is less than a non-central chi-squared random variable with 3 degrees of freedom, $\chi_{3,\kappa}^2$, as a function of the non-centrality parameter, κ.

peak on the false side. Ultimately, the question being asked is what is the probability that the peak on the true side is taller than the peak on the false side? This quantity is expressed as

$$P_t \triangleq P\left[\Phi(\mathbf{x}_t) > \Phi(\mathbf{x}_f)\right]. \tag{5.26}$$

Let

$$\alpha_b^a \triangleq \left[\tau_a - \bar{\tau}_a(\mathbf{x}_b)\right]^2, \qquad \begin{array}{l} a \in \{0, -, +\} \\ b \in \{t, f\} \end{array}. \tag{5.27}$$

Combining the terms of Eq. 5.20, it can be shown that Eq. 5.26 is equivalent to

$$P_t = P\left[\alpha_t^0 + \alpha_t^+ + \alpha_t^- < \alpha_f^0 + \alpha_f^+ + \alpha_f^-\right]. \tag{5.28}$$

To analyze each side of this inequality, note the measured values (τ_a) are normally distributed random variables. On the true side, their means equal the corresponding predictions $\bar{\tau}_a$, so $\tau_a - \bar{\tau}_a(\mathbf{x}_t)$ are zero-mean normally distributed random variables. However, on the false side, there may be some error between the mean of τ_a and the prediction $\bar{\tau}_a$, leading to $\tau_a - \bar{\tau}_a(\mathbf{x}_f)$ being normally distributed random variables but possibly having a non-zero mean.

Since each α_t^a is the square of a normal random variable, they are each χ_k^2 distributed with $k = 1$ degrees of freedom. Added together, the three α_t^a's, the left side of the inequality form a χ_k^2 random variable with $k = 3$ degrees of freedom. On the right side, the α_f^a random variables are the product of normal random variables that have the same variance, but do not necessarily have zero mean, so these have non-central

$\chi^2_{k,\kappa}$ distributions (Zelen and Severo, 1972). Since

$$\alpha_t^0 + \alpha_t^+ + \alpha_t^- \sim \chi_k^2 \quad \text{and} \quad \alpha_f^0 + \alpha_f^+ + \alpha_f^- \sim \chi_{k,\kappa}^2, \tag{5.29}$$

then

$$P_t = P[\chi_3^2 < \chi_{3,\kappa}^2]. \tag{5.30}$$

The parameter κ is the non-centrality parameter defined as

$$\kappa = \sum_{\forall a} \left(\frac{\tau_a - \bar{\tau}_a(\mathbf{x}_f)}{\sigma_{\tau_a}} \right)^2, \tag{5.31}$$

so the probability P_t depends entirely on κ. The parameter σ_{τ_a} is the standard deviation of τ_a, and depends on the accuracy with which the time delay can be measured. This relationship between central and non-central χ^2 distributions is shown in Fig. 5.4. Equation 5.30 can be rearranged as the distribution of the difference between two random variables, which can be computed by correlating their respective distribution functions, or alternatively conjugate-multiplying their characteristic functions. When $\kappa = 0$, the non-central $\chi^2_{k,\kappa}$ converges to a central χ^2_k and the probability that one is less than the other is 0.5, which would imply there is no time delay error on the false side. In short, Eq. 5.31 says that the accumulation of normalized squared time delay errors on the false side determine the array side prediction performance. This is analyzed with a simulation in the next section.

5.4.1 Simulation

Consider an environment with a planar but not flat bathymetry. In other words, the bottom lies in a plane but does not coincide with the $z = D$ plane. The plane has a constant slope such that one side of the array is deeper than the other. Array side discrimination is based on one side having different multipath characteristics than the other, so the gradient vector can not be perfectly parallel to array. Rather, it is the component of the gradient that is perpendicular to the array that is relevant. This is important because it says that to perform side discrimination, sensors deployed on a slope should be placed along bathymetric contours as opposed to an uphill-downhill arrangement. With an uphill-downhill arrangement, the bathymetry along the true and false bearings are identical, which eliminates differences in the multipath that are critical to this capability. In this simulation, the gradient vector is perpendicular to the array, and the sensors are on the bottom in 120 m of water. The simulation places the target on the deeper side at 1 km range at a bearing of 45°. The sound speed in the water is assumed to be constant with boundary reflections that are lossless.

In this case, the environment is known perfectly so the measured time delays are solely the result of correlating Gaussian noise signals in the presence of additive Gaussian noise. Hahn and Tretter (1973) derived the CRLB for the variance of the time delay in this scenario. Knapp and Carter (1976) further showed that this bound is reached with the generalized cross correlation algorithm, which is the same method used here. If the signal and noise have flat spectra in the band $[f_{\min}, f_{\max}]$ (using positive frequencies only), this bound is specified as

$$\sigma^2 = \frac{3f_s\left[1 + 2(\mathrm{SNR})\right]}{4\pi N\left(f_{\max}^3 - f_{\min}^3\right)(\mathrm{SNR})^2} \tag{5.32}$$

where N is the number of samples, f_s is the sample rate, and the SNR (in linear units, not decibels). The values used in the simulation are $f_s = 102400$ Hz, $N = 4096$, $f_{\min} = 0$ Hz, $f_{\max} = 3000$ Hz, and SNR = 10 dB.

The simulation is based on image theory. Since the seabed is a plane, the target is reflected over it to its image position. The receivers are then reflected over the surface to their image positions. Multipath time delays between the source and each receiver are computing using the distance between the image source and the image receivers. The ambiguity surface, Eq. 5.20, is computed on the shallow side to determine the location of false target. The false target position corresponds to a set of time delays that are close to, but do not necessarily match those of the true target position. These time delay errors exist for both direct-with-direct and direct-with-multipath correlations. From these errors, Eq. 5.31 is invoked along with Eq. 5.32 to determine κ, which maps directly to a probability value that specifies whether the peak at the true target position will be taller. The results of this simulation are shown in Fig. 5.5 for variations in the seabed slope and receiver separation.

As the seabed becomes more sloped, the side discrimination capability increases. Conversely, a flat seabed eliminates the unique propagation characteristics from each side of the array, which removes the array side discrimination capability. Consequently, the probability of choosing the correct side becomes 0.5. The same behavior is observed for the separation distance between receivers. As the receivers become closer together, even on a sloped seabed, the multipath rays from the target to each receiver become increasingly similar. Namely, the points on the seabed from which reflections occur start to converge, and the relative difference between pairs of rays on each side of the array correspondingly decrease. This quells the array side discrimination

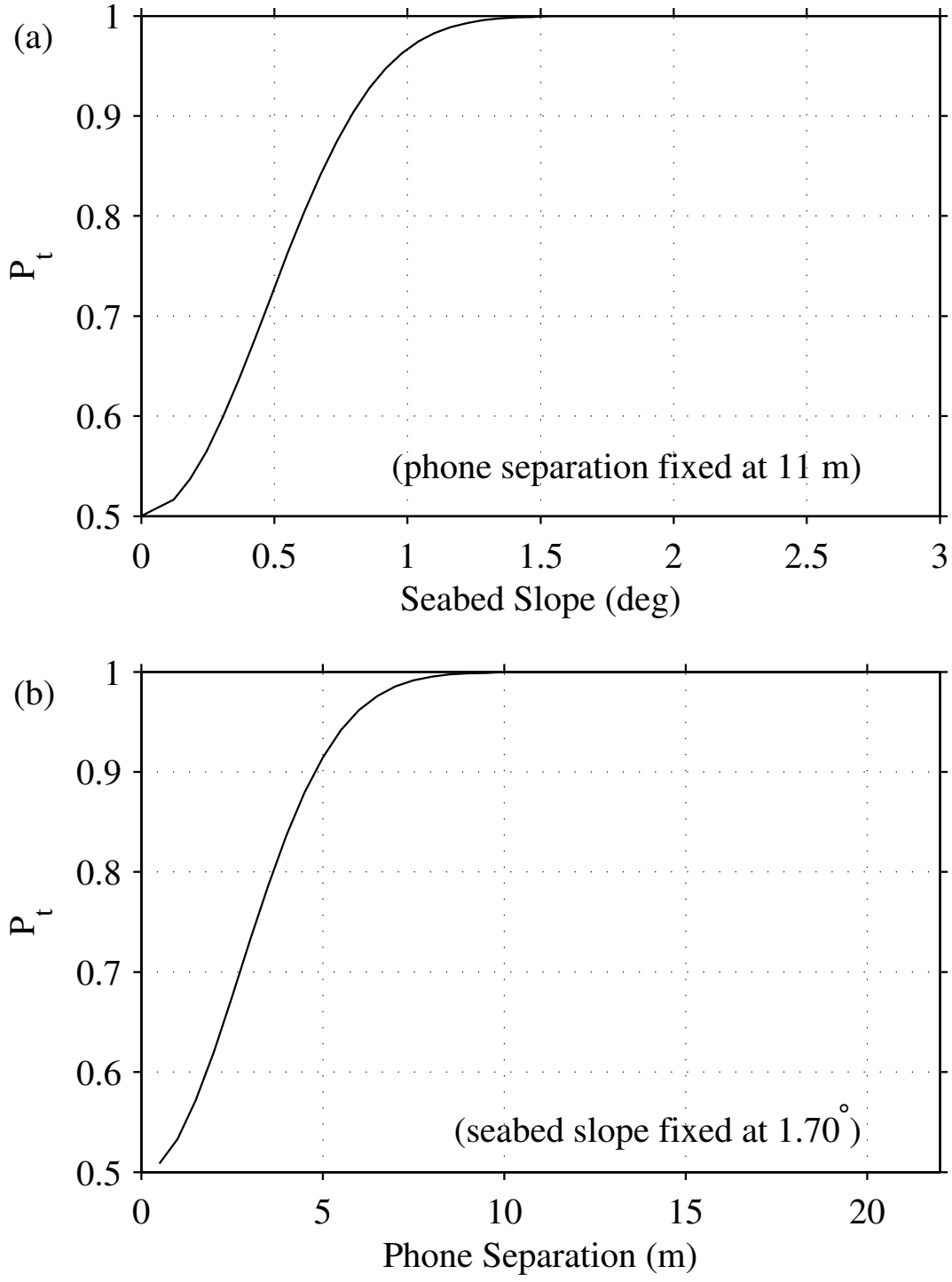


Figure 5.5: (a) Probability of choosing the true side for different perpendicular seabed slopes. (b) Same method but varying phone separation.

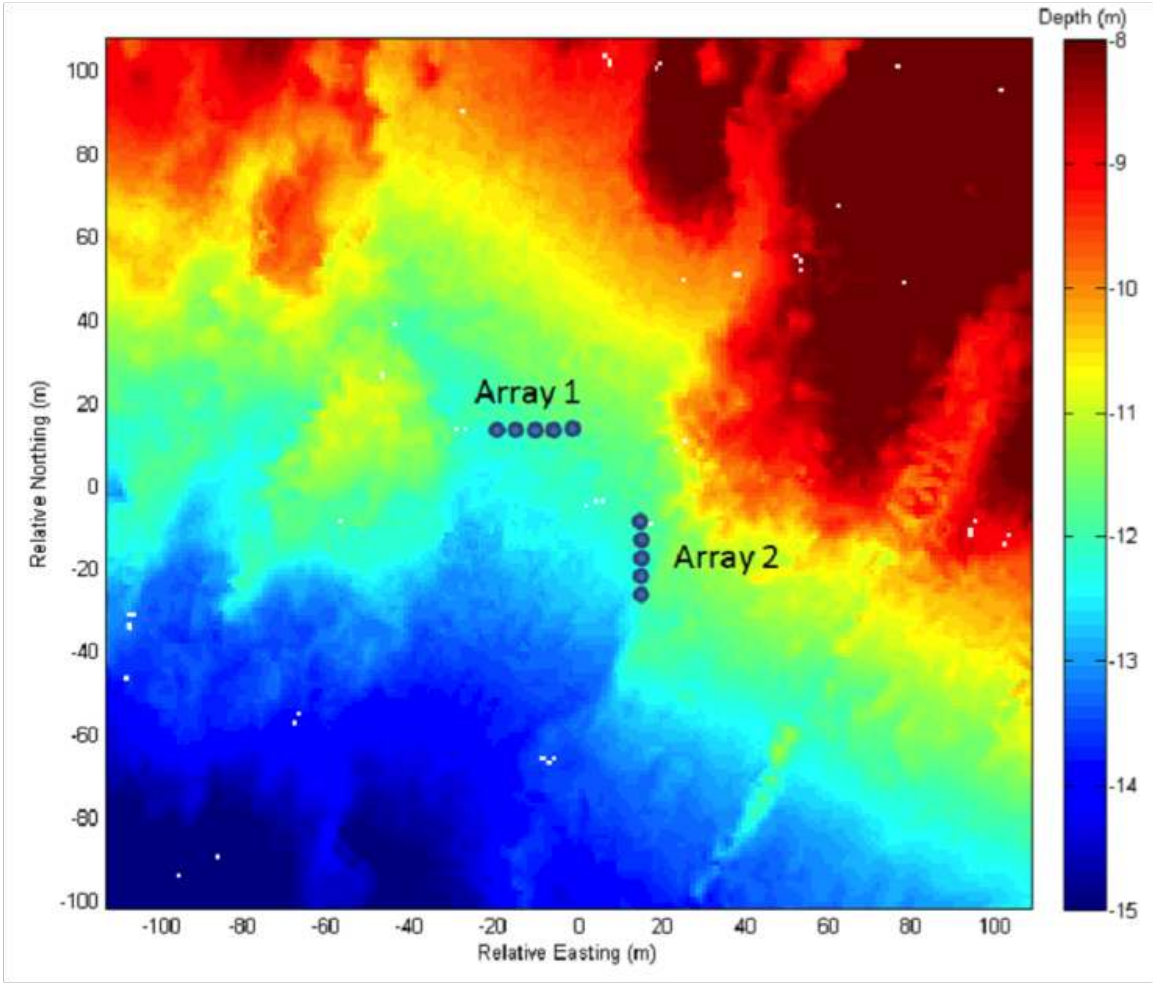


Figure 5.6: Arrays deployed in the Hawai‘i 2011 experiment.

capability.

5.5 Hawai‘i 2011 Experiment

The Hawai‘i 2011 experiment was conducted in August 2011 at the same site as the Hawai‘i 2010 experiment. The latter is described in Sec. 4.1. As before, this work was done in collaboration with the University of Hawai‘i. The Hawai‘i 2011

experiment was similar to the Hawai‘i 2010 experiment but with differences in the array configuration and the various targets and target tracks. Instead of one vertical array and one horizontal array, both arrays were laid out horizontally. They were orthogonal to each other, and were arranged in an “L” configuration on the seabed. This is shown in Fig. 5.6, along with the location of the horizontal array in the Hawai‘i 2010 experiment.

Only two elements of the HLA, spaced 11 m apart, were used in this study. The array was configured with a sample rate of 102.4 kHz, 24-bit dynamic range, 300 Hz low-cut filter, and 110 dB anti-aliasing filter set at 46.4 kHz. The hydrophones (HTI-92-WB) had a sensitivity of -160 dB re 1 V/μPa.

A rigid-hulled small boat with a single outboard engine was used as a target, and a handheld GPS device recorded its location. Time stamps in the GPS data and recorded acoustic data allowed for coarse-grained synchronization (on the order of 1 s) between the two sets of data. The boat executed several different maneuvers including driving in circles around the array deployment site. Spectral analysis indicated that the boat radiated noise in the 0-10 kHz band with the bulk of the energy below 3 kHz.

Bathymetry information for the local area was obtained from the SHOALS LIDAR bathymetry database at the University of Hawai‘i (University of Hawai’i at Manoa, Costal Geoglogy Group, 2012) which was ungridded data having roughly 1 m resolution. A gridding algorithm was used to convert this data to a square grid oriented the same way the local map coordinates (i.e. Northing and Easting). Grab samples near the deployment site indicated the seabed was composed of medium/coarse sand. The bathymetry, hydrophone locations, and track of the boat are shown in

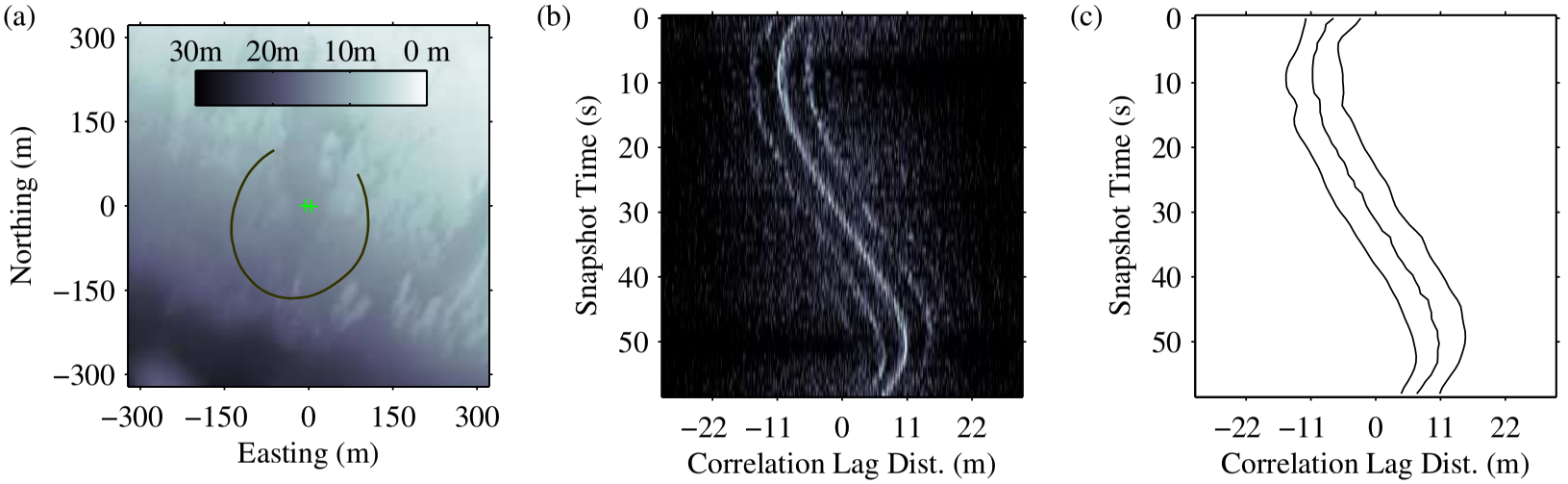


Figure 5.7: (a) Bathymetry and GPS boat track. The solid line shows the track of the small boat with a counter-clockwise trajectory. The "+" annotations indicate array element locations. (b) Correlogram showing $10\log_{10}|\Gamma(\tau)|^2$ evolving over snapshot time, plotted using 30 dB of dynamic range and with τ converted to wave travel distance. (c) Striation lines for τ_0 and $\tau_\pm$ output by the SIR particle filter.

Fig. 5.7(a). The bathymetry and hydrophone locations were used to compute the direct and bottom-surface eigenrays for a grid of points on the surface.

5.6 Striation Extraction Particle Filtering

This section describes the sequential Bayesian filtering methodology used to extract striation lines from a correlogram. The striation lines correspond to correlations of different pairs of arrivals at opposite hydrophones. As can be seen in the correlogram in Fig. 5.7(b), a prominent center striation line is flanked by two weaker striations. The center striation is caused by the direct arrival correlating with the direct arrival at the opposite hydrophone, and the flanking striations are due to a direct arrival at one hydrophone correlating with a multipath arrival at the opposite hydrophone. The offset of all three striations on the time delay (horizontal) axis at each time step (vertical axis) is the desired output. The extraction is divided into two trackers; the first tracks the center striation, which is then fed to the second tracker that tracks both flanking striations.

Given their ability to handle nonlinear problems, particle filters have been applied previously to the problem of acoustically tracking objects in space using multipath (Ward et al., 2003), and to directly analyze features of received data (Jain and Michalopoulou, 2011, Michalopoulou and Jain, 2012). The approach taken here is more akin to the latter by focusing on extracting time delays from the passive acoustic data. Sequential Importance Resampling (SIR) is among the simplest of the particle filter formulations (Ristic et al., 2004). In the tracking context, a sequence of observations are fed to the tracker to estimate hidden state variables at discrete time steps, denoted by k. State estimates at the previous time step, $k-1$, are used to

refine the probability at the current time step, giving this formulation its "Bayesian" nature.

In this context, the observations are the cross correlation at each time step, $\mathbf{\Gamma}_k = \Gamma_k(\tau)$ (Eq. 5.15 shows a noise-free measurement), and the state variables are the time delays, τ_k. Bold typeface indicates variables may be multidimensional. The prediction of the current state's (τ_k) probability distribution from a previous state's (τ_{k-1}) probability distribution is computed according to

$$p(\tau_k|\mathbf{\Gamma}_{1:k-1}) = \int p(\tau_k|\tau_{k-1})p(\tau_{k-1}|\mathbf{\Gamma}_{1:k-1})\ \mathrm{d}\tau_{k-1} \tag{5.33}$$

in which $\mathbf{\Gamma}_{1:k-1} = \mathbf{\Gamma}_1, \mathbf{\Gamma}_2, \ldots, \mathbf{\Gamma}_{k-1}$. Conceptually, this is based on a Markov state-space model in which the probability of the next state depends only on the previous state,

$$p(\tau_k|\tau_{1:k-1}) = p(\tau_k|\tau_{k-1}). \tag{5.34}$$

To incorporate a new measurement $\mathbf{\Gamma}_k$, the state is updated as

$$p(\tau_k|\mathbf{\Gamma}_{1:k}) \propto p(\mathbf{\Gamma}_k|\tau_k)p(\tau_k|\mathbf{\Gamma}_{1:k-1}), \tag{5.35}$$

for which the right hand side is then normalized to integrate to one. It is useful to think of $p(\mathbf{\Gamma}_k|\tau_k)$ as a function of τ_k that is parameterized by an actual (i.e. non-random) observation $\mathbf{\Gamma}_k$. Due to the proportionality relationship in Eq. 5.35, this expression can be represented as a likelihood function that need not integrate to one

$$L(\tau_k|\mathbf{\Gamma}_k) \propto p(\mathbf{\Gamma}_k|\tau_k). \tag{5.36}$$

Particle filters approximate the probability distributions as a sum of weighted delta functions,

$$p(\tau_k|\mathbf{\Gamma}_{1:k}) \approx \sum_{i=1}^{N} w^i \delta(\tau_k - \tau_k^i) \tag{5.37}$$

for which τ_k^i is the i^{th} particle and w^i is its weight, and N is the number of particles. The weights collectively sum to one, and provide a means to draw samples from a distribution other than $p(\tau_k|\mathbf{\Gamma}_k)$, for which there is often no closed form solution. This other distribution is called the importance density, and in the SIR algorithm is defined as the posterior probability distribution at the previous time step, $p(\tau_{k-1}|\mathbf{\Gamma}_{1:k-1})$.

