UNIVERSITY OF CALIFORNIA SAN DIEGO

Self-localization of a mobile swarm of underwater vehicles using ambient acoustic noise

A dissertation submitted in partial satisfaction of the
requirements for the degree
Doctor of Philosophy

in

Electrical Engineering (Intelligent Systems, Robotics and Control)

by

Perry W. Naughton

Committee in charge:

Ryan Kastner, Chair
Peter Gerstoft
John Hildebrand
William Hodgkiss
Jules Jaffe
Falko Kuester

2018

The dissertation of Perry W. Naughton is approved, and it is acceptable in quality and form for publication on microfilm and electronically:

_____

_____

_____

_____

_____

_____

Chair

University of California San Diego

2018

DEDICATION

To:

Pop and Oma for their focus on education

&

Grandpa and Grandma Naughton for their emphasis on faith

EPIGRAPH

Simple it's not, I'm afraid you will find,

for a mind maker-upper to make up his mind

-Theodor Seuss Geisel

TABLE OF CONTENTS

LIST OF FIGURES

ACKNOWLEDGEMENTS

At the culmination of a long, wandering, degree there are a lot of people to thank! And I will undoubtedly miss some...

My family has been a great support to me throughout my entire life, and I feel very fortunate to be part of a loving family that gets a long so well. My parents, Tim and Karina, always pushed me to be a better person, try hard, and *always* have fun. They consistently provide me with many opportunities to learn and grow. My sister, Ida, has always given me a different perspective on life through literature, music, and ants. My brothers, Chris and Phillip, keep me excited about hands on tinkering and getting outside to stay active.

My first academic thanks goes to my advisor, Ryan Kastner. As an undergraduate, Ryan, along with Albert Lin, rekindled my excitement for engineering research through the Engineers for Exploration program. This program showed me that engineering was an integral part of advancing scientific research. These experiences demonstrated the value of presentation and collaboration, which led to some of my most fulfilling pursuits of my graduate career (most of which did not make it into this dissertation, but will hopefully make it into some of the work of others). Ryan (almost) always left me with less headaches than I came to him with - even though his favorite answer is “what do you think you should do?”. I am grateful for his guidance.

Along the lines of collaborative research, I would also like to thank Falko Kuester and the Center for Interdisciplinary Science of Art, Architecture and Archeology for similar lessons and continued support in collaborative research. Finally, Curt Schurgers was also involved in E4E and this dissertation. He helped me to grow through his many conversations, edits and by challenging my thinking.

I owe Philippe Roux a huge thanks. Without Philippe this dissertation would be very different and may not have materialized at all. Philippe provided me with tremendous support, scientifically through discussions and lessons, professionally through edits and fellowships, and personally by hosting me in his lab at ISTerre, in Grenoble, France. His positive attitude was

always encouraging and made my research more enjoyable. Being hosted in France was an awesome experience and I am grateful for the many friends I met while at ISTerre.

This dissertation would also not exist without Jules Jaffe, who provided the infrastructure for every experiment presented in this dissertation, along with guidance and insight. Jules's emphasis on collaborative research immediately drew me to his lab and I am grateful that I was greeted with open doors. Paul Roberts and Devin Ratelle provided support with the data collection as well as advice for the experimental setup of Chapter 4 and calibrating the AUEs that were used throughout this dissertation.

I would like to that the rest of my committee: Peter Gerstoft, John Hildebrand, and William Hodgkiss for laying the foundation for much of this work, for the lessons, and for the feedback.

I would also like to thank the people who kept me sane. Early on, I was lucky to have many distractions provided by Brian Shughrou, Will Eddy and Katy Eddy (Beaver). I have had a strong group of climbing friends who kept me outside: Jason Oberg, Sonia Rackelmann, Wilson Lian, Dimitri Schreiber, and Jessica Carriere-Garwood. The members of the Kastner lab made research feel less like an isolated journey, and I am particularly grateful for Dustin Richmond who I got to share many gripes, growing pains, recipes, a garden, and a kegerator with.

I thank Joanne Shin for our cherished Internet relationship, as well as the companionship over the last two years. She has always been around to catch me when I fall, and I am excited for our next chapter.

I am grateful for all of the financial support that I received during my tenure. This work was supported by "INSPIRE Track I: Distributed Sensing Collective to Capture 3D Soundscapes" supported by the NSF under Grant No. 1344291. This material is also based upon work supported by NSF Grant No. OCE 09- 27449, the National Science Foundation Graduate Research Fellowship under Grant No. DGE-1144086, the National Science Foundation under IGERT Award DGE-0966375, "Training, Re- search and Education in Engineering for Cultural Heritage

Diagnostics, the NSF Graduate Research Opportunities Worldwide program, the Chateaubriand STEM Fellowship, the San Diego chapter of the ARCS Foundation, and the Friends of the International Center Scholarship at UCSD.

Chapter 3, in part, is a reprint of the material as it appears in The Journal of the Acoustical Society of America: P. Naughton, P. Roux, C. Schurgers, R. Kastner, J. S. Jaffe, and P. L. D. Roberts, (2016) "Ambient noise correlations on a mobile, deformable array". The dissertation author was the primary investigator and author of this paper.

Chapter 4, in part, has been submitted for publication of the material as it may appear in The Journal of the Acoustical Society of America: P. Naughton, P. Roux, C. Schurgers, R. Kastner, J. S. Jaffe, and P. L. D. Roberts, (2016) "Self-localization of a mobile swarm using noise correlations with local sources of opportunity". The dissertation author was the primary investigator and author of this paper.

Chapter 5, in part, is a reprint of the material as it appears in IEEE Access: P. Naughton, P. Roux, C. Schurgers, R. Kastner, J. S. Jaffe, and P. L. D. Roberts, (2018) "Self-localization of a deforming swarm of underwater vehicles using impulsive sound sources of opportunity". The dissertation author was the primary investigator and author of this paper.

Section 5.5, in part, has been submitted for publication of the material as it may appear in the fourth Underwater Communications and Networking Conference: P. Naughton, T. Salam, P. Roux, C. Schurgers, R. Kastner, J. S. Jaffe, and P. L. D. Roberts, (2018) "Self-synchronization of multiple vehicles using ambient impulsive noise". The dissertation author was the primary investigator and author of this paper.

VITA

2012	B. S. in Electrical and Computer Engineering, *Mangna Cum Laude*, University of California San Diego
2014	M. S. in Electrical and Computer Engineering, University of California San Diego
2012-2013	Graduate Teaching Assistant, University of California San Diego
2012-2018	Graduate Research Assistant, University of California San Diego
2016-2017	Visitng Research Assistant, Institut des Science de la Terre, Université Grenoble Alpes, France
2018	Ph. D. in Electrical and Computer Engineering, University of California San Diego

PUBLICATIONS

P. Naughton, P. Roux, C. Schurgers, R. Kastner, J. S. Jaffe, and P. L. D. Roberts, "Self-localization of a mobile swarm using noise correlations with local sources of opportunity," *The Journal of the Acoustical Society of America*, Submitted, 2018.

P. Naughton,T. Salam, P. Roux, C. Schurgers, R. Kastner, J. S. Jaffe, and P. L. D. Roberts,"Self-synchronization of multiple vehicles using ambient impulsive noise," *Underwater Communications and Networking*, Submitted, 2018.

P. Naughton, P. Roux, C. Schurgers, R. Kastner, J. S. Jaffe, and P. L. D. Roberts,"Self-localization of a deforming swarm of underwater vehicles using impulsive sound sources of opportunity," *IEEE Access*, vol. 6, pp. 1635-1646, 2018.

P. Naughton, P. Roux, R. Yeakle, C. Schurgers, R. Kastner, J. S. Jaffe, and P. L. D. Roberts, "Ambient noise correlations on a mobile, deformable array," *The Journal of the Acoustical Society of America*, vol. 140, no. 6, pp. 4260–4270, 2016.

R. Yeakle, P. Naughton, C. Schurgers, and R. Kastner, "Inter-node distance estimation from ambient acoustic noise in mobile underwater sensor arrays," in *Proceedings of OCEANS 2016 MTS/IEEE Monterey*, Monterey, CA, USA,2016

P. Naughton, C. Edwards, V. Petrovic, R. Kastner, F. Kuester, and S. Sandin "Scaling the Annotation of Subtidal Marine Habitats," in *Proceedings of the 10th International Conference on Underwater Networks & Systems*, Arlington, VA, USA, 2015

D. Mirza, P. Naughton, C. Schurgers, and R. Kastner, "Real-time collaborative tracking for underwater networked systems," *Ad Hoc Networks*, vol. 34, no. 3, pp. 196–210, 2015.

P. Bojakowski, K. C. Bojakowski, and P. Naughton, "A comparison between structure from motion and direct survey methodologies on the warwick," *Journal of Maritime Archaeology*, vol. 10, no. 2, pp. 159–180, 2015.

T. G. Garrison, D. Richmond, P. Naughton, E. Lo, S. Trinh, Z. Barnes, A. Lin, C. Schurgers, R. Kastner, and S. E. Newman, "Tunnel Vision: Documenting Excavations in Three Dimensions with Lidar Technology," *Advances in Archaeological Practice*, vol. 4, no. 2, pp. 192–204, 2016.

A. Kargar, K. Sun, Y. Jing, C. Choi, H. Jeong, Y. Zhou, K. Madsen, P. Naughton, S. Jin, G. Jung, and D. Wang, "Tailoring n-ZnO/p-Si branched nanowire heterostructures for selective photo-electrochemical water oxidation or reduction," *Nano Letters*,vol. 13, no. 7, pp. 3017-3022–204, 2013.

K. Sun, A. Kargar, N. Park, K. Madsen, P. Naughton, T. Bright, Y. Jing, D. Wang "Compound semiconductor nanowire solar cells," *IEEE Journal of Selected Topics in Quantum Electronics*, vol. 17, no. 4, pp. 1033–1049, 2011.

ABSTRACT OF THE DISSERTATION

Self-localization of a mobile swarm of underwater vehicles using ambient acoustic noise

by

Perry W. Naughton

Doctor of Philosophy in Electrical Engineering (Intelligent Systems, Robotics and Control)

University of California San Diego, 2018

Ryan Kastner, Chair

There is increasing interest in deploying swarms of underwater vehicles for marine surveys. In such surveys, the cost, endurance and utility of the vehicle swarm needs to be carefully optimized. One of the main challenges when designing these systems is coming up with an appropriate way to localize each vehicle in relation to the rest of the swarm. Current methods for underwater localization are restrictive in either cost, power consumption, or range.

This dissertation considers the self-localization of a deforming swarm of subsurface vehicles using ambient acoustic noise in the ocean. The experiments presented consider a group of independent underwater vehicles that passively record ambient sounds in the ocean with a single hydrophone while they float with subsurface currents. Three different self-localization

approaches are considered. The first involves estimating the acoustic impulse response between moving vehicles using cross-correlations of ambient noise, a known first step towards a self-localization estimate. Accurate estimates of the acoustic impulse response are shown between moving vehicle pairs. However, motion between the receivers limits the amount of time averaging that can be done, making the estimation susceptible to anisotropies in the ambient noise field. To overcome these anisotropies, the next approach jointly estimates the vehicle geometry and the directionality of the ambient noise field, without prior knowledge of either estimate. This creates a viable method for estimating the vehicle geometry on short time scales using correlations of low frequency noise in the ocean. Results are shown for a deployment carried out off of the coast of La Jolla. Last, a self-localization approach using impulsive noise from snapping shrimp is considered. Impulsive sound sources provide high intensity, broadband signals that facilitate accurate arrival time detections by each vehicle. However, the similarity between different signals presents a significant correspondence problem, which must be solved to provide accurate estimates of the changing geometry of the swarm. A geometric solution to this correspondence problem is shown and an estimation procedure is proposed to track the geometry of a swarm as it changes. The self-localization estimates are compared to estimates from an accurate acoustic localization system with good agreement.

Chapter 1

Introduction

1.1 Motivation

Our understanding of the ocean at different temporal and spatial scales is determined by the resolution of our technology. Traditional approaches to ocean sampling can be expensive and are rarely dense in both space *and* time. For example, infrastructure such as piers [1] or moorings [2] collect data at near continuous rates but are only able to sample from one location, making them temporally dense but not spatially dense. Surveys using ships [3, 4] or autonomous underwater vehicles [5, 6, 7] are able to cover large(r) swaths of area, providing spatial density. But, their costs limit the frequency of these surveys, making them sparse in time. The last few decades have seen many advances in oceanographic instrumentation and have increased the endurance, mobility, and spatial resolution of oceanographic surveys. However, some subsurface phenomena are still hard to measure with current technology and the cost of current surveys prevent many studies that are technologically possible.

Small scale subsea phenomena (< 1 km) that evolve in both space and time are particularly challenging to measure [8]. When studying small scale ocean phenomena, the data collected must have high temporal and spatial resolution compared to the processes being studied. For example,

studying subsea phenomena at scales below 1km requires a spatial sampling much finer than 1km. A swarm of coordinated vehicles or sensors can simultaneously sample the ocean at different locations, making them ideal tools to study these small scale processes that evolve in both space and time. However, designing an affordable and capable collection of coordinated vehicles is challenging.

One of the largest challenges in developing a coordinated sampling system is providing an accurate spatial reference between vehicles without dramatically decreasing the endurance or increasing the cost of deployments. Determining appropriate localization methodologies are key to understanding phenomena at these scales by creating affordable and effective coordinated surveys. Current localization protocols may not be sufficient for some studies.

This dissertation proposes methods to estimate the geometry of a swarm of underwater vehicles using ambient acoustic noise in the ocean. This passive approach can mitigate some of the drawbacks of current localization methods by decreasing the cost and power consumption of each vehicle in a coordinated sampling system. These advantages would improve the capability of many ocean surveys that rely on estimating the relative location of vehicles, especially those focused on studying small scale phenomena, by allowing high endurance, infrastructure free deployments. Examples of such surveys include: measuring submesoscale ocean dynamics to understand the redistribution of oil, plastics and marine organisms [8, 9], optically imaging the seafloor [10], and using the vehicles as an *ad hoc* acoustic array [11, 12, 13].

1.2 An overview on ocean sampling, and the move towards coordinated vehicles

A variety of vehicles have been built to study undersea environments at different scales. At the largest scale, subsurface floats have the endurance to provide an array of sensors that spatially sample the entire ocean. The largest effort in persistent, multi-vehicle ocean monitoring

at this scale is the Argo program [14, 15], which contains around 3,800 (in mid-2018) floats in the world's oceans to measure properties of the water (temperature and salinity) and large scale subsurface currents. Argo floats control their buoyancy and float with subsea currents, a principle first described by John Swallow [16] and later used by the predecessors to the current Argo floats: Autonomous Lagrangian Circulation Explorers (ALACE) [17, 18] and the Profiling ALACE floats (PALACE) [19]. Argo floats surface every 10 days to relay data and estimate their position using GPS. The horizontal displacement that the Argo floats experience on their way to the surface affects the accuracy (and therefore spatial resolution) of the resulting subsurface current measurements, and their temporal resolution is limited by infrequent surface intervals.

SOund Fixing And Ranging (SOFAR) floats were deployed in the SOFAR channel, a waveguide in the ocean where low frequency signals can propagate long distances. These floats were tracked by emitting a low frequency signal that was recorded by fixed infrastructure [20]. Having sound sources on the SOFAR floats is expensive, so RAFOS (SOFAR spelled backwards) floats were designed to listen for acoustic signals from fixed infrastructure instead of emitting signals [21]. These floats were tracked approximately once per day, with an accuracy of 0.5 km [22].

The infrequent updates and large positioning errors of the Argo, SOFAR, and RAFOS floats are adequate to study phenomena at large spatial scales, but cannot resolve smaller phenomena.

Floats, gliders and autonomous underwater vehicles have all been deployed to study the ocean at smaller scales. EM-APEX floats were deployed in a 2km area to measure subsurface currents by surfacing every hour to record their GPS positions [23], providing more accurate and frequent spatial estimates than Argo floats. Gliders, which control their buoyancy and use wings to move horizontally as they descend in depth, are another class of persistent vehicles that have been used to collect data at spatial scales on the order of kilometers. Examples of these gliders are the Slocum glider [24], the Spray [25], and the Seaglider [26]. Gliders typically move

around 35 cm/s, can be deployed for longer than a month, and surface to measure their GPS position, telemeter data, and receive instructions for their next profile. Coordinated gliders have been deployed to estimate both scalar (e.g. temperature) and vector fields (e.g. currents) [27, 28]. Like the floats, localizing gliders is typically limited to GPS measurements on the surface, which limits the spatial resolution of the data they collect.

Autonomous underwater vehicles (AUVs) are the most mobile scientific instruments to study subsea environments [29] and are commonly deployed to collect data on the order of kilometers. Generally torpedo-shaped, AUVs can efficiently move through the water at speeds around 1.5 m/s. There have been many different AUVs with different payloads, prices and endurances: Autonomous Benthic Explorer (ABE) [30], REMUS [31], OceanServer Iver3, MBARIs Dorado [32], and Starbug AUV [33]. To date, the AUVs deployed in the ocean mostly focus on reaching environments that would otherwise be impossible (deep sea and under ice) and have resulted in vehicles that are expensive, and can operate on the order of a day [34, 35]. AUVs typically localize themselves using fixed acoustic infrastructure, and onboard sensors [36].

The sophisticated localization systems of many AUVs are perfect for collecting data at high spatial resolutions, but they are expensive to deploy and maintain. While gliders and AUVs provide mobile instruments that are able to accurately record dense spatial data, their costs prevent more than a few vehicles from being deployed simultaneously, if they are able to be deployed at all. The consequences of this are two fold. First, only small numbers of these vehicles can be deployed simultaneously because of the cost of each vehicle. Small groups of these vehicles struggle to measure subsea phenomena that evolve in space and time (i.e. 4 dimensions). Studying these 4D phenomena requires simultaneous sampling from multiple positions [27, 37, 8]. In these coordinated surveys, the strength of the survey is proportional to the number of vehicles deployed. Second, many domains that could benefit from autonomous vehicle surveys are priced out of them [38].

Future trends

To provide increased spatial and temporal resolution, another class of AUVs are being developed that are designed for coordinated sampling. These vehicles are small and relatively inexpensive in order to maximize the number of vehicles that can be deployed. Mini - Autonomous Underwater Explorers (AUEs) were developed to provide a spatial swarm of of continuously tracked vehicles to measure internal waves, an example of a small scale 4D ocean phenomena that is impossible to sufficiently sample with other vehicles [8]. To achieve this, The AUE positions are estimated almost continuously from 5 surface pingers with known locations. While the AUEs were sufficiently tracked in this study, their localization performance relies on being within the range of acoustic infrastructure, which limits the areas that the vehicles can survey. Additionally, setting up the acoustic infrastructure requires extensive effort before a survey can start. The AUEs demonstrate the utility of a swarm of continuously tracked, mobile underwater vehicles and provides much of the motivation for this dissertation.

Keeping the vehicles inexpensive requires trade-offs in functionality, and many of the other examples of these vehicles developments are not field ready because of the difficulties of robust and affordable localization strategies. For example, the CoCoRo project [39] relies on optics to communicate and localize different vehicles, which is challenged in many underwater environments. Other example vehicles include Serafina [40], the HippoCampus [41], the MONSUN II [42], and the Vertex AUV [43]. Determining an affordable and robust spatial reference between vehicles is one of the main challenges holding swarms back from being adopted as an oceanographic instrument.

Newer vehicles are becoming smaller, more affordable, and have more endurance [8, 37]. These advantages offer new possibilities for sampling the ocean. By lowering the cost, vehicle platforms can be integrated into more surveys and collect data for domains that have not yet benefited from these technological advances. More affordable, high endurance vehicles increase the temporal and spatial resolution through coordinated sampling with multiple vehicles.

This increased resolution can measure gradients in the ocean environment and other physical phenomena that was previously difficult to measure using single vehicle surveys or surveys with infrequent localization estimates. However, there are technical challenges preventing persistent spatial sampling using underwater swarms, one of the main ones being affordable methods to robustly localize each underwater vehicle, which is described in the next section.

1.3 Underwater localization and its challenges

Tracking vehicles underwater is an active area of research [44, 45, 46, 47] and the implementation of a specific underwater localization protocol dictates important trade-offs in a marine survey including: the cost of the survey, the power consumption, duration, and the area covered. Almost all subsea localization strategies involve acoustics, the difference between localization strategies is where the acoustic signal is generated and how it is processed. The most accurate systems use georeferenced acoustic infrastructure to base their position from. This acoustic infrastructure sends an acoustic signal that is received by underwater vehicles to give a distance estimate between the vehicle and infrastructure. In this case, each acoustic measurement is an independent reference to a global coordinate frame [8, 48, 49, 50, 51, 52, 53, 54, 55]. These include long baseline systems [56, 48, 54, 55, 57], short baseline systems [58] and ultrashort baseline systems [59, 60, 49, 50]. While effective, these systems require significant effort to set up and limit the range of a deployment to be within acoustic range of the deployed infrastructure. Other systems leverage the return of a signal they generate (e.g. Doppler shifts [61, 62]) to measure the movement of each node in their environment. Successive measurements can be integrated into a trajectory estimate. These systems are less restricted in their movement and require less infrastructure, but they force the signal to be generated by each node, usually with an expensive sensor, making this method power intensive and expensive. Additionally, the integration over successive time samples means this method suffers from drift, usually requiring vehicles

to surface so that this drift can be corrected with a GPS measurement. In multi-node systems, positioning information can be exchanged between nodes through acoustic communication [63, 64, 65, 52] to increase the localization accuracy of each individual node. These systems can also suffer from large power consumption, poor localization performance, or both.

While active acoustic systems provide the most accurate loclization methods, their drawbacks prevent them from being used on other systems (such as the smaller, low cost vehicles that were described). The combination of being either power intensive or requiring vehicles to be within range of fixed infrastructure make active acoustic localization methods infeasible for many high endurance deployments. The costs associated with active acoustic systems can prevent them from being used on systems that are focused on low cost vehicles. Unfortunately, this means that some processes in the ocean, especially ones that happen on small scales, are left out of our understanding.

1.4 Dissertation Contributions

This dissertation proposes methods to estimate a spatial reference between independent underwater vehicles in a swarm using ambient acoustic noise that is already present in underwater environments. By using passive signals, many of the trade-offs of using active sources are mitigated: there could be less set up time, longer deployments, and no spatial restrictions on the vehicles. Additionally, this method only requires one hydrophone per vehicle, making this localization approach inexpensive compared to other solutions.

These improvements come at the cost of performing the localization offline and only being able to recover the swarm geometry in a relative coordinate frame because the vehicles cannot communicate and the locations of the noise sources are unknown, respectively. This self-localization solution may not be applicable to all classes of vehicles, especially those that require a real-time, global estimate. However, these advantages could improve the capability

of many ocean surveys that rely on estimating the relative location of vehicles, especially those focused on studying small scale phenomena, by allowing high endurance, infrastructure free deployments. Examples of such surveys include: measuring submesoscale ocean dynamics to understand the redistribution of oil, plastics and marine organisms [8, 9], optically imaging the seafloor [10], and using the vehicles as an *ad hoc* acoustic array [11, 12, 13]. Additionally, the relative coordinate frame could be augmented with more information, such as GPS positions on the surface, or a reference to a more expensive vehicle with an accurate navigation system, to provide a global estimate. While promising, localizing mobile vehicles using ambient signals presents many challenges. These challenges stem from the fact that the signals used cannot be designed, methods must use what is already present in the ambient soundscape. These challenges are explored throughout this dissertation.

1.5 Dissertation Outline

This dissertation examines different cases to the self-localization problem that are representative of what might be observed in different ocean environments. The details of the different cases are described per chapter:

Chapter 2 describes related work. Previous work on ambient noise processing and localization is introduced, which frames the contributions of this dissertation and introduces some of the applications that have been demonstrated using ambient acoustic noise. The applications demonstrate the potential of completely passive underwater sensing systems, and give examples of the applications that could be implemented by passive swarms of underwater vehicles given more development on processing with irregular and changing vehicle geometries.

Chapter 3 describes results of ambient noise correlations on a mobile, deformable array. The possibility of recovering the Time Domain Green's Function (also known as the acoustic impulse response) from short time correlations of ambient acoustic noise is examined as well as

the difficulty of choosing the time-window for the noise correlations. Recovering an estimate of the acoustic impulse response is a known method for a self-localization solution [66]. This chapter shows that there are salient correlation peaks that result from ambient noise correlations, and that for many pairs of floats, these peaks match what is expected from the acoustic impulse response. However, for other pairs, the arrival structure of the noise correlation is shaded by the directionality of the noise field.

Chapter 4 demonstrates a self-localization solution when there are dominating noise directions in the environment. This chapter proposes to jointly estimate the directionality of the noise field in addition to the geometry of the deformable swarm. Simulations are provided to demonstrate how different parameters effect the results including: inter-vehicle movement, azimuthal variance of the noise field, and the vehicle geometry. The proposed method is tested using data that was collected in San Diego when a boat was circling the drifting vehicles, and the self-localization solution was compared to an accurate active acoustic localization system.

Chapter 5 considers the self-localization of a deforming swarm of subsurface floating vehicles using impulsive sources of opportunity, such as the sounds of snapping shrimp that are present in warm coastal waters. Impulsive sound sources provide high intensity, broadband signals that facilitate accurate arrival time detections across each vehicle. This makes them useful references for a self-localization solution. However, the similarity between different signals presents a significant correspondence problem, which must be solved to provide accurate estimates of the changing geometry of the swarm. A geometric solution to this correspondence problem is shown and an optimization procedure is proposed to track the geometry of a swarm as it changes. The method is verified using a swarm of 17 self-ballasting subsurface floats that independently drifted with currents off of the coast of San Diego, California. The changing geometry of the floats was estimated using both an acoustic localization system and the proposed approach. The two estimates show good agreement, validating the method. Chapter 5 also shows how the proposed correspondence solution can be extended to estimate the clock offsets between

receivers. Clock-synchronization between vehicles is an important first step before any processing across vehicles can be performed, such as a self-localization solution. The results of the clock offset estimation are compared to linear drift estimates from GPS measurements at the surface. The two clock offset estimates show good agreement, validating our method.

Finally, Chapter 6 concludes this dissertation. The contributions are summarized and promising directions for future work are discussed.

Chapter 2

Background

This dissertation provides a self-localization estimate for mobile vehicles using only the acoustic noise already present in the ocean. This chapter describes some of the general characteristics of the ambient noise field and introduces past work on ambient noise processing and self-localization. Almost all of the prior work on ambient noise processing was performed between receivers or arrays with constant separation. This cannot be assumed for mobile vehicles. The challenges of processing ambient noise using mobile receivers will be echoed throughout the chapters in this dissertation.

Proven applications of ambient noise processing are also introduced in this chapter. These applications may be implemented on a swarm of independent mobile vehicles given more development, highlighting the potential for passive acoustic swarms. The applications also introduce methods that can be leveraged to extract useful information from the ambient soundscape.

2.1 Processing ambient acoustic noise in the ocean

The ambient acoustic noise in the ocean is the accumulation of many different sources including: anthropogenic activity (shipping traffic, airgun surveys, sonar systems), biologic noises

(snapping shrimp, vocalizations from cetaceans and fish) and environmental noises (earthquakes, rain, sea state). The noise levels have been rising for the last few decades in many frequency bands and geographic locations due to increasing anthropogenic activity in the oceans (e.g. shipping traffic, sonar systems and seismic exploration), and this trend is expected to continue [67, 68, 69]. Some of the typical noise levels and their origins are illustrated in Figure 2.1. The ambient noise levels are significant, especially in the lower frequencies. This is a problem for most systems that are built to study, navigate, or communicate in the ocean because they must compete with this large noise floor [70, 71, 72, 73]. Not surprisingly, there has been a significant amount of research to use this noise to study the ocean, instead of trying to constantly compete with it. Compared to traditional techniques that inject an active source into the environment, the possible benefits of using ambient acoustic noise are: it is less disruptive, requires less power, and it does not have to compete with the large acoustic noise floor (because the noise is the signal). Despite the clear advantage of using ambient acoustic noise, there are many challenges in extracting useful information from noise. These stem from the fact that noise is not a controlled source. The techniques, requirements, successes and challenges of ambient noise processing are discussed here. While not all of these are directly related to underwater localization, they either demonstrate techniques that could be helpful in extracting useful constraints for a localization solution, or they demonstrate applications that a passive acoustic swarm of vehicles may achieve.