Particle degeneracy happens when a significant portion of the particle weights become very small. These particles are essentially lost as they do not contribute meaningfully to the probability distribution. Resampling is used to re-draw the samples from regions of state space having greater probability density. This method follows from Ristic et al. (2004).

The first particle filter extracts the center striation using a random-walk state model defined as $\tau_{0_k} = \tau_{0_{k-1}} + v$, in which $v \sim \mathcal{N}(0, \sigma_{\tau_0}^2)$. The log-likelihood function is defined as

$$\log L(\tau_{0_k}|\mathbf{\Gamma}_k) = \int_{-\infty}^{\infty} \Gamma_k(\tau) g(\tau, \tau_{0_k}, \sigma)\, \mathrm{d}\tau. \tag{5.38}$$

The second particle filter extracts both flanking striations based on τ_{0_k} output by the first tracker. It uses a two-variable random walk state model, but instead of tracking the time offsets, it tracks two auxiliary variables. The flanking striations share roughly

the same absolute offset from τ_{0_k}, but there will be some small deviation of these absolute offsets on either side based on slight propagation differences. This deviation is important because it contains the information needed for array side discrimination. To model a varying absolute offset and a smaller deviation between sides, the state consists of two auxiliary variables, ζ_{1_k} and ζ_{2_k}. The state update is $[\zeta_{n_k} = \zeta_{n_{k-1}} + w_n]_{n\in\{1,2\}}$, in which $[w_n \sim \mathcal{N}(0, \sigma^2_{\zeta_n})]_{n\in\{1,2\}}$ such that $\sigma^2_{\zeta_2} < \sigma^2_{\zeta_1}$. The desired quantities, τ_{-_k} and τ_{+_k}, are defined in terms of these auxiliary variables as

$$\tau_{-_k} = \tau_{0_k} - \zeta_{1_k}, \qquad \tau_{+_k} = \tau_{0_k} + \zeta_{1_k} + \zeta_{2_k}. \tag{5.39}$$

The log-likelihood function is defined as

$$\log L(\tau_{+_k}, \tau_{-_k} | \mathbf{\Gamma}_k, \tau_{0,k}) = \int_{-\infty}^{\infty} \Gamma_k(\tau)[g(\tau, \tau_{-_k}, \sigma) + g(\tau, \tau_{+_k}, \sigma)]\, \mathrm{d}\tau. \tag{5.40}$$

To obtain point estimates from each posterior probability distribution, a Gaussian kernel smoother is applied to the particles to estimate the maximum a priori (MAP) value. The output of the tracker after both filters are the measured values τ_{0_k}, τ_{-_k}, and τ_{+_k} at each time interval k. In Sec. 5.5, the parameters were set to $N = 500$, $\sigma_{\tau_0} = 0.413$ ms, $\sigma_{\zeta_1} = 0.098$ ms, and $\sigma_{\zeta_2} = 0.001$ ms.

5.7 Experimental Results

A correlogram is shown in Fig. 5.7(b) in which multipath effects are evident. The strong, center striation is the correlation of direct arrivals. This is supported by the

fact that as the target circles around the array, this striation stays between ± 11 m, which are the limits for the correlation lag distance for the configured hydrophone spacing of 11 m. The multipath-with-direct correlations are visible as "shadow" striations that run adjacent to the main striation. All these striation lines were automatically traced using a SIR particle filter (see Sec. 5.6) and are shown in Fig. 5.7(c). Shadow striations from higher-order eigenrays are also faintly visible throughout the entire run, but are not used in this processing.

A comparison of localization using only the direct arrivals (corresponding to using just the first term of Eq. 5.20) with all three arrivals (all terms of Eq. 5.20) is shown in Fig. 5.8(a) and (b). The effects of using a flat seabed versus actual bathymetry are shown in Fig. 5.8(b) and (c), in which the peak near true target location is amplified with respect to the false target location. The full ambiguity function, $\Phi(\mathbf{x})$, is shown for several snapshots throughout the boat track in Fig. 5.9(a) for an 11 m spacing. The same processing was applied to phones separated by 2 m for the same times, and the results are shown in Fig. 5.9(b). Two things are evident with the shorter separation: each peak is less sharply defined in bearing due to the smaller aperture, and the false peaks on the near side are more prominent indicating a decrease in the array side discrimination capability. Poorer bearing resolution for shorter array lengths is consistent with a greater CRLB for bearing, described in Sec. 5.3. The image data in Figs. 5.8 and 5.9 were post-processed with a 2-D Hann filter to aid visualization of narrow features. Range errors at the start of the tracks Fig. 5.8(a) and (b) correspond to the SIR particle filter locking onto the striation lines.

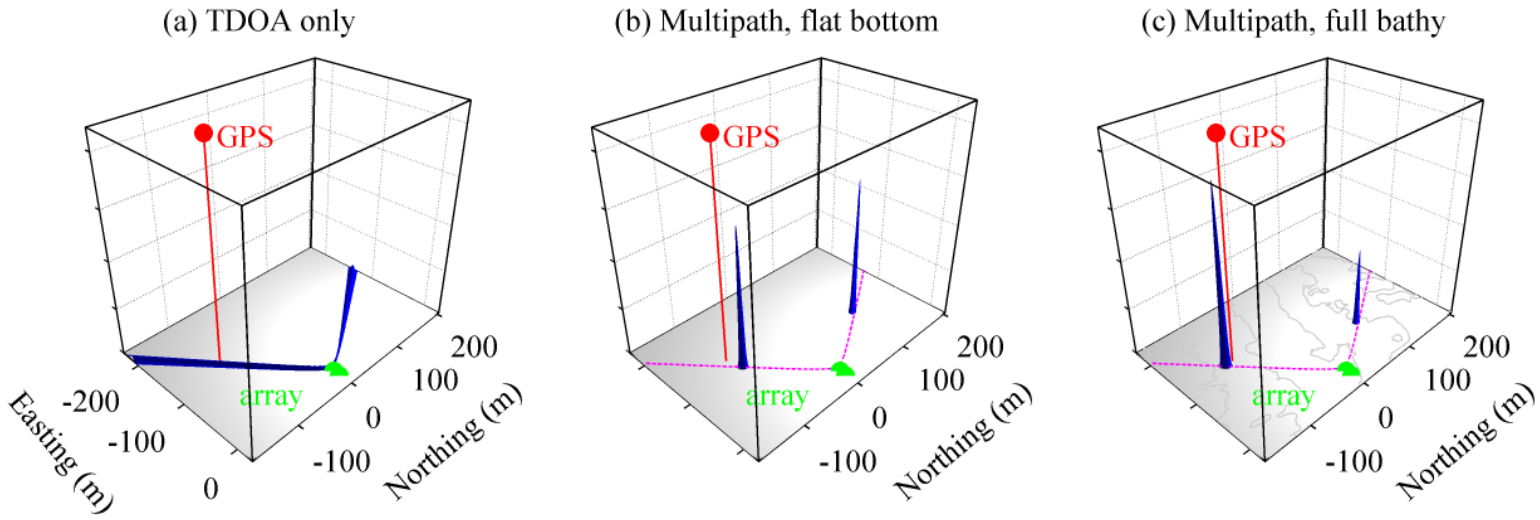


Figure 5.8: Comparison of ambiguity surfaces, $\Phi(\mathbf{x})$, for a single snapshot showing the effect of utilizing multipath and bathymetry information. Plots are normalized to unit volume to show the relative concentration of target location certainty. (a) A hyperbolic ambiguity is associated with only using the first term of Eq. 5.20. (b) Inclusion of multipath [the latter two terms of Eq. 5.20] and the assumption of a flat seabed cause the hyperbola to collapse to a single range, but a left-right ambiguity remains. (c) Using actual bathymetry improves the range estimate of the peak near the true target and also allocates a greater amount of target location certainty to it than the false target peak. Contour lines are shown at 2.5 m intervals. For (b) and (c), the hyperbola defined by τ_0 is also shown.

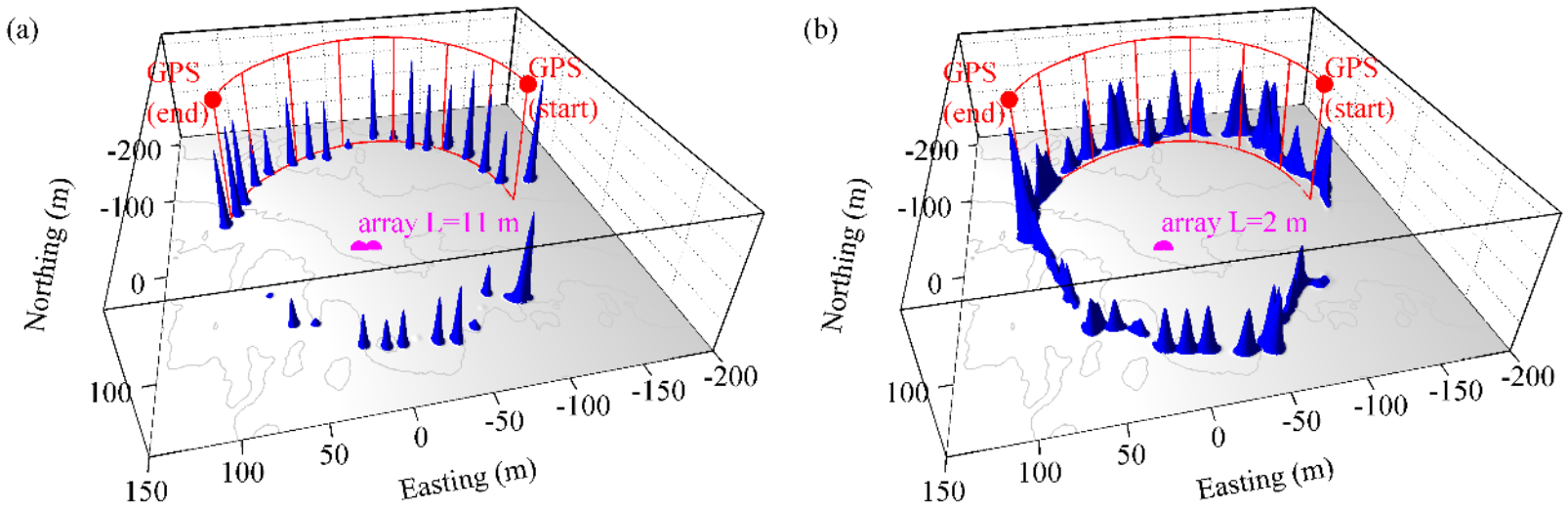


Figure 5.9: Ambiguity surfaces, $\Phi(\mathbf{x})$, over a sequence of snapshots comparing two array lengths, (a) $L = 11$ m (b) $L = 2$ m. The boat is driving in a circle around the array, and the data shown here is when it is on the south side. The camera is pointed mainly southward. The longer array length leads to better array side discrimination and better resolution in bearing. Within a few cycles, the particle filter locks onto the striation lines.

5.8 Chapter Summary

This chapter presents a simple technique for localizing a small boat using multipath arrivals recorded on two bottom-mounted hydrophones, and examines the factors that affect its performance. Hydrophones placed on the seabed constitute a favorable geometry for broadband target localization since it maximizes the time separation between the first and second arrivals, which facilitate extraction from a correlogram with a particle filter. Bathymetric variations can be exploited with ray tracing to improve range estimates, and can also support an array side discrimination capability for sloped seabeds. To expose this capability, array elements should be placed along contours of the bathymetry, rather than in an uphill-downhill arrangement. The probability of picking the correct side based on errors between predicted and actual multipath arrival times is reported. Theoretical limits of ranging accuracy are also given by modeling environmental and geometrical errors with the Bayesian Cramér-Rao lower bound.

One factor that is likely to impact performance of this technique is refraction caused by a non-constant vertical water sound speed profile. Errors in modeling this aspect of the environmental increase at greater ranges, which will likely increase the Bayesian CRLB. Performance was not observed to be impacted by seabed properties, which is attributable to the behavior of the envelope operation that removes frequency-independent phase changes. Analytic partial derivatives could be replaced with numeric methods, but this may require substantially greater computational resources. Regardless, one application for this technique might be monitoring small boat traffic in shallow water regions with relatively minimal deployment complexity and equipment.

Chapter 6

Volumetric Array Processing

This chapter presents a passive multi-target localization technique that uses a compact volumetric array to decompose the multipath structure arriving on a single bearing enabling direct measurement of multipath time delays that support target ranging. This differs from the two-hydrophone approach explained in Ch. 5 in that this algorithm employs array processing techniques more advanced than cross-correlating two hydrophones. The scale of the receiver arrays also differ, which is important because it critically affects how and where these types of systems can be deployed. The length of the POEMS-A systems used in Hawai'i 2010 were on the order of 10's of meters, whereas the volumetric array discussed in this chapter spans only 0.4 m.

The hardware and the signal processing requirements needed for this technique lend well to deployment from mobile platforms such as autonomous underwater vehicles (AUVs). Many common ocean waveguides support several ray paths between a source and receiver (known as eigenrays or arrivals), which are collectively referred to as the multipath structure. Since the multipath structure changes as a function of target range, measuring the interference pattern due to multipath can be informative of target range (Thode, 2000, Rakotonarivo and Kuperman, 2012). Multipath interference

is the result of time delays between arrivals causing constructive and destructive interference at different frequencies (Harrison, 2011). It has been the subject of much study because the pattern often has low sensitivity to waveguide parameters (Weston and Stevens, 1972). The technique presented in this chapter bypasses the interference pattern and directly measures the time delays between arrivals by cross correlating beams steered at multipath arrivals. It is based on a recent passive fathometry technique that measured multipath of surface noise. To relate this to the inference pattern, it shown that the most prominent spectrogram striations, and the waveguide invariant parameter β (Brekhovskikh and Lysanov, 2003), can be predicted using the measured multipath delays. While this approach is presented in the context of target localization, it also has implications for ocean tomography and geoacoustic inversion.

In order to utilize the multipath delays for target ranging, it is important that the delay is not so small as to preclude its accurate measurement. The approach in Ch. 5 was to place the hydrophones directly on the seabed thus maximizing the travel distance between the first and second eigenray arrivals. If it is possible to control the transmitted waveform, one approach is to use short-duration waveforms and source-receiver geometries that naturally separate rays (Holland and Osler, 2000). Many marine mammals also use short-duration waveforms. For example, whale clicks are impulsive and time delays between distinct multipath arrivals have been shown to provide estimates of animal locations (Nosal and Frazer, 2006, Tiemann et al., 2006, Mathias et al., 2013). The problem is less straightforward when the waveform is broadband noise, but a class of techniques built on the concept of "pulse compression" have been shown to be an effective method of estimating time-delays. Pulse compression is typically based on the generalized cross correlation algorithm (Carter and Knapp, 1976). It can also be used to estimate multipath time delays, and has

been studied extensively from a theoretical perspective for target range estimation (Hahn, 1975, Hamilton and Schultheiss, 1992, Badriasl et al., 2011).

Arrival angles also are functionally dependent on the environment, and are used in acoustic tomography experiments, such as Aulanier et al. (2013). This was based on a double-beamforming measurement concept introduced by Roux et al. (2008) involving a vertical receiver array and a vertical source array. The ambient noise field has also been shown to provide tomographic information, as recently demonstrated by Leroy et al. (2012) and Lani et al. (2013) extracting coherent wavefronts using multiple vertical line arrays.

Multipath coherence gives rise to an interference pattern that depends on range. Traditionally this has been treated as a nuisance factor, but recently it has been favorably utilized by a modified conventional beamformer to increase array gain (Rouseff and Zurk, 2011). However, multipath coherence can negatively affect adaptive beam forming algorithms that are based on eigendecomposition. A variety of methods have been developed for breaking the coherence [see Van Trees (2002d) for a summary]. An approach taken by Koch and Knobles (2005) for geoacoustic inversion using noise from a nearby ship of opportunity was to look for correlations between sub-apertures of a horizontal line array, thus avoiding the coherence problem. Later work by Stotts et al. (2010) used the same technique but to simultaneously invert for the ship location and geoacoustic properties.

While multipath coherence can cause problems for adaptive beamforming, this feature was exploited for geoacoustic inversion in the passive fathometer formulation that used surface-generated noise (such as from wind and breaking waves) as a source of opportunity. An algorithm based on conventional beamforming was introduced

by Siderius et al. (2006) and later refined by Gerstoft et al. (2008), Harrison and Siderius (2008); and Means and Siderius (2009). The two-way travel time from the array to the seabed was measured by beamforming at the end fire directions of a vertical line array to isolate the downward traveling noise and the seabed reflections. Interestingly, switching to adaptive beam forming (Siderius et al., 2010) improved time delay estimates by reducing interference from non-vertically-traveling waves. As described in the next section, a similar approach is taken to measure time delays between non-vertically traveling multipath arrivals originating from a distant surface noise source.

The conference paper Gebbie et al. (2010) looked at how cross-beamforming could be used to process multipath with the passive fathometer. This helped set the groundwork for the algorithm described here. This work was presented at the 21st International Congress on Acoustics Conference in Montreal Canada in June 2013 and published in the Proceedings of Meetings on Acoustics (Gebbie et al., 2013c). A similar work was also presented at the 1st International Conference & Exhibition on Underwater Acoustics, Corfu, Greece in June 2013 (Gebbie et al., 2013b). The multi-target capability and connection to waveguide invariance was presented at the Acoustical Society of America conference in Providence, RI in May 2014 (Gebbie et al., 2014d). Lastly, a full-length manuscript has been accepted for publication in the Journal of the Acoustical Society of America (Gebbie et al., 2014c) covering all but the waveguide invariance topic.

The rest of this chapter is organized as follows. Section 6.1 provides a description of the cross beam correlation methodology. Section 6.5 describes the GLASS'12 experiment. Section 6.6 presents the target localization results from that experiment.

Section 6.7 relates this work to the principles of waveguide invariance.

6.1 Problem Formulation

In the case of the passive fathometer, the vertical array was oriented in such a way that the ends of the array naturally pointed at the source (surface noise) and the multipath reflection (the seabed) while all other interference arrived mainly from broadside (i.e. horizontally). The vertical line array geometry is thus well suited to this type of distributed source, but applying it to a distant source requires forming beams at other elevation angles to match the multipath arrivals. While this is possible with a vertical line array, it also results in a conical beam pattern giving it ambiguity in bearing. This means that ambient noise and multipath from multiple targets will overlap in the beamformer output, making analysis of individual targets difficult. A solution is to use volumetric arrays that have aperture in all three spatial dimensions since they can form beams in any direction in bearing and elevation.

Cross beam correlation operates on two arrivals, so this discussion begins with a simplified model of a waveguide that includes only the first two arrivals. The spectral value at range r and frequency ω can be approximated with the first arrival normalized to one as

$$S \approx e^{-i\omega t_1} + Re^{-i\omega t_2}, \tag{6.1}$$

in which t_n is the time offset of the n^{th} arrival, $R = |R|e^{i\phi_R}$ is the complex amplitude of the second arrival relative to the first that accounts for propagation differences (i.e. additional reflections, losses due to spreading and attenuation, etc.). Note the

variables R and t_n depend on r; and we make the approximation that R is independent of frequency, such as would be the case for a half space seabed. Let the multipath time delay be defined as

$$\tau_{2,1} \triangleq t_2 - t_1. \tag{6.2}$$

Section 6.2 reiterates adaptive beamforming, and Sec. 6.3 discusses how to directly obtain $\tau_{2,1}$ from a noise-producing target by beamforming to isolate the individual terms of Eq. 6.1, which are then cross-correlated to expose $\tau_{2,1}$ in the time domain. The term ϕ_R is also measurable with cross-beam correlation, but it is ignored mainly because it is a property of the environment and does not contain much information about the target range.

6.2 Adaptive Beamforming

The first step is to determine the beams on which the multipaths are arriving. This is accomplished with adaptive beamforming that filters the acoustic data in bearing, elevation, and frequency. It results in a map showing where acoustic energy is being received at the array, and its spectral content. Thus, adaptive beamforming provides a means to determine the direction of a target in bearing, and its associated multipath arrivals in elevation. Adaptivity plays an important role in that it can focus in one direction while simultaneously rejecting noise arriving in other directions. As its name implies, adaptive beamforming adapts the beam pattern to the data being processed to improve overall gain in the steering direction. Conceptually, it does this by steering side lobes and nulls in directions that are most advantageous. For example, a loud

interferer arriving on a distinct bearing will have a null steered at it, whereas a relatively quiet bearing will get assigned a large side lobe.

As was described in Sec. 3.5.2, the minimum-power distortionless response (MPDR) adaptive beamformer is defined in Van Trees (2002c) by Eq. 3.156, and repeated here

$$\mathbf{w}_{\text{MPDR}} = \frac{\mathbf{R}^{-1}\mathbf{v}}{\mathbf{v}^H\mathbf{R}^{-1}\mathbf{v}}. \tag{6.3}$$

The term $\mathbf{R}$ is the sample-averaged cross-spectral density matrix (CSDM) defined in Eq. 3.125, and repeated here

$$\mathbf{R} = \frac{1}{M}\sum_{m=1}^{M}\mathbf{p}_m\mathbf{p}_m^H, \tag{6.4}$$

in which $\mathbf{p}_m$ is a vector of spectral values across all channels for the m^{th} snapshot. A snapshot consists of contiguous time-series data across all channels that is used to estimate spectral values. The total averaging time and snapshot length are important and will be discussed in Sec. 6.4. To eliminate strong tonal components (such as engine harmonics from a boat) pre-whitening is used to flatten the spectrum by $\mathbf{p}_m = \hat{\mathbf{p}}_m/|\hat{\mathbf{p}}_m|$ where $\hat{\mathbf{p}}_m$ is the measured spectral value. The plane wave array manifold is given by Eq. 3.117, and repeated here

$$\mathbf{v}(\mathbf{u}) = e^{-ik\mathbf{a}^T\mathbf{u}}, \tag{6.5}$$

for which $\mathbf{a}$ is the array geometry that specifies the position of each element in the columns of a matrix, $\mathbf{u}$ is a unit vector indicating the direction of wave propagation in a column vector, and k is the wavenumber in units of radians per meter. In the literature, MPDR is sometimes referred to as MVDR (where the V stands for

variance); however, strictly speaking the MVDR algorithm (Capon, 1969) requires *a priori* knowledge of the noise (N) plus interference (I) covariance matrix ($\mathbf{K}_{\mathrm{N+I}}$), whereas MPDR operates directly on the sample-averaged CSDM, $\mathbf{R}$. Equation 6.3 is a least-squares solution to

$$\underset{\mathbf{w}}{\operatorname{argmin}}\, \mathbf{w}^H \mathbf{R} \mathbf{w}, \tag{6.6}$$

which minimizes the power output of the beamformer while satisfying the distortionless constraint

$$\mathbf{w}^H \mathbf{v} = 1. \tag{6.7}$$

The formulation in Eq. 6.3 requires $\mathbf{R}$ to be full rank. It is often the case in practice that $\mathbf{R}$ has one or more very small eigenvalues, which cause the matrix inversion to become unstable. This can happen for a N element array with fewer than N snapshots, and is referred to as "snapshot deficiency" (Song et al., 2003). While techniques exist to compensate for this (Menon and Gerstoft, 2013), a common *ad hoc* technique for stabilizing the matrix inversion is to add a small amount of diagonal loading to $\mathbf{R}$; in this study a small amount (-30 dB relative to the mean spectral power across channels) is used (Van Trees, 2002b). Beyond stabilizing a rank deficient $\mathbf{R}$, diagonal loading can also increase the beamformer tolerance to mismatch errors, such as element location errors or wavefront curvature errors.

6.3 Cross Beam Correlation

Once multipath arrivals are spatially separated on different elevation beams, the next step is to determine the time delay of their respective signals. The noise signal from the source undergoes different delays based on unique ray trajectories through the waveguide. Cross-correlation can be used to "pulse compress" these coherent broadband noise signals into a broadband pulse in the time domain, which has an offset corresponding to the relative time delay. This shares some conceptual similarity to a matched filter commonly used in active sonar, but instead of a known transmitted waveform it uses a measured waveform.

In the time domain, cross correlation is computed by sliding a "reference" signal along the time axis while holding the "correlated" signal fixed, taking the product at each offset, and finally integrating. In the frequency domain this corresponds to conjugating the spectrum of the reference signal and multiplying it by the spectrum of the correlated signal. The cross spectrum of two arbitrary beams can thus be formulated as

$$C_{\mathrm{c,r}} = \left(\mathbf{w}_c^H \mathbf{p}\right) \left(\mathbf{w}_r^H \mathbf{p}\right)^* \tag{6.8}$$

$$= \mathbf{w}_c^H \left(\mathbf{p}\mathbf{p}^H\right) \mathbf{w}_r \tag{6.9}$$

in which $\mathbf{w}_c$ is the correlated beam and $\mathbf{w}_r$ is the reference beam. The quantity in parenthesis in Eq. 6.9 is the outer product of a single snapshot; if multiple snapshots are averaged, this is replaced with a sample-averaged CSDM, $\mathbf{R}$. Cross beam

correlation for arbitrary beams is thus defined as

$$C_{\mathrm{c,r}} = \mathbf{w}_c^H \mathbf{R} \mathbf{w}_r. \tag{6.10}$$

Inverse Fourier transforming (with the operator denoted as $\mathcal{F}^{-1}$) yields

$$\Gamma_{\mathrm{c,r}}(\tau) = \mathcal{F}^{-1}[\hat{C}_{\mathrm{c,r}}(\omega)]. \tag{6.11}$$

The envelope can then be used to eliminate constant phase offsets as

$$\Gamma_{\mathrm{c,r}}^{\mathrm{env}}(\tau) = \left| \Gamma_{\mathrm{c,r}}(\tau) + i\mathcal{H}\left[\Gamma_{\mathrm{c,r}}(\tau)\right] \right|, \tag{6.12}$$

in which $\mathcal{H}$ is the Hilbert transform. Time delays are then obtained by finding the offset of the largest peak as

$$\tau_{\mathrm{c,r}} = \underset{\tau}{\mathrm{argmax}} \left[\Gamma_{\mathrm{c,r}}^{\mathrm{env}}(\tau)\right] \quad \text{for} \quad \tau_{\min} < \tau. \tag{6.13}$$

The term $\tau_{\min}$ is an artifact due to white noise leakage and depends on the array geometry, the manifold origin, and the beams being correlated. It is defined as

$$\tau_{\min} = \max_n \left[\mathbf{a}_n^T(\mathbf{u}_r - \mathbf{u}_c)/c\right] + \frac{1}{B_s}, \tag{6.14}$$

in which B_s is the effective target bandwidth, and c is the sound speed at the array.

Equation 6.10 is a generalization of the passive fathometer that includes beams steered in arbitrary directions for an array of arbitrary geometry. For comparison, the vertical beams of the passive fathometer are $\mathbf{w}_{\mathrm{down}} = \mathbf{w}_r$ (the signal traveling down from the surface) and $\mathbf{w}_{\mathrm{up}} = \mathbf{w}_c$ (the signal traveling up from the seabed). It is useful to denote

the beam having the delayed signal as the correlated beam, which renders time-delays as positive quantities. However, this is merely a matter of convention.

6.4 Target Localization Procedure

This procedure uses array acoustic data and knowledge of the receiver position in the waveguide to estimate target range. In this scenario, the target is assumed to be at the surface. If additional information about the sound speed profile (SSP) of the waveguide is available, this can be used to improve range estimates. The total acoustic averaging time should be enough to produce a full rank $\mathbf{R}$, but short enough so the effects of target motion are minimal; this may be on the order of hundreds of milliseconds depending on the array size.

The first step in the procedure is to use adaptive beamforming to find all arrivals from all targets. Coherence between multipath arrivals can cause problems with adaptive beamforming since coherent signals arriving on distinct beams are represented by a single eigenvector. A way to mitigate this is to reduce the snapshot length in order to break the coherence. As the snapshot length approaches the time delay between multipath arrivals, each arrival will be increasingly mapped to distinct eigenvectors. This has the effect of stabilizing the adaptive beamformer for coherent arrivals originating from a single target. The output of the beamformer is a full three-dimensional map of all arrivals on all bearings and all elevations. For a single target on a single bearing, the multipath arrivals will be distributed on the elevation axis.

The second step is to cross correlate beams steered at different elevation angles on a single bearing to determine the time delay between arrivals. Unlike the previous

step that attempts to eliminate multipath coherence, this step re-averages $\mathbf{R}$ using snapshots that are longer than the multipath time delay, but without changing the total averaging time. The length of the snapshot should be on the order of several times the delay to be measured.

With a surface source in shallow water, the first arrival will usually be traveling downward and have the highest amplitude, so a method to find the delay between just the first two arrivals is to to steer the reference beam at the strongest arrival traveling downward and sweep the correlated beam over all elevation angles on the same bearing. If the correlated beam corresponds to an arrival that is delayed with respect to the arrival in the reference beam, a peak will appear at a positive time offset in the time series corresponding to the delay.

The third step is to use the receiver position in the waveguide, and possibly environmental data as well, to estimate the target range. For a surface target, image theory can be used to compute this time delay in a closed form (Jensen et al., 2011b). This assumes a constant water sound speed over depth. However, if data pertaining to the SSP or bathymetery exists, a ray model can be employed to compute the first two eigenray travel times. This same calculation is performed at all ranges for a constant receiver depth, and compared to the measured time delay. Ranges at which there is agreement imply a greater likelihood the target is at that range. This is demonstrated with experimental results in Sec. 6.6. Note that ray travel times can have varying degrees of sensitivity to the water SSP. Recent work has been done on travel time sensitivity kernels for range and depth dependent variations (Sarkar et al., 2011), but in this study we assume range independence.

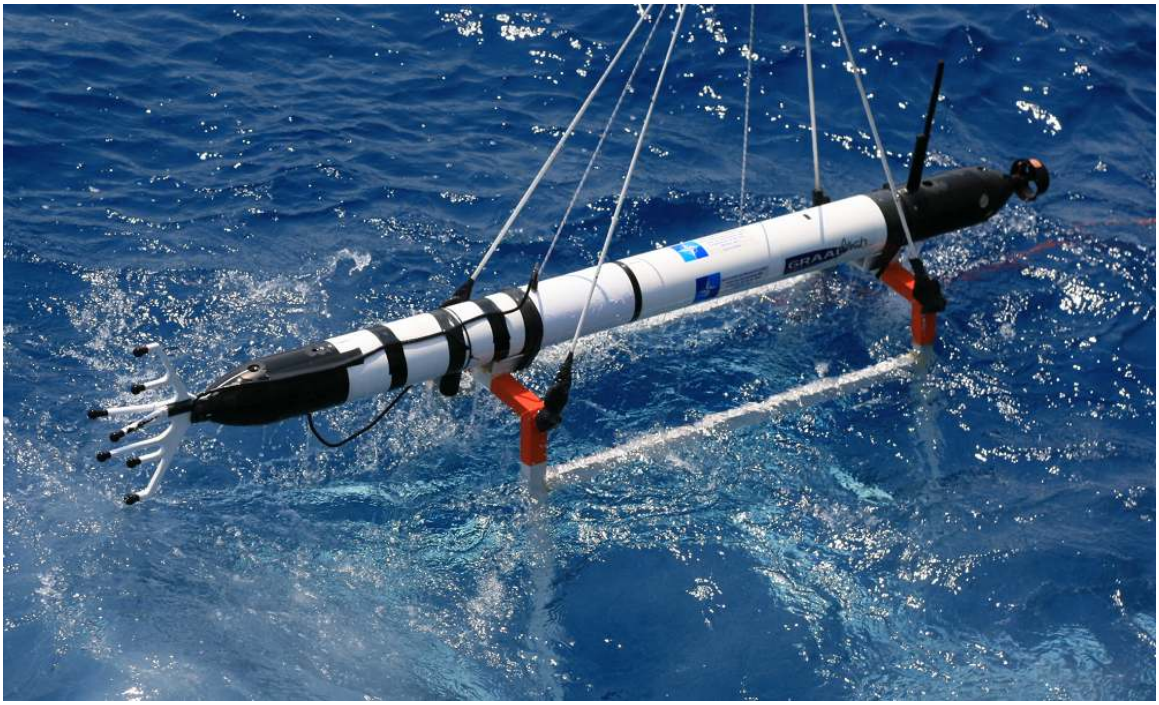


Figure 6.1: This photo shows the eFolaga AUV and fixed mooring frame submerging.

6.5 GLASS'12 Experiment

This section describes the experimental design used to validate the proposed technique. The GLASS'12 experiment was conducted by the Center for Maritime Research and Experimentation (CMRE) from the NRV *Alliance* in July 2012 off the coast of northern Italy near La Spezia. The purpose was to investigate potential uses of an autonomous underwater vehicle (AUV) affixed with a nose-mounted array.

The CMRE AUV eFOLAGA was chosen as the "autonomous" platform, and is shown in Fig. 6.1. It is capable of both propeller-driven and glider-driven locomotion. Despite these capabilities, there were technical challenges to getting the AUV to have the correct buoyancy, causing it to continually crash into the seabed. Since the pri-

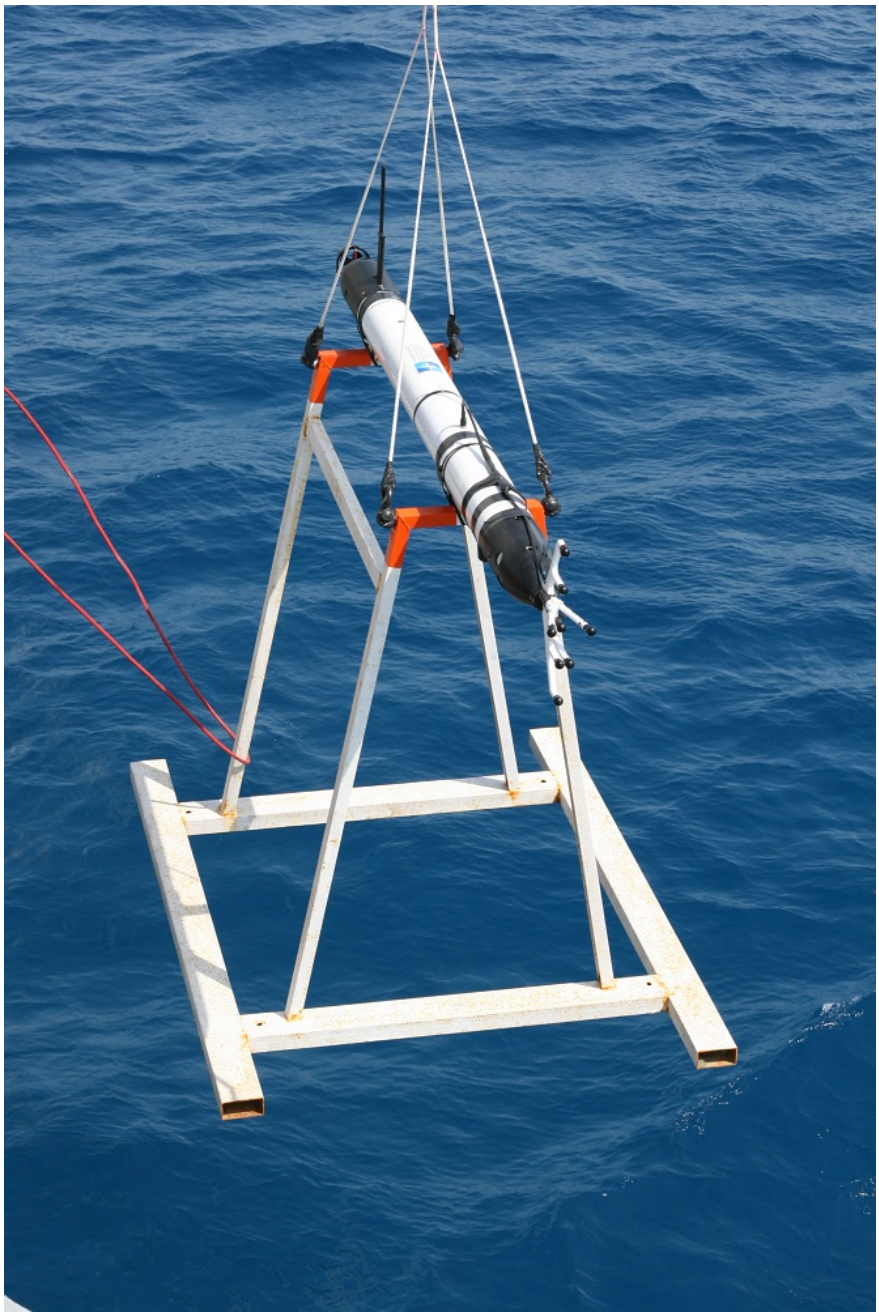


Figure 6.2: This photo shows the mooring to which the eFolaga AUV was attached.

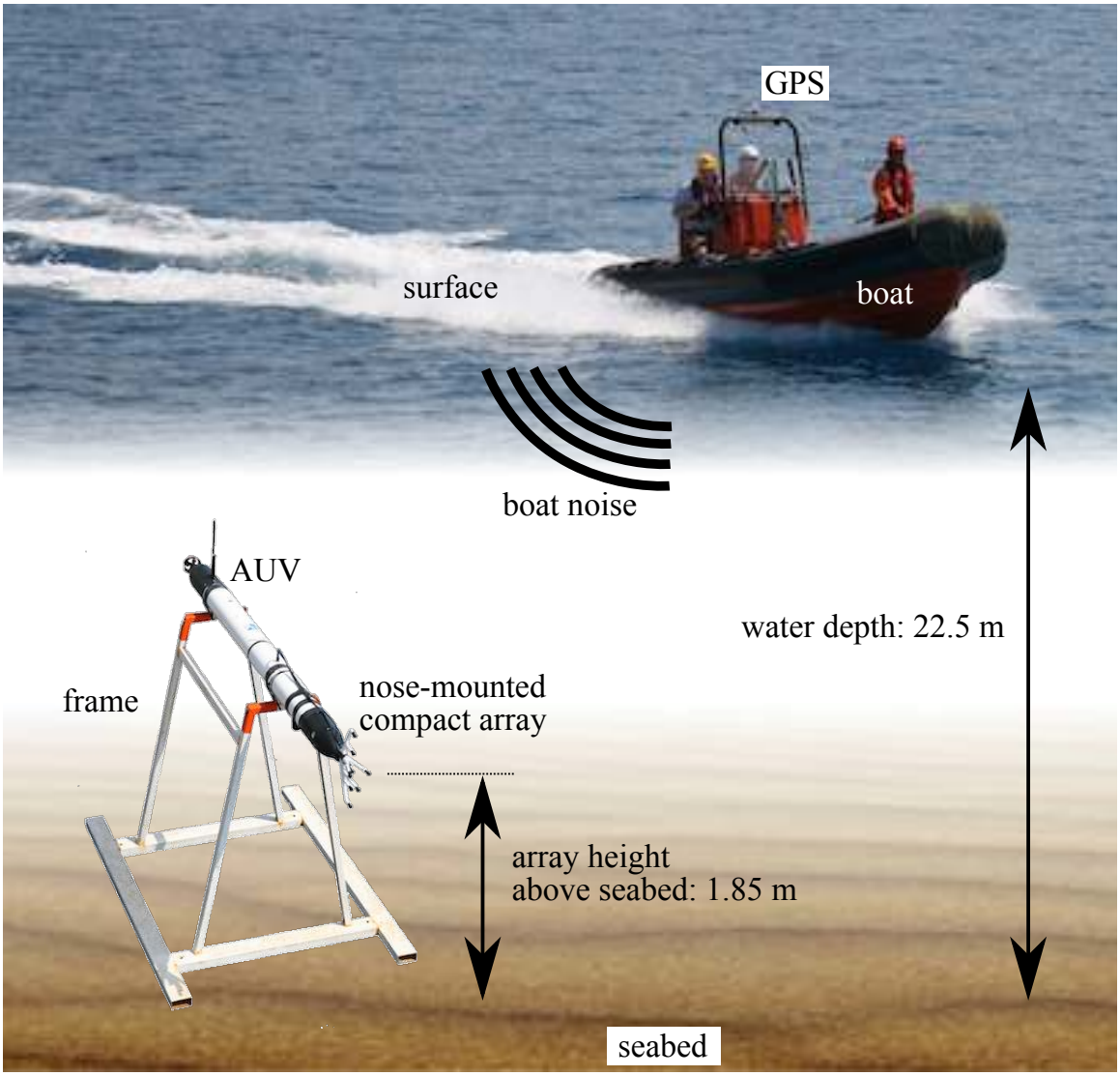


Figure 6.3: Experimental setup. The AUV is located on the seabed mounted on a fixed frame, and a small boat outfitted with a GPS collection device is operating on the surface. Noise from the boat is received by the hydrophone array mounted on the nose of the AUV.

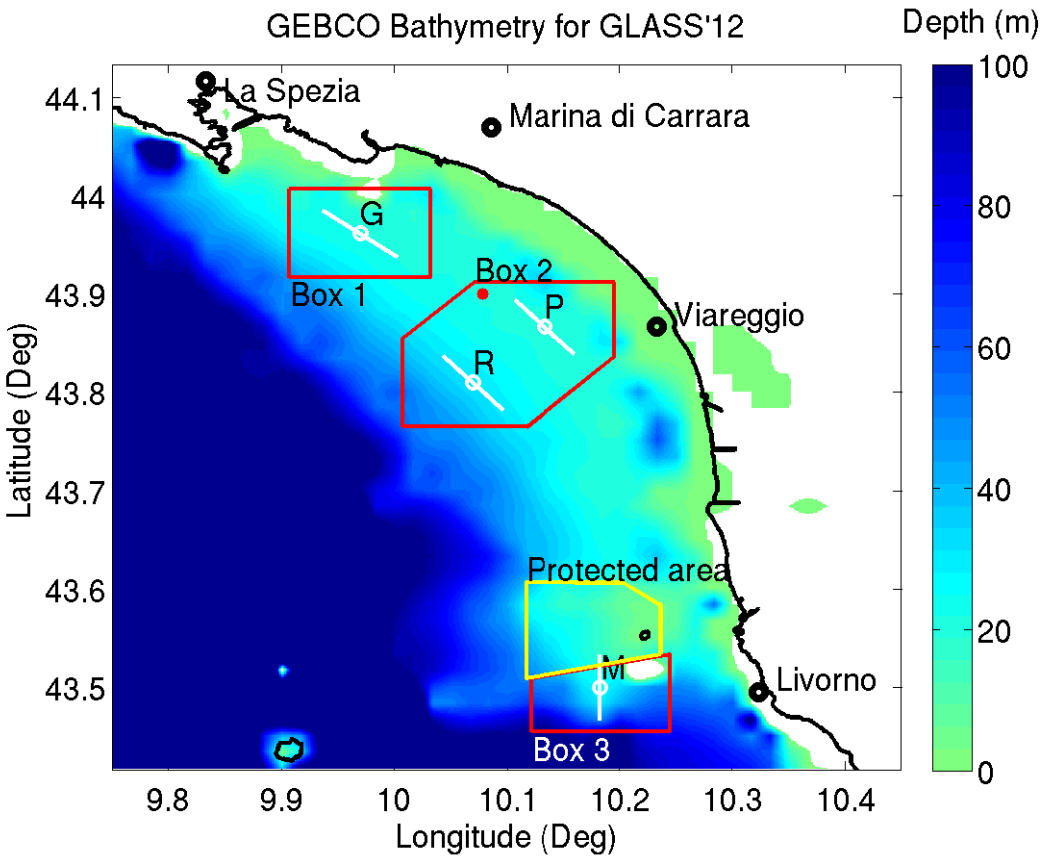


Figure 6.4: Bathymetry and overview map of experiment site (Nielsen et al., 2012). Data used in this chapter was taken at site G.

mary interest surrounded collection of data from the nose-mounted array, the entire AUV was attached to a fixed apparatus and moored to the seabed. This apparatus is shown in Fig. 6.2.

The experiment design is depicted in Fig. 6.3. The mooring placed the array approximately 1.85 m above the seabed, and the overall water depth was approximately 22.5 m, as measured by a towed EdgeTech SB216S sub-bottom profiler. An overview map of the deployment site and regional bathymetry is shown in Fig. 6.4. Results presented in the paper were recorded at point G. The SSP was measured with a

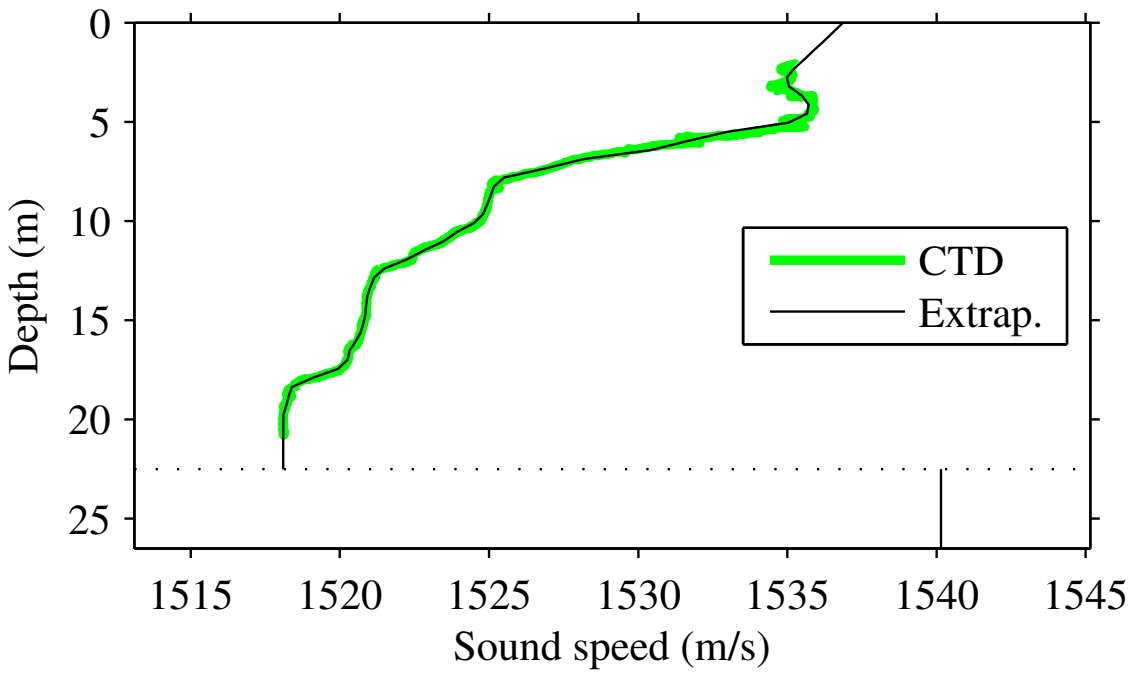


Figure 6.5: Water sound speed profile (SSP) showing measured and extrapolated parts.