2.1.1 Extracting coherent wavefronts from ambient acoustic noise

Much of the development in ambient noise processing stems from the result that an estimate of the impulse response (also known as the Time Domain Green's Function (TDGF)) can be recovered between two points by examining the noise received at those two points (see Figure 2.2 for a short explanation). The first experiments that demonstrated this result were performed by Lobkis and Weaver in ultrasonics [76, 77]. Their experiments showed that the impulse response between two receivers could be estimated by cross-correlating the ambient thermal fluctuations

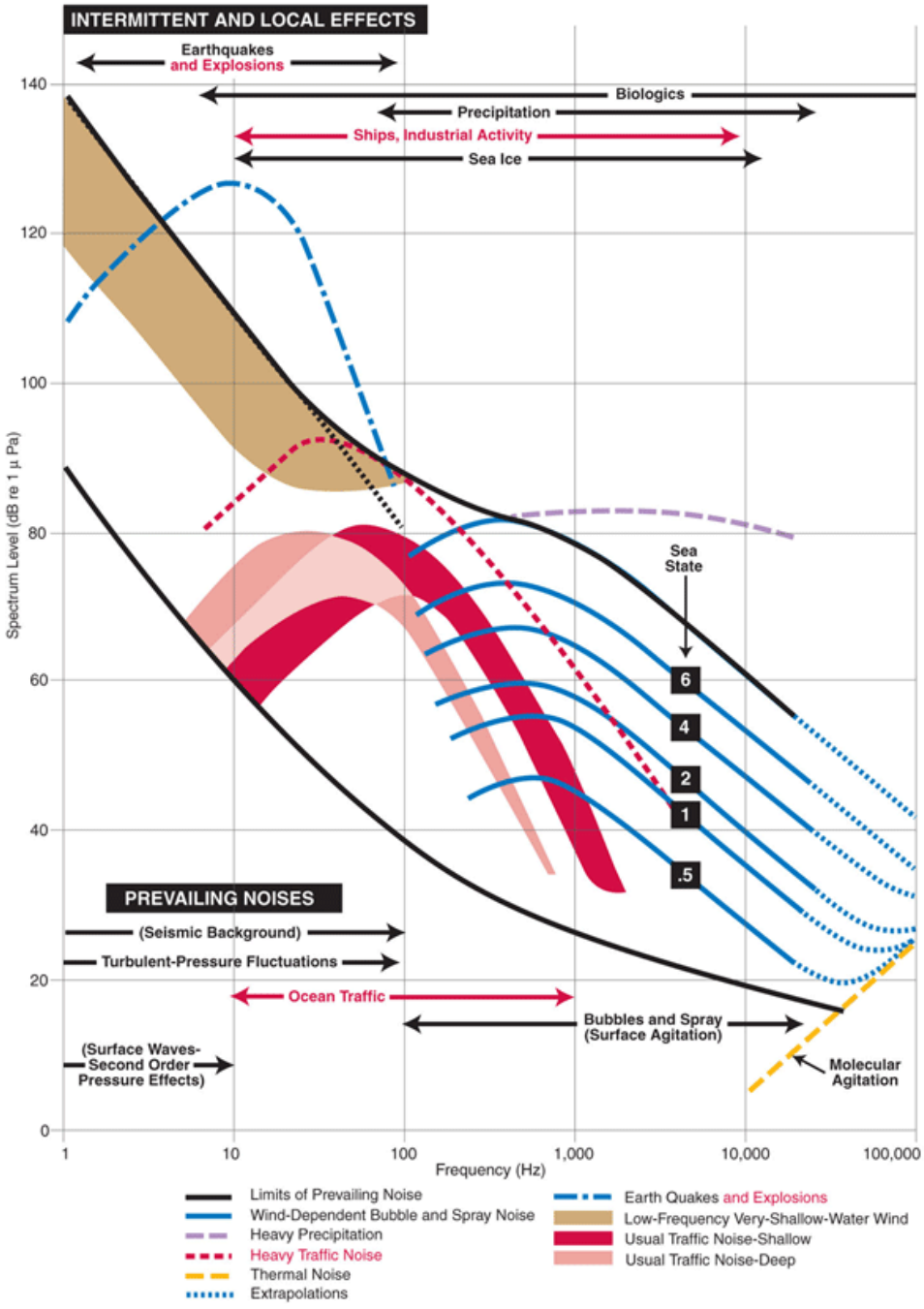

Figure 2.1: Typical ambient noise spectra in the ocean (first published by Wenz [74] but this version is from National Research Council, 2003 [75]).

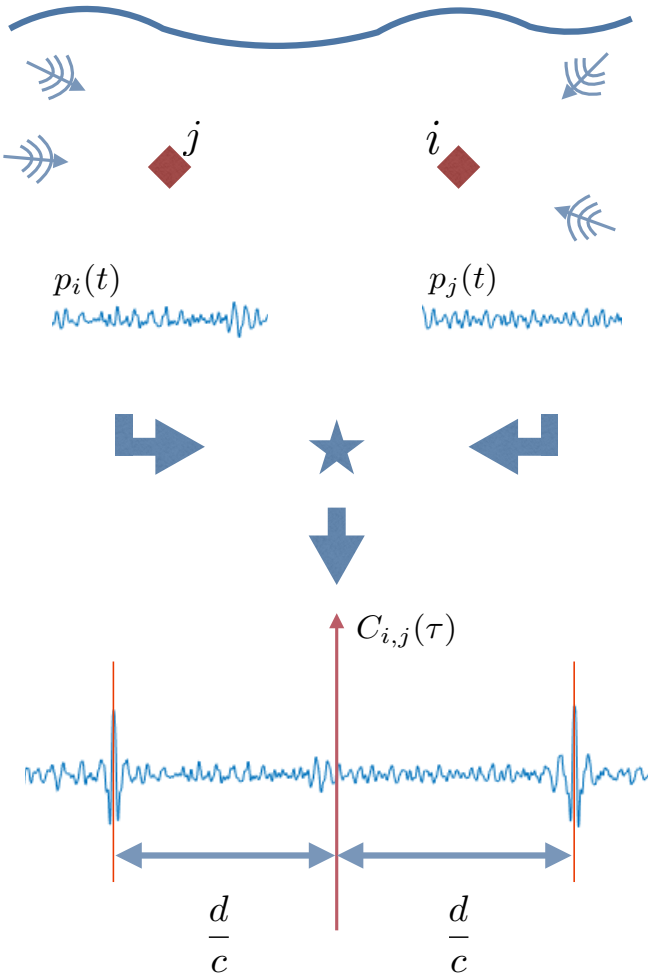

Figure 2.2: Noise correlation result: By correlating the noise received at two points (shown by red diamonds), the noise correlation function produces and double sided estimate of the acoustic impulse response (Time Domain Green's Function). One side of the noise correlation function (e.g. in positive correlation time) describes the propagation path from receiver i to receiver j and the other side of the noise correlation (e.g. in negative correlation time) describes the propagation path from receiver j to receiver i. In the figure, d represents the distance between the receivers and c represents the propagation speed of the acoustic signal. The peaks in the noise correlation function will be centered around the clock bias between the receivers. If the clocks are synchronized, then the peaks will be centered around 0 correlation time.

recorded by two transducers on an aluminum block. In other words, the impulse response could be estimated between two transducers by using the noise recorded at each transducer. Before this realization, an estimate of an impulse response from one point to another required an active signal to be sent from the starting point and measured at the receive point.

There are two main benefits to using noise correlations instead of an active source. The first is associated with the drawbacks of using an active source: it can be disruptive, power intensive, expensive and requires additional infrastructure. The second benefit is an increase in the number of ray paths that can be measured. Given n receivers, it is possible to estimate the impulse response between each receiver pair using noise (i.e. n choose 2 possible combinations). When using an active source, only the paths from the active source to each receiver can be estimated (n paths). This increase in ray paths has been especially beneficial in seismology [78].

The implications of this experiment have been profound. The drawbacks of using an active source are so large in some cases they cannot be used, and ambient noise has provided opportunity to sense different environments that would otherwise not be able to be understood. This result has been studied theoretically [79, 80, 81, 82, 83, 84] and applications have found their way into many different disciplines including seismic studies [78, 85, 86, 87, 88], structural health monitoring [89, 90, 91], medical imaging, [92], and ocean acoustics [93, 94, 66, 95, 96, 97, 98, 99, 100, 101, 102, 103]. The focus of this review will be on the works related to ocean acoustics.

2.1.2 Estimating the acoustic impulse response from ambient noise in the ocean

In the last decade, coherent processing of ambient acoustic noise developed into a new branch of ocean acoustic processing [93]. The first demonstration of extracting coherent information using ambient noise was performed by Roux et al. [94]. Since then, the spatial and temporal variation of the acoustic noise in the ocean encouraged experiments spanning a diverse set of

acoustic environments. Most of the literature on TDGF estimation in the ocean relies on frequency bands that are dominated by shipping noise (10Hz<f<1kHz) [94, 99, 100, 101, 102, 103]. These bands suffer the least from attenuation and contain the most power, which makes them ideal for estimating the TDGF over long distances. Biological sounds were considered (from a species of croaker fish) [66, 104] using higher frequencies (350-700Hz) and closer distances (1-50m). Low frequencies (<100Hz) were considered in the deep ocean [105, 106, 107]. In general, tens of minutes of noise has been used to recover the *arrival structure* (the time values) of the TDGF over distances from hundreds of meters to a few kilometers. Currently, only the arrival structure of the TDGF can be retrieved [94, 97]; the amplitude of the TDGF is often shaded by anisotropies in the ambient noise field as well as the spectra of the different noise sources in the ocean.

There are two competing forces in the emergence of the TDGF. There is an inherent variance in the Noise Correlation Function (NCF) resulting from the random distribution of noise sources contributing to the NCF. This variance is inversely proportional to the time-bandwidth product of the noise records under the assumption of an isotropic noise distribution [80]. Given a stationary environment, a longer time window results in a higher signal to noise ratio by decreasing the variance of the NCF. Similarly, longer time windows result in more coherent contributions that build the TDGF [100]. Unfortunately, the acoustic environment in the ocean is not stationary. Changes in the propagation of sound between two receivers degrades the emergence of the TDGF [104]. These temporal changes in the acoustic environment are the limiting factor in the successful recovery of the TDGF. To combat this, beamforming to enhance the emergence of horizontal coherent arrivals was proposed by Leroy et al. [102, 101] and shown to enhance the signal to noise ratio of the noise correlation function using frequency bands dominated by horizontally traveling shipping noise [100]. This beamforming procedure enabled short range tomography by reinforcing the most stable horizontal ray between the receivers for short distances [100, 101]. There has also been work to estimate the TDGF in a rapidly fluctuated medium by formulating the fluctuations as a constrained optimization algorithm [108]. The main challenge considered in this

dissertation is that the vehicle movement changes the propagation path between the receivers at a quicker rate than the changes in the environment, and the *ad hoc* geometry of an array of vehicles does not follow any optimal receiver spacing for a beamforming solution. These challenges are summarized in Section 2.3 and examined in the remaining chapters of this dissertation.

2.1.3 Applications of coherent ambient noise processing

Many applications have followed the discovery that the TDGF could be estimated by correlating ambient acoustic noise. Most of these applications were born out of techniques using active acoustics, and enjoy the advantage that no additional signals need to be introduced into the ocean. While not all of these applications are relevant to localization, they provide potential applications that a moving *ad hoc* passive acoustic swarm could work towards once a relative geometry of the vehicles is estimated.

Passive fathometer

A passive fathometer was implemented to measure the depth of the ocean and seabed layering using vertically traveling, surface generated, ambient noise (50<f<5kHz). This technique works by matching the downward traveling surface noise with its upward traveling reflections from different layers of the seafloor that provide strong reflections [96, 97, 98]. The matching between the two paths yields a two way travel time estimate that can be converted to a depth measurement given a known propagation velocity. In the fathometer implementation, delay sum beamforming was used to leverage multiple elements in a vertical array to enhance the emergence of the TDGF in the vertical direction, and reject horizontally traveling noise [96]. More theoretical developments were made [97] and an adaptive beamforming procedure was proposed to enhance the result of the original formulation [98].

Passive thermometry

Passive acoustic thermometry was proposed to measure the changing temperatures in the deep ocean by estimating changes in the speed of sound at depth. Temperature is one of the main variables that drives changes in the speed of sound underwater [109, 110]. By measuring how the propagation time of an acoustic signal between two fixed points changes, the change in the ocean's temperature between those points can be estimated. The change in propagation time has been studied by retrieving the TDGF through the cross correlation of low frequency noise across large distances in the deep ocean [106, 105]. This application may prove to be instrumental in understanding the rising temperatures of our ocean, especially in the deep ocean, which is both difficult to survey and important. The deep ocean is an important factor in forecasting the climate and how it may change because most of the thermal energy of the planet is stored there.

Array element self-synchronization and self-localization using ambient acoustic noise

The most relevant application of passively recovering the TDGF to this dissertation is for array self-localization and array self-synchronization [66]. Using estimates of the direct acoustic path from the TDGF, both the clock bias and distance estimates between different receivers can be extracted. For example, Figure 2.2 shows that the distance between the two receivers listening to noise can be estimated using the time separation of the two peaks in the noise correlation function (assuming knowledge of the propagation speed, c). The middle point between these two peaks provides an estimate of the clock bias between the receivers [66, 87]. Using distance estimates between all pairs of receivers, the relative geometry of an array can be formulated as a nonlinear optimization problem and estimated [66]. This gives an estimate of the array shape. This dissertation describes the challenges of this approach for mobile arrays in Chapter 3.

2.2 Sources of Opportunity

There is another branch of ambient noise processing that deals with "sources of opportunity". While there may not be a clear distinction between sources of opportunity and the noise considered in the last section on recovering the TDGF ("sources of opportunity" such as ships have been proposed to estimate the TDGF [95]), there typically two differences that change the way this data is processed. First, sources of opportunity are typically salient sources in the ambient noise field, while this is not always true in applications recovering the TDGF. Second, there is sometimes *a priori* information about these sound sources (e.g. the location of the sound source) that changes the way the processing can be executed. Some of the scenarios that have been considered in this scenario are highlighted in this section.

2.2.1 Acoustic Daylight

A picture like image of an object can be formed by measuring the reflections of incoherent noise from that object. This acoustic illumination is analogous to how our eyes use the reflection of sunlight to form an image of an object and led to the name "Acoustic Daylight". Buckingham et. al. [111] where the first to propose this concept when they were able to image targets placed in the ocean using an 'acoustic lens' - a parabolic reflector that focused incoming ambient noise that was reflected off of an object. The fundamental challenge with acoustic daylight is the large aperture needed to achieve usable angular resolution due to the wavelengths of acoustic noise that can be considered (about 3cm @ 50kHz compared to $5x10^{-5}$cm for green light) [111]. Therefore, the work in acoustic daylight imaging has used special purpose hardware like the Acoustic Daylight Ocean Noise Imaging System (ADONIS) [112] and the second generation Remotely Operated Mobile Ambient Noise Imaging System (ROMANIS) [113] to provide a usable, albeit still low, level of angular resolution.

2.2.2 Source Localization

Verlinden et al. [114] leveraged the fact that certain classes of vessel are required to broadcast data, such as their location and heading, for collision avoidance via the Automatic Identification Systems (AIS). The AIS system records the location of these vessels which provide "vessels of opportunity" with known locations. These known trajectories were used in order to populate a library of data derived replica vectors by cross-correlating the noise recorded by VLAs from a ship with a known position recorded in the AIS system. Using this replica library, new ships were localized by comparing their correlation results across two elements on different arrays to the library of correlation vectors from previously measured ships with known positions from the AIS system. Library vectors that were similar to the observed correlation vector indicated a greater likelihood that the ship in question was at the same location as the ship that produced the library replica. A similar approach, also using the AIS system, was used to estimate the waveguide invariant [115], and these two results were combined so that the library tracks could be extrapolated and localization solutions could be found for locations where the library ships had not been. Similar work was proposed that extracted the normalized cross-covariance of the elements of a VLA and used GPS data from a ship to train classifiers (feed forward neural network, support vector machine, random forest) for range classification of new ships passing the same environment [116, 117].

2.2.3 Localization methods using sources of opportunity

Sources of opportunity were considered to estimate the shape of an array. Given a known environment (the sound speed profile, bottom parameters, and bathymetry) and a with known signal, the tilt of an array can be estimated by optimizing the output of a matched field processor [118, 119]. Localization using acoustic transponders, matched field processors and independent tilt measurements have also been compared [119]. Relative arrival times from a moving research

vessel were also used to estimate the positions of each element in an array as well as an updated estimate of the position of the boat by iteratively linearising the acoustic ray equations [120]. The authors report that the standard deviation of their relative uncertainties was 0.2 m.

Work was proposed that considered localizing sonobuys using impulsive noise sources (e.g. lightbulbs [121]) that do not occur simultaneously and happen in from uncertain positions [122]. The authors propose an under determined inversion problem where they estimate the source positions, the initial sonobuoy positions and the sonobuoy depths. The under determined inversion is solved using regularization based on *a priori* estimates of the locations of the sonobouys, the locations of the noise events and a motion model for the sonobuoys. Their inversion algorithm then looks to optimize the travel time measurements with these uncertain estimates. Their algorithm was experimentally demonstrated in [123]. Their algorithm estimates the location of the sonobouys in global coordinates, but requires knowing (and therefore deploying) acoustic sources with approximate positions. This estimate will also suffer from drift because the solution relies on velocity assumptions for the buoys to smooth the trajectory in the time domain.

Auto calibration of *ad hoc* terrestrial microphone arrays

In terrestrial wireless sensor networks, studies focused on ad hoc self-localization of microphone arrays using sources of opportunity. Many scenarios, constraints and sensor packages have been considered and a review can be found in Plinge et al. [124]. The studies that are most similar to this work consider only one receiver per node, assume no knowledge of the ambient sound positions or their emittance times and assume the receivers are time synchronized. Some of these are similar to the work of Sabra et al. [66] in that they either assume a diffuse noise field [125] or rely on sources being in the endfire beam [126] so that they can directly estimate the distance between receivers. Under a far field assumption, the work of Thrun [127] showed that the relative geometry of an ad hoc microphone array could be robustly estimated in an affine space and later upgraded to a euclidean space when the noise sources are parameterized by an angle.

This method only requires a nonlinear optimization on a space that is P^2 variables, where P is the dimension that the receivers occupy. Therefore, this method is less prone to local minima than solutions that try to solve a least squares problem over all sound and receiver positions. The work of Thrun [127] was extended to a three-dimensional case [128]. Typically, these works assume that the correspondence problem is solved, meaning there are unambiguous correspondences of events between audio tracks of the receivers. In a real life ocean environment, this assumption is impractical for most frequency bands. The correspondence problem is exasperated the more nodes one considers at the same time. An approach to solving the correspondence problem is to look at minimal problems, where the correspondence ambiguity is the smallest. In Wendeberg et al. [129], the authors consider a two-dimensional localization problem and describe a receiver triangle. Under a far field assumption, they show that the TDOA measurements between the three receivers with one receiver as anchor form an ellipse. They use this information to extract distance and angle measurements from the receiver triangle and estimate the geometry of an array of receivers. For the three-dimensional case, minimal cases have been examined both in the far field, [130] and without a far field assumption [131]. These approaches look at a minimum of 4 and 5 receivers respectively.

2.3 Contributions of this dissertation

To date, most of the research on self-localization and ambient noise processing is limited to arrays where the distance between each element is constant. This cannot be assumed for independent mobile vehicles. In fact, the goal of this dissertation is to measure vehicle motion in relation to other vehicles. This dissertation does not assume to know information about the propagation environment or information about the noise sources, so approaches that optimize the output of a matched filter or assume *a priori* estimates of the locations of sound sources are not applicable. Instead, self-localization solutions are considered that try to track the changing TDGF

as well as solutions that utilize more salient "sources of opportunity".

The literature focused on recovering the TDGF was performed on arrays where the element separation is stationary, such as moored vertical [94, 99, 100, 102] or horizontal line arrays [66, 105, 99, 104, 96, 97, 98]. For stationary arrays, recovering the TDGF is limited by fluctuations in the propagation environment (e.g. the sound speed), which usually occur on the order of tens of minutes [66] to hours [104]. Chapter 3 provides some of the first results of noise correlations where each vehicle is an independent, untethered system. In this case, the distance between each receiver pair is changing, which changes the propagation path between vehicles. The change in propagation path due to receiver motion happens on shorter times scales than changes due to the propagation medium (now on the order of 10s of seconds). The result of this motion is that recovering TDGF is challenged and the self-localization method proposed by Sabra et. al [66] may not be robust. An approach to overcome this limitation is presented in Chapter 4 where the directionality of the noise field is jointly estimated along with the receiver positions.

There are also salient biologic signals in the ocean that could be useful in estimating the receiver positions. Since the self-localization problem must be solved using short time scales, signals that can be considered must happen frequently in time. Often the signals in this category are similar in structure and are difficult to discern by the characteristics of an individual signal alone (e.g. using correlation based techniques for matching). When many signals occur almost simultaneously from different locations there is ambiguity in which signal on one receiver corresponds to the same signal on another receiver. This ambiguity must be solved before a self-localization solution is formulated. This problem is studied in Chapter 5 and a self-localization solution is demonstrated using impulsive noise from snapping shrimp.

Chapter 3

Ambient noise correlations on a mobile, deformable array

3.1 Introduction

Chapter 2 discussed many of the applications that were built around recovering the Time Domain Green's Funtion (TDGF), also known as the acoustic impulse response, between two receivers listening to ambient acoustic noise in the ocean. Previous work considered receivers in the ocean that were stationary. In the stationary case, estimating the TDGF was limited by environmental fluctuations. This chapter considers ambient noise correlations on independent subsea receivers that float freely with the currents. The motion between the recivers is the main change in the propagation path between them, dominating the environmental fluctuations. This motion and resulting deformation of the array severely restricts the length of the correlation window that can be used and makes recovering the TDGF difficult. Despite these challenges, we demonstrate that we can recover the TDGF between two receivers for some instances. Consistently recovering the TDGF between all pairs of receivers would allow the self-localization solution proposed by Sabra et al. [66] as well as provide the foundation of the other applications that were

discussed in Section 2.1.3.

In addition to recovering the TDGF, some of the potential uses of an mobile array of vehicles are explored. Using thousands of measurements from short time cross-correlations, we provide an analysis of how the amplitude of the recovered TDGF compares to the theoretical amplitude. Additionally, we show that we can use this deformable array to detect dominant source directions in the ambient noise field and describe how these detected directions are consistent with our analysis of being able to recover the arrival information of the TDGF. These results suggest the possibility of source localization using a mobile, deformable array.

3.2 Methods

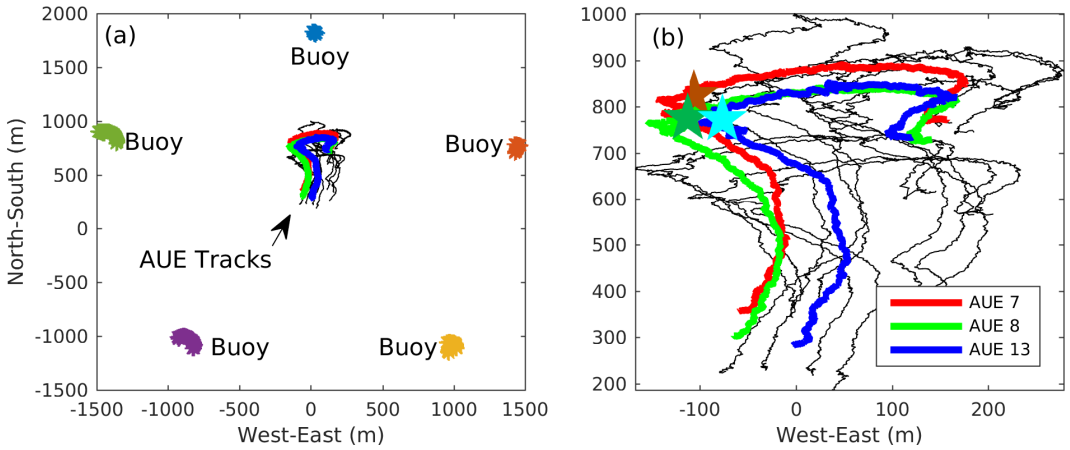

Figure 3.1: Experiment Setup: (a) 5 buoys are positioned on the surface of the ocean along the perimeter of a deployment. These buoys collect GPS data and take turns transmitting a linearly modulated chirp. The Autonomous Underwater Explorers (AUEs), positioned in the center of the buoy array, measure the time of arrival of these chirps and we compute estimates of their positions based on these signals. (b) A close up of the trajectories of the AUEs with 3 individual AUEs highlighted. We notice that the trajectories are rough because they are computed at discrete times with some noise. It is also difficult to deduce the time element of these trajectories. Where it looks like AUEs cross paths it is usually the case that an AUE is crossing the path of an older position of another AUE. To demonstrate this we plot the positions of the 3 highlighted AUEs with a star during one time in the deployment. From this we can see that AUE 7 is north of AUE 8 at this time even though it may look like AUE 7 and 8 are crossing paths if we were to only look at their trajectories.

3.2.1 Autonomous Underwater Explorers

The experiment leverages Autonomous Underwater Explorers (AUEs), designed and built at the Scripps Institute of Oceanography to increase underwater sampling resolution in both space and time. Each AUE is a buoyancy controlled unit that can track a depth profile in the ocean's water column by adjusting its buoyancy. The AUEs have no actuation in the horizontal direction, are fully at the mercy of the currents and can move substantially while deployed in the ocean. AUEs collect acoustic data from an HTI-96-MIN hydrophone, accelerometer data, temperature and pressure data.

The depth of the AUE is determined using the on board pressure sensor. To localize each AUE in the other two directions (latitude and longitude), an acoustic triangulation system is set up to act similar to a GPS system. Five buoys are positioned on the surface of the ocean to send a linear modulated chirp in the frequency range 8-15kHz to be received by the hydrophone of each AUE. Each of five buoys takes turns pinging and each packet of five pings occurs around every ten seconds. The AUEs have limited on-board processing and no communication infrastructure, so all localization is done offline after they are retrieved. Since the AUEs can move in the period that the acoustic signals are being received, we can either assume that they are stationary during the buoy signal acquisition or we can estimate their motion during the signal reception by employing a factor graph framework[52, 51]. We have performed both methods for this dataset and they yield similar trajectories. These estimated trajectories are used throughout the analysis given in this paper. To maintain their depth, each AUE uses a motor to drive a piston that changes its buoyancy. When the motor is on, the hydrophones are saturated and we have no usable acoustic signal for that time. We observe the motors to be on for at least 10% of the duration of the experiment. The clocks of the AUE are synchronized at the beginning and end of the experiment using a GPS receiver on each AUE, and a linear clock drift model is applied to each AUE clock before any processing on the acoustic data is performed.

3.2.2 Experiment

The experiment makes use of data that was collected off of the coast of San Diego in October of 2013. The purpose of this experiment was to validate the trajectory estimation of the AUEs as well as measure currents and internal waves in the ocean. For the experiment, 13 AUEs tracked a depth of 10m and drifted with the currents for 5 hours. A wirewalker[132] was deployed near the array which collected temperature and pressure data. From the wirewalker data, it was determined that the AUEs were sitting on a steep thermocline and the speed of sound was determined to be 1519 m/s at 10m depth (more discussion of the acoustic environment during the deployment is found in Figure 3.6 and in the propagation environment section). Figure 3.1 shows the trajectories of the AUEs and the position of each of the 5 buoys in relation to the trajectories. The trajectories of AUE 7, 8 and 13 are highlighted; we will be using these trajectories in our analysis. It is important to note that these trajectories have *no notion of time* meaning that it is difficult to differentiate when the AUEs crossed paths in the experiment from the times where the AUE crossed paths with an old position of another AUE.

3.3 Noise Correlation Function

3.3.1 Theory

We are interested in the feasibility of extracting the Time Domain Green's Function (TDGF) between a pair of *mobile* receivers through a cross-correlation of ambient ocean acoustic noise. Practically, the normalized noise correlation function (NCF) between signals $p_i(t)$ and $p_j(t)$ collected at receiver i and j, respectively, is computed as:

$$C_{i,j}(\tau) = \int_{t_c-\frac{T_r}{2}}^{t_c+\frac{T_r}{2}} p_i(t)p_j(t+\tau)dt \Bigg/ \sqrt{\int_{t_c-\frac{T_r}{2}}^{t_c+\frac{T_r}{2}} p_i(t)^2dt}\sqrt{\int_{t_c-\frac{T_r}{2}}^{t_c+\frac{T_r}{2}} p_j(t)^2dt} \tag{3.1}$$

Where t_c controls the center time of the signal and T_r the length of the correlation window.

With respect to ocean acoustics, it has been shown both experimentally [94] and theoretically [79] that the NCF yields an estimate of the arrival structure of the TDGF (the amplitude of the TDGF is generally not recovered). The intuition behind this estimate is that the wavefield components of the ambient noise field that travel through both receivers average coherently while the components of the noise field that only travel through one receiver average incoherently. Sufficient time averaging, either by controlling the time window parameter (T_r) or by averaging several time-gated windows, is needed to insure that there are components of the noise field that pass through both receivers. Otherwise the peaks of the TDGF will not emerge. In previous works, the length of the time window is generally on the order of 10-30 minutes for ocean noise [80, 66, 94, 99, 102]. The length of the time window is largely determined by the noise distribution in the array's environment and the frequency band chosen, which determines how isotropic the noise field is. For a specific acoustic environment with a horizontal array (element spacing between 1m and 120m) in shallow water, the optimal time average was determined to be 3 hours and 50 minutes before changes in the acoustic environment started to degrade the estimate of the TDGF[104].

The emergence rate of the TDGF has been studied in past work and it was determined that the variance of the normalized NCF is $(2B_\omega T_r)^{-1}$ for an isotropic distribution of noise sources[66]. This is the noise floor that we are trying to overcome to see an emergence of the direct path in the TDGF. The amplitude of the direct path of the TDGF between two receivers depends on their separation as well as the acoustic environment. Simplified models of the acoustic environment indicate that the amplitude of the TDGF is related to the receiver separation, R, either by $1/R$ (through cylindrical spreading) or $1/R^2$ (through spherical spreading). This gives us a rough rule of thumb for the relationship between the length of time, T_r, needed to see an emergence based on the receiver separation, R. Given twice the separation between receivers one must double or quadruple the time window in order to keep the signal to noise ratio constant, depending on the environment. Unfortunately, short time windows also make the NCF vulnerable to nearby

dominant sources because there is not enough time to average these sources out (see Figure 3 in Roux et al. 2004[94]). Dominant sources can bias both the amplitude and the arrival time of the peak in the NCF.