Figure 6.6: The passive acoustic array consisted of tetrahedral and line sub arrays. This was mounted to the nose of the eFolaga AUV.

conductivity-temperture-depth (CTD) instrument, and is shown in Fig. 6.5. Portions at the boundaries are extrapolated over depths for which no data was captured due to limitations of the instrument. The extrapolation is explained in Sec. 6.6.3. The local environment consisted of a primarily sandy bottom. Analysis of core samples from the seabed indicated a mean seabed density of 1.807 g/cm^3 and mean sound speed of 1540.1 m/s. Attenuation measurements for the seabed were not available, but a nominal value of 0.2 dB/λ was assumed.

Figure 6.7: The RHIB boat used as a broadband target for passive acoustic localization algorithms. The noise produced by this boat as it moves through the water is recorded by the hydrophones of the hydrophone array moored to the seabed.

The array shown in Fig. 6.6 was designed and built by CMRE. It consisted of eight elements; five in a vertical configuration, and three offset from the center element to form a tetrahedron with the center element. The elements are mounted on a rigid frame providing low element position error. The spacing between the vertical elements, as well as edge lengths of the tetrahedron, was 10 cm. The array frame itself was made of metal, so array elements were rigidly held in position giving the entire array a very low amount of mismatch error. The sample rate of the array was 100 kHz. The GPS coordinates of the array position were taken during the deployment.

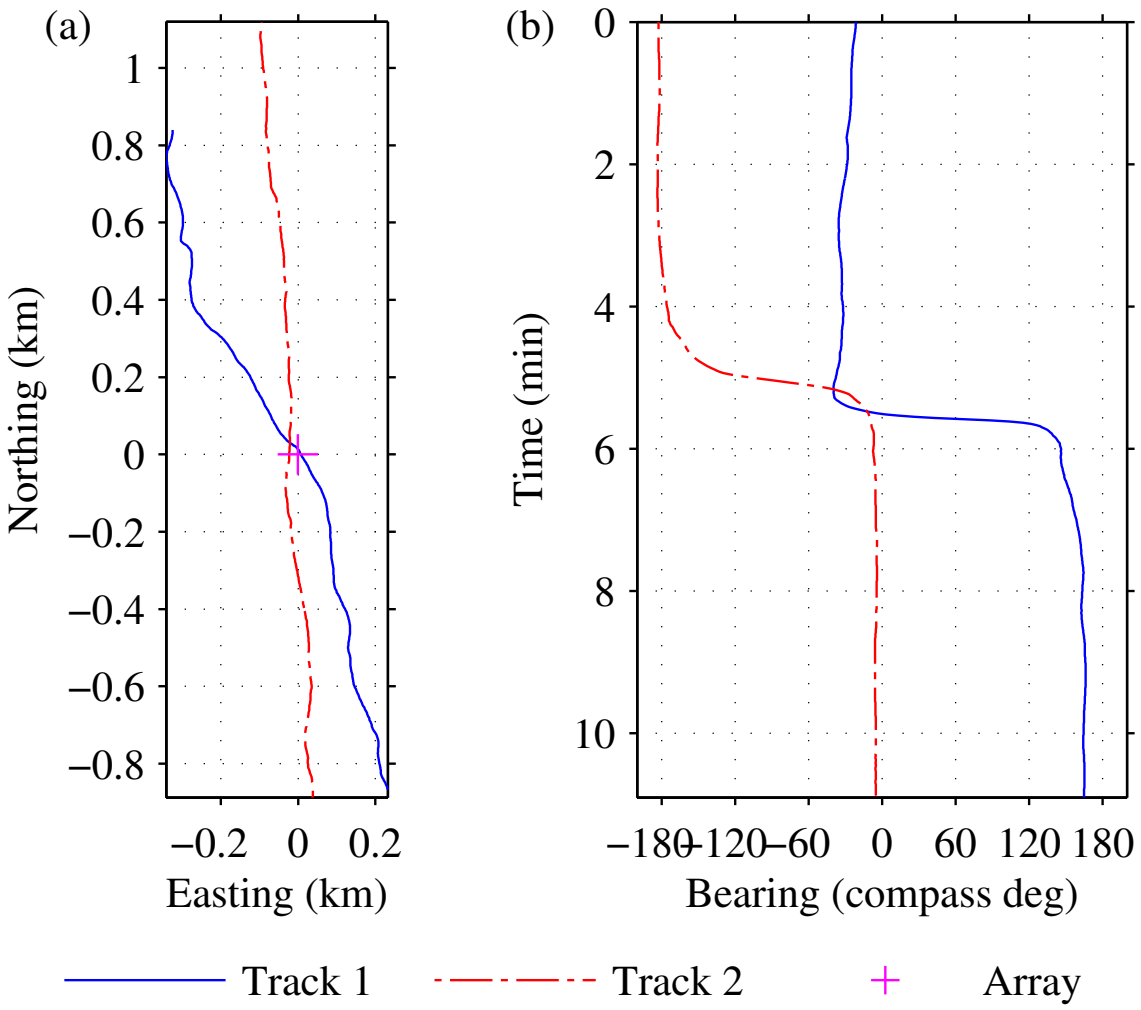


Figure 6.8: a) Tracks of the same boat for two passes by the array, taken from GPS records. b) Compass bearings of the two same two tracks. A multi-target scenario is simulated by adding the acoustic data from each track. This plot shows how they are aligned in time.

A small boat (the CMRE workboat) with an outboard engine outfitted with a portable GPS data recorder maneuvered in the vicinity of the array. The boat is shown in Fig. 6.7. The speed of the boat was roughly 3 m/s. Selected portions of the track of the boat are shown in Fig. 6.8(a). Besides the NRV *Alliance*, there were few vessels detected in the local region. To simulate a multiple target scenario, acoustic data from two boat passes by the array were added together. The time alignment of these two passes, showing the relative bearings of each pass are shown in Fig. 6.8(b).

6.5.1 Predicted Multipath Structure

This section describes how the multipath structure was determined from the measured environmental parameters. Bellhop (Porter and Bucker, 1987) is a popular ray tracing tool for analyzing acoustic propagation between two points in an ocean waveguide. Formulating this problem in a ray context has the advantage of being able to clearly see the contributions of each eigenray in terms of both space and time. Bellhop produces an estimate of the full path traversed by the eigenray. It also produces an estimate of the travel time for each eigenray, which depends on the SSP. As a matter of terminology for this section, eigenrays are labeled according to their sequence of boundary interactions; the direct eigenray is 'D', the bottom-reflected path is 'B', the path reflected first from the bottom then the surface is 'BS', etc.

Fig. 6.9 shows a set of eigenrays computed with Bellhop. The paths of these rays can be understood by noting the negative gradient of SSP shown in Fig. 6.5 (i.e. the trend shows decreasing sound speed with increasing depth). With the source placed at the surface, the first surface bounce is omitted following the same procedure in Holland and Osler (2000). The negative gradient causes rays to bend toward the seabed;

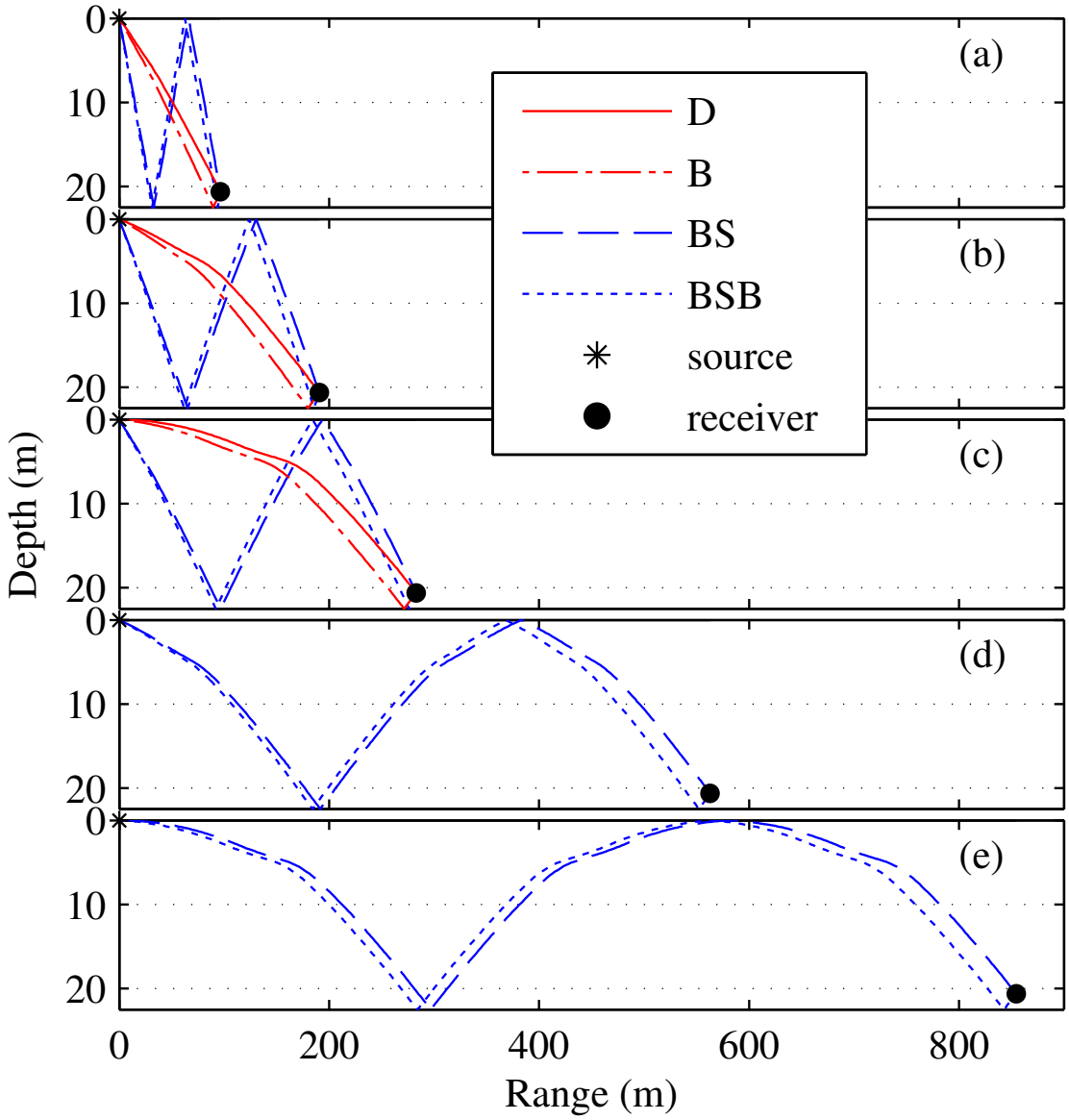


Figure 6.9: Ray trace using BELLHOP. Only D, B, BS, and BSB eigenrays are shown. Note the culling of the D and B eigenrays as the range approaches 300 m. Panel (c) represents the maximum range of the D eigenray, which corresponds to a 0° ray launch angle. Likewise, panel (e) shows the maximum range of the BS eigenray.

an effect that is amplified as rays are launched closer to the horizontal. Note that Fig. 6.9(c) is the last panel that shows a D and B eigenray. This corresponds to the maximum range for which the D eigenray exists. The launch angle for the D eigenray at that range is zero (exactly horizontal), meaning that just beyond that range the ray must first reflect off the seabed to be received by the array, which essentially culls the D eigenray. The same phenomenon happens for the BS ray; its limiting range is shown in Fig. 6.9(e).

6.6 Results and Analysis

This section demonstrates the proposed technique using data from the GLASS'12 experiment, and compares this against results derived from measured environmental and GPS data. Section 6.6.1 demonstrates how adaptive beamforming can be used to measure both the target bearing and elevation angles of individual multipath arrivals. Section 6.6.2 uses the beam directions with cross-beam correlation to measure time delays between the two dominant arrivals. Section 6.6.3 aggregates time delays over the entire boat run for single- and multi-target scenarios; and also illustrates range-dependent multipath features. Section 6.6.4 interpolates the range from measured time delays.

6.6.1 Beamforming to Find Multipath Arrivals

Volumetric arrays provide the capability of steering beams in any direction, which is used here to determine target bearings as well as the multipath arrival structure that is spread over the elevation angles. In this analysis, CSDMs are computed using

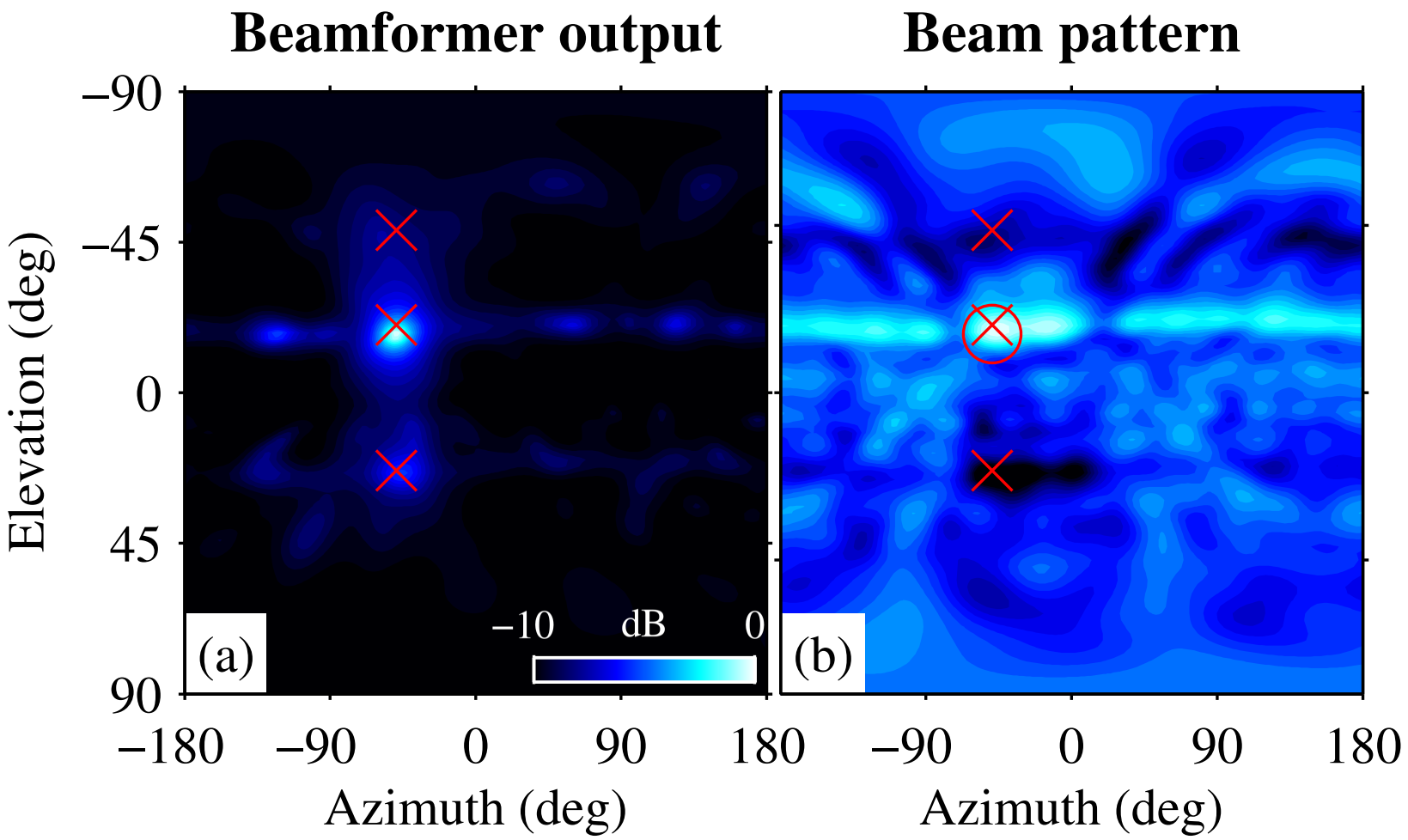


Figure 6.10: (a) Adaptive beamformer output averaged over 1–35 kHz. (b) Beam pattern steered at the strongest arrival traveling downward, averaged over the same band. The '×' annotations denote expected arrivals from ray tracing, and the '◯' annotation is the steering direction.

150 ms of total averaging time, and so the target motion is assumed to be negligible. This amount of time was chosen since it could be subdivided into enough snapshots to provide a good CSDM estimate, while also being short enough to avoid target motion smearing. The array shown in Fig. 6.6 is geometrically comprised of two sub arrays, however the full eight-element array is used in this analysis. Adaptive beamformer outputs are shown in Fig. 6.10(a) when the boat is at 58 m range and approaching the array. Snapshot lengths are 0.53 ms, and the predicted time delay between D and B eigenrays at this range is 0.904 ms. A Hann time window is applied to each snapshot during averaging, and snapshots overlap by 50%. Beamformer outputs are averaged over the 1–35 kHz frequency band. Note that the axes in Fig. 6.10(a) are oriented according to the wave propagation vector, for which negative angles correspond to downward-traveling waves and vice-versa. The expected arrival directions are predicted with a ray tracer; and these appear near strong beamformer outputs. Figure 6.10(b) is the adaptive beam pattern of the strongest beam traveling downward showing nulling of other beams, including multipath arrivals.

6.6.2 Measuring Time Delays with Cross Beam Correlation

Once the direction of the strongest arrival traveling downward is determined, that beam is then cross correlated with other elevations on the same bearing to find the beam containing the next multipath arrival to measure its delay. This step uses 10 ms snapshots in order to bring multipath coherence into the CSDM.

This processing follows the steps in Sec. 6.3 such that in Eq. 6.10 the reference beam is the strongest beam traveling downward and the correlated beam is steered in different elevation angles looking for correlations. If the correlated beam contains a delayed

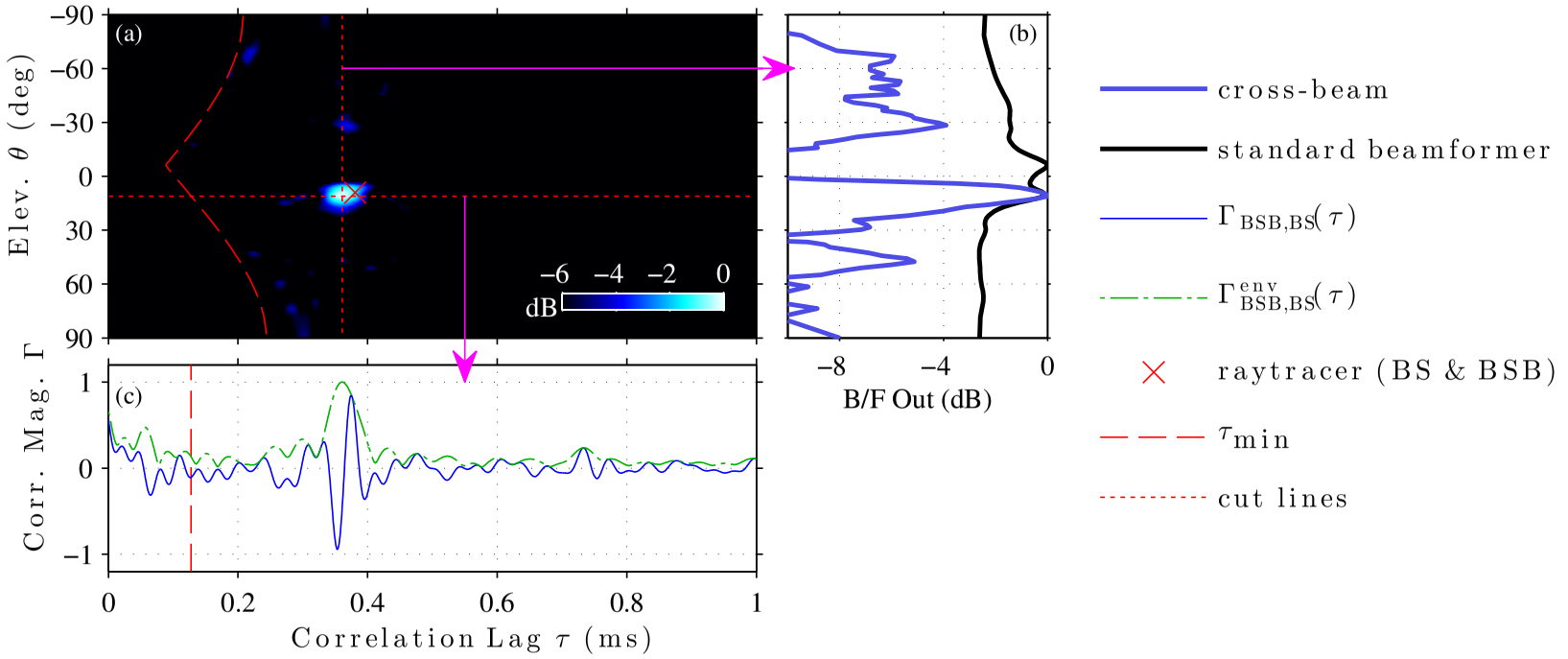


Figure 6.11: Cross beam correlation of the strongest arrival traveling downward with other elevation angles along the same azimuthal direction. (a) Correlation power envelope at each elevation angle and time delay. (b) A vertical slice of the cross beamformer taken at the measured multipath delay compared with a standard adaptive beamformer output. (c) Cross beamformer time series output at the measured multipath elevation angle showing the actual waveform and its envelope.

waveform with respect to the reference beam, the correlation delay will be positive in Eqs. 6.11 and 6.12. This is demonstrated in Fig. 6.11(a) in which the boat is at a distance of 797 m and is approaching the array. Each row is a separate cross-correlation between that elevation angle and the strongest beam traveling downward, measured at $-5.9°$. An annotation shows where the ray model predicts the BSB arrival to be; at $9.02°$ and 0.38 ms after the BS arrival. This agrees with the data in which a clear peak is visible on the $11.06°$ beam at a time delay offset of 0.36 ms.

Beyond validation against the ray model, the peak in Fig. 6.11(a) also matches a peak in the standard adaptive beamformer output, further indicating it is the multipath arrival from the seabed. This can be seen by cutting through the elevation angles at the measured delay, as shown in Fig. 6.11(b).

Further analyzing the actual time series on the $11.06°$ beam, the cross correlation and its envelope are shown in Fig. 6.11(c). The shape of the pulse in the non-envelope line is not a sinc function, as might be expected from a rectangular band of frequencies, but rather shows a significant amount of skewing. This is due to a phase change picked up from an additional seabed reflection, and is predictable from the ray tracer using measured environmental parameters and assuming a half space seabed. The envelope operation eliminates this constant phase term and yields a maximal value at the time delay between beams.

The minimum bound on time delay measurements, τ_{min}, varies as a function of the angular subtense between beams being cross correlated, and is computed according to Eq. 6.14. The observed peak width in Fig. 6.11(c) corresponds to an effective signal bandwidth of $B_s = 11.3$ kHz.

6.6.3 Time Evolution of Cross Beam Measurements

This section analyzes how measured time delays between the first two eigenray arrivals change as the boat maneuvers. By stacking the time series [i.e. Fig. 6.11(c)] at several steps during the boat pass, a cross beam correlogram is produced, as shown in Fig. 6.12. Note that this differs from the more typical usage of correlograms that are obtained by cross correlating two receivers. In contrast, this image shows the cross correlation of two beams steered at different elevation angles on a single bearing. The bearing corresponds to the direction of arrival of boat noise, and changes as the boat maneuvers. To analyze the multipath originating from a single track, prior knowledge of the boat's bearing was used to restrict the bearing search space. Overlaid on the plots of Fig. 6.12 is the expected time delay of different arrival pairs as estimated from Bellhop using the measured SSP. Note the ranges at which these lines terminate correspond to the ranges at which the different eigenray pairs are culled in Fig. 6.9. This also shows strong agreement with the ranges at which the correlation peaks in the acoustic data change visibility. This is true for both the D-with-B and BS-with-BSB correlations.

Similar results are obtained for Tracks 1 and 2 [comparing Fig. 6.12(a) and (b)]. Track 1 approaches the array from the north and Track 2 from the south. Yet, the BS-with-BSB correlation is visible in both only on the approach, indicating a strong aspect dependence on boat noise radiation. An artificial multi-target scenario is shown in Fig. 6.12(c) in which acoustic data from Track 1 and Track 2 are added together. The same processing is applied that focuses the cross beamformer on the arrivals from Track 1, and it is seen how adaptive beamforming effectively nulls the Track 2 interferer.

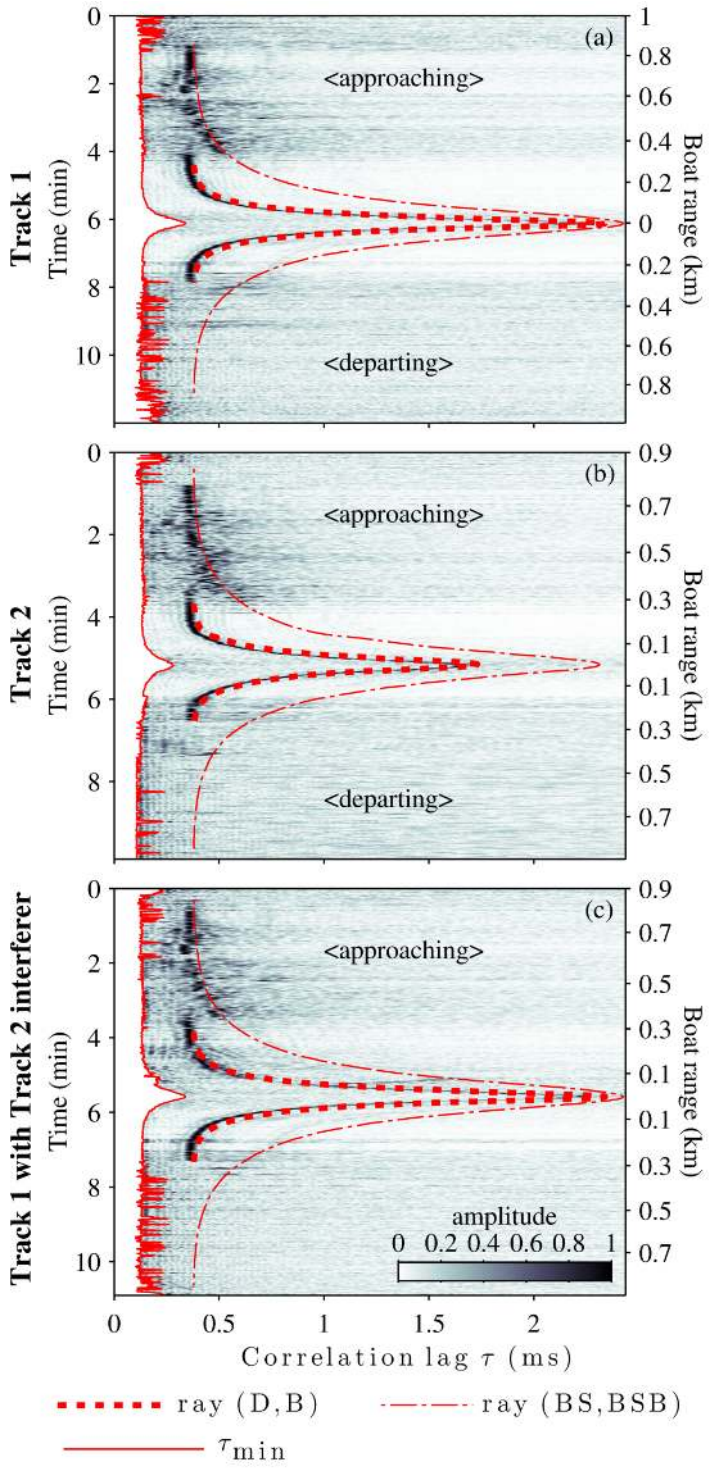


Figure 6.12: Correlograms of multipath-steered beams with overlays showing model predictions of multipath time delays. (a,b) individual boat tracks. (c) multi-target scenario (sum of acoustic data from both tracks) with cross-beamformer steered at Track 1.

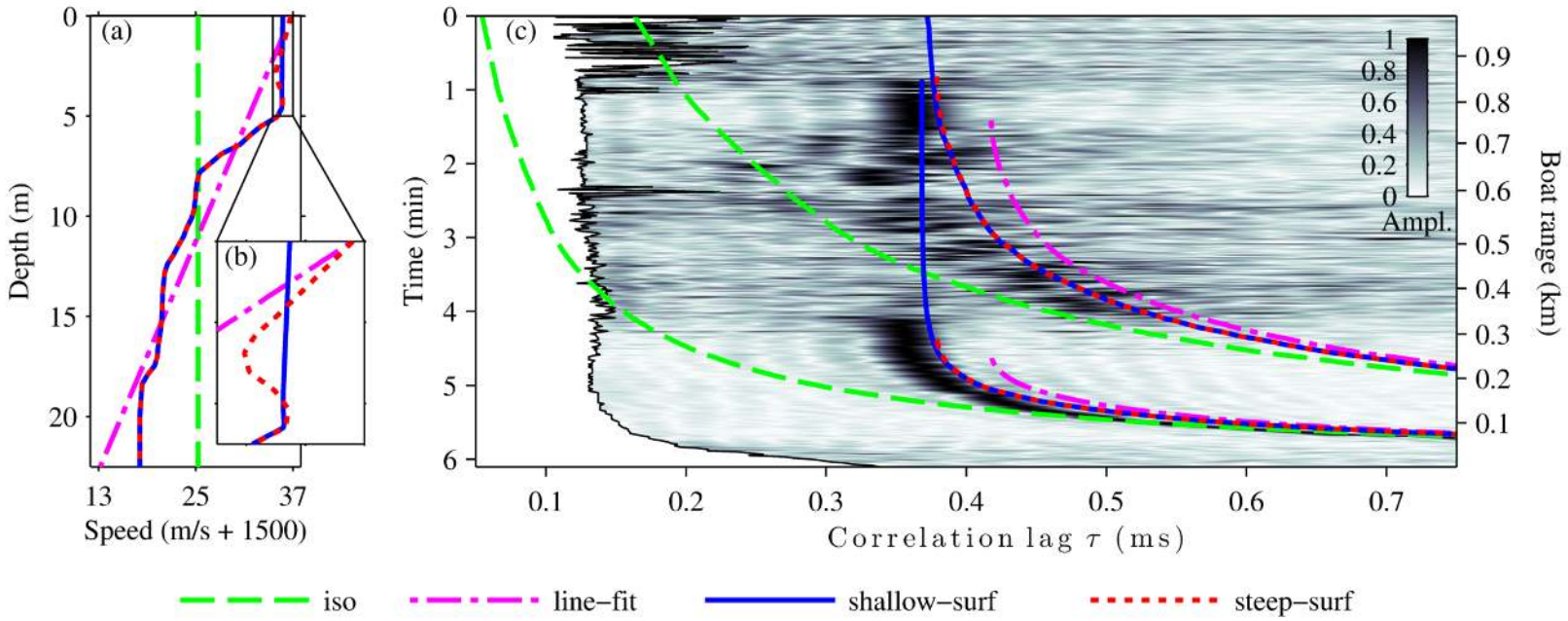


Figure 6.13: (a-b) Different SSPs used to configure the ray tracer. (c) Correlogram with overlays showing the multipath delays resulting from each SSP. There are two lines for each SSP that correspond to the delays between each of the two sets of eigenray pairs (D-with-B and BS-with-BSB). Note the ranges at which rays are culled match well with the measured SSP. When the near surface part of the SSP is changed to have a very small gradient, the range at which culling happens changes significantly.

Initially in this study, an isovelocity (constant SSP) model was adopted. This produced accurate predictions of the multipath delay out to about 100 m, but beyond that the range delay predictions were smaller than observed. These observations became explainable once a ray model was adopted and configured with the measured SSP. The downward refracting profile causes an increase in the delay, and its effect starts to become more pronounced at about 100 m range. This is evident by comparing the multipath delays computed using different SSPs to the acoustic data in Fig. 6.13. The isovelocity model clearly diverges from the acoustic data around this range whereas the other SSPs, which are downward-refracting, produce measurable delays out to farther ranges. The downward refraction causes rays launched near the surface to have steeper angles as they pass through the array depth. This implies that there is a minimum bound on the delay between a downward-traveling ray received at the array and the subsequent ray that bounces off the seabed. This is roughly at 0.36 ms.

The CTD data was processed using a linear fit, and is shown as the dot-dash line in Fig. 6.13. This line tracks better with the measured delays, and shows culling at comparable ranges. Culling happens when an eigenray path disappears due to refraction. For example, the range at which the horizontally-launched ray passes through the array depth corresponds to the maximum range at which the D eigenray (and thus its correlation with the B eigenray) exists.

As seen in Fig. 6.5, CTD data was recorded over the majority of the water column, but a few meters near the surface and seabed were not measured accurately. These sections were then populated by manually extrapolating to the boundaries using straight lines. The slope of these lines had a significant impact on the multipath structure. In

Fig. 6.13, the "shallow-surf" SSP uses a -0.03 m/s per m in the top 5 m, whereas the "slope-surf" uses a more pronounced -0.7 m/s per m slope in the top 2 m. A lower depth was used for the shallow-surf SSP to prevent surface ducting, as this was not a phenomenon under investigation. Using the measured SSP causes better agreement with the acoustic data. The main effect of flattening the profile near the surface is increasing the distance that rays travel near the surface before being pulled down by the steeper gradient, causing ray culling to occur at greater ranges. The value of -0.7 m/s per m was found through manual adjustment to match the culling behavior observed in the acoustic data. It was also observed that lowering the source depth by up to one meter had a minimal impact. It is evident that the cross beam output is highly sensitive to environment, and the water SSP in particular.

6.6.4 Range Estimation

The previous section demonstrated that the multipath delay is a stable measurement, and this section shows how that can be used for target ranging. The measured SSP indicates that the waveguide is downward refracting and that the first pair of arrivals are culled at specific ranges. This brings up an ambiguity as it is not known solely from the measured time delay which pair of eigenrays are being correlated. However, modeling can help with this. Assuming the correct pair of eigenrays is chosen, the range follows by matching the measured time delay to modeled time delays computed over range. The assumption being made here is that time delays for a single eigenray pair are monotonically decreasing as a function of range; and this is clearly observed in both modeled and measured curves in Fig. 6.12. However, this assumption may not be valid in regions with significant bathymetry variations.

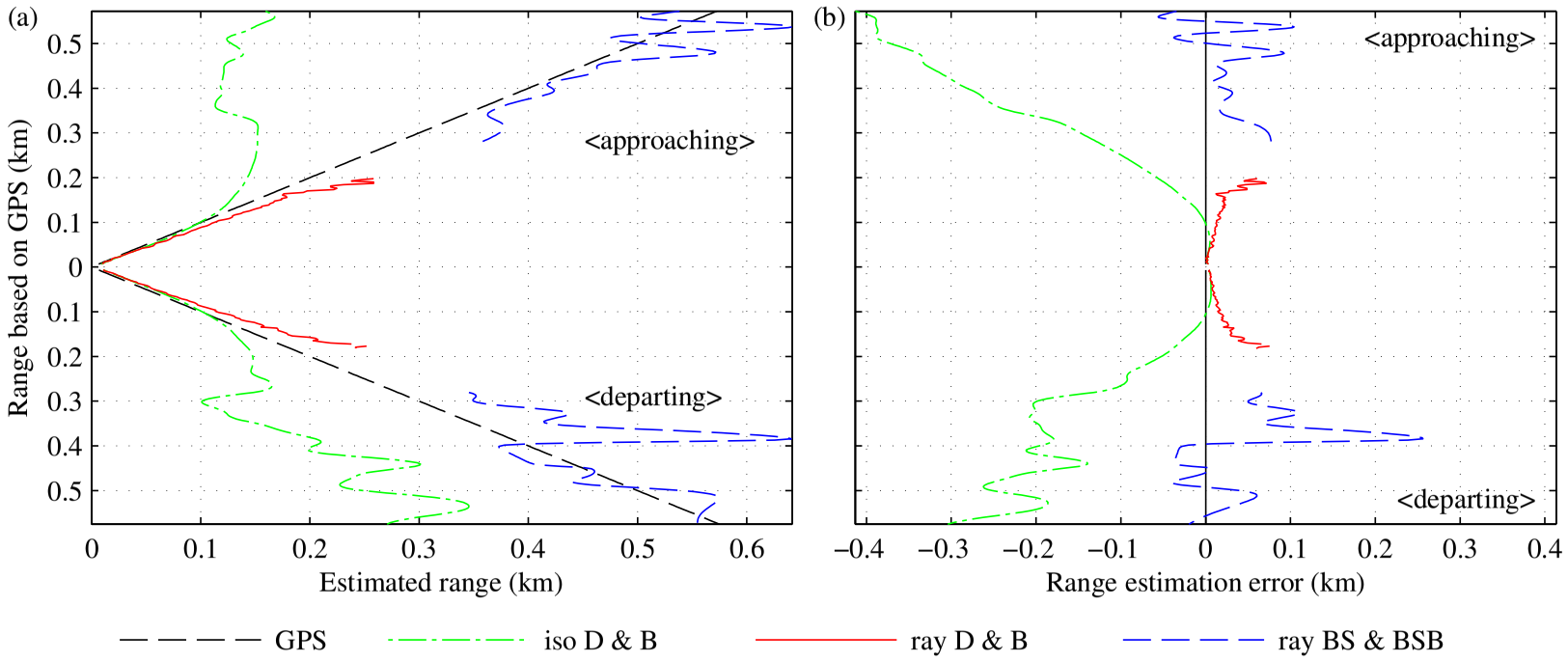


Figure 6.14: (a) Range estimates for different water SSPs. (b) Errors in range estimates relative to GPS records. The "iso" line assumes a constant SSP (isovelocity) while the "ray" lines use the measured SSP and a ray tracer to predict delays from different pairs of eigenrays. Refraction effects cause the isovelocity model to breakdown at closer ranges while a ray model provides a significantly improved detection range.

The range inversion results are shown in Fig. 6.14(a). Errors relative to the GPS records are shown in Fig. 6.14(b). Comparing isovelocity and ray models, it can be seen that the effects of refraction cause the isovelocity model to start to break down between 100 and 200 m, whereas the ray model provides reasonable estimates to over 500 m (over 22 water depths). Gaussian smoothing has been applied to the isovelocity and ray BS & BSB curves as a visual aid to mitigate the measured and modeled variance of $\tau_{2,1}$. Nevertheless, the variance of the range estimates is observed to increase at more distant ranges. This is because the time delay measurement becomes less sensitive to changes in range when the target is farther away. Visually, this can be seen in Fig. 6.12 by the flattening out of both the measured and predicted time delays at greater ranges.

6.7 Relationship to Waveguide Invariance

Multipath coherence gives rise to commonly observed striations in spectrograms, for which the slope has been shown to depend on the waveguide characteristics and source-receiver geometry (D'Spain and Kuperman, 1999). This has numerous uses such as passive determination of range and radial velocity from sources (Rakotonarivo and Kuperman, 2012, Kapolka, 2008), and improved active sonar beamforming (Zurk and Rouseff, 2012). The waveguide invariant (Chuprov, 1982), denoted by β, is the slope of striations in frequency-range space according to $\frac{\partial \omega}{\partial r} = \beta \frac{\omega}{r}$. While analysis has traditionally been formulated in a modal regime, Harrison (2011) showed that the relative eigenray travel times give the striations their shape, and computed β for several analytic sound speed profiles.

This section derives a simple form for both the spectrogram striations and for β in

the general scenario in which the geometry and environment support two dominant ray paths. First, the intensity at a receiver is obtained by multiplying out Eq. 6.1,

$$I = |S|^2 = SS^* \tag{6.15}$$

$$= \left(e^{-i\omega t_1} + Re^{-i\omega t_2}\right)\left(e^{-i\omega t_1} + Re^{-i\omega t_2}\right)^* \tag{6.16}$$

$$= 1 + |R|^2 + 2|R|\cos\left[\omega\tau_{2,1} - \phi_R\right] \tag{6.17}$$

using Eq. 6.2 that defines $\tau_{2,1}$ as the difference between t_2 and t_1. Let $\theta_I = \omega\tau_{2,1} - \phi_R$. The cosine term in Eq. 6.17 gives rise to intensity striations in range-frequency space that follow constant θ_I contours. Striation peaks correspond to $\theta_I = 2\pi N$ for $N = 0, 1, 2, \ldots$, and occur at frequencies

$$\omega\tau_{2,1} - \phi_R = 2\pi N \tag{6.18}$$

$$\omega = \frac{2\pi N + \phi_R}{\tau_{2,1}}. \tag{6.19}$$

Taking the derivative of Eq. 6.19 with respect to range and then substituting ω for the right side of Eq. 6.19 yields

$$\frac{\partial\omega}{\partial r} = -\frac{2\pi N + \phi_R}{\tau_{2,1}^2}\frac{\partial\tau_{2,1}}{\partial r} + \frac{1}{\tau_{2,1}}\frac{\partial\phi_R}{\partial r} \tag{6.20}$$

$$= -\left(\frac{2\pi N + \phi_R}{\tau_{2,1}}\right)\frac{1}{\tau_{2,1}}\frac{\partial\tau_{2,1}}{\partial r} + \frac{1}{\tau_{2,1}}\frac{\partial\phi_R}{\partial r} \tag{6.21}$$

$$= -\frac{\omega}{\tau_{2,1}}\frac{\partial\tau_{2,1}}{\partial r} + \frac{1}{\tau_{2,1}}\frac{\partial\phi_R}{\partial r}. \tag{6.22}$$

Extracting an $\frac{\omega}{r}$ term from the right side of Eq. 6.22, β for the N^{th} striation can be

defined as

$$\frac{\partial \omega}{\partial r} = \left(-\frac{r}{\tau_{2,1}} \frac{\partial \tau_{2,1}}{\partial r} + \frac{r}{\omega \tau_{2,1}} \frac{\partial \phi_R}{\partial r} \right) \frac{\omega}{r} \tag{6.23}$$

$$\beta = -\frac{r}{\tau_{2,1}} \frac{\partial \tau_{2,1}}{\partial r} + \frac{r}{\omega \tau_{2,1}} \frac{\partial \phi_R}{\partial r} \tag{6.24}$$

$$= -\frac{r}{\tau_{2,1}} \frac{\partial \tau_{2,1}}{\partial r} + \frac{r}{2\pi N + \phi_R} \frac{\partial \phi_R}{\partial r}. \tag{6.25}$$

Note that at higher striation numbers, or with a slowly changing ϕ_R (i.e. for reflections above critical angle), the first term dominates.

6.7.1 Computing Striations and β from Cross Beamformer Measurements

With *a priori* knowledge of range and radial velocity, β can be computed with Eq. 6.25 using the following measured quantities. The mulitpath time delay $\tau_{2,1}$ can be obtained from $\Gamma_{\mathrm{c,r}}^{\mathrm{env}}(\tau)$ in Eq. 6.12 using Eq. 6.13. Observing its change over time allows $\frac{\partial \tau_{2,1}}{\partial t_s}$ to be calculated. The envelope operation in Eq. 6.12 is needed to effectively eliminate ϕ_R that is present in $\Gamma_{\mathrm{c,r}}(\tau)$ (Eq. 6.11) which causes skewing of the compressed pulse. Strictly speaking, the change rate of ϕ_R is needed to predict the slope of a specific striation, however this is a secondary factor compared to the time delay. Nevertheless, a simple method of measuring ϕ_R from $\Gamma_{\mathrm{c,r}}(\tau)$ under the assumption that it does not vary with frequency is as follows.