The key challenge presented in this work is that the receivers are moving while they are deployed and the relative distances between the receivers are constantly changing. This prevents us from using long time averages of the cross-correlations that were used in previous literature to estimate the TDGF. This challenge has been mentioned before, where estimates of the TDGF were less reliable on top of a vertical array than on the bottom of the array because the anchor at the bottom of the array allowed less movement than at the top [102]. In our scenario, this problem is magnified by the fact that the receivers were designed to be Lagrangian and float freely with the currents. In fact, the AUEs were designed to study the small scale current fluctuations or the relative motion between the AUEs as opposed to the group motion (i.e. the average motion of the AUEs). In our formulation, we assume that the receivers are stationary for small segments of time so that we can use the theory developed in previous work. Even though there are local changes in the distances between the receivers, the movement between receivers is strongly correlated (driven by the currents). This helps keep the array together and we can assume that the receivers will be in the vicinity of each other during the entire deployment.

3.3.2 NCF Results

To compute the NCF, several preprocessing steps are executed before the correlation takes place. The first pre-processing step is to 'whiten' the signal's frequency spectrum, meaning equal power is assigned to each frequency band [104, 99, 80, 102]. This helps spread the contribution to the NCF across each frequency band instead of having the NCF be dominated by the frequency bands with the most power. Next, time series values with high amplitudes are truncated to the fourth standard deviation of the $p_i(t)$ and $p_j(t)$ signals. Last, each signal is normalized by the energy of the signal so that the peak of the auto-correlation of each signal would have amplitude 1.

The purpose of these preprocessing steps is to insure that peaks in the NCF come from coherent contributions of many sources instead of contributions from a few high energy sources which would likely bias the TDGF estimate, and to normalize each time series so that comparisons at different times can be made. More insight into these preprocessing steps can be found in References [104, 99, 80, 102].

Figures 3.2a and 3.2b show the NCF of receiver 8 and 13 and receiver 7 and 8, respectively, for the entire deployment time. These two pairs were chosen as examples to represent all possible permutations of AUE pairs in the array. For our experiments, we take the value of $T_r = 10s$ (more discussion on the choice of T_r and B_ω can be found in Section 3.3 C) in Equation 3.1 and t_c is taken to be integer multiples of 15 seconds. There is no overlap of data used between subsequent computations of the NCF. In each of these Figures, the y-axis represents the value of t_c, a measure of the deployment time (in hours), while the x-axis denotes the cross-correlation time (in seconds). For each correlation we used the frequency band $B_\omega = [100, 500]$Hz whitened according to the preprocessing procedure described above.

If we were to see the emergence of the TDGF we would expect a symmetry in the NCF around zero correlation time (it will be centered around zero only if the receivers are synchronized [66]). The peaks of the NCF would represent the causal and anti-causal solutions to the TDGF, meaning the peaks of the NCF would demonstrate the acoustic paths between the two receivers, including the direct and possibly reflected paths. In positive correlation time we would see the paths traveling from one receiver to the other, and in negative correlation time the paths would be traveling the opposite direction.

Figures 3.2c and 3.2d show the intensity of the NCF with the estimated travel time (from the computed positions of each receiver) plotted over the intensity to demonstrate the similarity of the NCF with the direct path between the two receivers. These distances were derived using the information shown in Figure 1. If the NCF was a representation of the TDGF, the correlation peaks would be related to the direct path shown. We see a difference in behavior between the

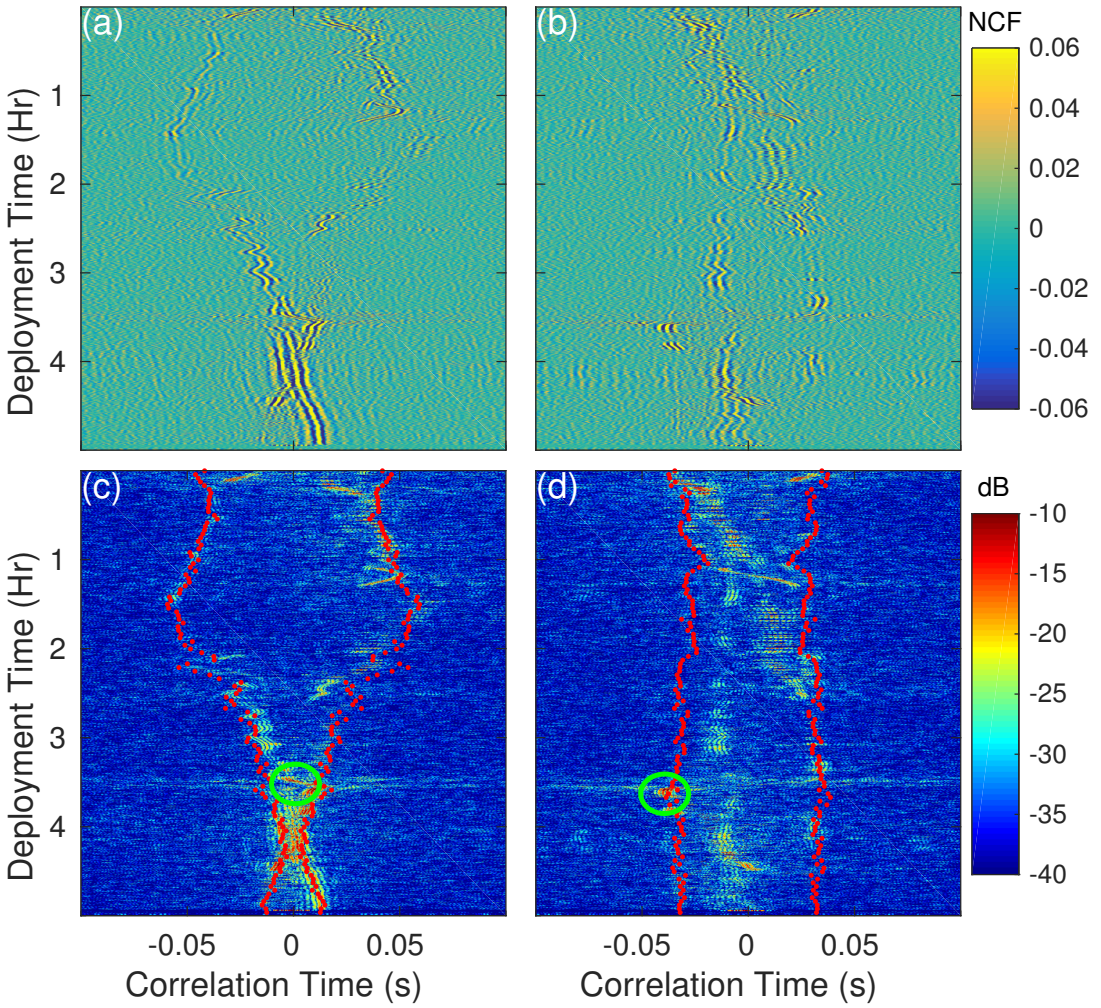

Figure 3.2: Noise Correlations: Left - AUE 8 and 13. Right - AUE 7 and 8. (a-b) Short Time Cross-Correlations of ambient acoustic noise between AUE pairs (x-axis) at different times during the AUE deployment (y-axis). The color represents the amplitude of the noise correlation function (NCF), defined in Equation 3.1. (c-d) The intensity (in dB) of the NCF with the estimated distances of the AUEs (generated from Figure 3.1) plotted over the intensity. From (a-d) we can see that for some pairs the peaks of the NCF align with the direct path between the pair of AUEs, for others the peaks do not show agreement with the distance between the pairs. We suggest that there is a directionality in the ambient noise field, and those pairs with their end-fire beams aligned with the dominant direction show the symmetry we would expect from the TDGF while those pairs whose endfire beam do not align do not show the symmetry. The green circles demonstrate a time when there is a change in the dominant noise direction of the ambient noise field, and we can see that the peaks match the distance for AUE 7 and 8 this time and not AUE 8 and 13 (This is described more in Figure 3.9). Parameters for correlations: time window length-10s, bandwidth-[100-500Hz]

two different pairs of receivers. The first pair, receiver 8 and 13, shows the symmetric correlation function that we would expect from the emergence of the TDGF. The symmetry follows the change in the direct path between the two receivers; as the receivers move closer together the peaks in the NCF move closer together. Not only are the trends similar but the time value of the peak frequently matches the travel time between the receivers.

In the frame of previous work, the results presented in Figure 3.2a and 3.2c are unique. We see a symmetric structure of the NCF that frequently matches the direct acoustic path. This is using a short correlation window, only ten seconds, compared to the tens of minutes or hours of previous work. This result shows promise that the first arrival time of the TDGF can be recovered using short time windows on receivers that are moving relative to each other.

The second pair (receiver 7 and 8) does not show the symmetric structure that we would expect from the TDGF and the peaks in the NCF do not match the direct path that was derived from the distance between the receivers. From a simple analysis of which pairs show the symmetric structure and which do not (for the pairs shown as well as pairs not shown) we attribute the difference in results to directionality in the noise field. Referring back to Figure 3.1, the relative geometry of the pairs of receivers shown by the stars are relatively consistent during the deployment. Pairs whose endfire beam (the endfire beam is defined by the ray that passes through both receivers) is mostly East-West demonstrates the symmetry described for the receiver pair 8 and 13, while receivers whose endfire beam aligns with North-South do not demonstrate the symmetry described. We can also see some directionality in the ambient noise field by examining pair 8 and 13. In this correlation, there is an imbalance in the magnitude of the peaks in the NCF. The peaks in positive correlation time are larger than the peaks in negative correlation time. From this, we can determine that most of the acoustic energy is being received from the East. The peaks in negative correlation time indicate that there are acoustic wavefronts that are propagating from the West, but the smaller magnitude of these peaks indicates less acoustic energy from this direction.

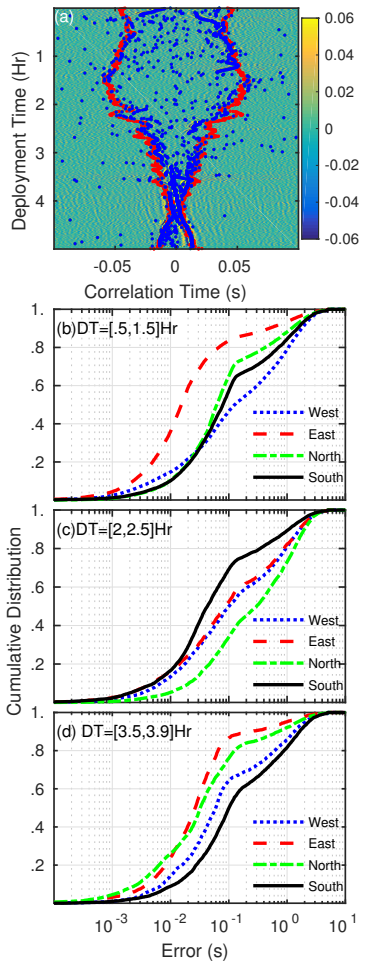

Figure 3.3: Arrival time accuracy analysis. Quantitative evaluation of the accuracy of the arrival time of the NCF peaks compared to the expected direct acoustic path (a) The maximum values of the NCF are shown for each center time, t_c, in both positive and negative correlation time (blue dots). The expected peaks of the NCF are also shown (red dots). (b-d) Cumulative distribution as a function of the difference in predicted direct path and the maximum value of each correlation peak (computed for all pairs of floats). The correlation peaks are separated so that a peak corresponds to an approximate direction of ambient noise sources that would generate that peak. These plots are shown for different Deployment Times (DT) (b) Deployment Times between .5 Hr and 1.5 Hr (c) Deployment Times between 2 Hr and 2.5 Hr (d) Deployment Times between 3.5 Hr and 3.9 Hr. These plots give a feel for both the accuracy of the arrival times as a function of distance as well as the time evolution of the ambient noise field during the 5 hour experiment.

We further quantify this observation by computing the residual describing the difference between the arrival time of the maximum of the NCF and the expected arrival time. We compute

these residuals for all pairs. Figure 3.3a shows an example for one pair. We are comparing the time value of the maximum (shown with blue dots) with the expected path (shown with the red dots). The maximum of the NCF is computed for both positive and negative correlation times for each time step and the orientation of the AUE pair is used to match the positive and negative correlation times with a direction of incidence noise. Figure 3.3b-d describe cumulative distribution of residuals for different directions of ambient noise. These plots are shown for different times in the deployment.

The results shown in Figure 3.3b-d demonstrate two important points about retrieving the arrival structure of the TDGF from ambient noise during this experiment. First, the noise field is anisotropic. Different source directions produce different distributions of residuals. For example, Figure 3.3b shows that the peaks that result from noise coming from the East match the expected TDGF the best (a greater percentage of peaks have smaller residuals) during the time period Hour .5 - Hour 1.5. The second point is that the ambient noise field is time varying. The residuals do not follow similar trends when viewed across different times during the deployment. The smaller residuals for the Easterly ambient noise of Figure 3.3b are not found in Figure 3.3c (Hour 2 - Hour 2.5). Also, Figure 3.3d shows a time in the deployment where ambient noise from the North provides accurate arrival times when compared to the other two times. Over all the deployment times, we can see that the smallest residuals occur during the time period Hour .5- Hour 1.5 with noise coming from the Easterly direction. The peaks resulting from Easterly noise at these times match the expected arrival structure with millisecond accuracy 10% of the time and match with 10 ms accuracy around 40% of the time. In evaluating these numbers it is important to understand that the positions of the AUEs are also estimates that are prone to some error. We do not expect to be able to estimate the direct arrival between the AUEs with less than millisecond accuracy.

The discrepancies in the arrival time during different deployment times points to a difficulty in ambient noise processing. We can see that in Figure 3.2a there are times where the NCF is single sided (see Figure 3.2a around 1 Hr) and even times when there seem to be no peaks at

all (see Figure 3.2a slightly after 2 Hours). This could either be a result of anisotropic noise fields; specifically non-stationary shipping events or it could be the result of interferers passing through the endfire region of a receiver pair. For lower frequencies that are typically dominated by shipping noise, coherent arrivals typically occur during discrete times (See Figure 2e in Lani et al. 2013 [100]). The long time windows of past works are chosen to make the noise fields act more stationary and more isotropic. There is a greater chance of coherent arrivals passing through both receivers when the NCF includes longer time periods. An important result of ambient noise processing is that arrival information of the TDGF can still be recovered even when the noise field is far from the isotropic assumption. In this work we are even further from the isotropic assumption because the receiver movement restricts the time window. We can see the effect of this in the results shown. Some pairs of floats are completely biased, such as the pair shown in Figure 3.2b and d. Even the pairs that do show the symmetric structure of the TDGF have times at which they are biased, such as the pair shown in Figure 3.2a and c.

Techniques have been investigated to overcome the difficulties of anisotropic noise that result in the incorrect arrival times of the NCF shown in Figure 3.2d. Specifically, in seismic studies [85, 78], a second-order correlation process is defined that uses surrounding receivers to improve the TDGF estimate of a given receiver pair. This second order correlation process benefits from scattering on seismic heterogeneities, and uses the other receivers in the array to act as secondary sources. This processes relies on significant reverberation and scattering in order to accumulate noise directions that are different from the dominant source directions. Unfortunately, the acoustic environment of the ocean does not permit enough scattering for this second-order correlation process to overcome anisotropy. We performed this second order correlation process on the data shown in Figure 3.2 and the results were similar to NCF.

3.3.3 Choosing T_r and B_ω

In the previous section, we presented results for computing the NCF for only one choice of T_r and B_ω. The choice of T_r and B_ω along with the relative velocity of the receivers determines the validity of our short time stationary assumption. In order for a wavefront with wavelength λ to contribute coherently to a peak in the NCF, the relative distance between receivers must change less than $\lambda/2$. This condition constrains both B_ω, through the wavelengths being correlated, and T_r through the amount of relative movement that can occur between the receivers (i.e. assuming a constant relative velocity between the receivers over the correlation window). This motivates using noise bands with longer wavelengths to make the NCF more robust to receiver movement. Longer wavelengths are also frequently used in noise correlations because they have the most energy and travel longer distances[93]. In our experiments the frequency band, $B_\omega = [100, 500]$Hz, was chosen by looking at the frequency content at the receivers and choosing bands with the largest energy. In this band we expect shipping traffic and noise from shore to be the dominating noise sources. Typically the choice of bandwidth is limited by the environmental conditions of the receivers, we do not want to correlate over bands which do not pass through both receivers.

Given a set bandwidth, the choice of T_r depends on the expected relative motion of the receivers. For our given wavelengths ($\lambda \approx [3, 15]m$) and rough estimates for relative motion (one to tens of centimeters per second) we are restricted to small values of T_r (on the order of seconds) before we see peak degradation of the NCF. In the results of Figure 3.2 and future figures we choose $T_r = 10s$, a conservative time window allowing for fairly large receiver motion (on the order of tens of centimeters per second).

Determining the optimal choice of B_ω and T_r is non-trivial; there are two competing effects resulting from the choice of these parameters. On one hand, the choice of B_ω and T_r defines a noise floor through the $(2B_\omega T_r)^{-1}$ relationship that must be overcome to see a peak in the NCF. On the other hand, choosing large B_ω and T_r will violate the stationary assumption and can cause interference in the correlation process.

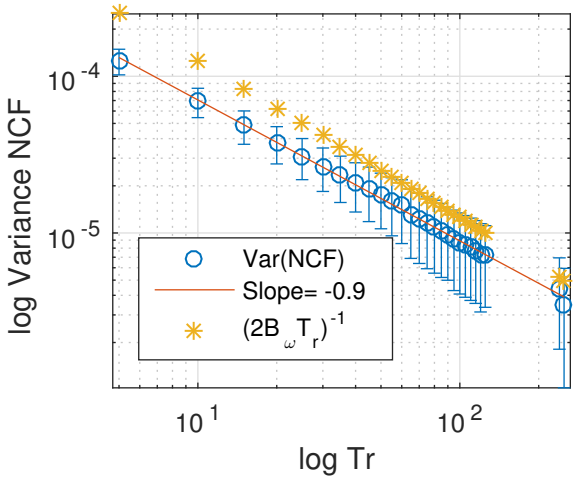

Figure 3.4: Variance of the NCF noise as a function of T_r. The variance of the NCF is computed between time lags .15 and 1s in both positive and negative time for all pairs and all center times, t_c, in the deployment. This is done for different values of T_r, the correlation window length. The mean of all of the computed variance measurements is shown along with one standard deviation of these measurements. A best fit line is drawn for the mean of the measured values. The slope of the best fit line is $-.9$, which is close to the theoretical value of -1, shown by (orange) asterisks, presented in Sabra et al [80]. This plot shows that the noise in the NCF decays close to the predicted model as we vary the window length. This information is important in determining the optimal parameter for T_r that depends on receiver movement and the noise floor.

To demonstrate this trade off we present two Figures. In Figure 3.4 we show the measured noise of the NCF compared to the theoretical model, $(2B_\omega T_r)^{-1}$. Here the noise of the NCF is defined by computing the variance of the NCF for $.15 > |\tau| > 1$s for all center times, t_c, during the deployment and across all pairs. The time lags, $.15 > |\tau| > 1$, were chosen because they are outside of all expected correlation peaks for all pairs. In Figure 3.4 we show the average of the measured variance along with the standard deviation of this measurement. We provide the values for the theoretical model, $(2B_\omega T_r)^{-1}$, for reference. We can see that the slope of the measured noise of the NCF is similar to the slope provided by the model, confirming that noise of the NCF decreases by a power law as T_r is increased.

Figure 3.5 is provided to show how the peak of the NCF evolves as T_r is adjusted. From our earlier analysis we expect the peak in the NCF to increase as we increase T_r until the distance between the receivers changes on the order of a half wavelength. In this example, we can see that

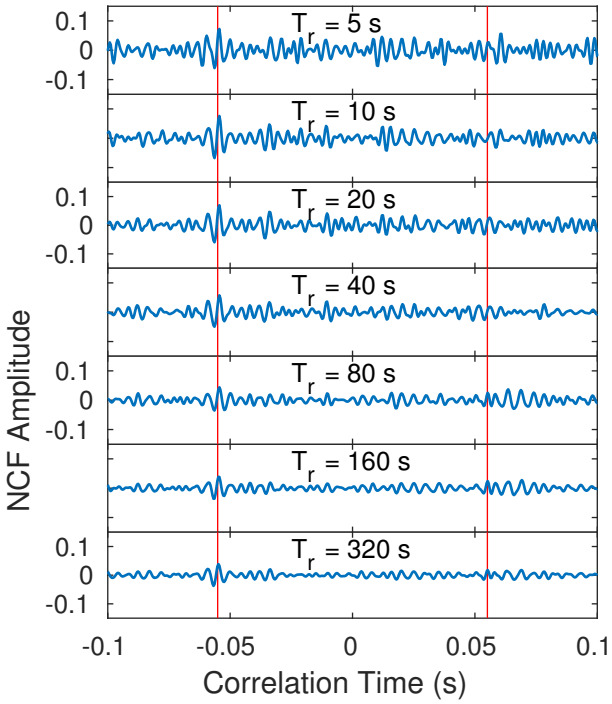

Figure 3.5: NCF peaks as a function of time window length (T_r). The NCF is plotted for different lengths of correlations starting with 5 seconds and ending at 320 seconds. All NCFs share a common center time of $t_c = 1$Hr 30s. A red vertical line is placed where we expect the correlation peak to be based on receiver separation. We can see that as we increase the time window length, the variance of the NCF (i.e. the amplitude of the "noise") decreases. However, the peak on the NCF does not behave as nicely. We can see the peak in negative correlation time become more pronounced as we increase the window length from 5 seconds to 20 seconds after which the amplitude of the peak starts to decline and become closer to the noise floor.

the NCF is one sided, meaning that there is only a peak in negative correlation time. As expected, the noise of the NCF decreases as T_r is increased from 5 seconds to 320 seconds, agreeing with Figure 3.4. Also, we can see the peak of the NCF in negative correlation time increase from 5 seconds until around 20 seconds before the peak begins to lose prominence. While 20 seconds may be close to an optimal choice in this example, it is difficult to determine this globally because it depends on the relative motion between the receivers as well as the noise field at that specific time. This motion is not constant and the noise field is not stationary. This is for one specific example in the deployment and while the intuition should translate to other cases, the optimal

value for T_r is expected to vary based on the noise distribution and relative movement.

Whether or not a peak will emerge in the NCF depends on the source distribution and noise coherence. Specifically, it will depend on whether or not there are noise sources whose components pass through both receivers and how much coherence is measured between the two receivers based on this propagation. We have already seen that this is not a constant process. There are times during the deployment where there is stronger coherence between the receivers and there are also directions where the noise propagation between the receivers results in stronger correlation peaks. In summary, the choice of B_ω and T_r depend on the source distribution, the source level and the relative movement between the receivers. All of these environmental conditions are time varying, and often unknown.

3.4 Propagation Environment

The acoustic environment of the AUEs during the deployment was not ideal for the propagation of the ambient sound field between the elements of the AUE array. The AUE's depth target, 10m, was on the boundary of the mixed layer and a steep thermocline, providing a strong downward refracting profile (Figure 3.6). The shallow depth target allowed surface reflections of the noise field to interfere with the direct path (i.e. the environment is a Lloyd's mirror). Figure 3.6 describes the propagation environment of the receivers based on time averaged temperature data collected every 15 minutes near the array. In panel (a) the sound speed profile is given. In panel (b) we see the transmission loss calculated for the center frequency (300 Hz) using the parabolic equation method [133]. This panel shows transmission loss colored in decibels, demonstrating the difficulty of horizontal transmission at 10m depth. Additionally, the eigenrays, which are valid propagation paths between receivers given the sound speed profile, are plotted for different values of receiver separation in the horizontal direction. The eigenrays show the surface reflected path as well as the downward diffracting ray paths. Both of these demonstrate

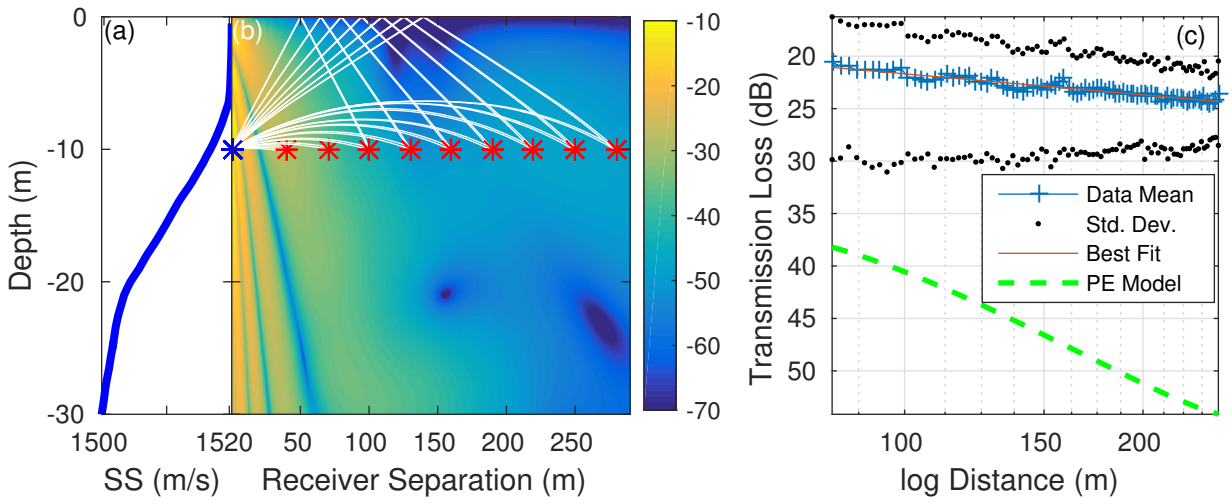

Figure 3.6: Environmental Conditions: (a): Sound Speed Profile averaged for 15 minute intervals during the deployment (b): eigenrays for a sample of horizontal distances from the receiver, we can see a surface reflection appear at distances close to 100m that may cause interference with the direct path. Beneath the eigenrays, the transmission loss is shown (in dB) as the solution to parabolic equations. Both the eigenray and the transmission loss are computed for a frequency of 300Hz. (c) transmission loss for both the measured data (along with the standard deviation for each distance bin) as well as the theoretical values from the parabolic equation. This plot demonstrates the challenges of horizontal propagation in the experiment's environment as well as the fact that on average, we recover a power law from the observed data.

the difficulty of ambient noise propagating between the receivers during this deployment and suggest that more favorable environmental conditions are possible.

Panel (c) of Figure 3.6 shows the transmission loss of the measured amplitude of the NCF from samples of correlation data (e.g. data from pairs like those shown in Figure 3.2), as well as the frequency averaged (for the $B_\omega = [100, 500]$Hz we are using) theoretical transmission loss generated from the parabolic equations. The measured amplitude was generated by taking the maximum value of the NCF for each short time cross correlation (i.e. the rows of Figures 3.2 (a) and (b)) indexed by binning the distance between the receivers at each time (the bin size used was 2m). All possible pairs of receivers were used in this analysis. Both the mean and the standard deviation (in log scale) are shown in this plot. The amplitude of the correlation function has large variance, highlighting the inconsistencies of the noise field during the deployment and the deviations from the isotropic noise field that theoretically recovers the amplitude of the

TDGF. Even though the variance of the amplitude is large, averaging over many samples of the amplitude retrieves a power law. This result is encouraging. On average, we are recovering amplitude information between the different receivers, that is, information of the transmission loss between receivers. While the complicated environment makes it difficult to say anything about the environment with certainty, there does seem to be an underlying physical process between the receivers.

We would expect the NCF to recover the amplitude of the TDGF only under an isotropic distribution of noise sources [79]. The theoretical transmission loss indicates greater attenuation than what is observed in the data. The smaller attenuation in the observed data indicates that the coherent contributions are coming from the far field, which is consistent with shipping and environmental noise sources that are expected in the lower frequency bands. Additionally, the frequency whitening and amplitude clipping that was performed in the computation of the NCF to spread the coherent arrivals over multiple sources, with the intention of making the noise field more "isotropic", may also alter the amplitude of the NCF. What is interesting is that we do see a power law for amplitude decay as a function of increased distance. We believe that this indicates that there is some physical information embedded in the amplitude of the NCF.