Starting with $C_{2,1}$ from Eq. 6.10, the pulse is first moved to the time origin by

$$\Gamma_{2,1}(\tau + \tau_{2,1}) = \mathcal{F}^{-1}\left[C_{2,1} e^{i\omega\tau_{2,1}} \right]. \tag{6.26}$$

This eliminates the linear phase term $e^{-i\omega\tau_{2,1}}$. Next, it is time-gated to remove all

but the correlation pulse, and then brought back into the frequency domain by

$$C_{2,1}^{\text{gate}} = \mathcal{F}^{-1}\left[\Gamma_{2,1}(\tau + \tau_{2,1})s(\tau)\right] = |R|e^{i\phi_R}, \tag{6.27}$$

in which $s(\tau)$ is a window function. The window function should be an even function having unity support of a minimum width $1/B_s$ (inverse of signal bandwidth); a Tukey window is used here. The result is an estimate of the spectra of the multipath correlation and has spectral levels that match the correlated signal. To estimate ϕ_R, the levels are averaged using a weighted mean over ω as

$$\phi_R = \angle\left[\left|C_{2,1}^{\text{gate}}\right|^{\alpha} C_{2,1}^{\text{gate}}\right]_{\omega}, \tag{6.28}$$

in which the $\angle$ operator returns the phase angle of a complex number. The variable α controls the degree of weighting which is useful to reduce corruption of the phase estimate by frequencies for which there is no signal (and thus random phase). A value of zero means no weighting, and a value of one weights according magnitude. The latter is a natural choice because Eq. 6.28 is being applied to a cross-correlated signal that already has units of intensity.

6.7.2 Experimental Prediction of Striations and β

This section demonstrates estimation of spectrogram striations and β using output from the cross beamformer. A single hydrophone spectrogram is shown in Fig. 6.15(a) for Track 1. Overlaid on this plot are striation lines predicted from $\tau_{2,1}$ and ϕ_R measured with the cross beamformer using Eq. 6.19 for $N = 1 \ldots 12$. Figure 6.15(b) shows a synthetic spectrogram obtained from Bellhop along with overlays showing

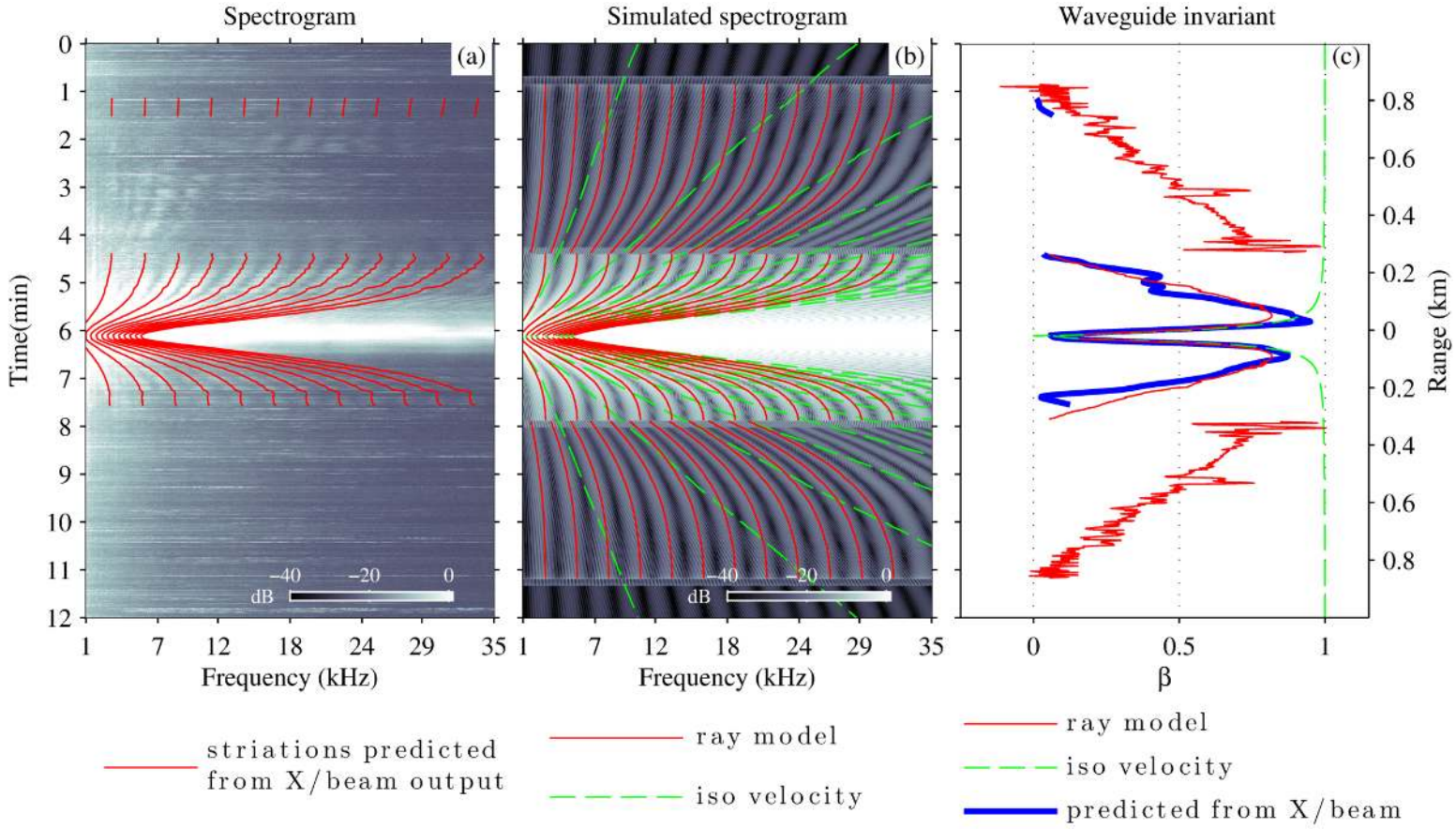


Figure 6.15: (a) Spectrogram of a single channel for the multi-target scenario. Overlays show striations predicted by Eq. 6.19 using $\tau_{2,1}$ and ϕ_R measured by the cross beamformer. (b) Simulation of acoustic data from the same receiver. Overlays use $\tau_{2,1}$ and ϕ_R from the eigenray outputs of Bellhop, and are shown for different SSPs. (c) Calculation of β using Eq. 6.25 from measured and modeled data, assuming a constant ω of 3 kHz. Noise in the outer lines of the ray model are likely due to floating point rounding errors.

different model predictions (also using Eq. 6.19). Striations are computed with eigenrays from BELLHOP, but using different SSPs. Sections of the data for which $\tau_{2,1}$ is not measurable or highly variant have been removed, as these produce meaningless output.

Using the derivatives of $\tau_{2,1}$ and derivatives of ϕ_R, β can be calculated according to Eq. 6.25; this is plotted against modeled results in Fig. 6.15(c) for a fixed frequency of 3 kHz. In order to estimate the derivatives of $\tau_{2,1}$ and ϕ_R, Gaussian smoothing was used to reduce the variance of both those quantities over target range. Note that noise in the outer lines of the ray model are likely due to floating point rounding errors, and would otherwise be smooth lines following the same trend.

6.8 Chapter Summary

The technique presented in this chapter is a generalization of the passive fathometer to non-vertically traveling waves using an array geometry other than a vertical line. While this has been formulated in the context of target localization, the fundamental approach of using cross-beam correlation to process noise from a distant source is important because it provides a new way to measure acoustic propagation through a waveguide. Target localization is essentially a matter of interpreting measured multipath time delays and phases while properly accounting for propagation effects. However, the propagation effects may themselves be the object of study in future work. For example, it was observed that measurements were sensitive to the water sound speed profile, implying this can be used for ocean acoustic tomography. Also, the phase of the seabed reflection coefficient is a separate piece of information contained in the output, which may serve as a useful basis for performing geoacoustic

inversion. Note that with this technique, the sound source is noise, such as from a passing boat.

Ray theory provides, perhaps, the most useful framework with which to understand the output of cross-beam correlation. It estimates a set of eigenrays that accurately describe and predict the measured quantities. In the experimental results, the strongest two eigenrays were measured out to about 35 water depths, but valid range estimates extended only to about 22 water depths. Environmental information proved to be important for accurately estimating more distant target ranges in which refraction effects were significant. The technique was demonstrated in a multi-target scenario (in which the acoustic data from two tracks were added) illustrating the ability of the compact volumetric array to handle off-bearing interference. The ray framework is also used to explain how this technique is related to the phenomenon of waveguide invariance. It is shown that measured time delays and phases can predict spectrogram striations and the parameter β (if range and radial velocity are known).

The array used in this study is highly versatile and well suited for mobile platforms. Hence, this processing technique opens new possibilities for experimental designs. For example, AUVs that transect the water column could leverage depth-dependent environmental measurements to enhance on-board passive target localization capabilities.

Chapter 7

Summary and Future Directions

This dissertation makes the following contributions to the field of underwater acoustics.

- Characterization of emissions from the propulsion system of an underway AUV in terms of aspect-dependent source level and spectral decomposition.
- A technique for localizing small boats in both range and unambiguous bearing using two fixed, bottom-mounted hydrophones by exploiting regional bathymetry. A rigorous performance analysis of target ranging and a probabilistic model for the array side discrimination capability.
- A technique for localizing small boats using a compact, AUV-mounted, volumetric hydrophone array that extends the core processing principle of the passive fathometer to a distant target. This represents a new method of measuring multipath, and its connection to commonly observed multipath-based phenomenon of spectrogram striations and waveguide invariance is elucidated theoretically and demonstrated experimentally.

This dissertation advances the capabilities of passive sonar technology for detecting and localizing small boats, and it lays essential groundwork that will be needed for developing detection and localization schemes for AUVs. This work is experimentally-driven and is offers practical information and robust solutions to pressing problems in this field.

7.1 Future Directions

The AUV noise characterization from Ch. 4 is useful information for building a localization system for a particular AUV. It is expected that different AUVs will produce different acoustic emissions. As such, further work might focus on comparing these emissions to see what features, if any, they have in common. While it is unlikely that propeller cavitation was a factor for the AUV in this study, this may not be the case for other platforms. Further analysis of the two tones that are emitted by the REMUS-100 should focus on their correlation. Ultimately, coherence between these tones might be exploited for improving range or depth localization.

The range localization performance of the two-hydrophone processing technique from Ch. 5 was formulated in the context of the Bayesian Cramér-Rao lower bound. While multiple environmental parameters were treated as "hidden," the vertical sound velocity profile was assumed to be constant. At distant target ranges, vertical gradients will likely have a significant impact on target location estimates. One question is how to parameterize the vertical sound velocity profile. However, a more challenging question is how to determine the partial derivatives of the received spectrum with respect to these parameters. A possible methodology may involve the use of Monte Carlo techniques to approximate these derivatives individually, or to approximate the

integral in Whittle's theorem (see Eq. 5.22) directly.

In Ch. 6, a volumetric array was used to localize a small boat in range and bearing using cross beamforming. To accurately estimate target ranges, it was necessary to predict the relative arrival times of the first two eigenray arrivals. At more distant ranges, the downward-refracting sound velocity profile increasingly affected range estimates. This sensitivity to the water sound speed suggests that cross beamforming may be used to study the water properties. Since cross-beamforming is a passive technique that processes noise from a moving target for which the source waveform need not be known, one potential use may be for passive ocean acoustic tomography with an opportunistic source. Another aspect of the cross beamformer output is the shape of the measured pulse. It was determined theoretically and experimentally that this shape encodes information about the seabed properties. A method was presented here to extract this information in order to demonstrate the relationship to spectrogram striations. However, future work may investigate whether it can be utilized for geoacoustic inversion.

7.2 Full List of Publications and Proceedings Related to this Research

Below is a comprehensive list of literature contributions and presentations by the author related to this research.

7.2.1 Journal Publications

- John Gebbie, Martin Siderius, and John S. Allen. Performance bounds of passively localizing a noise-producing target from two hydrophones by exploiting

bathymetry (submitted). *The Journal of the Acoustical Society of America*, 2014b

- John Gebbie, Martin Siderius, Peter L. Nielsen, and James Miller. Passive localization of noise-producing targets using a compact volumetric array (accepted). *The Journal of the Acoustical Society of America*, 2014c

- John Gebbie, Martin Siderius, Reid McCargar, John S. Allen, and Grant Pusey. Localization of a noisy broadband surface target using time differences of multipath arrivals. *The Journal of the Acoustical Society of America Express Letters*, 134(1):EL77–EL83, June 2013a. doi: 10.1121/1.4809771

- John Gebbie, Martin Siderius, and John S. Allen. Aspect-dependent radiated noise analysis of an underway autonomous underwater vehicle. *The Journal of the Acoustical Society of America Express Letters*, 132(5):EL351–EL357, October 2012b. ISSN 0001-4966. doi: 10.1121/1.4754419

- John Gebbie, Martin Siderius, and John S. Allen. Passive acoustic array harbor security applications. *Marine Technology Society Journal*, 45(3):103–110, May 2011a. ISSN 0025-3324. doi: 10.4031/MTSJ.45.3.3

7.2.2 Conference Proceedings

- John Gebbie, Martin Siderius, Peter L. Nielsen, James Miller, Steven Crocker, and Jennifer Giard. Small boat localization using adaptive 3-D beamforming on a tetrahedral and vertical line array. In *1st International Conference & Exhibition on Underwater Acoustics*, page 42, Corfu, Greece, June 2013b

- John Gebbie, Martin Siderius, Peter L. Nielsen, James H. Miller, Steven Crocker, and Jennifer Giard. Small boat localization using adaptive three-dimensional beamforming on a tetrahedral and vertical line array. In *International Conference on Acoustics*, page 070072, Montreal, Canada, June 2013c. doi: 10.1121/1.4800565

- John Gebbie, Martin Siderius, John Allen, and Grant Pusey. Small boat localization using time difference of multipath arrivals from two bottom mounted hydrophones. In *European Conference on Underwater Acoustics*, Edinburgh, UK, July 2012a

- John Gebbie, Martin Siderius, Lanfranco Muzi, and Joel Paddock. Extracting the rayleigh reflection coefficient from the passive fathometer. In *OCEANS*

2010, pages 1–10, Seattle, WA, September 2010. IEEE. doi: 10.1109/oceans.2010.5664434

7.2.3 Conference Talks

- John Gebbie, Martin Siderius, Peter L. Nielsen, and James Miller. Passive multi-target localization by cross-correlating beams of a compact volumetric array. In *167th Meeting of the Acoustical Society of America*, Providence, RI, May 2014d. doi: 10.1121/1.4877919
- John Gebbie, Martin Siderius, and John S. Allen. Aspect-dependent acoustic characterization of an underway autonomous underwater vehicle. In *167th Meeting of the Acoustical Society of America*, Providence, RI, May 2014a. doi: 10.1121/1.4877143
- John Gebbie, Martin Siderius, and John S. Allen. Target detection in a shallow water environment using a twodimensional array. In *161st Meeting of the Acoustical Society of America*, volume 129, pages 2602+, Seattle, WA, April 2011b. doi: 10.1121/1.3588631

7.2.4 Co-Authored Works

- Tyler H. Wai, John S. Allen, John Gebbie, and Martin Siderius. Spectral changes in snapping shrimp sounds between dusk and dawn (expected). *The Journal of the Acoustical Society of America Express Letters*, 2014
- Tyler H. Wai, John S. Allen, John Gebbie, and Martin Siderius. Nonlinear time series analysis of snapping shrimp sounds. In *165th Meeting of the Acoustical Society of America*, volume 134, pages 4147+, San Francisco, CA, December 2013a. doi: 10.1121/1.4831199
- Tyler H. Wai, John S. Allen, John Gebbie, and Martin Siderius. Spectral changes in snapping shrimp sounds between dusk and dawn. In *OCEANS 2013*, pages 1–6, San Diego, CA, 2013b. IEEE
- John S. Allen, Grant Pusey, John Gebbie, and Martin Siderius. Integrated approaches to tracking in cluttered environments. In *162nd Meeting of the Acoustical Society of America*, volume 130, pages 2378+, San Diego, California, USA, April 2011. doi: 10.1121/1.3654532

- L. Muzi, M. Siderius, J. Gebbie, and J. Paddock. On the use of adaptive beam forming techniques for geoacoustic inversion of marine ambient noise. In *OCEANS 2010*, pages 1–6, Seattle, WA, September 2010. IEEE. ISBN 978-1-4244-4332-1. doi: 10.1109/oceans.2010.5664300

7.2.5 Unpublished Works

- John Gebbie. Small boat bearing localization using Time-Difference of arrival. unpublished, May 2013
- John Gebbie. A discussion of units and the discrete fourier transform. unpublished, April 2011
- John Gebbie. Seabed layer tracking with a particle filter. unpublished, December 2009b
- John Gebbie. Beam-forming on a vertical line array. unpublished, June 2009a

Bibliography

Costa rica implementing surveillance system for cocos island, May 2014.

J. Abel and K. Lashkari. Track parameter estimation from multipath delay information. *IEEE Journal of Oceanic Engineering*, 12(1):207–221, January 1987. ISSN 0364-9059. doi: 10.1109/joe.1987.1145224.

Michael A. Ainslie. *Principles of Sonar Performance Modelling (Springer Praxis Books / Geophysical Sciences)*. Springer-Verlag, 1st edition. edition, November 2010. ISBN 3540876618.

John S. Allen, Grant Pusey, John Gebbie, and Martin Siderius. Integrated approaches to tracking in cluttered environments. In *162nd Meeting of the Acoustical Society of America*, volume 130, pages 2378+, San Diego, California, USA, April 2011. doi: 10.1121/1.3654532.

Antenna-Theory.com. Minimum Mean-Square error (MMSE) weights. http://www.antenna-theory.com/arrays/weights/mmse.php (last viewed 05/05/2014).

Paul T. Arveson and David J. Vendittis. Radiated noise characteristics of a modern cargo ship. *The Journal of the Acoustical Society of America*, 107(1):118–129, January 2000. ISSN 00014966. doi: 10.1121/1.428344.

Roland Aubauer, Marc O. Lammers, and Whitlow W. L. Au. One-hydrophone method of estimating distance and depth of phonating dolphins in shallow water. *The Journal of the Acoustical Society of America*, 107(5):2744–2749, May 2000. doi: 10.1121/1.428660.

Florian Aulanier, Barbara Nicolas, Jérôme I. Mars, Philippe Roux, and Romain Brossier. Shallow-water acoustic tomography from angle measurements instead of travel-time measurements. *The Journal of the Acoustical Society of America Express Letters*, 134(4):EL373–EL379, September 2013. doi: 10.1121/1.4820468.

Dan Bacher. Schwarzenegger's 'marine poaching areas' to go into effect on may 1. *California Progress Report*, April 2010.

L. Badriasl, K. Dogancay, and S. Arulampalam. 3D passive localization in shallow water using bearing and multipath time-delay measurements. In *Seventh International Conference on Intelligent Sensors, Sensor Networks and Information Processing (ISSNIP)*, pages 473–478. IEEE, December 2011. ISBN 978-1-4577-0675-2. doi: 10.1109/issnip.2011.6146564.

A. B. Baggeroer, W. A. Kuperman, and P. N. Mikhalevsky. An overview of matched field methods in ocean acoustics. *IEEE Journal of Oceanic Engineering*, 18(4): 401–424, 1993. ISSN 0364-9059. doi: 10.1109/48.262292.

Lenny Bernstein. U.S. rules would allow 'seismic air guns' in search for offshore oil, gas, February 2014.

P. Blanc-Benon. Implication of shallow waters for source localization: Time-delays estimation versus matched-field processing. In *OCEANS '95. MTS/IEEE. Challenges of Our Changing Global Environment. Conference Proceedings.*, volume 2, pages 826–831 vol.2. IEEE, October 1995. ISBN 0-933957-14-9. doi: 10.1109/oceans.1995.527314.

Brian Borowski, Alexander Sutin, Heui-Seol Roh, and Barry Bunin. Passive acoustic threat detection in estuarine environments. *Optics and Photonics in Global Homeland Security IV*, 6945(13), April 2008. doi: 10.1117/12.779177.

L. M. Brekhovskikh and Yu P. Lysanov. *Fundamentals of Ocean Acoustics (Modern Acoustics and Signal Processing)*, pages 140–146. New York: Springer-Verlag, 3rd edition, April 2003. ISBN 0387954678.

M. Bruno, Kil W. Chung, H. Salloum, A. Sedunov, N. Sedunov, A. Sutin, H. Graber, and P. Mallas. Concurrent use of satellite imaging and passive acoustics for maritime domain awareness. In *International Waterside Security Conference (WSS), 2010*, pages 1–8. IEEE, November 2010. ISBN 978-1-4244-8894-0. doi: 10.1109/wssc.2010.5730229.

Homer P. Bucker. Use of calculated sound fields and matchedfield detection to locate sound sources in shallow water. *The Journal of the Acoustical Society of America*, 59(2):368–373, February 1976. doi: 10.1121/1.380872.

Barry Bunin, Alexander Sutin, George Kamberov, Heui-Seol Roh, Bart Luczynski, and Matt Burlick. Fusion of acoustic measurements with video surveillance for estuarine threat detection. In *Defense + Security*, volume 6945, pages 694514+. The International Society for Optical Engineering., April 2008. doi: 10.1117/12.779176.

John F. Burns and Steven L. Myers. The warship explosion: The overview; blast kills sailors on U.S. ship in yemen. *The New York Times*, October 2000.

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

William M. Carey, Jason D. Holmes, and James F. Lynch. The applicability of a small autonomous vehicle towed array system to ocean acoustic measurements and signal processing. In *Proceedings of Meetings on Acoustics*, volume 4, page 070007, July 2009. doi: 10.1121/1.3196623.

Damian Carrington. Whales flee from military sonar leading to mass strandings, research shows, July 2013.

G. Carter and C. Knapp. Time delay estimation. In *Acoustics, Speech, and Signal Processing, IEEE International Conference on ICASSP 1976*, volume 1, pages 357–360. IEEE, April 1976. doi: 10.1109/icassp.1976.1169979.

G. C. Carter, A. H. Nuttall, and C. K. Yuen. On the weighted overlapped segment averaging method for power spectral estimation. *Proceedings of the IEEE*, 68(10):1352–1354, October 1980. ISSN 0018-9219. doi: 10.1109/proc.1980.11866.

G. Clifford Carter. *Coherence and time delay estimation : an applied tutorial for research, development, test, and evaluation engineers*. IEEE Press, New York, 1993. ISBN 0780310063.

Douglas H. Cato. Simple methods of estimating source levels and locations of marine animal sounds. *The Journal of the Acoustical Society of America*, 104(3):1667–1678, September 1998. doi: 10.1121/1.424379.

CBS News. Drug traffickers' vehicle of choice, June 2013.

Y. T. Chan and K. C. Ho. A simple and efficient estimator for hyperbolic location. *Signal Processing, IEEE Transactions on*, 42(8):1905–1915, August 1994. ISSN 1053-587X. doi: 10.1109/78.301830.

Kil W. Chung, Alexander Sutin, Alexander Sedunov, and Michael Bruno. DEMON acoustic ship signature measurements in an urban harbor. *Advances in Acoustics and Vibration*, 2011:1–13, March 2011. ISSN 1687-6261. doi: 10.1155/2011/952798.

S. D. Chuprov. Interference structure of a sound field in a layered ocean. In L. M. Brekhovskikh and I. B. Andreevoi, editors, *Ocean Acoustics: Current State*, pages 71–91. Nauka, Moscow, 1982.

James Churnside. Airborne lidar probes undersea waves. *SPIE Newsroom*, pages 1–2, March 2012. doi: 10.1117/2.1201203.004162.

C. S. Clay. Waveguides, arrays, and filters. *Geophysics*, 31(3):501–505, June 1966. ISSN 0016-8033. doi: 10.1190/1.1439789.

H. Cox, R. Zeskind, and M. Owen. Robust adaptive beamforming. *IEEE Transactions on Acoustics, Speech and Signal Processing*, 35(10):1365–1376, 1987. ISSN 0096-3518. doi: 10.1109/TASSP.1987.1165054.

W. C. Cummings and D. V. Holliday. Sounds and source levels from bowhead whales off pt. barrow, alaska. *The Journal of the Acoustical Society of America*, 82(3): 814–821, September 1987. ISSN 0001-4966. doi: 10.1121/1.395279.

Joe Cuschieri and Susan Frandsen. Acoustic signature of an AUV. In *146th Meeting: Acoustical Society of America*, volume 114, page 2333. Acoustical Society of America, ASA, October 2003. doi: 10.1121/1.4781051.

J. Dauwels. Computing bayesian Cramer-Rao bounds. In *Information Theory, 2005. ISIT 2005. Proceedings. International Symposium on*, pages 425–429. IEEE, 2005. ISBN 0-7803-9151-9. doi: 10.1109/isit.2005.1523369.

Mark de Berg. Delaunay triangulations. In *Computational geometry algorithms and applications*, chapter 9, pages 191–218. Springer, Berlin, third edition, March 2008. ISBN 3540779736.

Stan E. Dosso and Jan Dettmer. Bayesian matched-field geoacoustic inversion. *Inverse Problems*, 27(5):055009+, April 2011. ISSN 0266-5611. doi: 10.1088/0266-5611/27/5/055009.

G. L. D'Spain and W. A. Kuperman. Application of waveguide invariants to analysis of spectrograms from shallow water environments that vary in range and azimuth. *The Journal of the Acoustical Society of America*, 106(5):2454+, 1999. ISSN 00014966. doi: 10.1121/1.428124.

Christine Erbe. Underwater noise of small personal watercraft (jet skis). *The Journal of the Acoustical Society of America Express Letters*, 133(4):EL326–EL330, March 2013. doi: 10.1121/1.4795220.

James J. Faran and Robert Hills. Correlators for signal reception. Technical Report 27, Office of Naval Research, Contract N5 ORI-76 Project Order X, September 1952.

Paul G. Fernandes, Pete Stevenson, Andrew S. Brierley, Frederick Armstrong, and Simmonds. Autonomous underwater vehicles: future platforms for fisheries acoustics. *ICES Journal of Marine Science: Journal du Conseil*, 60(3):684–691, January 2003. doi: 10.1016/s1054-3139(03)00038-9.

Richard P. Feynman, Robert B. Leighton, and Matthew L. Sands. *The Feynman lectures on physics*. Basic Books, 2011. ISBN 0465023827.

L. Fillinger, P. de Theije, M. Zampolli, A. Sutin, H. Salloum, N. Sedunov, and A. Sedunov. Towards a passive acoustic underwater system for protecting harbours against intruders. In *International Waterside Security Conference (WSS), 2010*, pages 1–7. IEEE, November 2010. ISBN 978-1-4244-8894-0. doi: 10.1109/wssc.2010.5730243.

R. G. Fizell and S. C. Wales. Source localization in range and depth in an arctic environment. *The Journal of the Acoustical Society of America*, 78(S1):S57–S58, November 1985. doi: 10.1121/1.2022889.

E. R. Franchi and M. J. Jacobson. Ray propagation in a channel with DepthVariable sound speed and current. *The Journal of the Acoustical Society of America*, 52 (1B):316–331, February 1972. ISSN 0001-4966. doi: 10.1121/1.1913094.

B. Friedlander. Accuracy of source localization using multipath delays. *IEEE Transactions on Aerospace and Electronic Systems*, 24(4):346–359, July 1988. ISSN 00189251. doi: 10.1109/7.7176.

John Gebbie. Beam-forming on a vertical line array. unpublished, June 2009a.

John Gebbie. Seabed layer tracking with a particle filter. unpublished, December 2009b.

John Gebbie. A discussion of units and the discrete fourier transform. unpublished, April 2011.

John Gebbie. Small boat bearing localization using Time-Difference of arrival. unpublished, May 2013.

John Gebbie, Martin Siderius, Lanfranco Muzi, and Joel Paddock. Extracting the rayleigh reflection coefficient from the passive fathometer. In *OCEANS 2010*, pages 1–10, Seattle, WA, September 2010. IEEE. doi: 10.1109/oceans.2010.5664434.

John Gebbie, Martin Siderius, and John S. Allen. Passive acoustic array harbor security applications. *Marine Technology Society Journal*, 45(3):103–110, May 2011a. ISSN 0025-3324. doi: 10.4031/MTSJ.45.3.3.

John Gebbie, Martin Siderius, and John S. Allen. Target detection in a shallow water environment using a twodimensional array. In *161st Meeting of the Acoustical Society of America*, volume 129, pages 2602+, Seattle, WA, April 2011b. doi: 10.1121/1.3588631.

John Gebbie, Martin Siderius, John Allen, and Grant Pusey. Small boat localization using time difference of multipath arrivals from two bottom mounted hydrophones. In *European Conference on Underwater Acoustics*, Edinburgh, UK, July 2012a.

John Gebbie, Martin Siderius, and John S. Allen. Aspect-dependent radiated noise analysis of an underway autonomous underwater vehicle. *The Journal of the Acoustical Society of America Express Letters*, 132(5):EL351–EL357, October 2012b. ISSN 0001-4966. doi: 10.1121/1.4754419.

John Gebbie, Martin Siderius, Reid McCargar, John S. Allen, and Grant Pusey. Localization of a noisy broadband surface target using time differences of multipath arrivals. *The Journal of the Acoustical Society of America Express Letters*, 134(1): EL77–EL83, June 2013a. doi: 10.1121/1.4809771.

John Gebbie, Martin Siderius, Peter L. Nielsen, James Miller, Steven Crocker, and Jennifer Giard. Small boat localization using adaptive 3-D beamforming on a tetrahedral and vertical line array. In *1st International Conference & Exhibition on Underwater Acoustics*, page 42, Corfu, Greece, June 2013b.

John Gebbie, Martin Siderius, Peter L. Nielsen, James H. Miller, Steven Crocker, and Jennifer Giard. Small boat localization using adaptive three-dimensional beamforming on a tetrahedral and vertical line array. In *International Conference on Acoustics*, page 070072, Montreal, Canada, June 2013c. doi: 10.1121/1.4800565.

John Gebbie, Martin Siderius, and John S. Allen. Aspect-dependent acoustic characterization of an underway autonomous underwater vehicle. In *167th Meeting of the Acoustical Society of America*, Providence, RI, May 2014a. doi: 10.1121/1.4877143.

John Gebbie, Martin Siderius, and John S. Allen. Performance bounds of passively localizing a noise-producing target from two hydrophones by exploiting bathymetry (submitted). *The Journal of the Acoustical Society of America*, 2014b.

John Gebbie, Martin Siderius, Peter L. Nielsen, and James Miller. Passive localization of noise-producing targets using a compact volumetric array (accepted). *The Journal of the Acoustical Society of America*, 2014c.

John Gebbie, Martin Siderius, Peter L. Nielsen, and James Miller. Passive multitarget localization by cross-correlating beams of a compact volumetric array. In *167th Meeting of the Acoustical Society of America*, Providence, RI, May 2014d. doi: 10.1121/1.4877919.

Peter Gerstoft, William S. Hodgkiss, Martin Siderius, Chen-Fen Huang, and Chris H. Harrison. Passive fathometer processing. *The Journal of the Acoustical Society of America*, 123(3):1297–1305, March 2008. ISSN 00014966. doi: 10.1121/1.2831930.

Gwyn Griffiths, Peter Enoch, and Nicholas W. Millard. On the radiated noise of the autosub autonomous underwater vehicle. *ICES Journal of Marine Science: Journal du Conseil*, 58(6):1195–1200, December 2001. doi: 10.1006/jmsc.2001.1120.

P. E. Hagen, N. Storkersen, K. Vestgard, and P. Kartvedt. The HUGIN 1000 autonomous underwater vehicle for military applications. In *OCEANS 2003*, volume 2, pages 1141–1145. IEEE, September 2003. ISBN 0-933957-30-0. doi: 10.1109/oceans.2003.178504.

W. Hahn and S. Tretter. Optimum processing for delay-vector estimation in passive signal arrays. *IEEE Transactions on Information Theory*, 19(5):608–614, September 1973. ISSN 0018-9448. doi: 10.1109/tit.1973.1055077.

William R. Hahn. Optimum signal processing for passive sonar range and bearing estimation. *The Journal of the Acoustical Society of America*, 58(1):201–207, July 1975. doi: 10.1121/1.380646.

M. Hamilton and P. M. Schultheiss. Passive ranging in multipath dominant environments. i. known multipath parameters. *Signal Processing, IEEE Transactions on*, 40(1):1–12, January 1992. ISSN 1053-587X. doi: 10.1109/78.157176.

M. Hamilton and P. M. Schultheiss. Passive ranging in multipath dominant environments: Part II-unknown multipath parameters. *Signal Processing, IEEE Transactions on*, 41(1):1+, January 1993. ISSN 1053-587X. doi: 10.1109/tsp.1993.193122.

C. H. Harrison and D. G. Simons. Geoacoustic inversion of ambient noise: A simple method. *The Journal of the Acoustical Society of America*, 112(4):1377–1389, October 2002. doi: 10.1121/1.1506365.

Chris H. Harrison. Anomalous signed passive fathometer impulse response when using adaptive beam forming. *The Journal of the Acoustical Society of America*, 125(6): 3511–3513, June 2009. doi: 10.1121/1.3126345.

Chris H. Harrison. The relation between the waveguide invariant, multipath impulse response, and ray cycles. *The Journal of the Acoustical Society of America*, 129(5): 2863+, May 2011. ISSN 00014966. doi: 10.1121/1.3569701.

Chris H. Harrison and Martin Siderius. Bottom profiling by correlating beam-steered noise sequences. *The Journal of the Acoustical Society of America*, 123(3):1282–1296, March 2008. doi: 10.1121/1.2835416.

Christopher H. Harrison. Target detection and location with ambient noise. *The Journal of the Acoustical Society of America*, 123(4):1834–1837, April 2008. doi: 10.1121/1.2872516.

R. M. Heitmeyer, W. B. Moseley, and R. G. Fizell. Full field ambiguity function processing in a complex Shallow-Water environment. In Ronald A. Wagstaff and Arthur B. Baggeroer, editors, *High Resolution Spatial Processing in Underwater Acoustics*, pages 171–191. NSTL, Mississippi, 1985.

Charles W. Holland and John Osler. High-resolution geoacoustic inversion in shallow water: A joint time- and frequency-domain technique. *The Journal of the Acoustical Society of America*, 107(3):1263–1279, March 2000. doi: 10.1121/1.428415.

J. D. Holmes and A. Kukulya. Characteristics of an autonomous underwater vehicle with a towed hydrophone array. In *OCEANS 2006*, pages 1–5. IEEE, September 2006. ISBN 1-4244-0114-3. doi: 10.1109/oceans.2006.306932.

J. D. Holmes, W. M. Carey, J. F. Lynch, A. E. Newhall, and A. Kukulya. An autonomous underwater vehicle towed array for ocean acoustic measurements and inversions. In *Oceans 2005 - Europe*, volume 2, pages 1058–1061. IEEE, June 2005. ISBN 0-7803-9103-9. doi: 10.1109/oceanse.2005.1513204.

Jason D. Holmes. *Investigation of ocean acoustics using autonomous instrumentation to quantify the water-sediment boundary properties*. PhD thesis, Boston University College of Engineering, 2007.

Jason D. Holmes, William M. Carey, and James F. Lynch. An overview of unmanned underwater vehicle noise in the low to mid frequencies bands. In *POMA - 159th Meeting Acoustical Society of America/NOISE-CON 2010*, page 065007, October 2010. doi: 10.1121/1.3492795.

Chen-Fen Huang, Peter Gerstoft, and William S. Hodgkiss. Uncertainty analysis in matched-field geoacoustic inversions. *The Journal of the Acoustical Society of America*, 119(1):197–207, January 2006. doi: 10.1121/1.2139075.

Paul Hursky, Michael B. Porter, Martin Siderius, and Vincent K. McDonald. High-frequency (816 kHz) model-based source localization. *The Journal of the Acoustical Society of America*, 115(6):3021–3032, June 2004. ISSN 0001-4966. doi: 10.1121/1.1690078.

J. P. Ianniello. Large and small error performance limits for multipath time delay estimation. *Acoustics, Speech and Signal Processing, IEEE Transactions on*, 34(2): 245–251, April 1986. ISSN 0096-3518. doi: 10.1109/TASSP.1986.1164820.

Hiroshi Iwakami, Tamaki Ura, Kenichi Asakawa, Teruo Fujii, Yoshiaki Nose, Junichi Kojima, Yuichi Shirasaki, Teruyuki Asai, Senzo Uchida, Naoto Higashi, and Tetsuo Fukuchi. Approaching whales by autonomous underwater vehicle. *Marine Technology Society Journal*, 36(1):80–85, March 2002. ISSN 00253324. doi: 10.4031/002533202787914232.

Rashi Jain and Zoi-Heleni Michalopoulou. A particle filtering approach for spatial arrival time tracking in ocean acoustics. *The Journal of the Acoustical Society of America Express Letters*, 129(6):EL236–EL241, May 2011. doi: 10.1121/1.3574766.

Finn B. Jensen, William A. Kuperman, Michael B. Porter, and Henrik Schmidt. *Computational Ocean Acoustics*. Springer New York, New York, NY, second edition, 2011a. ISBN 978-1-4419-8677-1. doi: 10.1007/978-1-4419-8678-8.

Finn B. Jensen, William A. Kuperman, Michael B. Porter, and Henrik Schmidt. Ray methods. In William M. Hartmann, editor, *Computational Ocean Acoustics*, Computational Ocean Acoustics, chapter 3, pages 115–232. Springer New York, New York, NY, second edition, 2011b. ISBN 978-1-4419-8677-1. doi: 10.1007/978-1-4419-8678-8.

Daphne Kapolka. Equivalence of the waveguide invariant and two path ray theory methods for range prediction based on lloyd's mirror patterns. In *European Conference on Underwater Acoustics Paris*, pages 357–361, 2008. doi: 10.1121/1.2934751.

Steven M. Kay. *Fundamentals of statistical signal processing: Detection theory*, volume 2. Prentice Hall PTR, Upper Saddle River, New Jersey, 1998. ISBN 013504135.

Thomas H. Kean, Lee H. Hamilton, Richard Ben-Veniste, Bob Kerrey, Fred F. Fielding, John F. Lehman, Jamie S. Gorelick, Timothy J. Roemer, Slade Gorton, and James R. Thompson. The 9/11 commission report, July 2004.

Ronald T. Kessel and Reginald D. Hollett. Underwater intruder detection sonar for harbour protection: State of the art review and implications. Technical Report ADA457007, Nato Undersea Research Centre, La Spezia, Italy, October 2006.

Kilo Nalu Study Area. Kilo nalu study area. http://www.soest.hawaii.edu/OE/kilonalu/images/oahu_area_map.gif (last viewed: 2014/06/02).

Lawrence E. Kinsler, Austin R. Frey, Alan B. Coppens, and James V. Sanders. *Fundamentals of Acoustics*. Wiley, 4th edition, December 1999. ISBN 0471847895.

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

Robert A. Koch and David P. Knobles. Geoacoustic inversion with ships as sources. *The Journal of the Acoustical Society of America*, 117(2):626–637, February 2005. doi: 10.1121/1.1848175.

Marc O. Lammers, Russell E. Brainard, Whitlow W. L. Au, T. Aran Mooney, and Kevin B. Wong. An ecological acoustic recorder (EAR) for long-term monitoring of biological and anthropogenic sounds on coral reefs and other marine habitats. *The Journal of the Acoustical Society of America*, 123(3):1720–1728, March 2008. doi: 10.1121/1.2836780.

Shane W. Lani, Karim G. Sabra, William S. Hodgkiss, W. A. Kuperman, and Philippe Roux. Coherent processing of shipping noise for ocean monitoring. *The Journal of the Acoustical Society of America Express Letters*, 133(2):EL108–EL113, January 2013. ISSN 00014966. doi: 10.1121/1.4776775.

Christophe Laplanche. A bayesian method to estimate the depth and the range of phonating sperm whales using a single hydrophone. *The Journal of the Acoustical Society of America*, 121(3):1519–1528, March 2007. doi: 10.1121/1.2436644.

Christophe Laplanche, Olivier Adam, Maciej Lopatka, and Jean-François Motsch. Male sperm whale acoustic behavior observed from multipaths at a single hydrophone. *The Journal of the Acoustical Society of America*, 118(4):2677–2687, October 2005. doi: 10.1121/1.2033567.

Donald Launer. *Navigation through the ages*. Sheridan House, 2009. ISBN 1574092782.

Harry B. Lee. A novel procedure for assessing the accuracy of hyperbolic multilateration systems. *Aerospace and Electronic Systems, IEEE Transactions on*, AES-11 (1):2–15, January 1975. ISSN 0018-9251. doi: 10.1109/taes.1975.308023.

N. Lee, L. M. Zurk, and J. Ward. Evaluation of reduced-rank, adaptive matched field processing algorithms for passive sonar detection in a shallow-water environment. In *Signals, Systems, and Computers, 1999. Conference Record of the Thirty-Third Asilomar Conference on*, volume 2, pages 876–880. IEEE, October 1999. ISBN 0-7803-5700-0. doi: 10.1109/acssc.1999.831835.

S. H. Lee, C. S. Ryu, and K. K. Lee. Near-field source localisation using bottom-mounted linear sensor array in multipath environment. *IEE Proceedings - Radar, Sonar and Navigation*, 149(4):202+, 2002. ISSN 13502395. doi: 10.1049/ip-rsn:20020494.

Eric A. Lehmann and Robert C. Williamson. Posterior Cramer-Rao bound for acoustic source tracking in reverberant environments. Technical report, National ICT Australia Technical Reports, 2007.

Charlotte Leroy, Shane Lani, Karim G. Sabra, William S. Hodgkiss, W. A. Kuperman, and Philippe Roux. Enhancing the emergence rate of coherent wavefronts from ocean ambient noise correlations using spatio-temporal filters. *The Journal of the Acoustical Society of America*, 132(2):883–893, August 2012. ISSN 0001-4966. doi: 10.1121/1.4731231.

Zizheng Li, Lisa M. Zurk, and B. Ma. Vertical arrival structure of shipping noise in deep water channels. In *OCEANS 2010*, pages 1–8. IEEE, 2010. ISBN 978-1-4244-4332-1. doi: 10.1109/oceans.2010.5664539.

J. Lotz, Lisa M. Zurk, James McNames, T. Ellis, and J. L. Ecochard. Coral fish shoal detection from acoustic echograms. In *OCEANS 2007*, pages 1–7. IEEE, 2007. ISBN 978-0933957-35-0. doi: 10.1109/oceans.2007.4449220.

B. Magliozzi, D. B. Hanson, and R. K. Amit. Propeller and propfan noise. Technical Report N92-10599, NASA, Windsor Locks, Connecticut, 1992.

Irene Magyar, W. M. Schleidt, and B. Miller. Localization of sound producing animals using the arrival time differences of their signals at an array of microphones. *Experientia*, 34(5):676–677, May 1978. doi: 10.1007/bf01937030.

Nathan Marchand. Error distributions of best estimate of position from multiple time difference hyperbolic networks. *Aerospace and Navigational Electronics, IEEE Transactions on*, ANE-11(2):96–100, June 1964. ISSN 0096-1957. doi: 10.1109/tane.1964.4502170.

Delphine Mathias, Aaron M. Thode, Jan Straley, and Russel D. Andrews. Acoustic tracking of sperm whales in the gulf of alaska using a two-element vertical array and tags. *The Journal of the Acoustical Society of America*, 134(3):2446–2461, September 2013. ISSN 0001-4966. doi: 10.1121/1.4816565.

MathWorks. *Symbolic Math Toolbox version 5.11*. The MathWorks Inc., Natick, Massachusetts, matlab r2013b v8.2.0.701 edition, 2013.

MATLAB. *version 7.10.0 (R2010a)*. The MathWorks Inc., Natick, Massachusetts, 2010.

Steven L. Means and Martin Siderius. Effects of sea-surface conditions on passive fathometry and bottom characterization. *The Journal of the Acoustical Society of America*, 126(5):2234–2241, November 2009. ISSN 0001-4966. doi: 10.1121/1.3216915.

Christoph Mecklenbrauker. A bibliography of publications about matched field processing. Privately Published, August 2001.

Ravi Menon and Peter Gerstoft. High resolution beamforming using l1 minimization. In *Proceedings of Meetings on Acoustics*, volume 19, pages 1–4, 2013. doi: 10.1121/1.4799519.

Z. H. Michalopoulou and M. B. Porter. Matched-field processing for broad-band source localization. *Oceanic Engineering, IEEE Journal of*, 21(4):384–392, October 1996. ISSN 0364-9059. doi: 10.1109/48.544049.

Zoi-Heleni Michalopoulou and Rashi Jain. Particle filtering for arrival time tracking in space and source localization. *The Journal of the Acoustical Society of America*, 132(5):3041+, 2012. ISSN 00014966. doi: 10.1121/1.4756954.

B. Miller and S. Dawson. A large-aperture low-cost hydrophone array for tracking whales from small boats. *The Journal of the Acoustical Society of America*, 126(5): 2248–2256, November 2009. doi: 10.1121/1.3238258.

Leonard Moreavek and T. J. Brudner. USS asheville leads the way in high frequency sonar. *Undersea Warfare*, 1(3), April 1999.

Xavier Mouy, David Hannay, Mikhail Zykov, and Bruce Martin. Tracking of pacific walruses in the chukchi sea using a single hydrophone. *The Journal of the Acoustical Society of America*, 131(2):1349–1358, February 2012. ISSN 0001-4966. doi: 10.1121/1.3675008.

L. Muzi, M. Siderius, J. Gebbie, and J. Paddock. On the use of adaptive beam forming techniques for geoacoustic inversion of marine ambient noise. In *OCEANS 2010*, pages 1–6, Seattle, WA, September 2010. IEEE. ISBN 978-1-4244-4332-1. doi: 10.1109/oceans.2010.5664300.

Lanfranco Muzi and Martin Siderius. Synthetic-array beamforming for bottom-loss estimation using marine ambient noise. *Proceedings of Meetings on Acoustics*, 19 (1):070033, June 2013. ISSN 1939-800X. doi: 10.1121/1.4799120.