3.5 Incoherent Beamformer

In Section 3.3 we concluded that we could not recover the arrival structure of the TDGF from the NCF between some receiver pairs because of an imbalance of the noise distribution in the environment. In this section we assume that there is a dominant source direction and we provide a technique to detect its presence as well as its direction. Coherent beamforming is a typical solution to this problem, yet coherent beamforming would not work well here because: 1. The spacing between receivers is on the order of tens to hundreds of wavelengths whereas coherent beamforming typically requires spacing on the order of 1/2 wavelength to avoid aliasing

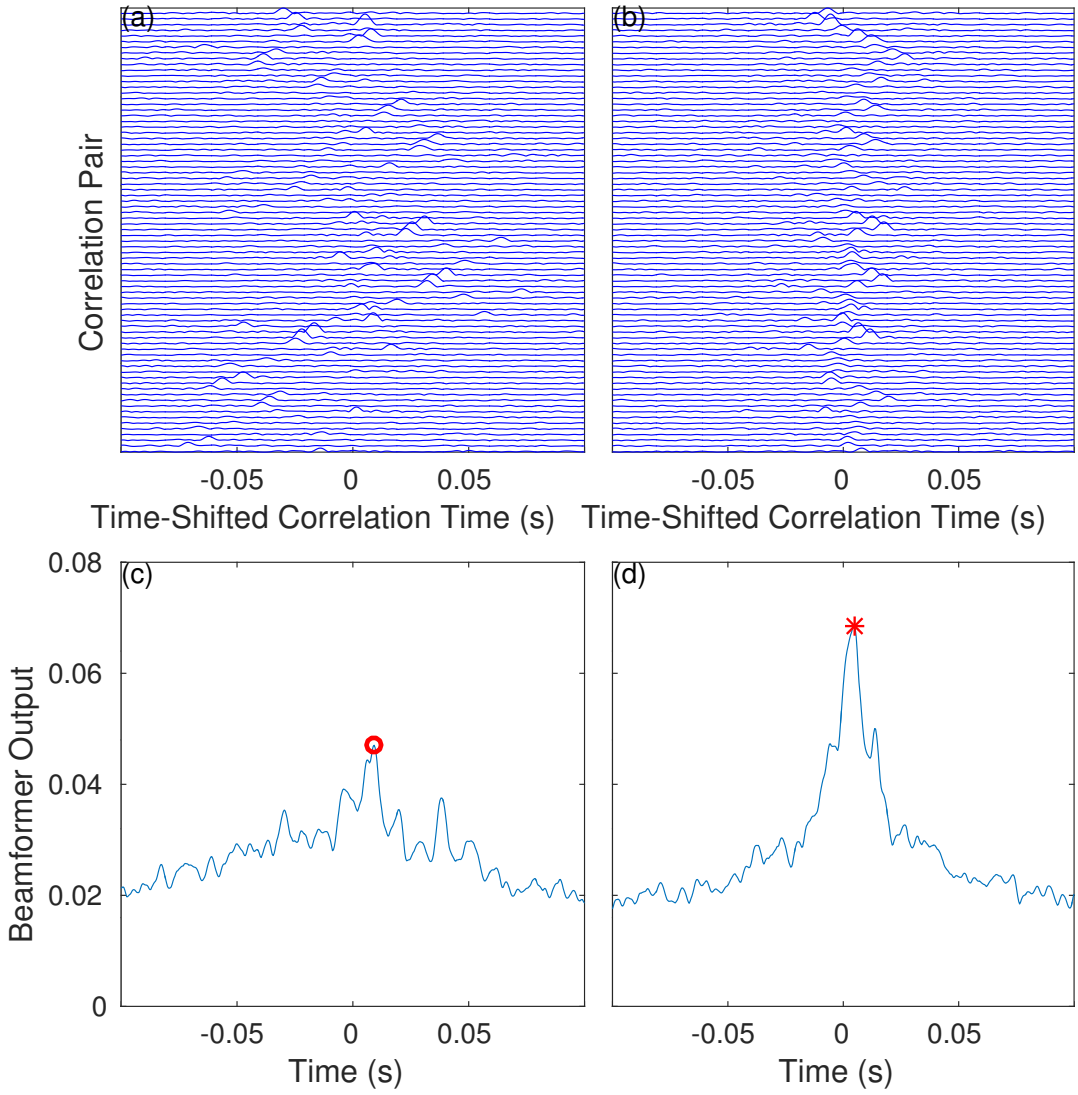

Figure 3.7: Illustrations of the incoherent beamformer procedure. A dominant noise direction is assumed ((a) 45°or Northeast, (b) East) and the NCF is time shifted so that the estimated location of the peak is at 0 correlation time. The locations of the receivers and an assumed noise direction are used to calculate this time shift. When the peaks align around 0 correlation time, such as in (b), their average at 0 correlation time is large. The average for (b) is shown in (d). When the peaks are not aligned, (a), the average of the peaks at 0 correlation time is small. The average of (a) is shown in (c). We take the max of the average at within 10ms of zero correlation time (to account for some error in the system) to be the output of the incoherent beamformer. When the output is large we suggest that the assumed dominant noise direction is supported (in this case directly East). When the dominant noise direction is small the dominant noise direction is not supported.

issues 2. The position of the receivers are only estimated and are not precise enough for any type of coherent processing 3. The configuration of the array is *ad hoc* and does not follow any optimal pattern. To overcome these limitations, we introduce an incoherent beamformer. We show that we are able to recover the dominant noise directions received by the *ad hoc* array of receivers using the noisy localization results from the high frequency pinging system, and we show how these detected directions are consistent with our analysis in Section 3.3.

To describe the incoherent beamformer, we start with the observation that given a plane wave representation of the dominant source, a peak will arise at time δ_p in the NCF between receiver i and receiver j according to:

$$\delta_p = \frac{P_\theta(S_j - S_i)}{c} + \nu \tag{3.2}$$

Where $P_\theta(\cdot)$ is the projection onto the direction of the dominant source (parameterized by θ), S_i and S_j are the positions of receiver i and j, c is the speed of sound underwater and ν is a noise term. The noise term, ν, comes from inaccuracies in the estimated positions, S_i and S_j, as well as violations of the plane wave assumption. Each NCF gives a measurement of the difference in distance of each receiver along the direction of propagation of the plane wave. We can build an incoherent beamformer by 1. assuming a dominant noise direction modeled by a plane wave 2. estimating where the peaks in the NCF would be based on Equation 3.2 3. time shifting the envelope of the NCF so that the expected peak would be at 0 correlation time for all pairs 4. average the NCFs from all pairs. Here we are working with the envelope of the NCF because of the challenges described with phase coherent processing. With this formulation, if there is a dominant source coming from the assumed direction *and* this dominant direction can be modeled by a plane wave then we would expect all of the peaks in the NCFs to average coherently at 0 correlation time. If the dominant source is not coming from the assumed direction, then time shifting the peaks will provide peaks that are not centered around zero and the peaks will not average coherently.

Using the estimated positions of each receiver (S_i and S_j) from the high frequency pinging system and assuming a dominant source direction, θ, (in this case 45°, or Northeast) Figure 3.7 a shows the time shifted envelopes using the procedure described. The peaks in this plot would line up close to zero time if the dominant direction was supported. We can see that this is not the case for the direction assumed in Figure 3.7a. In contrast, Figure 3.7b shows the same plot for a different assumed direction (East) and we can see the peaks align around zero. This represents an agreement of the time-shifted correlations. To further quantify this agreement, Figure 3.7c and 3.7d show the average of the peaks in 3.7a and 3.7b, respectively. In Figures 3.7c and 3.7d the maximum is chosen to be the maximum value of the average within 10 ms of 0 correlation time. We believe 10ms adequately accounts for the error in the system. We call this maximum the output of the incoherent beamformer.

We can perform this calculation for many assumed dominant noise directions to get an idea of the distribution of dominant noise sources for a given time during the deployment. Figure 3.8 shows this result for all directions during the same time that was used in the analysis of Figure 3.7. Here we can see that the dominant noise direction is coming directly from the East because this gives us the largest output for the incoherent beamformer. Additionally, East looks like the only direction of dominant source during this time in the deployment because there is only one direction that produces a large output. Also highlighted are the angles that were shown in Figure 3.7, East corresponds to Figure 3.7b and 3.7d, while 45 degrees (Northeast) corresponds to Figures 3.7a and 3.7c, respectively.

The output shown in Figure 3.8 describes the environment of the array. In this case, the largest response from the incoherent beamformer is East, which is pointing directly towards shore. This is a typical response for this deployment and this supports our hypothesis that dominant directions in the noise field is the reason for symmetry being present in the NCF for some pairs, and not for others. Specifically, it supports the observation that the receivers with their endfire beam along the East-West direction would exhibit the arrival information of the TDGF because

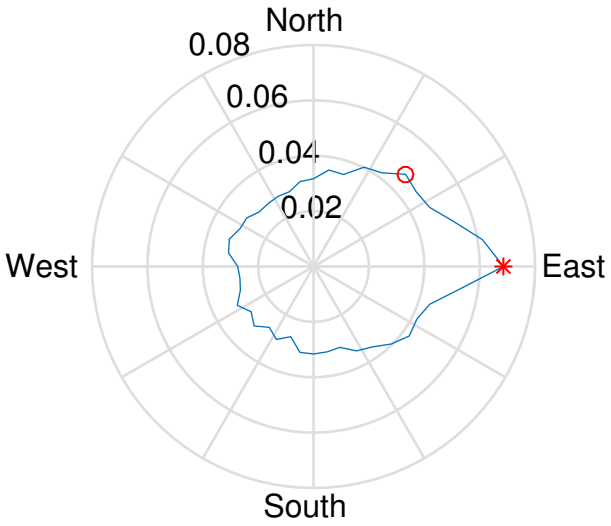

Figure 3.8: The Incoherent Beamformer output for one time during the deployment. We repeat the procedure of assuming a dominant noise direction and computing the output of the incoherent beamformer for all possible directions. This provides intuition of the dominant noise directions at a specific time. For example, at this particular time during the deployment the majority of the noise field is coming from directly East of the array. This result supports our hypothesis that the noise field is not isotropic and is biased in the East-West direction. This is consistent with the observation in Figure 2 that receiver pairs whose endfire beam are in the East-West directions demonstrate fair estimates of the arrival structure of the TDGF and receivers whose endfire is directed more in the North-South direction do not.

these noise directions pass through both receivers while the pairs whose endfire beams where North-South do not exhibit this relationship. This is consistent with the analysis of Figures 3.1 and Figures 3.2 that were provided.

We can show the plot of Figure 3.9 as a function of deployment time so that we can understand the time evolution of dominant signals in the ocean. Each row of Figure 3.9 is the same information shown in Figure 3.8 unrolled with the color showing the output amplitude of the incoherent beamformer. Like the NCF function presented in Figure 3.2, these are also computed in 15 second intervals during the deployment. For this plot we chose a period of time (approximately Hour 3.5 to Hour 3.8) where we have verified that there is an audible boat in the vicinity of the receivers. This boat provides dominant noise directions that are different from the

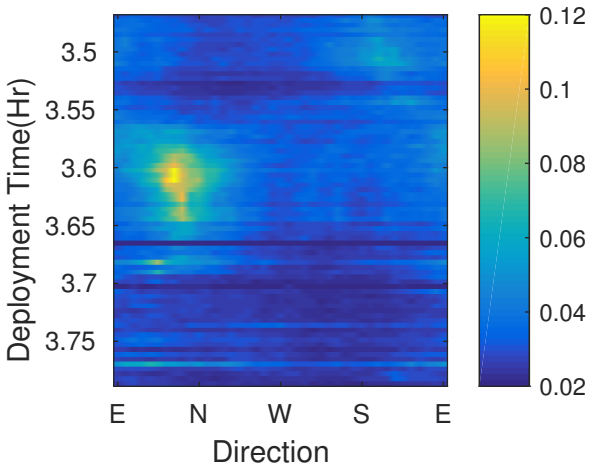

Figure 3.9: Time evolution of the beamformer output. Here each row is computed similar to Figure 5 and is unrolled with the amplitude represented by the color shown on the color bar. We can see that there is a dominant direction of the ambient sound field just North of East. This is different than the dominant directions usually detected. We have verified that there is an identifiable boat in the recordings during this time so the presence of a directional bias in the ambient noise field is expected. This plot demonstrates the time varying nature of dominant noise directions during the deployment in $B_\omega = [100, 500]Hz$ band and this result is consistent with the arrival structure of the noise correlation function during this time in the deployment (shown by a green circle in Figure 2c and 2d).

rest of the deployment. We can see from Figure 3.9 the direction of the incident noise coming from the boat and we can see that this lasts for less than 10 minutes.

For the boat case, it is interesting to see how our assumption of a plane wave representation for the ambient noise holds up. We can see that the peak in this figure is spread over many values of θ which may indicate that we are not computing the appropriate time shifts from Equation 3.2 because the source is too close to the array. This may hint that there may be able to be some improved detection by refining the estimate to include nearby sources (that cannot be modeled by a plane wave). Nonetheless, we see that the incoherent beamformer is able to detect the presence of dominant noise signals even though the spacing between the receivers is large, the geometry of the array is not optimal for coherent processing and the positions of the receivers are not precisely known.

Figure 3.9 shows a dominant noise direction from the north as opposed to the typical easterly

direction. If we revisit the NCF in Figure 3.2 we can see the effect of this boat, highlighted by the green circles in the intensity plot. We can see that for a brief period of time, the arrival information matches with the pair whose endfire beam is in the North-South direction and not for the pair whose endfire is in the East-West direction. This is also consistent with the arrival analysis of Figure 3.3 where we can see that the arrival information is more accurate for the North direction during Hour 3.5 to Hour 3.9 (Figure 3.3) than during other times in the deployment (Figures 3.3b and c). This again shows consistency of the incoherent beamformer observations with the results we obtain in the NCF.

3.6 Conclusion

We presented a new application of ambient acoustic noise processing using mobile receivers whose geometry deforms over time. In this challenging environment, we were able to extract information related to the travel time between pairs of receivers using only 10 second cross-correlations. The amplitude of the TDGF is generally not retrieved in ambient noise processing because of complications in the noise distribution. The average amplitude of the NCF follows a power law when tracked over varying distances of receiver pairs demonstrating that the amplitude may follow some physical model. We developed an incoherent beamformer that we used for source localization. We also used the incoherent beamformer to detect the direction of anisotropy, which was consistent with our analysis of the noise correlation function.

In terms of a self-localization solution there is some promise in the results presented. The TDGF is not consistently retrieved for all pairs of receivers. For example, Figure 3.2 shows that for some geometry of receivers the NCF is not double sided and symmetric. Without being able to extract the TDGF for all receivers the self-localization approach from Sabra et. al. [66] will be challenged. However, there are some arrivals (peaks in the NCF), even with the short time windows, and for some pairs of receivers these arrivals are double sided. This demonstrates that

there are usable constraints that result from low frequency noise correlations, but these constraints are biased by anisotropies in the noise field. The next chapter of this dissertation looks at using locally dominant sources for a self-localization solution.

3.7 Acknowledgements

Chapter 3, in part, is a reprint of the material as it appears in The Journal of the Acoustical Society of America: P. Naughton, P. Roux, C. Schurgers, R. Kastner, J. S. Jaffe, and P. L. D. Roberts, (2016) “Ambient noise correlations on a mobile, deformable array”. The dissertation author was the primary investigator and author of this paper.

Chapter 4

Self-localization of a mobile swarm using noise correlations with local sources of opportunity

4.1 Introduction

In the last chapter, the challenges of ambient noise processing on a mobile array were considered. These challenges stem from the fact that movement between vehicles changes the propagation path between them and restricts the time windows that can be used in the noise correlations. Despite the tough restrictions on the length of time that can be used, we saw that noise correlations across mobile vehicles still exhibit arrivals from different noise directions. Without noise coming from all directions, this arrival structure is biased from the dominant noise directions in the ocean environment. Consequently, accurate distance estimates between individual vehicles cannot be estimated and only the receiver separation projected onto a dominant noise direction can be retrieved. This projected information is used in the proposed self-localization solution.

For mobile underwater vehicles, a self-localization solution is required that jointly estimates the vehicle geometry and the dominant noise directions of the ambient noise field when recovering the acoustic impulse response between all pairs is unrealistic. In this chapter, a robust method to perform this estimation is proposed and evaluated using data collected from a multi vehicle survey where each vehicle independently drifts with subsea currents. The vehicle geometry as well as the direction of arrival of a moving boat are accurately estimated and compared to estimates from an accurate, independent localization system. Simulations are provided to examine three different factors that affect the proposed solution: inter-vehicle motion, vehicle geometry, and the azimuthal variance of the noise field.

4.2 Theory and Definitions

4.2.1 Problem Description

We consider a group of underwater vehicles that have a single hydrophone on board each vehicle and are close enough to each other that they record common noises in the ambient soundscape. The deployments considered in Refs [11, 12, 8] fit this requisite. We are able to determine their depth using onboard pressure sensors, and are interested in estimating their horizontal positions *relative* to one another. We can only solve for the positions of the vehicle up to a rigid transformation (i.e. a translation and a rotation) because we do not assume to know the locations of any noise sources or vehicle positions. This is useful for deployments that aim to study spatial data in a relative sense (e.g. chemicals, currents, or optical images), or to measure the time difference of arrivals of acoustic signals that are received by the elements of the swarm. Additionally, relative localization solutions can be augmented with a global reference when it is available, such as a GPS position when the vehicle surfaces.

We estimate the relative geometry of the vehicle swarm using cross-correlations of low frequency noise in the ocean, and assume that this noise is coming from the far field. In our

experiment, the low frequency noise is dominated by a moving vessel, so the developments in this section reflect a single moving sound source. However, our approach is not limited to this scenario and can be modified to include multiple dominating sources. Examples of sources that could be used are low frequency noise from shore or noise from biological sources [66]. Naughton et al. [11] shows examples of noise correlations that result from multiple dominating sources (e.g. shipping noise and noise from the coast), that are common in coastal environments.

The outline for our procedure is shown in Figure 4.1. We start with acoustic data from independent vehicles and use noise correlations to extract time difference of arrival (TDOA) constraints between vehicles. Our experiment and TDOA extraction methods are reported in Section 4.3 and 4.4, respectively. With these constraints, the estimation described in Section 4.2.3 and 4.2.4 is performed and then refined by the procedure described in Section 4.2.5.

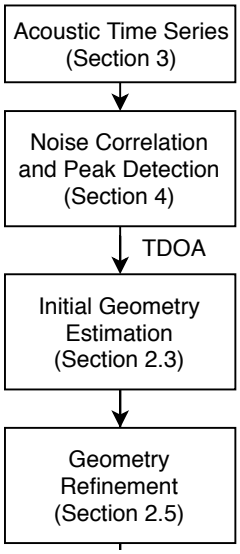

Figure 4.1: Outline of our processing pipeline: We start with raw acoustic data from each individual vehicle. The vehicles and the deployment are described in Section 4.3. We compute noise correlations between the vehicles and extract TDOA values from the noise correlations, as described in Section 4.4. Finally, we perform the estimation described in Section 4.2 by first estimating the relative positions using the formulation of Section 4.2.3 and 4.2.4 and then refining the positions using the procedure described in Section 4.2.5.

4.2.2 Definitions

We define the 2 dimensional position of vehicle i, at time t, as $R_{i,t}$, the collection of vehicle positions at time t as a matrix, $\mathbf{R}_t$, and the 2 dimensional position of the dominant sound source at time t as $S_t = [S_{x,t}, S_{y,t}]$. We can arbitrarily assign the x, y position of any vehicle (in this case the first vehicle), to $R_{1,t} = [0,0]$, because the solution can only be solved in a relative coordinate frame. Throughout the text we will refer to the vehicle defined at the origin as the anchor vehicle. With this assignment, we define a relative position matrix for n vehicles as:

$$\overline{\mathbf{R}}_t = \begin{bmatrix} x_{2,t} - x_{1,t} & y_{2,t} - y_{1,t} \\ x_{3,t} - x_{1,t} & y_{3,t} - y_{1,t} \\ \vdots & \vdots \\ x_{n,t} - x_{1,t} & y_{n,t} - y_{1,t} \end{bmatrix} = \begin{bmatrix} x_{2,t} & y_{2,t} \\ x_{3,t} & y_{3,t} \\ \vdots & \vdots \\ x_{n,t} & y_{n,t} \end{bmatrix}$$

We assume that S_t is in the far field so that the propagation of noise from S_t to every vehicle $R_{i,t}$ can be modeled by a plane wave and parameterized by a single angle, α_t. We also assume that S_t is moving so that $S_t \neq S_{t+\varepsilon}$ for some small value of ε. This gives us two collections of variables we wish to estimate, the time evolution of the angles from the dominant noise source, α_t, and the time evolution of the relative position matrix, $\overline{\mathbf{R}}_t$.

4.2.3 Solution

The estimation is based on time difference of arrival (TDOA) measurements. The time difference of arrival between vehicle i and j for the noise source at time t is defined as:

$$\delta_{i,j,t} = \frac{\|R_{i,t} - S_t\|_2}{c} - \frac{\left\|R_{j,t} - S_t\right\|_2}{c} \tag{4.1}$$

Here $\|\cdot\|_2$ denotes Euclidean distance, and c is the speed of sound underwater. By only looking at the cases where $j = 1$, remembering that we set $R_{1,t} = [0,0]$, and incorporating the far

field assumption, Equation 4.1 can be rewritten as:

$$\delta_{i,1,t} \approx \frac{[\overline{R}^x_{i,t}\ \overline{R}^y_{i,t}][\cos(\alpha_t)\ \sin(\alpha_t)]^T}{c} \tag{4.2}$$

Equation 4.2 describes a constraint on the relative geometry, $\overline{\mathbf{R}}_t$, but it is only a constraint along the angle α_t. In order to estimate the direction perpendicular to α_t, the estimation must consider angles that describe the component perpendicular to α_t. In other words, our estimation must include at least one more angle that is sufficiently different from α_t to provide constraints along the component perpendicular to α_t. This requires assumptions on both the motion of the sound source, S_t, and the relative positions of the vehicles, $\overline{\mathbf{R}}_t$. Specifically, there is an assumption that the relative geometry remains unchanged ($\overline{\mathbf{R}}_{t_o} \approx \overline{\mathbf{R}}_{t_o \pm t_w}$) for some time window, t_w, so that only the relative positions of the swarm at the center time, $\overline{\mathbf{R}}_{t_o}$ need to be estimated. This is an assumption on the *relative* movement between the vehicles, which is usually much smaller than the vehicle movement on a global scale, especially for vehicles that are drifting with currents or moving in a formation. We have seen this assumption on the vehicle geometry holds true for the presented experiment as well as in other deployments [11, 12]. The second assumption is that the source, S_t, will provide a collection of angles $\alpha_{t_0-t_w:t_0+t_w}$ that are sufficiently different from the center angle, α_{t_0}. Together these two assumptions provide what is needed for a self-localization solution: a vehicle geometry that is stationary for long enough to estimate the geometry in two dimensions.

This defines the approach. Given a time window, $t = t_0 - t_w : t_0 + t_w$, the estimation will produce an estimate of the vehicle geometry at the center time, $\overline{\mathbf{R}}_{t_o}$ as well as the collection of angles, $\alpha_{t_0-t_w:t_0+t_w}$, which we will denote $\alpha_{\mathbf{t_0},\mathbf{t_w}}$ for simplicity. The constraints of Equation 4.2 set up a minimization problem to estimate the angle of arrival of the source at each time, and the relative geometry of the vehicles [127]:

$$\underset{\overline{\mathbf{R}}_{t_o},\Lambda_{t_0,t_w}}{\arg\min}\|\overline{\mathbf{R}}_{t_o}\Lambda_{t_0,t_w} - c\mathbf{D}_{t_0,t_w}\|^2 \tag{4.3}$$

Here, $||\cdot||$ denotes the Frobenius norm, Λ_{t_0,t_w} describes the constraints placed by the angles, α_t:

$$\Lambda_{t_0,t_w} = \begin{bmatrix} \cos(\alpha_{t_0-t_w}) & \cos(\alpha_{t_0-t_w+1}) & \dots & \cos(\alpha_{t_0}) & \dots & \cos(\alpha_{t_0+t_w-1}) & \cos(\alpha_{t_0+t_w}) \\ \sin(\alpha_{t_0-t_w}) & \sin(\alpha_{t_0-t_w+1}) & \dots & \sin(\alpha_{t_0}) & \dots & \sin(\alpha_{t_0+t_w-1}) & \sin(\alpha_{t_0+t_w}) \end{bmatrix}$$

and $\mathbf{D}_{t_o,t_w}$ is a collection of the TDOAs compared to vehicle 1, $\delta_{i,1,t}$. Since we are comparing to the same vehicle, we drop the subscript of 1 (i.e. $\delta_{i,1,t} = \delta_{i,t}$) and define $\mathbf{D}_{t_o,t_w}$ as:

$$\mathbf{D}_{t_o,t_w} = \begin{bmatrix} \delta_{2,t_o-t_w} & \delta_{2,t_o-t_w+1} & \dots & \delta_{2,t_o} & \dots & \delta_{2,t_o+t_w} \\ \delta_{3,t_o-t_w} & \delta_{3,t_o-t_w+1} & \dots & \delta_{3,t_o} & \dots & \delta_{3,t_o+t_w} \\ \vdots & \vdots & \ddots & \vdots & \ddots & \vdots \\ \delta_{n,t_o-t_w} & \delta_{n,t_o-t_w+1} & \dots & \delta_{n,t_o} & \dots & \delta_{n,t_o+t_w} \end{bmatrix}$$

In the noise free case, $\mathbf{D}_{t_o,t_w}$ should be at most rank 2 because it is the product of two rank 2 matrices ($\overline{\mathbf{R}}_{t_o}$, Λ_{t_0,t_w}). Therefore, an affine solution to Equation 4.3 can be found using a singular value decomposition of the matrix $\mathbf{D}_{t_o,t_w}$:

$$UVW^T = \text{svd}(\mathbf{D}_{t_o,t_w}) \tag{4.4}$$

and only considering the first two singular values and vectors for the estimation of $\overline{\mathbf{R}}_{t_o}$ and Λ_{t_o,t_w}:

$$\overline{\mathbf{R}}_{t_o} = U'V' \qquad \text{and} \qquad \Lambda_{t_o,t_w} = W'^T \tag{4.5}$$

In other words, U' is an $(n-1)$ x 2 matrix consisting of the first two left singular vectors (where n is the number of vehicles), W' is an $(2*t_w)$ x 2 matrix consisting of the first two right singular vectors (where t_w controls the number of time samples used in the estimation) and V' is a diagonal matrix consisting of the two largest singular values. The product of $\overline{\mathbf{R}}_{t_o}$ and Λ_{t_o,t_w} from Equation 4.5 give an optimal rank 2 approximation to $\mathbf{D}_{t_o,t_w}$, and an optimal affine solution to the minimization problem of Equation 4.3. However, this solution is not unique and given any invertible matrix, C,

$$\overline{\mathbf{R}}^*_{t_o} = U'V'C^{-1} \qquad \text{and} \qquad \Lambda^*_{t_o,t_w} = CW'^T \tag{4.6}$$

is also a solution to 4.3. In general, an affine solution will not uphold the Euclidean constraint that $\cos(\alpha_i)^2 + \sin(\alpha_i)^2 = 1$, so we choose C such that this discrepancy is minimized [127]:

$$C^* = \arg\min_C \| [1\ 1]\Lambda^*_{t_o,t_w} \cdot \Lambda^*_{t_o,t_w} - [1\ 1\ \ldots\ 1]\|^2 \tag{4.7}$$

where $\cdot$ denotes the dot product. Once an optimal value of C is computed, the optimal value of the relative distance matrix, $\overline{\mathbf{R}}^*_{t_o}$, and the optimal angles of arrival, $\Lambda^*_{t_o,t_w}$, are given by Equation 4.6. The nonlinear least squares problem of Equation 4.7 is only a 4-parameter estimation problem and can be estimated using methods like gradient descent (our implementation uses Levenberg-Marquardt). This is less complex than trying to directly estimate the relative positions of all of the nodes as well as the angle of arrival of all of the noise sources, and makes the estimation more robust to local minima.

A variety of angles are needed in order for this estimation procedure to be well conditioned. When all α_{t_0,t_w} are similar, the position of the vehicles will not be constrained along the dimension that is perpendicular to the angles α_{t_0,t_w}, and the optimization procedure of Equation 4.3 will be unstable. This introduces a fundamental trade-off when considering a single moving source:

increasing the time window, t_w, increases the opportunity for more azimuthal variety, but it also means that the assumption of a constant swarm geometry is less likely to hold, $\overline{\mathbf{R}}_{t_o} = \overline{\mathbf{R}}_{t_o \pm t_w}$.

4.2.4 Condition number of C

C provides a "Euclidean upgrade" to the affine solution based on the angles of arrival computed in Equation 4.6 [127]. When the angles of arrival are similar, the estimation of Equation 4.6 becomes ill-conditioned and will be more sensitive to noise on the input values, $\mathbf{D}_{t_o,t_w}$. A way to measure this is to look at the condition number of C after it is estimated by Equation 4.7, which estimates the worst case loss of precision from the Euclidean upgrade. If the condition number of C is large, the final solution to Equation 4.6 will be more sensitive to noise on the input, $\mathbf{D}_{t_o,t_w}$, and the solution is less likely to produce a good estimate.

The choice of the anchor unit when computing $\overline{\mathbf{R}}_t$ affects the condition number of C. Since the swarm presented in Section 4.3 is relatively small, Equation 4.6 is computed for all possible values of the anchor unit and the solution that yields the smallest condition number for C is chosen. This chooses a solution that is least sensitive to input noise.

4.2.5 Refinement

The problem with the solution outlined in the last section is that it does not utilize all of the information between all vehicle pairs. Since the relative distance matrix is relative to an arbitrary anchor unit, we are only considering the TDOA values compared to that anchor unit, i.e. $\delta_{i,1,t}$. This means that the solution is susceptible to poor detections on the anchor unit and that there is information received by other pairs that could be useful for the self-localization solution.

To overcome this, the solution to Equation 4.6 is refined by minimizing the following nonlinear optimization problem using the Levenberg-Marquardt algorithm:

$$\overline{\mathbf{R}}^*_{t_o}, \alpha^*_{\mathbf{t_o},\mathbf{t_w}} = \underset{\overline{\mathbf{R}}_{t_o},\alpha_{\mathbf{t_o},\mathbf{t_w}}}{\arg\min} \sum_{h=1}^{N} \sum_{i=h+1}^{N} \sum_{t=t_o-t_w}^{t_o+t_w} ([R_{h,t_o} - R_{i,t_o}][\cos(\alpha_t)\ \sin(\alpha_t)]^T - c\delta_{h,i,t})^2 \tag{4.8}$$

We use the solution computed in Equation 4.6 as the starting point for this optimization problem. Equation 4.8 refines the solution of Equation 4.6 to include the information from new pairs of vehicles.