National Defense Research Committee. *Principles and Applications Of Underwater Sound - Summary Technical Report of Division 6*, pages 175–199. Number 7. 1946.

National Research Council (U.S.). *An evaluation of the U.S. Navy's extremely low frequency communications system ecological monitoring program*. National Academy Press, Washington, DC, 1997. ISBN 9780309055901.

National September 11 Memorial & Museum. *USS Cole Bombing*, 2014. http://www.911memorial.org/uss-cole-bombing (last viewed 04/28/2014).

Peter L. Nielsen, James H. Miller, Martin Siderius, Steven Crocker, and Jennifer Giard. GLASS'12 - the GLider acoustics sensing of sediments experiment 2012. Technical Report CMRE-FR-2012-XX, Center for Maritime Research and Experimentation, 2012.

Harbor and Coastal Security. Northrop Grumman, Herndon, VA, 2011.

E. Nosal and L. Neil Frazer. Track of a sperm whale from delays between direct and surface-reflected clicks. *Applied Acoustics*, 67(11-12):1187–1201, November 2006. ISSN 0003682X. doi: 10.1016/j.apacoust.2006.05.005.

Eva-Marie Nosal and L. Neil Frazer. Sperm whale three-dimensional track, swim orientation, beam pattern, and click levels observed on bottom-mounted hydrophones. *The Journal of the Acoustical Society of America*, 122(4):1969–1978, October 2007. doi: 10.1121/1.2775423.

Tobias G. Oesterlein, Chensong He, Jorge E. Quijano, Richard L. Campbell, Lisa M. Zurk, and Martin Siderius. Extraction of time-frequency target features. In *2010 44th Asilomar Conference on Signals, Systems and Computers*, pages 2156–2163. IEEE, November 2010. ISBN 978-1-4244-9722-5. doi: 10.1109/ACSSC.2010.5757933.

George L. Ogden, Lisa M. Zurk, Mark E. Jones, and Mary E. Peterson. Extraction of small boat harmonic signatures from passive sonar. *The Journal of the Acoustical Society of America*, 129(6):3768–3776, June 2011. doi: 10.1121/1.3583500.

Helen H. Ou, Pasang Sherpa, and Lisa M. Zurk. Tracking of small vessels with passive acoustic sensors. *The Journal of the Acoustical Society of America*, 130(4):2450, October 2011. ISSN 0001-4966. doi: 10.1121/1.3654841.

Michael B. Porter and Homer P. Bucker. Gaussian beam tracing for computing ocean acoustic fields. *The Journal of the Acoustical Society of America*, 82(4):1349–1359, October 1987. doi: 10.1121/1.395269.

Michael B. Porter and A. Tolstoy. The matched field processing benchmark problems. *Journal of Computational Acoustics*, 02(03):161–185, September 1994. ISSN 0218-396X. doi: 10.1142/s0218396x94000129.

Jorge E. Quijano, Stan E. Dosso, Jan Dettmer, Lisa M. Zurk, Martin Siderius, and Chris H. Harrison. Bayesian geoacoustic inversion using wind-driven ambient noise. *The Journal of the Acoustical Society of America*, 131(4):2658–2667, April 2012. ISSN 0001-4966. doi: 10.1121/1.3688482.

Jorge E. Quijano, Stan E. Dosso, Jan Dettmer, Lisa M. Zurk, and Martin Siderius. Trans-dimensional geoacoustic inversion of wind-driven ambient noise. *The Journal of the Acoustical Society of America*, 133(1):EL47–EL53, January 2013. ISSN 0001-4966. doi: 10.1121/1.4771975.

S. T. Rakotonarivo and W. A. Kuperman. Model-independent range localization of a moving source in shallow water. *The Journal of the Acoustical Society of America*, 132(4):2218+, 2012. ISSN 00014966. doi: 10.1121/1.4748795.

M. J. D. Rendas and J. M. F. Moura. Cramer-Rao bounds for passive range and depth in a vertically inhomogeneous medium. In *Acoustics, Speech, and Signal Processing, 1990. ICASSP-90., 1990 International Conference on*, volume 5, pages 2779–2782. IEEE, April 1990. doi: 10.1109/icassp.1990.116202.

M. J. D. Rendas and J. M. F. Moura. Cramer-Rao bounds for passive range and depth in a vertically inhomogeneous medium. *Signal Processing, IEEE Transactions on*, 39(12):2593–2610, December 1991. ISSN 1053-587X. doi: 10.1109/ICASSP.1990.116202.

Branko Ristic, Sanjeev Arulampalam, and Neil Gordon. A tutorial on particle filters. In *Beyond the Kalman filter : particle filters for tracking applications*, chapter 3, pages 35–66. Artech House, Boston, 2004. ISBN 158053631.

Donald Ross. *Mechanics of underwater noise*. Pergamon Press, 1976. ISBN 0080211828.

Daniel Rouseff and Lisa M. Zurk. Striation-based beamforming for estimating the waveguide invariant with passive sonar. *The Journal of the Acoustical Society of America Express Letters*, 130(2):EL76–EL81, July 2011. doi: 10.1121/1.3606571.

Philippe Roux, Bruce D. Cornuelle, W. A. Kuperman, and W. S. Hodgkiss. The structure of raylike arrivals in a shallow-water waveguide. *The Journal of the Acoustical Society of America*, 124(6):3430–3439, December 2008. ISSN 00014966. doi: 10.1121/1.2996330.

Jit Sarkar, Bruce D. Cornuelle, and W. A. Kuperman. Information and linearity of time-domain complex demodulated amplitude and phase data in shallow water. *The Journal of the Acoustical Society of America*, 130(3):1242+, 2011. ISSN 00014966. doi: 10.1121/1.3613709.

H. Schmidt. *OASES User Guide and Reference Manual*. Department of Ocean Engineering at Massachusetts Institute of Technology, Cambridge, MA http://acoustics.mit.edu/faculty/henrik/oases.html, 3.1 edition, October 2004.

Michael S. Schmidt and Thom Shanker. To smuggle more drugs, traffickers go under the sea. *The New York Times*, September 2012.

Ralph Schmidt. A new approach to geometry of range difference location. *Aerospace and Electronic Systems, IEEE Transactions on*, AES-8(6):821–835, November 1972. ISSN 0018-9251. doi: 10.1109/taes.1972.309614.

He Shaoyin and Liu Yudong. Matched multipath source localization in deep ocean. In *Signal Processing Proceedings, 1998. ICSP '98. 1998 Fourth International Conference on*, volume 2, pages 1431–1434 vol.2. IEEE, 1998. ISBN 0-7803-4325-5. doi: 10.1109/icosp.1998.770889.

Pasang Sherpa. Passive acoustic vessel localization. Master's thesis, Portland State University, 2012.

Martin Siderius. Using practical supergain for passive imaging with noise. *The Journal of the Acoustical Society of America Express Letters*, 131(1):EL14–EL20, December 2012. ISSN 00014966. doi: 10.1121/1.3663282.

Martin Siderius and Chris Harrison. High-Frequency geoacoustic inversion of ambient noise data using short arrays. In *High Frequency Ocean Acoustics Conference*, volume 728, pages 22–31. AIP, 2004. doi: 10.1063/1.1842993.

Martin Siderius, Chris H. Harrison, and Michael B. Porter. A passive fathometer technique for imaging seabed layering using ambient noise. *The Journal of the Acoustical Society of America*, 120(3):1315–1323, September 2006. doi: 10.1121/1.2227371.

Martin Siderius, Heechun Song, Peter Gerstoft, William S. Hodgkiss, Paul Hursky, and Chris Harrison. Adaptive passive fathometer processing. *The Journal of the Acoustical Society of America*, 127(4):2193–2200, April 2010. doi: 10.1121/1.3303985.

Martin Siderius, Lanfranco Muzi, Chris H. Harrison, and Peter L. Nielsen. Synthetic array processing of ocean ambient noise for higher resolution seabed bottom loss estimation. *The Journal of the Acoustical Society of America*, 133(3):EL149–EL155, March 2013. ISSN 0001-4966. doi: 10.1121/1.4774074.

E. K. Skarsoulis, A. Frantzis, and M. Kalogerakis. Passive localization of pulsed sound sources with a 2-Hydrophone array. In *Proceedings of the Seventh European Conference on Underwater Acoustics, ECUA 2004*, Delft, The Netherlands, July 2004.

Cristiano Soares and Sérgio M. Jesus. Broadband matched-field processing: Coherent and incoherent approaches. *The Journal of the Acoustical Society of America*, 113 (5):2587–2598, May 2003. doi: 10.1121/1.1564016.

H. Song, W. A. Kuperman, W. S. Hodgkiss, P. Gerstoft, and Jea S. Kim. Null broadening with snapshot-deficient covariance matrices in passive sonar. *IEEE Journal of Oceanic Engineering*, 28(2):250–261, April 2003. ISSN 0364-9059. doi: 10.1109/joe.2003.814055.

E. Sorensen, H. H. Ou, L. M. Zurk, and M. Siderius. Passive acoustic sensing for detection of small vessels. In *OCEANS 2010*, pages 1–8. IEEE, September 2010. ISBN 978-1-4244-4332-1. doi: 10.1109/oceans.2010.5664542.

John L. Spiesberger. Hyperbolic location errors due to insufficient numbers of receivers. *The Journal of the Acoustical Society of America*, 109(6):3076–3079, June 2001. doi: 10.1121/1.1373442.

John L. Spiesberger. Geometry of locating sounds from differences in travel time: Isodiachrons. *The Journal of the Acoustical Society of America*, 116(5):3168–3177, November 2004. doi: 10.1121/1.1804625.

John L. Spiesberger and Magnus Wahlberg. Probability density functions for hyperbolic and isodiachronic locations. *The Journal of the Acoustical Society of America*, 112(6):3046–3052, December 2002. doi: 10.1121/1.1513648.

Operational Research Staff. The range reliability and accuracy of a low frequency loran system. Technical Report ORS-P-23, Office of Chief Signal Officer, Pentagon, Washington, D.C., January 1946.

John Staples. Nuclear activist swims into faslane. *The Scotsman*, April 2001.

Kenneth Steele, Bobby Ulich, and Anthony Dietz. Lidar as a complementary sensor technology for harbor security. In *SPIE Defense & Security Symposium*, March 2005.

Rustam Stolkin, Alexander Sutin, Sreeram Radhakrishnan, Michael Bruno, Brian Fullerton, Alexander Ekimov, and Michael Raftery. Feature based passive acoustic detection of underwater threats. *Proceedings of SPIE*, 6204:620408–620408–10, May 2006. doi: 10.1117/12.663651.

S. A. Stotts, R. A. Koch, S. M. Joshi, V. T. Nguyen, V. W. Ferreri, and D. P. Knobles. Geoacoustic inversions of horizontal and vertical line array acoustic data from a surface ship source of opportunity. *Oceanic Engineering, IEEE Journal of*, 35(1):79–102, January 2010. ISSN 0364-9059. doi: 10.1109/joe.2009.2032256.

United S. Supreme Court. Winter, secretary of the navy, et al. v natural resources defense council, inc., et al. In *Supreme Court Case*, Washington, DC, October 2008.

A. Sutin, B. Bunin, A. Sedunov, N. Sedunov, L. Fillinger, M. Tsionskiy, and M. Bruno. Stevens passive acoustic system for underwater surveillance. In *International Waterside Security Conference (WSS), 2010*, pages 1–6. IEEE, November 2010. ISBN 978-1-4244-8894-0. doi: 10.1109/wssc.2010.5730286.

A. Sutin, H. Salloum, M. DeLorme, N. Sedunov, A. Sedunov, and M. Tsionskiy. Stevens passive acoustic system for surface and underwater threat detection. In *Technologies for Homeland Security (HST), 2013 IEEE International Conference on*, pages 195–200. IEEE, November 2013. ISBN 978-1-4799-3963-3. doi: 10.1109/ths.2013.6698999.

John D. Sutter. Blind man uses his ears to see, November 2011.

Aaron Thode. Tracking sperm whale (physeter macrocephalus) dive profiles using a towed passive acoustic array. *The Journal of the Acoustical Society of America*, 116(1):245–253, July 2004. doi: 10.1121/1.1758972.

Aaron M. Thode. Source ranging with minimal environmental information using a virtual receiver and waveguide invariant theory. *The Journal of the Acoustical Society of America*, 108(4):1582+, 2000. ISSN 00014966. doi: 10.1121/1.1289409.

Christopher O. Tiemann, Aaron M. Thode, Janice Straley, Victoria O'Connell, and Kendall Folkert. Three-dimensional localization of sperm whales using a single hydrophone. *The Journal of the Acoustical Society of America*, 120(4):2355–2365, October 2006. doi: 10.1121/1.2335577.

Alexandra Tolstoy. *Matched field processing for underwater acoustics*. World Scientific, 1993. ISBN 9810210590.

Caitlyn Toropova, Imèn Meliane, Dan Laffoley, Elizabeth Matthews, and Mark Spalding. Global ocean protection : present status and future possibilities, November 2010.

Jonathan P. Towle, David Herold, Ryan Johnson, and Harold Vincent. Low-cost acoustic sensors for littoral anti-submarine warfare (ASW). In Edward M. Carapezza, editor, *Proc. SPIE 6538*, pages 653814–653814–6, April 2007. doi: 10.1117/12.724242.

James Traer and Peter Gerstoft. Coherent averaging of the passive fathometer response using short correlation time. *The Journal of the Acoustical Society of America*, 130(6):3633–3641, December 2011. doi: 10.1121/1.3654026.

James Traer, Peter Gerstoft, H. C. Song, and William S. Hodgkiss. On the sign of the adaptive passive fathometer impulse response. *The Journal of the Acoustical Society of America*, 126(4):1657–1658, October 2009. ISSN 0001-4966. doi: 10.1121/1.3206696.

James Traer, Peter Gerstoft, and William S. Hodgkiss. Ocean bottom profiling with ambient noise: A model for the passive fathometer. *The Journal of the Acoustical Society of America*, 129(4):1825–1836, April 2011. doi: 10.1121/1.3552871.

University of Hawai'i at Manoa, Costal Geoglogy Group. *SHOALS LIDAR Bathymetry Database*, June 2012. www.soest.hawaii.edu (date last viewed 04/23/2014).

R. J. Urick. Generalized form of the sonar equations. *The Journal of the Acoustical Society of America*, 34(5):547–550, May 1962. doi: 10.1121/1.1918166.

Robert J. Urick. *Principals of Underwater Sound for Engineers*. McGraw-Hill, Inc., 1967. ISBN 07-066085-9.

Svein Vagle and Holly Burch. Acoustic measurements of the sound-speed profile in the bubbly wake formed by a small motor boat. *The Journal of the Acoustical Society of America*, 117(1):153–163, January 2005. doi: 10.1121/1.1819502.

Harry L. Van Trees. *Detection, Estimation, and Modulation Theory, Part I*. John Wiley and Sons, Inc., New York, first edition, September 2001a. ISBN 0471095176.

Harry L. Van Trees. Multiple parameter estimation. In *Detection, Estimation, and Modulation Theory, Part I*, chapter 2.4.3, pages 74–85. John Wiley and Sons, Inc., New York, 2001b. ISBN 0471095176.

Harry L. Van Trees. *Optimum Array Processing (Detection, Estimation, and Modulation Theory, Part IV)*. John Wiley & Sons, Inc., New York, 1 edition, March 2002a. ISBN 0471093904.

Harry L. Van Trees. Mismatched MVDR and MPDR beamformers. In *Optimum Array Processing (Detection, Estimation, and Modulation Theory, Part IV)*, chapter 6.6, pages 488–513. John Wiley & Sons, Inc., New York, first edition, March 2002b. ISBN 0471093904.

Harry L. Van Trees. Sample matrix inversion (SMI). In *Optimum Array Processing (Detection, Estimation, and Modulation Theory, Part IV)*, chapter 7.3, pages 728–752. John Wiley & Sons, Inc., New York, first edition, March 2002c. ISBN 0471093904.

Harry L. Van Trees. Correlated and coherent signals. In *Optimum Array Processing (Detection, Estimation, and Modulation Theory, Part IV)*, chapter 9.6, pages 1233–1243. John Wiley & Sons, Inc., New York, first edition, March 2002d. ISBN 0471093904.

Tyler H. Wai, John S. Allen, John Gebbie, and Martin Siderius. Nonlinear time series analysis of snapping shrimp sounds. In *165th Meeting of the Acoustical Society of America*, volume 134, pages 4147+, San Francisco, CA, December 2013a. doi: 10.1121/1.4831199.

Tyler H. Wai, John S. Allen, John Gebbie, and Martin Siderius. Spectral changes in snapping shrimp sounds between dusk and dawn. In *OCEANS 2013*, pages 1–6, San Diego, CA, 2013b. IEEE.

Tyler H. Wai, John S. Allen, John Gebbie, and Martin Siderius. Spectral changes in snapping shrimp sounds between dusk and dawn (expected). *The Journal of the Acoustical Society of America Express Letters*, 2014.

D. L. Waldron and L. Mullen. Underwater optical ranging: A hybrid LIDAR-RADAR approach. In *OCEANS 2009, MTS/IEEE Biloxi - Marine Technology for Our Future: Global and Local Challenges*, pages 1–7. IEEE, October 2009. ISBN 978-1-4244-4960-6.

Shane C. Walker. A model for the spatial coherence of arbitrarily directive noise in the depth-stratified ocean. *The Journal of the Acoustical Society of America*, 131 (5):EL388–EL394, May 2012. ISSN 0001-4966. doi: 10.1121/1.3701707.

Darren B. Ward, E. A. Lehmann, and R. C. Williamson. Particle filtering algorithms for tracking an acoustic source in a reverberant environment. *Speech and Audio Processing, IEEE Transactions on*, 11(6):826–836, November 2003. ISSN 1063-6676. doi: 10.1109/tsa.2003.818112.

William A. Watkins and William E. Schevill. Sound source location by arrival-times on a non-rigid three-dimensional hydrophone array. *Deep Sea Research and Oceanographic Abstracts*, 19(10):691–706, October 1972. ISSN 00117471. doi: 10.1016/0011-7471(72)90061-7.

William A. Watkins, Peter Tyack, Karen E. Moore, and James E. Bird. The 20Hz signals of finback whales (balaenoptera physalus). *The Journal of the Acoustical Society of America*, 82(6):1901–1912, December 1987. ISSN 0001-4966. doi: 10.1121/1.395685.

Michelle J. Weirathmueller, William S. D. Wilcock, and Dax C. Soule. Source levels of fin whale 20hz pulses measured in the northeast pacific ocean. *The Journal of the Acoustical Society of America*, 133(2):741–749, January 2013. ISSN 0001-4966. doi: 10.1121/1.4773277.

Lauren Wenzel, Terry McTigue, and Mimi D'Iorio. *Marine Reserves in the United States*. Office of Ocean and Coastal Resource Management, NOAA Ocean Service, 1305 East West Hwy (N/ORM), Silver Spring, MD 20910, U.S.A., July 2011.

D. E. Weston and K. J. Stevens. Interference of wide-band sound in shallow water. *Journal of Sound and Vibration*, 21(1):57–64, March 1972. ISSN 0022460X. doi: 10.1016/0022-460x(72)90205-2.

Edward C. Whitman. SOSUS: The "secret weapon" of undersea surveillance. *Undersea Warfare*, 7(1), October 2005.

P. Whittle. The analysis of multiple stationary time series. *Journal of the Royal Statistical Society. Series B (Methodological)*, 15(1), 1953.

J. P. Wilkinson and P. Wadhams. Modeling the flow of oil under sea ice: a role for the AUV. In *Autonomous Underwater Vehicles, 2008. AUV 2008. IEEE/OES*, pages 1–5. IEEE, October 2008. ISBN 978-1-4244-2939-4. doi: 10.1109/auv.2008.5347601.

F. G. Wood and W. E. Evans. Adaptiveness and ecology of echolocation in toothed whales. In René-Guy Busnel and James F. Fish, editors, *Animal Sonar Systems*, volume 28 of *NATO Advanced Study Institutes Series*, pages 381–425. Springer US, 1980. doi: 10.1007/978-1-4684-7254-7_16.

T. C. Yang and T. Yates. Matched-beam processing: Application to a horizontal line array in shallow water. *The Journal of the Acoustical Society of America*, 104(3): 1316–1330, September 1998. doi: 10.1121/1.424341.

T. C. Yang, Kwang Yoo, and T. Yates. Matched-beam processing: Range tracking with vertical arrays in mismatched environments. *The Journal of the Acoustical Society of America*, 104(4):2174–2188, October 1998. doi: 10.1121/1.423730.

Y. X. Yuan, C. Carter, and J. E. Salt. Near-optimal range and depth estimation using a vertical array in a correlated multipath environment. *IEEE Transactions on Signal Processing*, 48(2):317–330, February 2000. ISSN 1053-587X. doi: 10.1109/78.823960.

Yun X. Yuan and J. E. Salt. Range and depth estimation using a vertical array in a correlated multipath environment. *IEEE Journal of Oceanic Engineering*, 18(4):500–507, October 1993. ISSN 0364-9059. doi: 10.1109/48.262300.

Marvin Zelen and Norman C. Severo. Probability functions. In Milton Abramowitz and Irene A. Stegun, editors, *Handbook of Mathematical Functions With Formulas, Graphs, and Mathematical Tables*, 55, pages 925–964. U.S. Government Printing Office, Washington, D.C., 10 edition, December 1972.

R. Zimmerman, G. L. D'Spain, and C. D. Chadwell. Decreasing the radiated acoustic and vibration noise of a mid-size AUV. *IEEE Journal of Oceanic Engineering*, 30(1):179–187, January 2005. ISSN 0364-9059. doi: 10.1109/joe.2004.836996.

Lisa Zurk, Joe Lotz, Tim Ellis, James McNames, and JeanLouis Ecochard. Sonar mapping for coral reef conservation. In *151st Meeting of the Acoustical Society of America*, volume 120, page 3060, November 2006. doi: 10.1121/1.4787314.

Lisa M. Zurk and Daniel Rouseff. Striation-based beamforming for active sonar with a horizontal line array. *The Journal of the Acoustical Society of America Express Letters*, 132(4):EL264–EL270, September 2012. doi: 10.1121/1.4748281.

Lisa M. Zurk, N. Lee, and J. Ward. 3D adaptive matched field processing for a moving source in a shallow water channel. In *OCEANS '99 MTS/IEEE. Riding the Crest into the 21st Century*, volume 2, pages 728–731 vol.2. IEEE, 1999. ISBN 0-7803-5628-4. doi: 10.1109/oceans.1999.804877.

Lisa M. Zurk, Nigel Lee, and James Ward. Source motion mitigation for adaptive matched field processing. *The Journal of the Acoustical Society of America*, 113(5):2719–2731, May 2003. doi: 10.1121/1.1561817.