4.3 Experiment

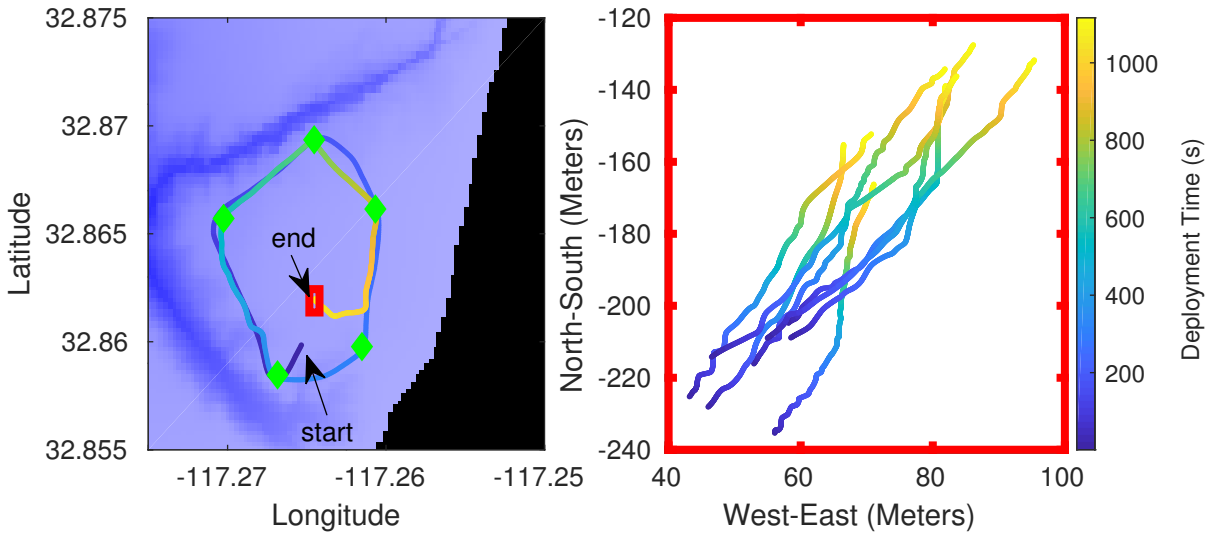

Figure 4.2: Experiment Setup: (left) 5 acoustic pingers, shown by green diamonds, are moored along the perimeter of the deployment, which took place off of the coast of La Jolla. The bathymetry as well as the shoreline are shown by colors (darker blue is deeper, black denotes the shoreline). These pingers collect GPS data and take turns transmitting a linearly modulated chirp (8-15kHz). 8 Autonomous Underwater Explorers (AUEs), deployed inside the red box, hold a depth of 6 meters and float with the subsea currents. The trajectories of the AUEs are estimated using time of flight measurements from the 5 acoustic pingers. While the AUEs were deployed, a boat circled the swarm twice and the GPS trajectory of the boat is shown with the start and end positions indicated. The right panel shows a close up of the AUE trajectories where the red bounding box matches the box on the left panel. Deployment times for both the boat and AUE trajectories are shown by the colorbar on the right.

We described the core estimation of the relative vehicles positions in Section 4.2, and we

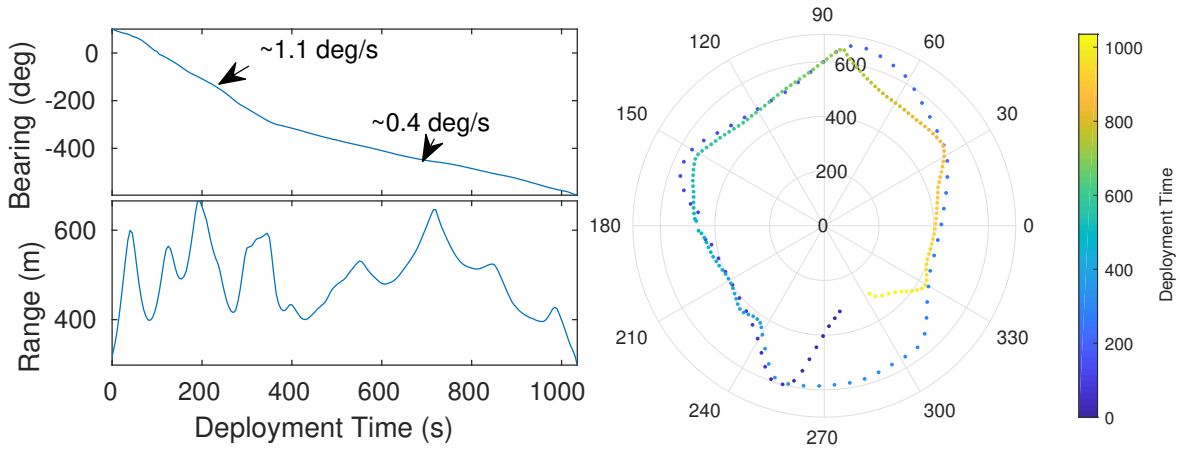

Figure 4.3: Range and bearing of the boat: (left-top) The bearing of the boat in relation to the center of the AUE swarm. (left-bottom) The distance between the boat and the center of the AUE swarm. The bearing plot shows a change in the derivative of the bearing around 400 seconds into the deployment as the boat slows its speed (from 1.1 deg/s to 0.4 deg/s). (left-bottom) The range plot shows that the boat is in the far field of the AUE swarm for the entire deployment. (right) The same information is shown in polar form.

will describe how we extract the constraints for our estimation in Section 4.4. In this section we describe our experiment, which will make the description of Section 4.4 more clear.

Our self-localization approach was tested using data from a multi-vehicle deployment. Acoustic data was collected using Autonomous Underwater Explorers (AUEs), designed and built at the Scripps Institution of Oceanography[8]. Each AUE is a self-contained sensor that collects acoustic data from an HTI-96-MIN hydrophone, in addition to accelerometer, temperature, and pressure data. The AUEs track a depth in the ocean's water column by adjusting their buoyancy. When the motor is on, the hydrophones are saturated. For an individual AUE, the motors ran at least 10% of the duration of the experiment and are independent of the other AUEs. The AUEs have no actuation in either horizontal direction and currents move the AUEs substantially while deployed in the ocean. The clocks of the AUEs are synchronized at the beginning and end of the experiment using a GPS receiver on each AUE, and a linear drift correction is applied to each AUE clock before any processing on the acoustic data is performed. The AUEs measure

their depth using an onboard pressure sensor. In order to provide a 'Ground Truth' estimate for our vehicle geometry, an acoustic triangulation system is set up to act similar to a GPS system. Five acoustic buoys are positioned on the surface of the ocean to send a linearly modulated chirp (8-15kHz) to be received by the hydrophone of each AUE. Each of five buoys takes turns pinging and each packet of five pings occurs every ten seconds. The AUEs have limited on-board processing and no communication infrastructure, so all localization is done offline after they are retrieved. We believe that these localization estimates are accurate to about 1 meter.

The experiment was performed off of the coast of the Scripps Institution of Oceanography and is described in Figure 4.2. Eight AUEs and the five pinger localization system were deployed close to shore, in water that was approximately 20 meters deep. The pingers were separated between 0.5 and 1km apart. The AUEs were programmed to hold a depth of 6 meters and were accurate to ± 1 meter. While the AUEs were deployed, a small boat circled the acoustic pinging array twice while recording its GPS position at 1 Hz. In the first loop around the array, the boat traveled approximately 11 m/s and circled the array in approximately 400 seconds. During the second loop, the boat was slowed to 4 m/s and traveled around the array for approximately 700 seconds. Both the bearing and range of the boat from the center of the swarm is shown in Figure 4.3. In this figure, we can see the change in speed of the boat in the bearing plot. The range plot also shows that the boat is in the far field of the AUE swarm.

4.4 Noise Correlation Function

In order to provide physical constraints from the moving boat, arrival information must be extracted between AUEs. This is done by correlating low frequency noise between pairs of vehicles to extract TDOA information from dominant noise directions (i.e. Equation 4.1). These TDOA measurements provide the input to the self-localization solution described in Section 4.2. This section describes implementation details regarding these low frequency correlations.

There are conditions that make the boat difficult to detect for some deployment times. First, there is missing data in the acoustic recordings resulting from the vehicle running its motor to adjust its buoyancy (described in Section 4.3). The boat can also be difficult to detect when it is broadside to the vehicles, and when there are other noise sources in the vicinity of the swarm broadcasting in the same frequency band. To combat these issues, we preprocess the acoustic data before correlating each pair of audio tracks. Our process is described in Section 4.4.1. We also employ temporal smoothing to reject spurious arrivals caused by missed detections or other noise sources in the vicinity of the array, described in Section 4.4.2.

4.4.1 Preprocessing and noise correlations

Each AUE, indexed by i, records a pressure signal that we denote as $\tilde{p}_i(t)$. Before correlations are performed between different vehicles, the noise is preprocessed similar to work focused on estimating the acoustic impulse response between receivers [104, 99, 80, 102]. These preprocessing steps help suppress spurious arrivals in the noise correlation function, allowing the noise correlations to focus on the dominant source. The first pre-processing step is to whiten the signal's frequency spectrum as well as defining the bandwidth, B_ω, that will be correlated over. The whitening weighs each frequency band equally across the defined bandwidth, B_ω, in the computation of the noise correlations instead of having a few high energy bands dominate. Next, time series values with high amplitudes are truncated to the fourth standard deviation of each $\tilde{p}_i(t)$ signal. More insight into these preprocessing steps can be found in the literature [104, 99, 80, 102]. We define the preprocessed version of $\tilde{p}_i(t)$ as $p_i(t)$ and compute the normalized noise correlation between two signals, $p_i(t)$ and $p_j(t)$ as:

$$C_{i,j,t_c}(\tau) = \int_{t_c-\frac{T_r}{2}}^{t_c+\frac{T_r}{2}} p_i(t)p_j(t+\tau)dt \Bigg/ \sqrt{\int_{t_c-\frac{T_r}{2}}^{t_c+\frac{T_r}{2}} p_i(t)^2 dt}\sqrt{\int_{t_c-\frac{T_r}{2}}^{t_c+\frac{T_r}{2}} p_j(t)^2 dt} \tag{4.9}$$

Where t_c controls the center time of the signal and T_r the length of the correlation window. The normalized cross correlation allows us to make comparisons between different correlation times by removing the total energy at each time step. The noise correlations effectively have two user defined parameters: T_r and B_ω. For our experiments, we took T_r=3s and $B_\omega = [0.1\text{-}1.0$ kHz].

The envelope of the noise correlation between two different pairs is shown by the colorbar in Figure 4.4. Each row is a short time noise correlation centered at the center time, t_c, indicated on the y axis. The center times are spaced one second apart. As expected, the noise correlations are dominated by the circling boat. The noise correlation changes as the boat moves around the vehicles.

4.4.2 Temporal smoothing

While the SNR of the boat is large for most correlations, the short time windows and hydrophone dropouts from buoyancy adjustment create peaks in the noise correlations that are not from the boat or eliminate peaks in the correlations entirely. To improve the estimation, the detections are smoothed along the deployment time axis to enforce the constraint that the peak in the noise correlation must vary continuously in time. To do this, the Viterbi algorithm[134] was implemented to estimate the sequence of TDOA values, $\hat{\delta}_{i,j,t}$ for $t_c = [1,2,\ldots,t_{end}]$. The implementation is heuristic in that the choices of probability distributions are not rigorously justified, but were chosen so that the estimates visually tracked the peaks in the correlation well.

The Viterbi algorithm requires two probabilities to be defined, one describing transition probabilities between $\hat{\delta}_{i,j,t-1}$ and $\hat{\delta}_{i,j,t}$ and the other the probability of the noise correlation having the observed structure given the true TDOA value, $\delta_{i,j,t-1}$. We define the transition probabilities as Gaussians centered around the last state:

$$P(\delta_{i,j,t} = \beta | \delta_{i,j,t-1} = \gamma) = \mathcal{N}(\gamma, \sigma^2) \tag{4.10}$$

For the observation probability, we use a heuristic that the probability of the noise correlation observation given that the TDOA measurement is equal to β is related to the envelope of the normalized noise correlation evaluated at β:

$$P(C_{i,j,t}|\delta_{i,j,t} = \beta) \propto \text{env}(C_{i,j,t}(\beta)) \tag{4.11}$$

where $\text{env}(\cdot)$ denotes the envelope operator.

For our purposes, this algorithm tries to accumulate the most correlation energy while only allowing small variation in the time value of consecutive correlation peaks. This temporal constraint on the peaks of the correlation function can be adjusted by the standard deviation choice of the transition probability. We chose $\sigma = 0.3ms$, which visually tracked the peaks well for the pairs examined. In our case, this variable was not optimized, but could be chosen given *a priori* knowledge of expected movement of a source in relation to the vehicles. Given these probabilities, the Viterbi algorithm produces the most likely sequence of TDOA values, $\hat{\delta}_{i,j,t}$, based on the time evolution of the noise correlations. This smoothing is performed for noise correlations between all pairs of vehicles.

The accuracy of the detection for every pair of vehicles is summarized in Figure 4.5. For this Figure, the value of t_c is shown on the x-axis and multiple statistics about the discrepancy between $\hat{\delta}_{i,j,t}$ and $\delta_{i,j,t}$ (i.e. the '+' and '.' of Figure 4.4) are shown for all pairs. The residuals do not have a mean of 0 for much of the deployment time, but this bias is usually under 1ms. The spike that was visible in the correlations given in Figure 4.4 (see the right panel at $t_c = 700s$) is also visible in Figure 4.5. What is difficult to deduce in Figure 4.5 but is visible in Figure 4.4 is that the errors are not distributed around the expectation (plotted by '+' in Figure 4.4). Figure 4.4 shows that the detected correlation peaks frequently have a constant offset from the estimated correlation peaks. This could be a result of slight inaccuracies in the "ground truth" positions of the AUE, which would effect the value of $\delta_{i,j,t}$. This is important because this error is clearly not a result of a white Gaussian process and cannot be modeled as such when we try to understand

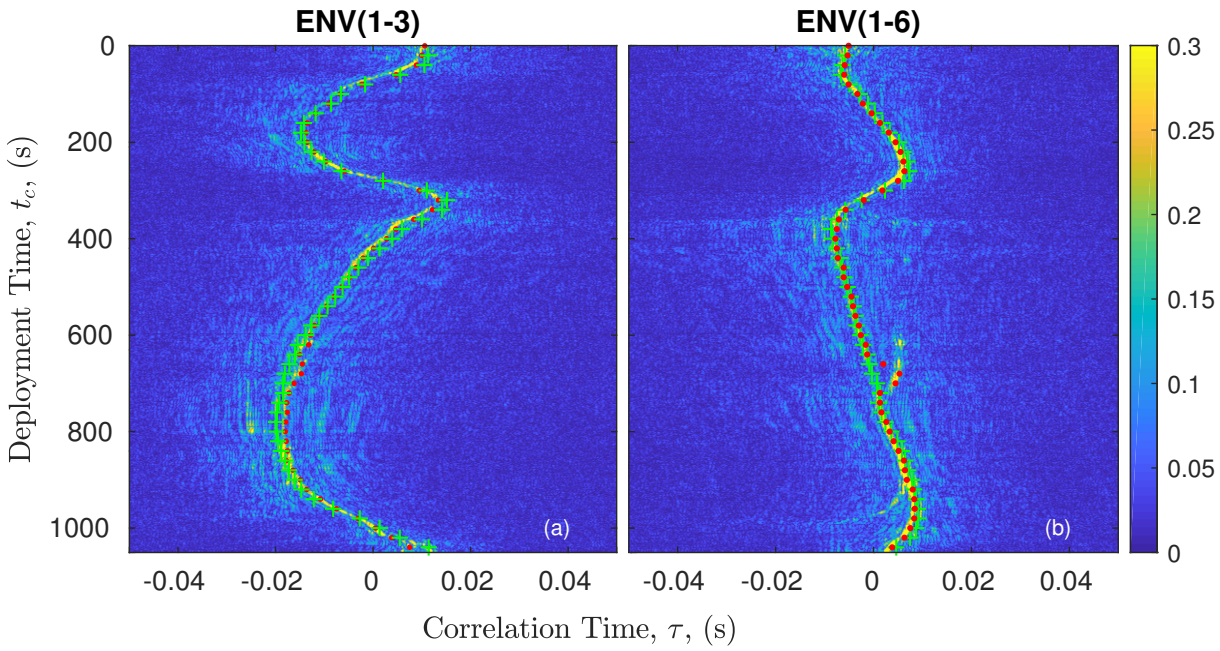

Figure 4.4: Noise correlations: (a-b) The envelope of the noise correlations are shown between two different pairs of AUEs while the boat is circling the array. Each row is the envelope of a short time cross correlation with the amplitude shown by the colorbar. The center time of the cross correlation is shown by the deployment time on the y-axis. The expected peaks of each noise correlation are shown by green '+' and were computed using the GPS position from the boat and the estimated locations of the AUEs from the 5 acoustic pingers. The red dots show the estimated location of the peaks, $\hat{\delta}_{i,j,t}$, for each deployment time after the smoothing operation of Section 4.4.2 is performed. The detected peaks in the noise correlation match the expected peaks, meaning that the boat is the main dominating source in the noise correlation. This provides confidence that the estimated signal shown by dots can be used to self-localize the swarm. Parameters: T_r=3s and B_ω = [0.1-1.0 kHz].

the robustness of the solution using simulations in Section 4.6. In fact, the results in Section 4.5 have better outcomes than what would be expected if the error is modeled as a Gaussian process with a standard deviation as high as what is reported in Figure 4.5. This indicates that while the estimates of the time difference of arrival, $\hat{\delta}_{i,j,t}$, do not exactly match the expected time difference of arrival, $\delta_{i,j,t}$, they still provide a consistent model that generates an answer similar to what we expect from the "ground truth" locations of the AUEs.

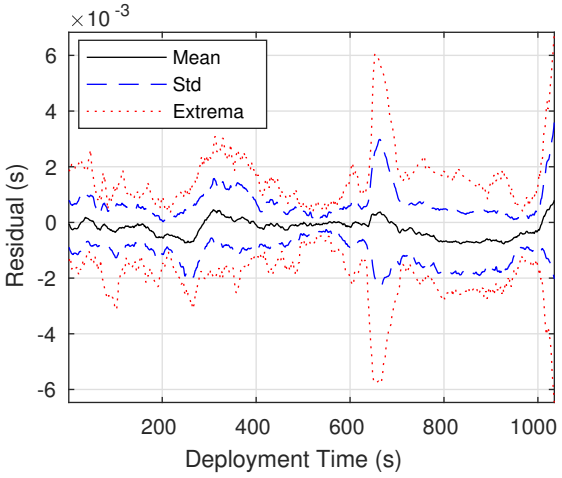

Figure 4.5: Discrepancy between the detected, $\hat{\delta}_{i,j,t}$, and expected noise correlation peaks, $\delta_{i,j,t}$, for all pairs: The mean (black solid), standard deviation (blue dashed) and the maximum and minimum (red dotted) discrepancies are shown as a function of deployment time for all combination of pairs of vehicles. The expected noise correlation peaks do not perfectly match the measured correlation peaks, and there is at least one pair that is off by more than 1ms at every deployment time. There is a spike close to deployment time 700s where many of the noise correlations jumped to another peak, shown in Figure 4.4 (b). These discrepancies may adversely affect the comparison between the estimated relative trajectories from the boat noise and the measured ones from the high frequency pinging system.

4.5 Results

The estimation procedure described by Equations 4.6 and 4.8 was applied to the data described in Figure 4.4 and Figure 4.5. Examples of two solutions with different center times, $t_o = 160$ and $t_o = 945$, and a time window of 30s, (i.e. $t_w = 15$s) are shown in Figure 4.6. The left panel of Figure 4.6 compares the estimate of the vehicle geometry (shown by '*' markers) to the ground truth estimate (shown by 'o' markers for the center time). The estimated ground truth motion of each AUE around the center time is shown by black dots. Since we only estimate the geometry in a relative sense, the comparison is given after an optimal rigid transformation is computed between the estimates using the algorithm described in Arun et al.[135]. The root mean square error (RMSE) in meters is reported after this fit is applied. The results show good agreement, on the order of one or two meters for the two cases shown. The results show promise for the method proposed in Section 4.2, without any knowledge of the position of the noise source,

or any prior information on the relative geometry of the vehicles, we are able to estimate the relative geometry, $\overline{\mathbf{R}}_{t_o}$, as well as the angle of arrivals of the noise source $\alpha_{\mathbf{to},\mathbf{t_w}}$.

Another way to evaluate the estimation is to compare the estimated angles of arrival with the true angles of arrival. The RMSE is reported, in degrees, for the discrepancy between the estimated angles of arrival and the known angles of arrival from the GPS measurements of the boat. This RMSE is reported after the rotation from the rigid transformation was applied to match coordinate frames. We can see that the angle estimates match well to what they are expected to be, with an RMSE of 2.9 degrees and 0.48 degrees for the top and bottom figures, respectively. In this frame of reference, the estimation recovers the geometric interpretation of the soundscape, which can be useful for passive sensing in ocean environments.

It is important to put the results in Figure 4.6 into the context of Figure 4.5, which showed that the detection results do not perfectly agree with their expectation. In fact, Figure 4.5 shows that a 2ms error between the expected time value of the noise correlation function and the measured time value is not uncommon. This could translate to an RMSE on the order of a few meters. The source of the discrepancies in Figure 4.5 are difficult to understand. As previously mentioned, the positions of the AUEs from the high frequency pinging system are also estimates of the AUE positions and are expected to have an accuracy around 1 meter. A static sound speed of $c = 1500m/s$ is assumed in the calculation of Equations 4.3 and 4.8 and this may have a slight effect on the accuracy. In reality, most of the error probably results from the small angles that are used for the estimation. We can see that the error is perpendicular to the angle of arrivals, especially in the bottom panel of Figure 4.6. This is the direction that is the most difficult to resolve given the angle of arrivals of that deployment time.

The estimation results of Equation 4.8 are not similar across the entire deployment time. Figure 4.7 shows a time ($t_o = 110$s) that resulted in a solution that was not close to the geometry of the swarm. In the top solution of Figure 4.7 there is greater variation in the true angles of arrival than what we see in Figure 4.6. Despite this extra time window (here $t_w = 30s$) and azimuthal

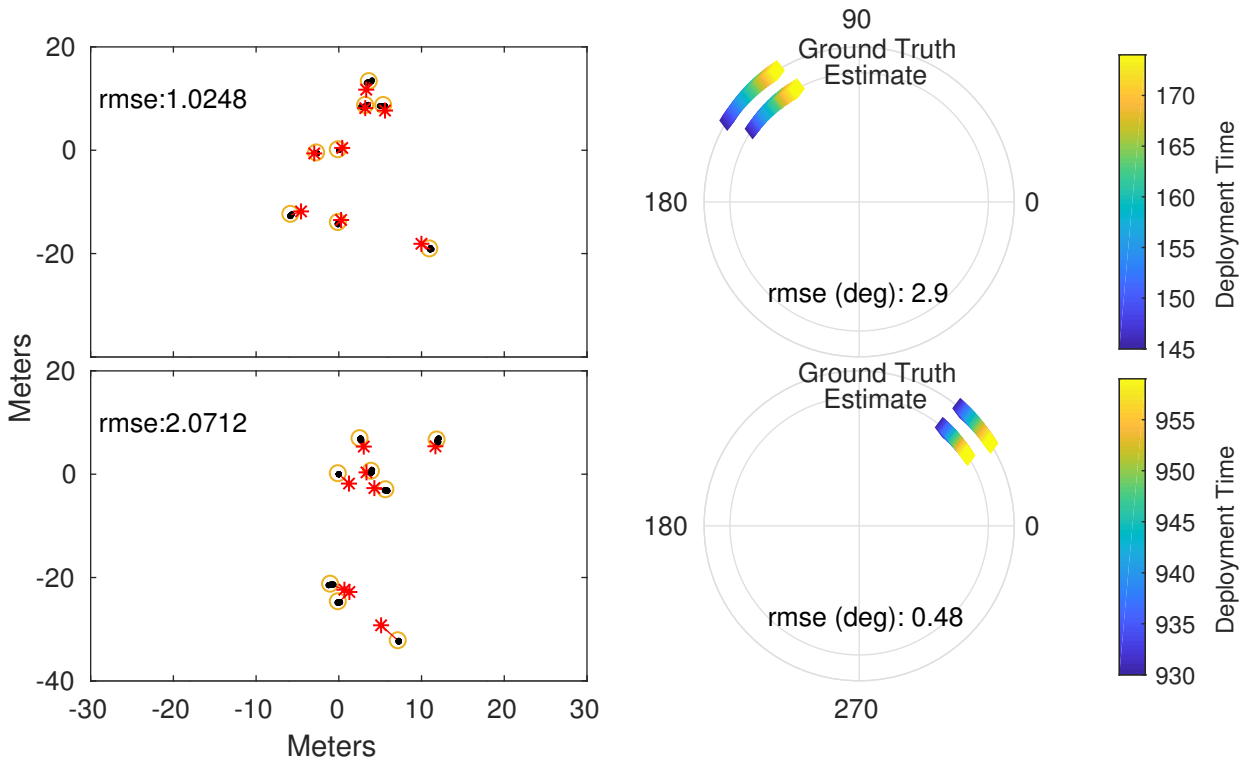

Figure 4.6: Example Solutions: The solutions to Equation 4.8 are shown for two different deployment times, $t_o = 160s$ and $t_o = 945s$. In the left panels, small black dots represent the estimated position of each vehicle for each second in the chosen 30 second window ($t_w = 15s$), the positions at the center time, t_o, are marked by a 'o'. The swarm geometry estimated from the noise is shown by '*' after an optimal translation and rotation is applied to put the estimate in the same coordinate frame as the ground truth estimates. The root mean square error in meters is reported between the position at the center time and the estimate after the fit is performed. On the right panels, the angles of arrival are reported for both the ground truth (on the outside ring) and the angle of arrival estimates (on the inside ring) after the optimal rotation is applied. The RMSE error is shown, in degrees, for the angles at each time step. From these two examples, we can see that the method is accurately able to estimate the relative geometry of each vehicle and the angle of arrival of the dominant source.

variation, the solution provided in the top panel of Figure 4.7 is not a good estimate, both in terms of estimating the relative geometry of the swarm and also the angle of arrival of the noise. However, as we increase the window length to add more variation to the angle of arrivals, we eventually get a solution that is more comparable to the solutions reported in Figure 4.6. This example shows that increasing the angle of arrivals provides more stability even if it increases the total magnitude of the vehicle motion during the time window (the vehicle motion for the time

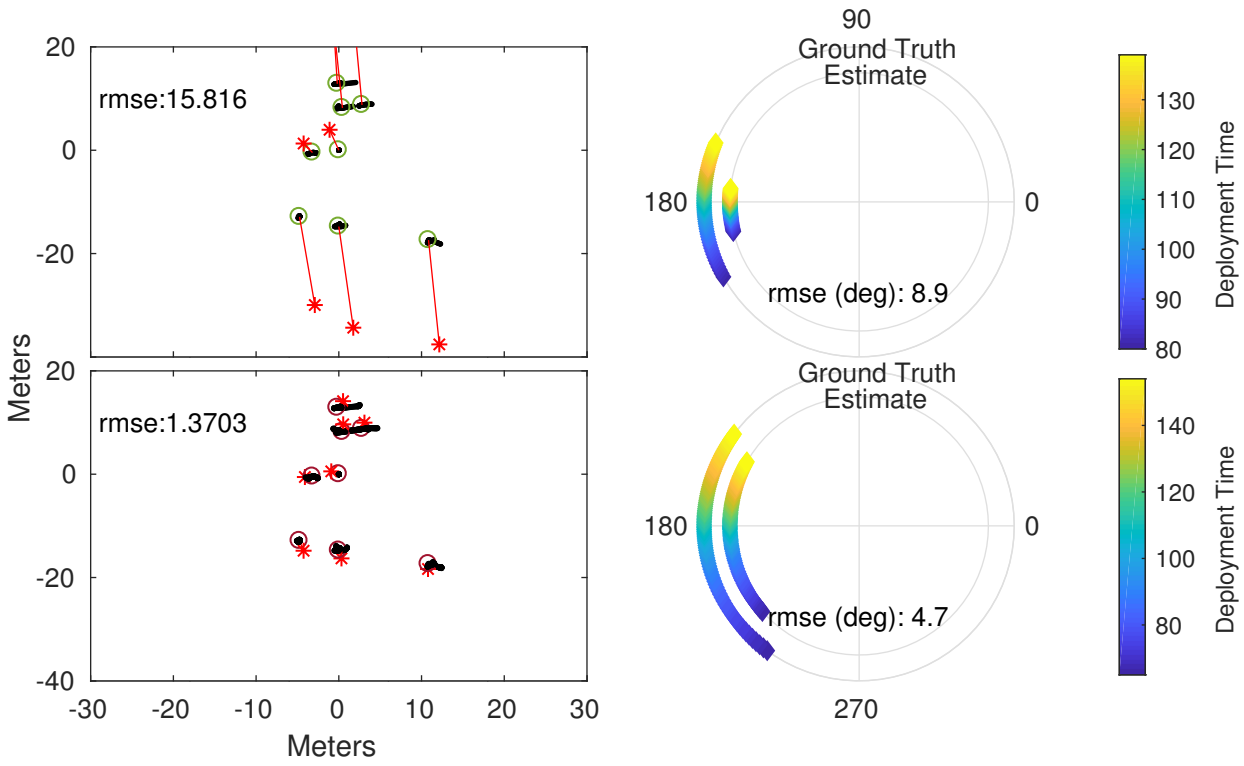

Figure 4.7: Challenges: Similar plots to Figure 4.6 but here the time window length is varied (top $t_w = 30$, bottom $t_w = 45s$) for the same deployment time, $t_o = 110s$. This figure demonstrates that some angles of arrival are more challenging to the estimation procedure. The estimation is poor in the top panel using a 60 second time window. By increasing the time window to 90 seconds ($t_o = 45s$), we can achieve the levels of accuracy that were reported in Figure 4.6. More insight into the interplay between the angle of arrival, vehicle motion, and vehicle geometry is investigated through simulations in Section 4.6.

window is shown by black dots). It also highlights the difficulty of understanding the interplay of all the different factors that could influence the performance the estimation. These factors are examined through simulations in the next section.

4.6 Simulations

4.6.1 Simulation description

This section aims to decouple some of the effects that influence the results shown in Figure 4.6 and 4.7 through simulations. Simulations are first shown for two different vehicle geometries

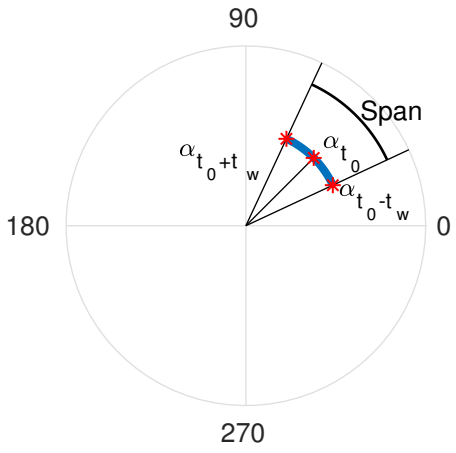

Figure 4.8: Simulation Explanation: A visual definition of the parameters used in the simulations (Figure 4.9 and 4.10). The span dictates the magnitude of the angles used, while α_{t_0} describes the orientation of angles in relation to the vehicles.

to demonstrate that the vehicle geometry influences the solution and are later shown for random realizations of the vehicle geometry to show that some results generalize to many geometries.

The first set of simulations use two different geometries for the swarm, $\mathbf{R}$, which are shown in Figures 4.9 (a) and (d). In order to understand the effects of vehicle motion on our solution, simulations when the stationary assumption is exact ($\mathbf{R}_{t_o} = \mathbf{R}_{t_o,t_w}$) are compared to simulations when the stationary assumption is only approximate ($\mathbf{R}_{t_o} \approx \mathbf{R}_{t_o,t_w}$). In Figures 4.9 (a) and (d) the total movement for the vehicles is shown by small black dots for each of 60 time instances (an arbitrarily number of time instances, but similar to the examples shown in Figures 4.7). The center time is chosen as the position of the vehicles when the stationary assumption is exact, which is also considered the ground truth position when the vehicles have motion. Figures 4.9 (b) and (e) show results when there is movement and (c) and (f) show times when the vehicles are completely stationary. 60 source positions, S_t, are defined in polar coordinates with $r = 1000m$ and various values of $\alpha_{\mathbf{t_0},\mathbf{t_w}}$.

We define $\alpha_{\mathbf{t_0},\mathbf{t_w}}$ as 60 source positions evenly spaced in angle from $\alpha_{t_0-t_w}$ to $\alpha_{t_0+t_w}$, and describe $\alpha_{\mathbf{t_0},\mathbf{t_w}}$ using two parameters: the center angle, α_{t_0}, and the 'span' of the angles, ($\alpha_{t_0+t_w} - \alpha_{t_0-t_w}$). Figure 4.8 shows a visual explanation of these values. The span refers to the total

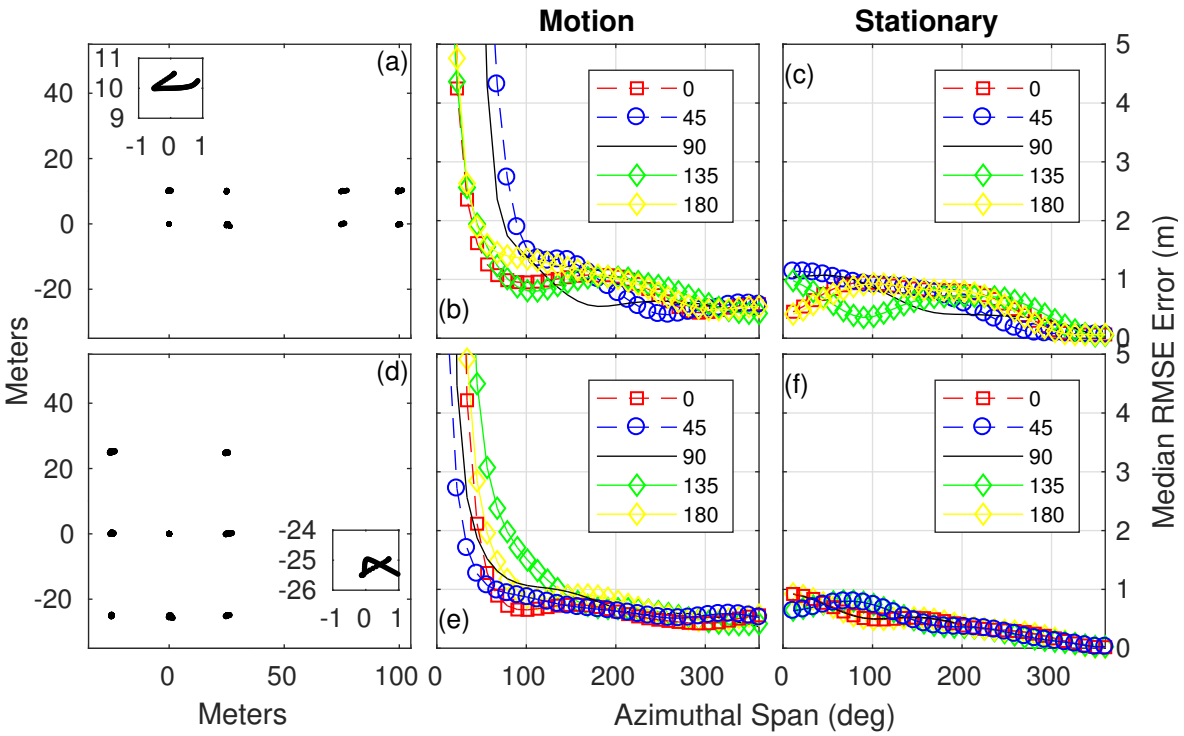

Figure 4.9: Simulations are provided for two different swarm geometries, (a) and (d). Figures (a) and (d) show the position of the floats for each of 60 seconds with black dots. A close up of one of the vehicles is provided in the inset to show an example of the vehicle movement. The vehicle movement is similar to what is seen in the real data. The simulations are computed for vehicles that are moving, (b) and (e), as well as for vehicles that are stationary, (c) and (f). The stationary case uses the center time of the defined trajectory for each vehicle. The simulations are computed similar to what was shown in Figures 4.6 and 4.7 where the swarm geometry and angle of arrival are computed using noisy TDOA measurements. In Figures (b-c,e-f), the center angle of arrival, α_{t_o}, for each line is given in the legend. The total variation of the angles is shown on the x-axis and the median RMSE of 200 iterations between the ground truth and the estimated geometry are shown. These simulations show the effects of swarm geometry, motion, and angle of arrival on the total accuracy. More details are described in the text.

azimuthal range, in degrees, and is a measure of the amount of azimuthal information that is given by the source. For example, a span of 40° with a center angle of $\alpha_{t_0} = 0^\circ$ means that the 60 source positions are evenly spaced from -20° to 20°. A large azimuthal span means that there is adequate angles for the estimation problem to be well conditioned. A small azimuthal span means that all of the angles are similar, and the estimation will be challenged. In order to understand the interaction between the swarm geometry, $\mathbf{R}$, and the angles of arrival, $\alpha_{\mathbf{t_0},\mathbf{t_w}}$, the choices of $\alpha_{\mathbf{t_0},\mathbf{t_w}}$ are varied while keeping the relative motion constant. For each of the lines in Figure 4.9 (b-c,e-f)

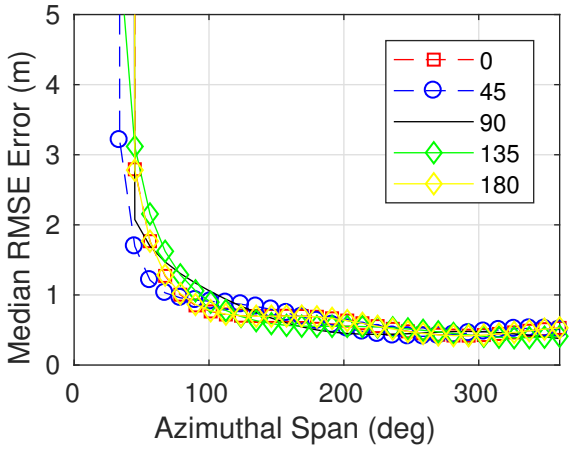

Figure 4.10: Simulation with random positions: In order to show that the trends in Figure 4.9 are typical of most deployments, this figure shows the same set up as Figure 4.9 except that for each of 200 iterations a random geometry is chosen in the box.

the center angle, α_{t_0}, is shown for increments of 45° as indicated by the legend. The center angles are defined with respect to the vehicle geometry in (a) and (d). The x-axis describes the span of angles (i.e. $\alpha_{t_0+t_w} - \alpha_{t_0-t_w}$). With the defined parameterizations of the source positions and vehicle geometry, TDOA values are calculated using Equation 4.1 ($c = 1500 m/s$). The relative geometry of the vehicles was estimated using these TDOA values in Equations 4.3 and 4.8 after zero mean Gaussian noise with a standard deviation of 0.8 meters was added to the $c\Delta$ terms of Equations 4.3 and 4.8. This process was repeated 200 times with different noise realizations, and the median results are shown.

4.6.2 Simulation analysis

There are some emerging trends from the simulation results. The first trend is that movement degrades the self-localization estimate. This is especially apparent when the span of $\alpha_{\mathbf{t_0},\mathbf{t_w}}$ is small. The median RMSE values of the solution are large when there is small azimuthal variation and movement (Figures 4.9 (b) and (e)), but are small when the vehicles are stationary, (c) and (f). More subtly, the stationary cases always have a smaller median value of the RMSE than the cases with movement. This result is expected, as the model assumes that the relative geometry of the

vehicles is stationary.

We can also see that the angle of arrival with respect to the geometry of the swarm plays a role in the results from the different curves shown in Figure 4.9. Some of the curves have lower RMSE for the same azimuthal span (compare $\alpha_{t_0} = 0^\circ$ which is better than $\alpha_{t_0} = 45^\circ$ in Figure 4.9 (b)). Again, this is especially apparent in the simulations with movement, where we can see that there is a separation in the median RMSE curves as a function of the center angle of arrival. The angle of arrivals that result in the best RMSE curve are not consistent across the different swarm geometries. Additionally, the RMSE curve is also not symmetric in angles even though the swarm geometries are symmetric. This last fact is likely due to the small scale relative motion in the vehicles, which is constant across the noise realizations and geometries - the relative motion favors some angles of arrival more than others.

In order to show that the trends of Figure 4.9 (b) and (e) generalize to geometries that are not shown, Figure 4.10 is given. Figure 4.10 shows similar information to Figure 4.9 (b) and (e) but this time the vehicle positions are chosen randomly in a box ($0 < x < 50, 0 < y < 50$). For each realization of the vehicle geometry, the same motion vectors and noise levels are assigned that were given in Figure 4.9 (a) and (d). Figure 4.10 shows that the trends are similar to what was reported in Figure 4.9 except the effects of the center angle α_{t_0} are not as apparent after averaging over the different geometries.

4.6.3 Comparing the simulations and data

The simulations of Figure 4.9 support the findings that were presented in the data analysis (Figures 4.6 and 4.7). Figure 4.9 (b) and (e) show that when the relative geometry of the vehicles is not stationary, there are certain angles of arrival that produce worse results (e.g. $\alpha_{t_o} = 45^o$ for Figure (b) and $\alpha_{t_o} = 135^o$ for Figure (e)). The angles that perform worse are dictated by the relative geometry of the swarm and the motion of the vehicles. This result is also seen in the experimental results. Figure 4.7 shows an example of a challenged angles of arrival that are

coming from approximately 180°. For these angles of arrival, the solution required a larger span of angles to obtain a good solution than the solutions in Figure 4.6. Conversely, the simulations show that for some center angles, only a small span of angles is needed to consistently recover a solution (e.g. 0° for both plots), which is consistent with the data presented in Figure 4.6. The simulations also show that when there is no movement between the vehicles, small angles can be used to successfully recover the swarm geometry. Inter-vehicle movement could be a factor in the difference between the results presented in Figure 4.6 and Figure 4.7, especially since the vehicle movement in the simulations is not large (it does not stray further than one meter away from the position at the center time).

4.7 Conclusion

This work demonstrated a self-localization solution for moving vehicles using low frequency noise from an anisotropic ambient noise field. The movement between vehicles changes the propagation path and forces the solution to consider short time scales. A solution was proposed that was able to jointly estimate the array geometry as well as the angle of arrival of the noise sources using short time windows. Experimental validation was provided using a moving boat with known location, which allowed us to verify that our solution was able to produce accurate direction of arrival estimates without any *a priori* knowledge of the source location. Simulations were provided that showed that small azimuthal variation, coupled with vehicle movement degrades the solution. The stability of the solution was also dependent on the interaction between array geometry and the angles of arrival of the source; different center angles resulted in different accuracies for the solution in both data and simulations. Ultimately, the self-localization solution provides a new and passive way to track the movement of a swarm of underwater vehicles, providing a spatial estimate to the data they collect. This is useful for many coordinated surveys, especially those that require high endurance missions.

4.8 Acknowledgments

Chapter 4, in part, has been submitted for publication of the material as it may appear in The Journal of the Acoustical Society of America: P. Naughton, P. Roux, C. Schurgers, R. Kastner, J. S. Jaffe, and P. L. D. Roberts, (2016) "Self-localization of a mobile swarm using noise correlations with local sources of opportunity". The dissertation author was the primary investigator and author of this paper.

Chapter 5

Self-localization of a deforming swarm of underwater vehicles using impulsive sound sources of opportunity

5.1 Introduction

In the last chapter, a self-localization procedure was proposed that used time difference of arrival (TDOA) constraints from a local source of opportunity to recover the geometry of independent underwater vehicles. Recovering TDOA measurements between vehicles was relatively easy because of the loud dominating source present in the soundscape. However, recovering these constraints can become difficult when considering moving vehicles for two reasons. The first reason is that it can be difficult to match the arrival time of an individual signal across the recordings of each vehicle. This happens when signals have similar structure and occur frequently in time. We saw an example of this in Figure 4.4. As another example, consider trying to find matching pairs between two identical sound sources across two vehicles. If the sounds are transmitted far apart in time, then the signals with the closest arrival times on the recordings of

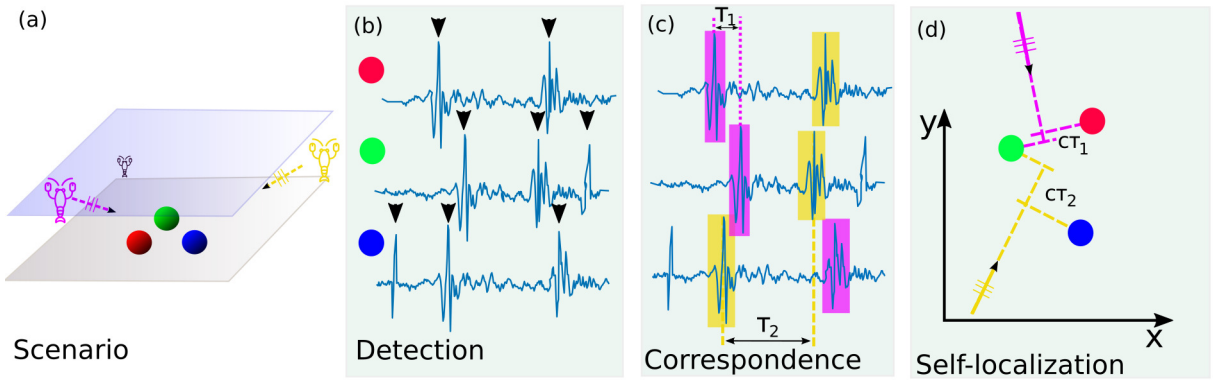

Figure 5.1: Overview of our solution: (a) We are trying to track the changing geometry of moving vehicles (shown by red, green, and blue spheres) using ambient sounds in the ocean, such as the pictured snapping shrimp. (b) Our first step is to determine accurate arrival times for these signals on each recording (shown by the black arrows); we discuss our approach in Section 5.3.1. (c) Our next step is to determine the appropriate correspondence between the detected signals. This is a difficult problem that we examine in Section 5.2.1, and discuss our implementation and challenges in Section 5.3.2 and 5.3.2, respectively. (d) The final step is to perform the self-localization estimate using the difference of arrival times (τ_1 and τ_2 from panel (c)). This is discussed in Section 5.2.2 and implementation details are described in Section 5.3.3.

each vehicle yield the matching pairs. However, if they transmit at the same time, it is difficult to determine the matching pairs without knowing the geometry between the vehicles and sound sources. This problem, typically known as a correspondence problem, must be solved before further processing can proceed. A second challenge is that not all vehicles recover the same information. A self-localization solution must incorporate a sparse set of constraints to estimate the whole swarm geometry.

Section 5.5. This approach only requires one hydrophone per vehicle, which is less expensive than solutions using real time clocks and less power intensive than applications that require communication. It can be used in conjunction with the passive self-localization solution presented in this chapter to self-localize vehicles that have imprecise clocks.

5.2 Problem description and background

We consider a set of moving vehicles (shown by red, green, and blue spheres in Figure 5.1 (a)) that are recording ambient sounds in the water column. Most of the sounds are from the far field, and some sounds are salient in each recording. Our goal is to track the deforming geometry of the vehicles using the salient signals in the ambient acoustic noise, such as the shrimp pictured in Figure 5.1 (a). It is easy to determine accurate arrival times for salient signals on each recording (shown by the black arrows in Figure 5.1 (b)); we discuss our approach in Section 5.3.1. However, it is difficult to determine the correspondence between signals across vehicles. Consider the two shrimp in Figure 5.1 (a), a yellow one and a purple one. If they produce similar sounds, at similar times, from dissimilar positions, the order of arrival of each may be recorded differently at each vehicle. This is shown in Figure 5.1 (c), where the order of arrival times between the purple and yellow shrimp is different for one of the vehicles. This makes picking out the same signal in all recordings difficult. Additionally, choosing correspondences based on metrics such as correlation fail because of the similarity between the signals[113]. We propose a geometric argument to estimate correspondences between recordings in Section 5.2.1, and discuss our implementation and challenges in Section 5.3.2 and 5.3.2, respectively. The difference of arrival times (τ_1 and τ_2 from Figure 5.1 (c)) place constraints on the final geometry of the vehicles (they must be separated by $c\tau_1$ and $c\tau_2$ along the direction of travel from the purple and yellow sound sources, respectively, where c is the propagation speed of the signal) shown in Figure 5.1 (d). Not all constraints are shown so the figure does not become overloaded. Using these constraints to estimate the geometry of the vehicles and locations of the noise sources is discussed in Section 5.2.2 and implementation details are described in Section 5.3.3.

We can only estimate the relative geometry because we do not assume global information for either the source or vehicle positions. This means that our solution is only unique up to a rigid rotation, translation, and mirroring. We are concerned about estimating only the 2D position of the

vehicles because the depth of the vehicles can be accurately estimated using a pressure sensor. We denote the 2 dimensional position of vehicle i, $R_i = (R_{xi}, R_{yi})$, and the two-dimensional position of the sound source j, $S_j = (S_{xj}, S_{yj})$. Throughout this work we also denote the $N \times 2$ dimensional vector $\mathbf{R} = [R_{x1}, R_{y1}; R_{x2}, R_{y2}; \ldots R_{xN}, R_{yN}]$ which is the collection of the N vehicle positions. We define a $M \times 2$ dimensional vector $\mathbf{S} = [S_{x1}, S_{y1}; S_{x2}, S_{y2}; \ldots S_{xM}, S_{yM}]$ for the positions of M different sound sources. Our estimation makes use of the time each sound, j, is received at each vehicle, i, denoted $t_{i,j}$. Since we do not know the emission time of each sound source, we will work with time difference of arrival (TDOA) measurements to negate the unknown emission time. The TDOA is defined as:

$$\tau_{i,h,j} = \frac{\|R_i - S_j\|_2 - \|R_h - S_j\|_2}{c} \tag{5.1}$$

where $\tau_{i,h,j} = t_{i,j} - t_{h,j}$, $\|\cdot\|_2$ is the Euclidean distance, and c is the speed of sound in water. This gives us a system of equations that we can use to estimate the unknown positions of the swarm, $\mathbf{R}$, as well as the unknown positions of the sound sources, $\mathbf{S}$:

$$\arg\min_{\mathbf{R},\mathbf{S}} \sum_{h=1}^{N} \sum_{i=h+1}^{N} \sum_{j=1}^{M} \upsilon_{h,i,j} \left(\|R_i - S_j\|_2 - \|R_h - S_j\|_2 - c(\tau_{i,h,j}) \right)^2 \tag{5.2}$$

where $\upsilon_{h,i,j}$ is an indicator variable that takes the value of 1 if noise source j is detected by both vehicles i and h and 0 if such a correspondence cannot be made. This minimization problem is sensitive to two main pitfalls that need to be addressed to find an accurate estimate for $\mathbf{R}$, which are described throughout the rest of this section. First, correspondences must be correct for this minimization function to minimize a meaningful objective function. Finding appropriate values for both $\tau_{i,h,j}$ and $\upsilon_{h,i,j}$ by looking at minimal sets of vehicles is described in Section 5.2.1. Another difficulty with equation 5.2 is that it is prone to get stuck in local minima. Developing a robust solution to solving equation 5.2 is discussed in Section 5.2.2.

5.2.1 Correspondence solution

The first challenge is to determine correct correspondences of noise sources across the vehicles of the swarm. This challenge is illustrated in Figure 5.1 (a) and (c). If two noise events happening at similar times are recorded by two vehicles, they may not be recorded in the same order due to the geometry of the vehicles and noise sources. Furthermore, it is known that correlation based metrics for determining the correspondence between impulsive noises fail because different impulsive noises have similar signal characteristics [113]. This problem needs to be solved to estimate an accurate geometry of our swarm; incorrect correspondences will degrade our geometry estimation.

In order to solve this correspondence problem, we make three simplifying assumptions. The first assumption is on the maximum inter vehicle distance within the swarm, which can either be defined in a deployment, estimated or set conservatively. This assumption limits the number of possible correspondences by defining a window, δ_w, in time that we allow correspondences to be accepted from. The second assumption is that we can model our localization problem as a 2D problem, meaning that we are considering vehicles that are mostly planar. Generally, the depth of the vehicles can be cheaply and accurately determined in the ocean by measuring pressure. Our assumption here also extends to parametrizing the noise source locations. This assumption means that the horizontal distance between a vehicle and sound source is a good approximation for the total distance traveled. In other words, the vertical distance between the sound source and vehicle is far smaller than the horizontal distance. The final assumption is that the noise locations are in the far field. This is related to the previous assumption. The second and third assumptions allow us to parameterize the noise source locations by an angle instead of a three-dimensional coordinate. The validity of this assumption is based on the ratio between the inter-vehicle spacing and the distances between the vehicles and noise sources. Typically, for the far field assumption to hold, the vehicle separation should be 2-4 times smaller than the distance of the noise sources to the swarm [129, 127]. This assumption is justified for shallow water cases where the bathymetry

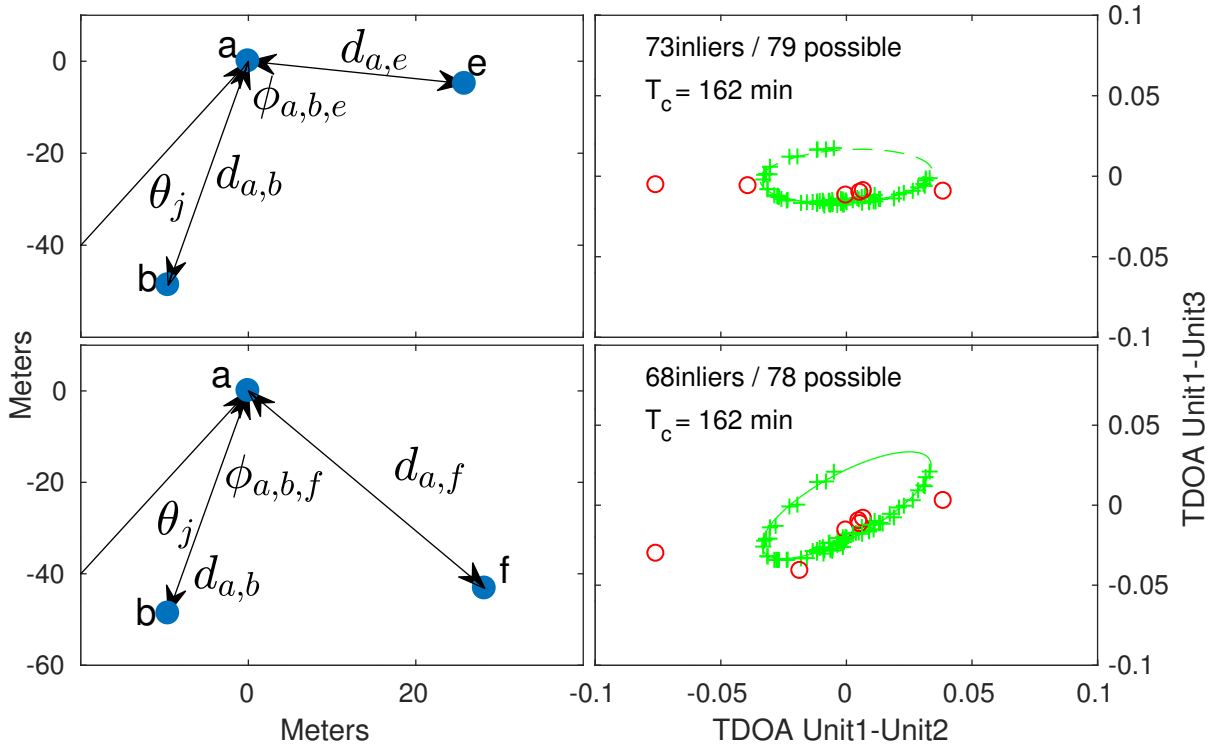

Figure 5.2: Finding correct correspondences: (left) the relative geometry of two different vehicle triplets at $T_c = 162$ min. (right) the corresponding output to the RANSAC framework for the geometry on the left. We can see that the geometry of the vehicles puts constraints on the TDOA measurements between the arrivals of the impulses. For correct correspondences, the TDOA measurements lie on an ellipse parameterized by the angle of arrival, θ_j, of each noise source. Comparing the top and bottom panels shows that changing the geometry of the vehicles changes the geometry of the ellipse. We can also see that the RANSAC framework does a good job at finding an optimal inlier set (green '+'), while rejecting outliers (red 'o'). The correspondences that lie on each ellipse are accepted for further processing while the correspondences in red are labeled as incorrect and rejected. Some outliers are not shown to focus more on the ellipse ($\delta_w = 0.2s$, $\varepsilon_t = 3ms$).

is smaller than the horizontal distance that the signal must travel. These assumptions minimize the number of vehicles that need to be considered to solve the correspondence problem. This minimal set is discussed in the following paragraphs and is important because the correspondence problem quickly becomes intractable as we consider more vehicles. The addition of another vehicle increases the combinations of possible correspondences that must be considered. While it is possible to consider a solution that is not planar, it would require considering more vehicles at a time in our correspondence solution. Our results in Section 5.4 show experimental justification

for these assumptions.

A key insight in solving the correspondence problem is that there are constraints on the TDOA measurements that are imposed by the geometry of the vehicles. With the far field assumption we can approach the correspondence problem by looking at two TDOA measurements between three vehicles. Under a far field assumption these TDOA measurements must lie on an ellipse [129]. We can see this mathematically by examining the TDOA between two pairs of vehicles: pair a and e, and pair a and b (this configuration is shown in top left panel of Figure 5.2, which will be discussed in full in Section 5.3.2):

$$\tau_{a,b,j} = \frac{d_{a,b}}{c} cos(\theta_j) + o_{a,b} \tag{5.3}$$

$$\tau_{a,e,j} = \frac{d_{a,e}}{c} cos(\theta_j + \phi_{a,b,e}) + o_{a,e} \tag{5.4}$$

Here $d_{a,b}$ is the distance between vehicle b and a. θ_j is the angle of arrival of sound source j in relation to the line intersecting vehicles a and b and $\theta_j + \phi_{a,b,e}$ is the angle of arrival in relation to vehicles a and e. Clock offsets between the vehicles are also included as $o_{a,e}$ and $o_{a,b}$. While discussing our self-localization solution, the clocks are assumed to be synchronized (Section 5.3 describes our justification of the clock synchronization). This means that $o_{a,e} = o_{a,b} = 0$. We discuss a self-synchronization solution in Section 5.5 where the offsets are estimated.

$\tau_{a,b}$ and $\tau_{a,e}$ trace an ellipse, centered at $(\tau_{a,b} = o_{a,b}, \tau_{a,e} = o_{a,e})$ parameterized by the angle of arrival of the sound source, θ_j. There are two interesting things about this result. The first is that with the knowledge of the ellipse parameters we can estimate the 2D geometry of the 3 vehicle swarm (e.g. parameters $d_{a,e}$, $d_{a,b}$ and $\phi_{a,b,e}$). Likewise, given the geometry of the swarm we can also estimate the parameters of the ellipse. Most importantly, given that the sound sources are coming from the far field, they must lie on this ellipse[129].

Our approach is to evaluate correspondences three vehicles at a time. We determine the

correct ones by computing the two TDOA measurements of each correspondence and examining which TDOA values lie on the ellipse. Correspondences that result in a point that lies off of the ellipse we reject as incorrect and we retain the correspondences on the ellipse for further processing. The problem with this approach is that the ellipse is not known. Unlike Wendeberg et al. [129] which assumed perfect correspondence, we are using the ellipse result to determine the correct correspondences without knowledge of what the ellipse should be. We do this by determining the ellipse with the most "votes" in a Random Sampling and Consensus (RANSAC) framework [136]. The RANSAC algorithm estimates an ellipse using a minimal set, N_{min} (in our case it takes $N_{min} = 3$ points to estimate an ellipse with a known center, and $N_{min} = 5$ points if the clocks are not synchronized and we must estimate the clock offsets), counting how many candidate correspondences "agree" with that minimal set and iterating until we find the set with the most "agreement". Typically, the set with the most agreement is called the "inlier set". Here agreement is defined as being within a threshold distance from the ellipse, ε_t. The RANSAC algorithm is summarized below for a collection of candidate correspondences, $\chi_{a,b,e}$.

Algorithm 1 RANSAC

Input: $\chi_{a,b,e}$ collection of possible correspondences
Output: True correspondences S_{true}, parameterization of ellipse E_{final}

1: $maxInliers \leftarrow 0$
2: **for** every combination of N_{min} correspondences (3 for synchronized clocks, or 5 for unsynchronized clocks) in $\chi_{a,b,e}$, or until max iterations **do**
3: $S_{rand} \leftarrow$ random N_{min} points from $\chi_{a,b,e}$
4: $E \leftarrow$ parameterization of ellipse estimated from S_{rand}
5: $N \leftarrow$ number of entries in S within ε_t from E.
6: **if** $N > maxInliers$ **then**
7: $maxInliers \leftarrow N$
8: $E_{final} \leftarrow E$
9: $S_{true} \leftarrow$ points in S that are within ε_t from E
10: **end if**
11: **end for**
12: $E_{final} \leftarrow$ ellipse estimated from final S_{true} in line 9

The output of the RANSAC algorithm gives us information on the correct correspondences

which we can use to estimate the geometry of the vehicles.

5.2.2 Estimation

Once we have a collection of arrival times for each vehicle and we know which arrivals from one vehicle correspond to the arrivals in another vehicle we can use this information to estimate a geometry for the swarm.

Given a swarm, $\mathbf{R}_\rho$, and noise sources, $\mathbf{S}_\sigma$, where every vehicle has a known time of arrival for each noise source, we define two new variables:

$$\overline{\mathbf{R}}_\rho = \begin{bmatrix} x_{\rho_2} - x_{\rho_1} & y_{\rho_2} - y_{\rho_1} \\ x_{\rho_3} - x_{\rho_1} & y_{\rho_3} - y_{\rho_1} \\ \vdots & \vdots \\ x_{\rho_n} - x_{\rho_1} & y_{\rho_n} - y_{\rho_1} \end{bmatrix}$$

$$\overline{\mathbf{D}}_{\rho,\sigma} = \begin{bmatrix} t_{\rho_2,\sigma_1} - t_{\rho_1,\sigma_1} & t_{\rho_2,\sigma_2} - t_{\rho_1,\sigma_2} & \cdots & t_{\rho_2,\sigma_m} - t_{\rho_1,\sigma_m} \\ t_{\rho_3,\sigma_1} - t_{\rho_1,\sigma_1} & t_{\rho_3,\sigma_2} - t_{\rho_1,\sigma_2} & \cdots & t_{\rho_2,\sigma_m} - t_{\rho_1,\sigma_m} \\ \vdots & \vdots & \ddots & \vdots \\ t_{\rho_n,\sigma_1} - t_{\rho_1,\sigma_1} & t_{\rho_n,\sigma_2} - t_{\rho_1,\sigma_2} & \cdots & t_{\rho_n,\sigma_m} - t_{\rho_1,\sigma_m} \end{bmatrix}$$

Matrix $\overline{\mathbf{R}}_\rho$ is a relative distance matrix to vehicle R_{ρ_1} and matrix $\overline{\mathbf{D}}_{\rho,\sigma}$ is a matrix of the TDOA of each vehicle compared to vehicle R_{ρ_1}. Again, we are not able to estimate the global location of the vehicles because we do not know the locations of any sound sources or vehicles. Making our parameterization relative to the location of one unit does not change our result because we are only able to estimate the geometry of the swarm in a relative sense. With this in mind, we can arbitrarily set $R_{\rho_1} = (0,0)$ so that $\overline{\mathbf{R}}_\rho = \mathbf{R}_\rho$ and drop the $\overline{\mathbf{R}}_\rho$ notation.

With the above parameterizations, the angle of arrival of each sound source and the relative

geometry of the vehicles can be estimated by minimizing the following equation [127]:

$$\arg\min_{\mathbf{R}_\rho, \Lambda_\sigma} \|\mathbf{R}_\rho \Lambda_\sigma - c\overline{\mathbf{D}}_{\rho,\sigma}\|^2 \tag{5.5}$$

Here Λ_σ is a parameterization of the matrix $\mathbf{S}$ in angles instead of 2D coordinates:

$$\Lambda_\sigma = \begin{bmatrix} cos(\alpha_{\sigma_1}) & cos(\alpha_{\sigma_2}) & \dots & cos(\alpha_{\sigma_m}) \\ sin(\alpha_{\sigma_1}) & sin(\alpha_{\sigma_2}) & \dots & sin(\alpha_{\sigma_m}) \end{bmatrix}$$

Equation 5.5 can be trivially solved in an affine space using a singular value decomposition of the matrix $\overline{\mathbf{D}}_{\rho,\sigma}$, meaning that the constraint $cos(\alpha_i)^2 + sin(\alpha_i)^2 = 1$ is not upheld. We refer the reader to Thrun [127] for these details. A Euclidean solution can then be estimated by solving a nonlinear optimization problem with P^2 variables, where P is the dimension of the space. In our case P is 2. This number is independent of the number of vehicles, N, and noise sources, M, that we are considering.

In the general case, every vehicle in $\mathbf{R}_\rho$ may not know every arrival time in $\mathbf{S}_\sigma$. In this case, our approach is to estimate different subsets of the swarm in which all vehicles do have an arrival time for every sound, then merge and refine these results using equation 5.2.

5.3 Self-localization Experiment and Implementation

Our experiment consisted of 17 subsurface floats, which we call Autonomous Underwater Explorers (AUEs). Jaffe et al. [8] and Naughton et al. [11] have more details about the deployment, including both the AUE specifications and environmental considerations. Some key points are summarized here. The AUEs were programmed to hold a depth of 10 meters and drifted with the subsea currents for 5 hours while collecting acoustic data. The vehicles were spaced such that neighboring vehicles were within 10s of meters of each other, with the whole swarm spanning approximately 300m in each direction. The trajectories of the AUEs were recovered using 5

surface pingers that collected timing and positioning information from onboard GPS receivers and transmitted linearly frequency modulated chirps (7-15kHz) that were received by the AUEs. These trajectories are used to validate our estimates of the swarm geometry from the impulsive noises. In addition to the chirps from the surface pingers, the AUEs recorded impulsive noise from snapping shrimp. The shrimp produced a signal from an unknown location, but are known to spend most of their time on the seafloor, which was approximately 50 meters below the oceans surface during this deployment. The clocks of the AUEs are corrected by comparing GPS measurements at the beginning and end of the AUE deployments from onboard GPS sensors. The acoustic environment of the AUEs during the deployment was not ideal for the propagation of the noises around the elements of the AUE swarm. The AUE's depth target, 10m, was on the boundary of the mixed layer and a steep thermocline, providing a strong downward refracting profile. The shallow depth target allowed surface reflections of the noise field to interfere with the direct path (i.e. the environment is a Lloyd's mirror) cf Figure 3 in Naughton et al. [11]. Therefore, shadow zones are expected between the vehicles depending on vehicle distance. Between the shadow zones and the hydrophone dropouts due to the buoyancy adjustment (dropouts account for about a ten percent loss of audio), we cannot expect all vehicles to hear all sources.

To estimate the changing geometry of the vehicles we assume that the geometry of the swarm is unchanged for a period of time, $\delta_s = 4$ min, centered at a center time, T_c. This time period was chosen empirically and is based on the expected relative movement between the vehicles. Notice that the relative movement is different from the absolute motion of the vehicles. For the slowly drifting vehicles in our experiment, the relative movement between the vehicles, driven by variations in the current experienced at each vehicle, is smaller than the rate that the swarm moves through the ocean, which is driven by the mean flow of the current. This allows δ_s to be large compared to the average current speed. The center time of each window, T_c, is chosen so that each window overlaps by $\delta_s/2$ with the next and previous window. Our approach is to solve for the relative geometry independently for each time window by performing the following

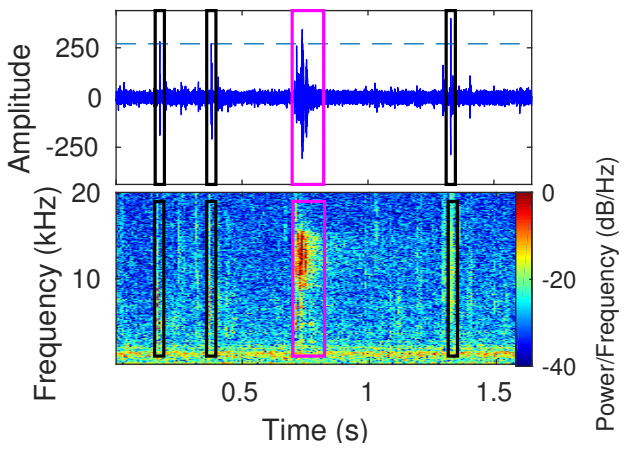

Figure 5.3: Examples of The signals used in our analysis: The frequency content and time series are plotted to show the two types of signals utilized in this work. The high energy signal enclosed by a pink box is transmitted by 5 different pingers with known GPS positions and transmit times. We use these signals to estimate the ground truth geometry of our moving swarm. The high amplitude impulses enclosed by black boxes are assumed to be snapping shrimp whose positions and emission time are unknown. We estimate the geometry of our swarm using these noise sources and compare our solution to the more controlled measurements from the pingers. Also, shown is the amplitude threshold that the snap must exceed to be considered for further processing.

steps (which are outlined in Figure 5.1 (b-d)):

1. For each T_c in the deployment, select a window of data δ_s seconds long, centered at T_c, from each vehicle.

2. Detect the impulsive sources of opportunity for each vehicle (Figure 5.1 (b)).

3. Determine the correct correspondences between vehicles using the developments of Section 5.2.1 (Figure 5.1 (c)).

4. Jointly solve for the vehicle positions, $\mathbf{R}$, and noise positions, $\mathbf{S}$, using the development of Section 5.2.2 (Figure 5.1 (d)).

5.3.1 Detection

The first step in our pipeline is to extract the arrival time of signals that could be detected across vehicles of the swarm. For this demonstration we focus on impulsive noise from snapping shrimp, a common signal in the soundscape of warm coastal waters [137]. There are three main considerations for our snapping shrimp detector: we only classify an 'impulse' as something we can accurately determine the arrival time of, we do not classify signals from the GPS pingers as snaps (the pings are also a broadband pulse), and we only classify the first arrival of a snap. To ensure the first requirement, we only consider impulses that are larger than 10 standard deviations of the acoustic time series. From these candidate snaps, we examine the spectrum of the signal to make sure that there is adequate energy in the bands 4-17 kHz. In order to make sure that we do not classify pings from the buoys as a snap, we detect the pings using a matched filter and define a dead zone in time around each ping where we do not allow snaps to be accepted. The window of the dead zone is 25 ms before the detected ping and 100 ms after the detected ping. The width of this dead zone was chosen empirically by looking at the received reflections from the pinger and is conservative. Similarly, to make sure that we only detect the first arrival of the impulsive sound we define a dead zone around detected snaps so that we do not allow another snap to be detected 10 ms after one of the snaps. Figure 5.3 shows examples of the impulsive snaps and the surface pingers that we use to estimate the trajectories of the vehicles.

5.3.2 Correspondence solution

The first step in determining the correct correspondences is to determine all the possible correspondences. We start with the vehicle that will be computed in both TDOA measurements, vehicle a. For each detected noise event, $t_{a,j}$, from a we look for corresponding events in both vehicle b and e that are within $\pm\delta_w = 0.2$s away from the arrival time, $t_{a,j}$. There can be more than one noise event that corresponds to $t_{a,j}$ from either vehicle b or e. The outcome of this first

pass is a collection of L different possible correspondences between the three vehicles for the given time window centered at T_c. We call this collection χ_{a,b,e,T_c}. To make the processing easier, we throw away entries in χ_{a,b,e,T_c} where there are ambiguous correspondences. Meaning if there are entries in χ_{a,b,e,T_c} that contradict each other, we discard them.

We then run the candidate correspondences through the RANSAC procedure described in Section 5.2.1. Figure 5.2 shows two examples of the swarm geometry of a vehicle triangle (left) and the corresponding output of our RANSAC algorithm (right). The comparison between the top pair and the bottom pair shows that changing geometry of the vehicles changes the geometry of the ellipse, where each correspondence on the ellipse is related to the angle of arrival of a sound source. The right panels show that the algorithm does a good job accepting the candidate matches (green '+') and rejecting the outlier correspondences (red 'o'). This analysis is performed for all possible combinations of a, b, and e.

Failure Modes

Unfortunately there are configurations when the ellipse estimation degenerates and becomes unstable. The first case is when the far field assumption is violated. For our model, this is a function of two different physical components of the environment: the inter-vehicle distances, $d_{a,b}$ and $d_{a,e}$ as well as the source positions, $\mathbf{S}$. This means our system will only find correspondences for subsets of the swarm where the vehicle triplets are close together. Another failure mode is when the vehicle triplet becomes linear, i.e. $\phi_{a,b,e}$ in equation 5.4 approaches 0. In this case, $\tau_{a,b,j}$ and $\tau_{a,e,j}$ trace a line instead of an ellipse and our estimation will become unstable. We throw away the information χ_{a,b,e,T_c} that produces a degenerate ellipse.

These two failure modes mean that our system only accepts correspondences that are local compared to the size of the swarm. It also means that for certain choices of a, b, and e, χ_{a,b,e,T_c} will not produce a valid ellipse and some vehicle triplets will not have any matching correspondences. Our next step is to combine the information from the good vehicle triplets to

construct correspondences across as much of the swarm as possible.

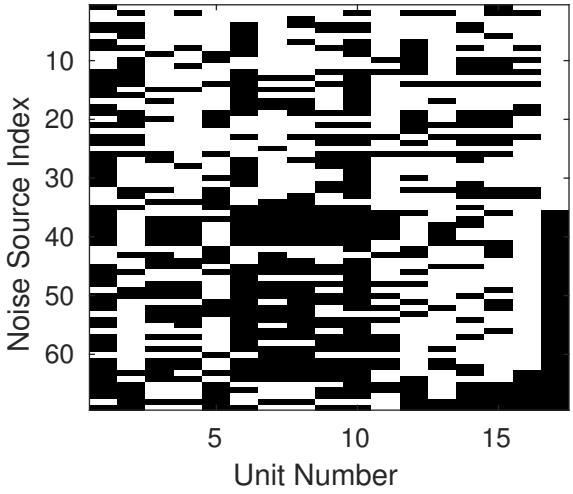

Figure 5.4: An example of the sparsity of the constraints for one time step: The x-axis labels the 17 vehicles and the y-axis the noise indices that were detected across a minimum of 4 vehicles for a specific choice of T_c. A white box indicates that the time of arrival is known for the vehicle on the x-axis and the noise index on the y-axis. These time of arrivals are used to solve our final optimization problem for the location of the vehicles and sound sources. The structure of this problem makes the optimization difficult.

Merging Correspondences

Once we have all the information from the sets of vehicle triplets, we combine the information to look for a full correspondence set across the entire swarm. We do this by going back to correspondence matches between pairs of vehicles instead of the vehicle triplets that we were working with before. Notice that there is redundant information when trying to find correspondences between vehicle pairs when we are using information from vehicle triplets. That is, when considering correspondences between vehicle a and b, the ellipse estimation can fail for many choices of e and we can still have some idea of correct correspondences between a and b from the values of e in χ_{a,b,e,T_c} that produced valid results.

In order to accept a correspondence between a pair of vehicles we ensure that it was accepted in at least 70 percent of the valid ellipses that considered the correspondence, and we ensure that there were at least 2 ellipses that were possible between the pairs. This protects us from an

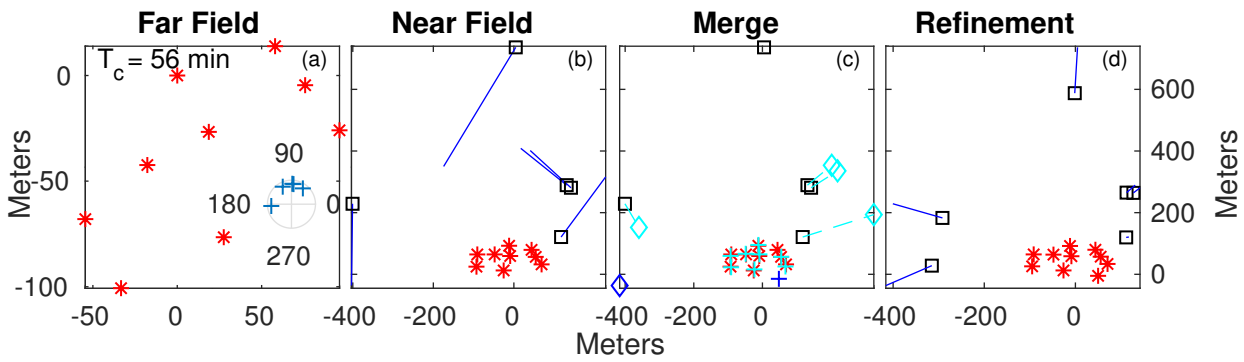

Figure 5.5: Steps of our self-localization solution: (a) A far field approximation is used to solve for an initial geometry of a subset of the swarm, $\mathbf{R}_\rho$, shown by red '*' as well as the angle of arrival of a subset of the noise sources, Λ_σ. The angle of arrival of the noise sources relative to the swarm geometry is shown in the inset. (b) The solution is refined by relaxing the far field approximation. Noise sources that were parameterized by an angle are now parameterized by a 2D position, so we are relaxing the assumption that the noise sources are in the far field. Lines indicate the initial and final locations of each entry in $\mathbf{R}_\rho$ and $\mathbf{S}_\sigma$. (c) A new subset is chosen with different, but overlapping, values of ρ and σ. The steps in (a) are computed and the solution is fit to the previous solution. In (c), vehicles that overlap with the original subset are shown by cyan '+' whereas new ones are a dark blue '+'. Likewise, new sound sources are shown by a dark blue diamond and the corresponding noise sources are shown by a cyan diamond. (d) The initial guess of step (c) is refined with equation 5.2 to get a new, refined value of the swarm geometry and the noise sources. Steps (c)-(d) are iterated until there are no more values of ρ and σ.

incorrect single ellipse. Finally, if there is any contradiction between any pairs of correspondences, we throw away all of those possible correspondences.

Once the pairwise correspondences are made we assume that they are associative so that correspondences can be drawn that were not accepted by the ellipse estimation. This may accept correspondences that could violate the far field assumption. We address this in Section 5.3.3 by allowing noise sources to be parameterized by a two-dimensional coordinate instead of only an angle. Our last constraint is that we throw away noise events that we cannot track across a minimum of 4 vehicles. An example of correspondences for a specific choice of T_c is shown in Figure 5.4, where a white box indicates that the time of arrival is known for the unit indicated on the x-axis and the snap index located on the y-axis. We can see the sparsity of the optimization problem we are left with, a single noise source is rarely tracked across the entire swarm. This

makes finding a global solution to the optimization problem difficult.

5.3.3 Self-localization

After determining the correct correspondences, $t_{i,j}$, we construct a joint estimate of the geometry of the swarm, **R**, and the relative locations of the sound sources, **S**. This involves solving the minimization problem of equation 5.2. We do this by breaking the swarm into subsets, estimating each subset independently, then incrementally piecing subsets together to estimate a solution for the entire swarm. We choose a subset of the swarm $\rho \subseteq \{1,2,...,N\}$ of length k and a subset of the noise sources $\sigma \subseteq \{1,2,...,M\}$ of length l so that every vehicle in ρ has an arrival time for every sound in σ. Subsets are chosen so that $k,l \geq 5$. For each noise source, we choose a collection of 5 noise sources that have the greatest overlap in detected arrival times for the same vehicles (a minimum of 5 vehicles for the subset to be kept). This criterion ensures every valid noise source is included in at least one subset, it also creates redundancy in the collection of subsets. In the next section, we evaluate the stability of different subsets and discard some subsets if they do not meet specific criterion, so this redundancy helps reduce the number of noise sources that are excluded. Examples of two different subsets of noise sources and vehicles that were chosen by these criteria are shown in Figure 5.5. At the end of collecting subsets, we verify that every vehicle is included in at least one subset. If there are vehicles that cannot be included, we decrease the constraint of $k,l \geq 5$ to 4 and 3 but we add these subsets to the global solution last, and if there are still vehicles that cannot be included, they are not included in the final result for that choice of T_c. For each subset we start with a far field assumption and leverage the developments of Thrun [127]. From the far field approximation, we refine the positions of the sound sources by allowing the entires in **S** to be parameterized by (x,y) coordinates instead of angles. Once we have a solution for each subset, we piece them together and run through the optimization of equation 5.2 again. The details of each step are explained in the following subsections.

Far field approximation

Given a subset of the swarm, ρ, and noise sources, σ, we estimate their far field approximations $\mathbf{R}_\rho$ and Λ_σ solving the minimization procedure described in Section 5.2.2. Each time we compute a solution to equation 5.5 we evaluate its stability. In order to keep this estimate for further processing it must meet two criteria. First we make sure that there is an adequate variety of angles that produced the result. If all the angles of arrival are similar then a small change in matrix $\overline{\mathbf{D}}_{\rho,\sigma}$ can produce a very different estimate of $\mathbf{R}_\rho$. We also check to make sure that the residuals of equation 5.5 are reasonable before we accept the result. An example of a solution to equation 5.5 can be found in Figure 5.5 (a) where the estimates of the geometry of a subset are shown by red asterisks and the angle of arrival is shown by the polar plot inset.

Near field refinement

Once we have estimates for $\mathbf{R}_\rho$ and Λ_σ we use these values as an initial estimate for equation 5.2 to refine the positions of the noise sources, $\mathbf{S}_\sigma$, and make small modifications to the geometry of the swarm $\mathbf{R}_\rho$, by allowing the positions of the noise sources to be parameterized by a two-dimensional position instead of an angle. While not always necessary, this can improve our estimate of the swarm geometry if the detected noise sources are not accurately parameterized by an angle after the merging correspondences step of Section 5.3.2. For our initial estimate of $\mathbf{S}_\sigma$ we choose $\mathbf{S}_\sigma = r\Lambda_\sigma$ where r is a constant that is an estimate of how far the noise sources are from the swarm (we choose $r = 400m$). This refinement step is shown in Figure 5.5 (b) where the blue lines indicate the difference between the initial and final estimates of the noise sources (black squares) and the geometry of the swarm (red '*'). The refinement is done using the Levenberg-Marquardt algorithm.

Adding different subsets to the estimate

From a refined estimate of the swarm, our next step is to add more vehicles to the estimate. We choose a different value of ρ and σ that has the most vehicles in common with our current estimate of the swarm geometry. We then perform the far field approximation and refinement on this new subset of data. Next we find an optimal rigid body transformation[135] between the overlapping units of the new subset and current estimate to position the two estimates in the same coordinate frame. If there is not enough overlap in vehicles to determine an appropriate transformation, we use the estimated location of the noise sources to help with the rotation estimate (we do not add subsets that do not have enough combined overlap between the vehicles and noise sources). This step is shown in Figure 5.5 (c) where we are adding a new subset of vehicles (dark blue '+') and noise sources (dark blue diamonds) by using the overlapping vehicle estimates (shown by the red '*' and cyan '+'). Once everything is in the same coordinate frame we use this new estimate as an initial estimate for further refinement in equation 5.2 (Figure 5.5 (d)). We iterate this step until all of predetermined subsets have been added.

5.4 Self-localization Results

The final results of our optimization procedure are shown in Figure 5.6 for four different times during the deployment. In this figure, the estimated geometry from the high frequency pinging system is shown by a blue circle. The self-localization estimate, $\mathbf{R}$, from the impulsive noises is shown by red asterisks. Additionally, some noise source locations, $\mathbf{S}$, are shown by a yellow box, although most are outside of the frame to focus on the swarm estimate. The plots show the overlay after an optimal rigid transformation is applied between the two estimates [135], since we cannot place the estimate from the noise sources in a global coordinate frame. Also reported is the number of snaps that were used in the estimate and the root mean square error (RMSE) in meters, which measures the discrepancy between the geometry of the two estimates.

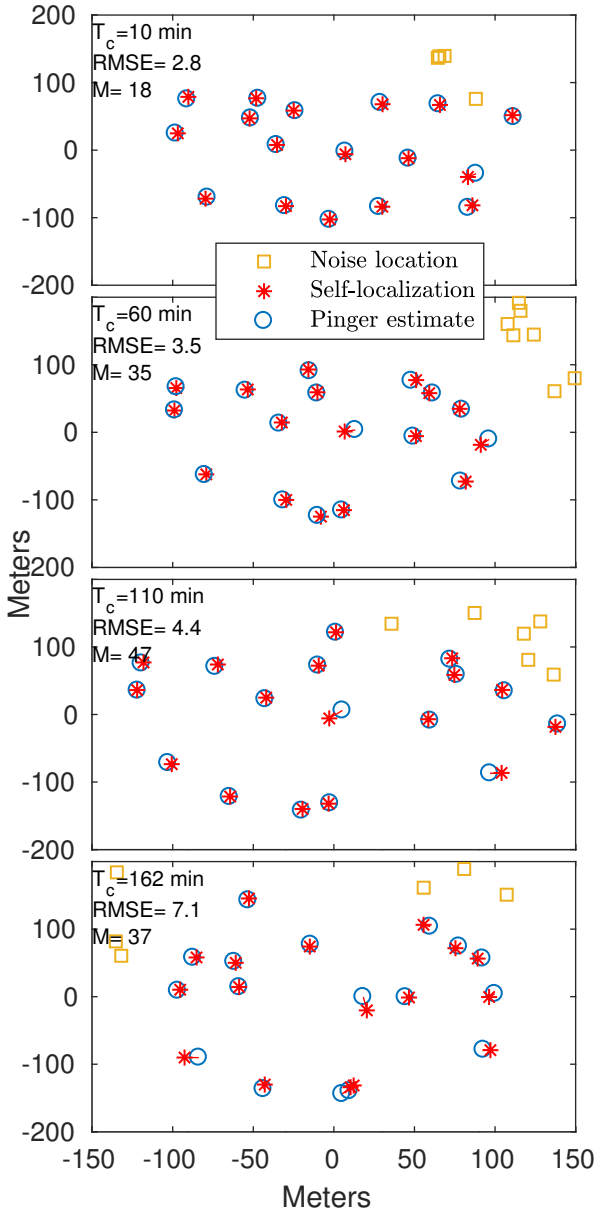

Figure 5.6: Full solution for multiple deployment times: The full solution of our method is shown for different values of T_c (a: $T_c = 10$ min b: $T_c = 60$ min c: $T_c = 110$ min d: $T_c = 162$ min). The blue circles represent the geometry estimate from the high frequency GPS pingers and the red '*' represents our estimate using the impulsive noise sources. Yellow squares represent the estimated location of noise sources that are close to the swarm. Many noise source locations are out of frame to focus on the swarm. The number of noise sources that are estimated, M, is also shown. Also reported is the RMSE in meters between the swarm geometry estimated by the impulsive noise sources and the acoustic localization system. This Figure shows good agreement between our self-localization solution and the estimate provided by the acoustic localization system.

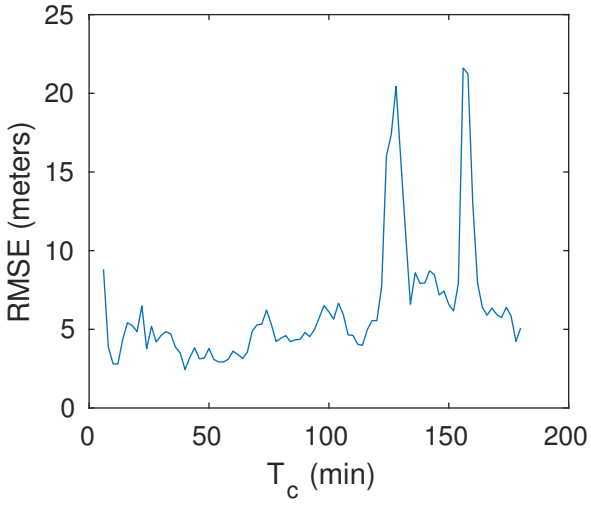

Figure 5.7: RMSE as a function of deployment time: RMSE, in meters, as a function of center time T_c, between a median filtered version of the self-localization solution and the estimates from the acoustic localization system. The RMSE measures the discrepancy between the two solutions. The RMSE shows good agreement for most of the deployment, validating our proposed method.

For many cases, all the units are estimated and the estimate matches what is reported by the pingers. When the estimate of the swarm has a higher RMSE (see Figure 5.6 b-d) it is typically the same units that do not match the estimate provided by the pingers. An example is the unit that is closest to the origin, (0,0). This result may point to an inaccurate estimate from the pinging system (e.g. that a reflected arrival was detected as the direct arrival for a ping).

Figure 5.7 shows the RMSE, in meters, as a function of deployment time for the first 3 hours of the deployment after a median filter of length 5 is applied to the position estimates. The median filter ensures that there is an estimate of the vehicle locations for every time step and smooths solutions that arrived at a suboptimal solution. Only the first three hours are shown. After the third hour, the swarm drifted over the impulsive noises and the bottom section of the swarm moved away from the top section. These events significantly violated our far field model, described in Section 5.3.2 and 5.3.3, degrading the results. The RMSE is usually around 5 meters, and for periods of time it stays below 5 meters. When the RMSE does jump, it usually results from weak connections between the bottom four vehicles that are far away from the rest of the units. The "ground truth" that we are comparing our measurements to are also estimates. The

expected accuracy of the pinger estimates are on the order of a few meters [8]. In this frame of reference, the results presented show the effectiveness of the proposed method.

5.5 Self-Synchronization Experiment

In our self-localization experiment we assumed that the clocks were synchronized between the vehicles (as described in Section 5.3), which is important to derive accurate spatial constraints from the TDOA data. In this section, we estimate the clock offsets, $o_{a,b}$ and $o_{a,e}$ of Equations 5.3 and 5.4, as a function of deployment time by using the correspondence method of Section 5.2.1. In order to evaluate our results, we compare our estimated clock bias to a linear interpolation of the clock drift based on the GPS measurements. This linear interpolation has been used in other analyses [11, 8], as well as the self-localization solution, and we believe that the clock estimates are accurate to 1ms for the entire 5 hour deployment. Therefore, if our clock bias estimation is within 1ms of the linear interpolation, we believe our estimation is reasonable. We call this linear interpolation our 'Ground Truth', but stress that it is also an estimate.

5.5.1 Results and Discussion

For our experiments, we split the deployment time into time windows that were $\delta_s = 240s$ long and ran the method described in Section 5.2.1 and Algorithm 1 for each individual window. Our windows are spaced $\delta_s/2$ seconds apart, so there is overlap between consecutive windows.

The development of Section 5.2.1 and Algorithm 1 (with $N_{min} = 5$) was applied to a triplet of vehicles (units 7, 8, and 17). Examples of the RANSAC portion of Algorithm 1 for the estimation of ellipses are shown in Figure 5.8 for different deployment times. Most of the estimations of the bias between clocks take the form of Figure 5.8 (a), where the ellipse is stable and the center of the ellipse, shown by an '*', matches the expected value of clock bias from our ground truth estimate, shown by a diamond. These results show we are able to use impulsive noise to accurately estimate

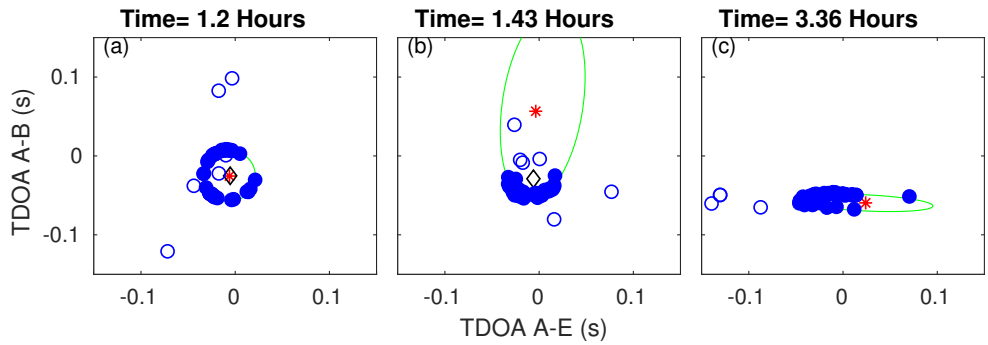

Figure 5.8: Examples of our estimation for different times during the deployment. The candidate correspondences are shown by 'o' markers. A filled marker denotes that the candidate correspondence was accepted by the RANSAC algorithm, while an open marker denotes that it was rejected as an outlier. The final ellipse is shown as a least squares fit to the inlier set. The middle marker for the ellipse is shown with a '*' and the expected center of the ellipse is shown by a black diamond. (a) shows an example of a good estimate (b) shows an example of a challenged estimate, because of the azimuthal variety of the angle of arrival of the snaps (c) shows an example of a challenged estimate because of the geometry of the vehicles; when the vehicles approach a line (i.e. when $\phi_{a,b,e} \approx 180\,\text{deg}$) the ellipse degenerates and the estimation becomes unstable.

the clock bias between two pairs of vehicles. Our RANSAC algorithm robustly finds the TDOA values that originate from the same sound source.

While the estimation worked for most of the deployment, there were times when it did not. During the time window of Figure 5.8 (b), most of the angle of arrivals of the snaps are similar. Under this condition, only one side of the ellipse is well constrained and the ellipse is more sensitive to noise in the time of arrival measurements. An interesting thing to note about Figure 5.8 (b) is that even though the ellipse is far from the optimal – the ellipse should look similar to Figure 5.8 (a) – the center is still constrained along the TDOA A-E axis. The resulting ellipse gives a biased estimate for the TDOA A-B pair, but an accurate estimate for the A-E pair because of the angles of arrivals of the snaps.

Another difficulty in the ellipse estimation is shown in Figure 5.8 (c). This is the result of the vehicles geometry becoming linear (i.e. $\phi_{a,b,e}$ in Equation 5.4 approaches 180°). In this case, the ellipse parameterized by Equation 5.3 and 5.4 approach a line and an ellipse is not a valid model for this vehicle geometry. For the vehicle triplet that was chosen for this analysis, the

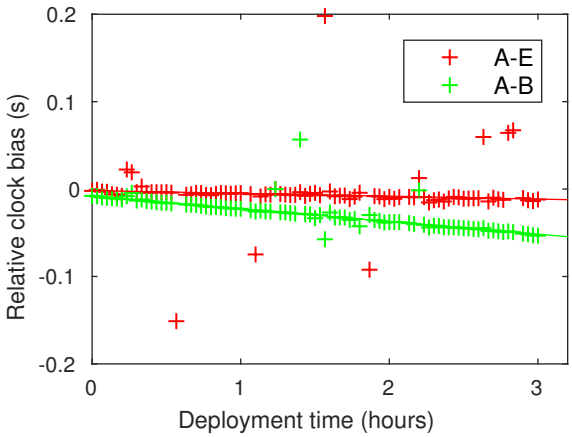

Figure 5.9: RANSAC Results: The clock synchronization results are shown for 3 hours during the deployment. The Ground Truth estimate is shown by a line. The '+' markers show the estimation of clock bias that comes from the center of the ellipse, examples of which are shown in Figure 5.8.

geometry was close to linear for the last 2 hours, making the estimation difficult.

Our final, unfiltered estimates of the clock bias for each time step are shown in Figure 5.9 for the first 3 hours of the deployment. After the first 3 hours, the geometry of the vehicle triplet made the ellipse estimation unstable (see Figure 5.8 (c)). In Figure 5.9, each point is taken from an estimation of the ellipse center, such as the ones shown in Figure 5.8. The red '+' (A-E) markers are taken from the time value of the ellipse on the x-axis and the green '+' (A-B) values are taken from the y-axis, respectively. Also shown on this plot are the Ground Truth estimates (marked by a line). We see that by using our technique we are able to accurately estimate the clock biases between the vehicles. These estimates are frequently close to what we expect from the linear drift estimates. There are some noticeable jumps from the linear clock drift, especially for the A-E pair. These result from estimations similar to the ones shown in Figure 5.8 (b). Fortunately, we can identify and filter out these outliers.

Figure 5.10 is given to provide a quantitative comparison of self-synchronization method to the Ground Truth estimate. A Cumulative Distribution Function is given to describe the absolute distance from the two estimates for pairs A-E (red) and A-B (green). The y-axis gives the percentage of estimates that differ from the Ground Truth by less than the value given on the

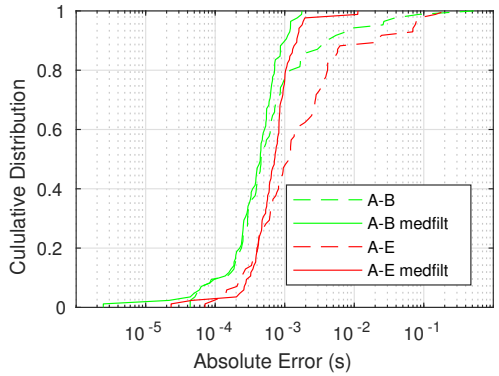

Figure 5.10: The plotted CDF describes the difference between the self-synchronization estimates and the Ground Truth estimates. Pair A-B is shown in green and A-E in red. Dashed lines show the CDF of the raw data and solid lines show the data after a median filter is used. The median filtered estimates show that over 90% of our estimates are within 1.1ms of the Ground Truth estimate.

x-axis. Results for both the raw measurements (dashed) (i.e. those shown in Figure 5.9), and after a median filter of length 9 is applied to the raw data. This figure validates our self-synchronization method - for the median filtered version, over 90% of the estimates are within 1.1ms of the Ground Truth estimates. Given that we expect the Ground Truth estimates to be within around 1ms of GPS time, our results show good agreement.

While this analysis was only performed between one vehicle triplet, the results show promise that impulsive noise can be used to synchronize the clocks between vehicles. There are challenges when the vehicles become linear. With a dense swarm of underwater vehicles, the clock offsets between all units could be estimated by taking advantage of vehicle triplets that have favorable geometry and realizing that the clock offsets are transitive. We leave this analysis as future work.

5.6 Conclusion

In this work we looked at the self-localization problem of an untethered deforming swarm of subsurface vehicles. We assumed that the swarm was stationary for short periods of time and estimated the relative geometry at each time step. Our sources of opportunity were impulsive,

giving us good time resolution but providing a significant correspondence problem. We proposed a method to solve this correspondence problem by assuming that the noise sources could be modeled by a plane wave and examining the correspondences 3 vehicles at a time. We accepted correspondences that agreed with a geometric model. From these correspondences, an incremental solution was developed that first solved local subproblems and then merged those subproblems into a global estimate of the swarm geometry. Our approach was able to accurately track the deformation of the swarm as a function of time, verified by an independent, accurate localization procedure.

We proposed a self-synchronization method that used ambient impulsive noise to estimate the clock drift between pairs of receivers. The self-synchronization matched our Ground Truth estimate of the clock bias based on GPS measurements. The self-synchronization solution does not require any assumptions, such as a linear clock drift, making it useful for long deployments when the clock drift will not be linear. The self-synchronization can mitigate the expense of using real time clocks and acoustic communication for synchronization which in turn would lower the price per vehicle and allow more vehicles to be deployed.

5.7 Acknowledgements

Chapter 5, in part, is a reprint of the material as it appears in IEEE Access: P. Naughton, P. Roux, C. Schurgers, R. Kastner, J. S. Jaffe, and P. L. D. Roberts, (2018) “Self-localization of a deforming swarm of underwater vehicles using impulsive sound sources of opportunity”. The dissertation author was the primary investigator and author of this paper.

Section 5.5, in part, has been submitted for publication of the material as it may appear in the fourth Underwater Communications and Networking Conference: P. Naughton, T. Salam, P. Roux, C. Schurgers, R. Kastner, J. S. Jaffe, and P. L. D. Roberts, (2018) “Self-synchronization of multiple vehicles using ambient impulsive noise”. The dissertation author was the primary

investigator and author of this paper.

Chapter 6

Conclusion

This dissertation demonstrated self-localization procedures for a swarm of independent underwater vehicles using ambient acoustic noise. The ambient noise in the ocean depends on the geographic location and also varies temporally. Therefore, different environments afford different solutions to the self-localization problem. The fundamental physical challenge considered in this dissertation is unambiguously detecting constraints from a variety of azimuthal directions to estimate the geometry of the vehicles in two dimensions. Different approaches to the self-localization were given in order to encompass different scenarios that might be encountered in underwater environments.

Recovering an estimate of the acoustic impulse response using correlations of low frequency noise was examined. This requires an isotropic distribution of noise sources over the correlation window to work between all vehicle pairs. While this technique showed promise - for pairs of vehicles the expected double sided structure of the noise correlation was recovered - the short time windows required by mobile vehicles makes this method susceptible to anisotropic noise fields. For some pairs of receivers, the noise correlation function did not recover an accurate estimate of the acoustic impulse response. More research could provide insight into evaluating the possibility of correlating low frequency noise to estimate the acoustic impulse response between

moving vehicles. Particularly, it would be interesting to look at environments where the noise field may be expected to be more isotropic than what is considered in this dissertation.

In response to this strict constraint on the time windows, a method was formulated that jointly estimated the changing geometry of the moving swarm along with the direction of ambient noise sources in the ocean. This approach assumed the vehicles were stationary for short periods of time in order to perform this joint estimation. Simulations were provided that examined the effect of violating this stationary assumption as well as the affects of the swarm geometry and the azimuthal variance of the noise. The proposed method was shown to accurately estimate the geometry of the vehicles as well as the angles of arrival of a boat as it circled the vehicles.

One of the challenges of this estimation procedure is that the constraints must be derived from the same signal across each vehicle. When the vehicles are spaced far from each other, and the signals of interest happen frequently in time and are similar in structure, it can be difficult to choose the same signal in the recordings of different vehicles. This correspondence problem was considered and a method to estimate the correct correspondences was proposed. A self-localization solution was demonstrated using the estimated correspondences between snapping shrimp signals. The self-localization solution was also able to handle missing data, where each vehicle recorded a different subset of snapping events than the rest of the vehicles. The correspondence solution was extended to a self-synchronization solution. Results showed promise for being able to determine the clock drift between triplets of vehicles by examining the time difference of arrival of impulsive signals received by the receivers.

Bibliography

[1] D. M. C. Jr. and M. Lindegren, "Sea surface temperature variability at the scripps institution of oceanography pier," *Journal of Physical Oceanography*, vol. 44, no. 11, pp. 2877–2892, 2014.

[2] L. D. Talley, G. L. Pickard, W. J. Emery, and J. H. Swift, "Instruments and methods," in *Descriptive Physical Oceanography*, ch. 16, pp. 1–83, Elsevier, 2011.

[3] A. Gartman, M. Hannington, J. W. Jamieson, B. Peterkin, D. Garbe-Schnberg, A. J. Findlay, S. Fuchs, and T. Kwasnitschka, "Boiling-induced formation of colloidal gold in black smoker hydrothermal fluids," *Geology*, vol. 46, no. 1, p. 39, 2017.

[4] P. Khanna, A. W. Droxler, J. A. Nittrouer, J. W. Tunnell Jr, and T. C. Shirley, "Coralgal reef morphology records punctuated sea-level rise during the last deglaciation," *Nature communications*, vol. 8, no. 1, p. 1046, 2017.

[5] M. Johnson-Roberson, O. Pizarro, S. B. Williams, and I. Mahon, "Generation and visualization of large-scale three-dimensional reconstructions from underwater robotic surveys," *Journal of Field Robotics*, vol. 27, no. 1, pp. 21–51, 2010.

[6] A. Kim and R. M. Eustice, "Real-time visual slam for autonomous underwater hull inspection using visual saliency," *IEEE Transactions on Robotics*, vol. 29, pp. 719–733, June 2013.

[7] R. Stokey, T. Austin, B. Allen, N. Forrester, E. Gifford, R. Goldsborough, G. Packard, M. Purcell, and C. von Alt, "Very shallow water mine countermeasures using the remus auv: a practical approach yielding accurate results," in *OCEANS, 2001. MTS/IEEE Conference and Exhibition*, vol. 1, pp. 149–156, IEEE, 2001.

[8] J. S. Jaffe, P. J. S. Franks, P. L. D. Roberts, D. Mirza, C. Schurgers, R. Kastner, and A. Boch, "A swarm of autonomous miniature underwater robot drifters for exploring submesoscale ocean dynamics," *Nature Communications*, vol. 8, no. 14189, 2017. Article.

[9] E. A. D'Asaro, A. Y. Shcherbina, J. M. Klymak, J. Molemaker, G. Novelli, C. M. Guigand, A. C. Haza, B. K. Haus, E. H. Ryan, G. A. Jacobs, H. S. Huntley, N. J. M. Laxague, S. Chen, F. Judt, J. C. McWilliams, R. Barkan, A. D. Kirwan, A. C. Poje, and T. M. Özgökmen,

"Ocean convergence and the dispersion of flotsam," *Proceedings of the National Academy of Sciences*, 2018.

[10] C. Roman, G. Inglis, and B. McGilvray, "Lagrangian floats as sea floor imaging platforms," *Continental Shelf Research*, vol. 31, no. 15, pp. 1592 – 1598, 2011.

[11] P. Naughton, P. Roux, R. Yeakle, C. Schurgers, R. Kastner, J. S. Jaffe, and P. L. D. Roberts, "Ambient noise correlations on a mobile, deformable array," *The Journal of the Acoustical Society of America*, vol. 140, no. 6, pp. 4260–4270, 2016.

[12] P. Naughton, P. Roux, C. Schurgers, R. Kastner, J. S. Jaffe, and P. L. D. Roberts, "Self-localization of a deforming swarm of underwater vehicles using impulsive sound sources of opportunity," *IEEE Access*, vol. 6, pp. 1635–1646, 2018.

[13] H. Schmidt, J. G. Bellingham, and J. W. Bales, "Mobile underwater arrays," Apr. 13 1999. US Patent 5,894,450.

[14] "Argo." http://www.argo.ucsd.edu/. Accessed: March, 2018.

[15] D. Roemmich, G. C. Johnson, S. Riser, R. Davis, J. Gilson, W. B. Owens, S. L. Garzoli, C. Schmid, and M. Ignaszewski, "The argo program: Observing the global ocean with profiling floats," *Oceanography*, vol. 22, no. 2, pp. 34–43, 2009.

[16] J. C. Swallow, "A neutral-buoyancy float for measuring deep currents," *Deep Sea Research (1953)*, vol. 3, no. 1, pp. 74–81, 1955.

[17] R. E. Davis, "Observing the general circulation with floats," *Deep Sea Research Part A. Oceanographic Research Papers*, vol. 38, pp. S531–S571, 1991.

[18] R. Davis, L. Regier, J. Dufour, and D. Webb, "The autonomous lagrangian circulation explorer (alace)," *Journal of atmospheric and oceanic technology*, vol. 9, no. 3, pp. 264–285, 1992.

[19] R. Davis, J. Sherman, and J. Dufour, "Profiling alaces and other advances in autonomous subsurface floats," *Journal of atmospheric and oceanic technology*, vol. 18, no. 6, pp. 982–993, 2001.

[20] T. Rossby and D. Webb, "Observing abyssal motions by tracking swallow floats in the sofar channel," *Deep Sea Research and Oceanographic Abstracts*, vol. 17, no. 2, pp. 359 – 365, 1970.

[21] T. Rossby, D. Dorson, and J. Fontaine, "The rafos system," *Journal of Atmospheric and Oceanic Technology*, vol. 3, no. 4, pp. 672–679, 1986.

[22] J. Swallow, B. McCartney, and N. Millard, "The minimode float tracking system," in *Deep Sea Research and Oceanographic Abstracts*, vol. 21, pp. 573–595, Elsevier, 1974.

[23] A. Y. Shcherbina, M. A. Sundermeyer, E. Kunze, E. DAsaro, G. Badin, D. Birch, A.-M. E. Brunner-Suzuki, J. Callies, B. T. Kuebel Cervantes, M. Claret, *et al.*, "The latmix summer campaign: submesoscale stirring in the upper ocean," *Bulletin of the American Meteorological Society*, vol. 96, no. 8, pp. 1257–1279, 2015.

[24] D. C. Webb, P. J. Simonetti, and C. P. Jones, "Slocum: An underwater glider propelled by environmental energy," *IEEE Journal of oceanic engineering*, vol. 26, no. 4, pp. 447–452, 2001.

[25] J. Sherman, R. E. Davis, W. Owens, and J. Valdes, "The autonomous underwater glider" spray"," *IEEE Journal of Oceanic Engineering*, vol. 26, no. 4, pp. 437–446, 2001.

[26] C. C. Eriksen, T. J. Osse, R. D. Light, T. Wen, T. W. Lehman, P. L. Sabin, J. W. Ballard, and A. M. Chiodi, "Seaglider: A long-range autonomous underwater vehicle for oceanographic research," *IEEE Journal of oceanic Engineering*, vol. 26, no. 4, pp. 424–436, 2001.

[27] N. E. Leonard, D. A. Paley, R. E. Davis, D. M. Fratantoni, F. Lekien, and F. Zhang, "Coordinated control of an underwater glider fleet in an adaptive ocean sampling field experiment in monterey bay," *Journal of Field Robotics*, vol. 27, no. 6, pp. 718–740, 2010.

[28] N. E. Leonard, D. A. Paley, F. Lekien, R. Sepulchre, D. M. Fratantoni, and R. E. Davis, "Collective motion, sensor networks, and ocean sampling," *Proceedings of the IEEE*, vol. 95, no. 1, pp. 48–74, 2007.

[29] J. Yuh, "Design and control of autonomous underwater robots: A survey," *Autonomous Robots*, vol. 8, pp. 7–24, Jan 2000.

[30] D. R. Yoerger, A. M. Bradley, and B. B. Walden, "The autonomous benthic explorer (abe): A deep ocean auv for scientific seafloor survey," *Oceanogi. aphic*, p. 79, 1991.

[31] B. Allen, R. Stokey, T. Austin, N. Forrester, R. Goldsborough, M. Purcell, and C. von Alt, "Remus: a small, low cost auv; system description, field trials and performance results," in *OCEANS'97. MTS/IEEE Conference Proceedings*, vol. 2, pp. 994–1000, IEEE, 1997.

[32] D. Thompson, D. Caress, D. Clague, D. Conlin, J. Harvey, E. Martin, J. Paduan, C. Paull, J. Ryan, H. Thomas, *et al.*, "Mbari dorado auv's scientific results," in *Oceans-San Diego, 2013*, pp. 1–9, IEEE, 2013.

[33] M. Dunbabin, J. Roberts, K. Usher, G. Winstanley, and P. Corke, "A hybrid auv design for shallow water reef navigation," in *Robotics and Automation, 2005. ICRA 2005. Proceedings of the 2005 IEEE International Conference on*, pp. 2105–2110, IEEE, 2005.

[34] J. G. Bellingham and K. Rajan, "Robotics in remote and hostile environments," *Science*, vol. 318, no. 5853, pp. 1098–1102, 2007.

[35] J. G. Bellingham, C. A. Goudey, T. R. Consi, J. W. Bales, D. K. Atwood, J. J. Leonard, and C. Chryssostomidis, "A second generation survey auv," in *Autonomous Underwater Vehicle Technology, 1994. AUV '94., Proceedings of the 1994 Symposium on*, pp. 148–155, 1994.

[36] J. Vaganay, J. J. Leonard, and J. G. Bellingham, "Outlier rejection for autonomous acoustic navigation," in *Proceedings of IEEE International Conference on Robotics and Automation*, vol. 3, pp. 2174–2181 vol.3, Apr 1996.

[37] T. B. Curtin, J. G. Bellingham, J. Catipovic, and D. Webb, "Autonomous oceanographic sampling networks," *Oceanography*, vol. 6, no. 3, pp. 86–94, 1993.

[38] P. Naughton, C. Edwards, V. Petrovic, R. Kastner, F. Kuester, and S. Sandin, "Scaling the annotation of subtidal marine habitats," in *Proceedings of the 10th International Conference on Underwater Networks & Systems*, WUWNET '15, (New York, NY, USA), pp. 31:1–31:5, ACM, 2015.

[39] T. Schmickl, R. Thenius, C. Moslinger, J. Timmis, A. Tyrrell, M. Read, J. Hilder, J. Halloy, A. Campo, C. Stefanini, *et al.*, "Cocoro–the self-aware underwater swarm," in *Self-Adaptive and Self-Organizing Systems Workshops (SASOW), 2011 Fifth IEEE Conference on*, pp. 120–126, IEEE, 2011.

[40] F. S. Schill, "Distributed communication in swarms of autonomous underwater vehicles," tech. rep., The Australian National University, 2007.

[41] A. Hackbarth, E. Kreuzer, and E. Solowjow, "Hippocampus: A micro underwater vehicle for swarm applications," in *Intelligent Robots and Systems (IROS), 2015 IEEE/RSJ International Conference on*, pp. 2258–2263, IEEE, 2015.

[42] C. Osterloh, T. Pionteck, and E. Maehle, "Monsun ii: A small and inexpensive auv for underwater swarms," in *Robotics; Proceedings of ROBOTIK 2012; 7th German Conference on*, pp. 1–6, VDE, 2012.

[43] F. Schill, A. Bahr, and A. Martinoli, "Vertex: A new distributed underwater robotic platform for environmental monitoring," in *Distributed Autonomous Robotic Systems*, pp. 679–693, Springer, 2018.

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

[45] G. Han, J. Jiang, L. Shu, Y. Xu, and F. Wang, "Localization algorithms of underwater wireless sensor networks: A survey," *Sensors*, vol. 12, no. 2, pp. 2026–2061, 2012.

[46] V. Chandrasekhar, W. K. Seah, Y. S. Choo, and H. V. Ee, "Localization in underwater sensor networks: survey and challenges," in *Proceedings of the 1st ACM international workshop on Underwater networks*, pp. 33–40, ACM, 2006.

[47] L. Stutters, H. Liu, C. Tiltman, and D. J. Brown, "Navigation technologies for autonomous underwater vehicles," *IEEE Transactions on Systems, Man, and Cybernetics, Part C (Applications and Reviews)*, vol. 38, pp. 581–589, July 2008.

[48] M. M. Hunt, W. M. Marquet, D. A. Moller, K. R. Peal, W. K. Smith, and R. C. Spindel, "An acoustic navigation system," tech. rep., Woods Hole Oceanographic Institution, 1974.

[49] N. R. Rypkema, E. M. Fischell, and H. Schmidt, "One-way travel-time inverted ultra-short baseline localization for low-cost autonomous underwater vehicles," in *2017 IEEE International Conference on Robotics and Automation (ICRA)*, pp. 4920–4926, May 2017.

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

[51] J. Yi, D. Mirza, R. Kastner, C. Schurgers, P. Roberts, and J. Jaffe, "Toa-ts: Time of arrival based joint time synchronization and tracking for mobile underwater systems," *Ad Hoc Networks*, vol. 34, pp. 211–223, 2015.

[52] D. Mirza, P. Naughton, C. Schurgers, and R. Kastner, "Real-time collaborative tracking for underwater networked systems," *Ad Hoc Networks*, 2014.

[53] S. E. Dosso, M. R. Fallat, B. J. Sotirin, and J. L. Newton, "Array element localization for horizontal arrays via occams inversion," *The Journal of the Acoustical Society of America*, vol. 104, no. 2, pp. 846–859, 1998.

[54] B. J. Sotirin and J. A. Hildebrand, "Acoustic navigation of a large-aperture array," *The Journal of the Acoustical Society of America*, vol. 87, no. 1, pp. 154–167, 1990.

[55] B. J. Sotirin and J. A. Hildebrand, "Large aperture digital acoustic array," *IEEE Journal of Oceanic Engineering*, vol. 13, pp. 271–281, Oct 1988.

[56] L. Techy, K. A. Morganseny, and C. A. Woolseyz, "Long-baseline acoustic localization of the seaglider underwater glider," in *Proceedings of the 2011 American Control Conference*, pp. 3990–3995, June 2011.

[57] R. M. Eustice, H. Singh, and L. L. Whitcomb, "Synchronous-clock, one-way-travel-time acoustic navigation for underwater vehicles," *journal of field robotics*, vol. 28, no. 1, pp. 121–136, 2011.

[58] B. Bingham, D. Mindell, and T. Wilcox, "Integrating sharps ii precision navigation into jason/medea two-vehicle operation," in *OCEANS, 2005. Proceedings of MTS/IEEE*, pp. 1603–1609, IEEE, 2005.

[59] K. G. Kebkal, O. G. Kebkal, R. Bannasch, and S. G. Yakovlev, "Performance of a combined usbl positioning and communication system using s2c technology," in *2012 Oceans - Yeosu*, pp. 1–7, May 2012.

[60] B. Allotta, F. Bartolini, A. Caiti, R. Costanzi, F. Di Corato, D. Fenucci, J. Gelli, P. Guerrini, N. Monni, A. Munafò, *et al.*, "Typhoon at commsnet 2013: experimental experience on auv navigation and localization," *IFAC Proceedings Volumes*, vol. 47, no. 3, pp. 3370–3375, 2014.

[61] R. McEwen, H. Thomas, D. Weber, and F. Psota, "Performance of an auv navigation system at arctic latitudes," *IEEE Journal of Oceanic Engineering*, vol. 30, pp. 443–454, April 2005.

[62] B. Allen, R. Stokey, T. Austin, N. Forrester, R. Goldsborough, M. Purcell, and C. von Alt, "Remus: a small, low cost auv; system description, field trials and performance results," in *OCEANS '97. MTS/IEEE Conference Proceedings*, vol. 2, pp. 994–1000 vol.2, Oct 1997.

[63] L. Freitag, M. Grund, S. Singh, J. Partan, P. Koski, and K. Ball, "The whoi micro-modem: an acoustic communications and navigation system for multiple platforms," in *Proceedings of OCEANS 2005 MTS/IEEE*, pp. 1086–1092 Vol. 2, Sept 2005.

[64] A. Bahr, J. J. Leonard, and M. F. Fallon, "Cooperative localization for autonomous underwater vehicles," *The International Journal of Robotics Research*, vol. 28, no. 6, pp. 714–728, 2009.

[65] L. E. Freitag, M. Grund, J. Partan, K. Ball, S. Singh, and P. Koski, "Multi-band acoustic modem for the communications and navigation aid auv," in *Proceedings of OCEANS 2005 MTS/IEEE*, pp. 1080–1085 Vol. 2, Sept 2005.

[66] K. G. Sabra, P. Roux, A. M. Thode, G. L. D'Spain, W. S. Hodgkiss, and W. A. Kuperman, "Using ocean ambient noise for array self-localization and self-synchronization," *IEEE Journal of Oceanic Engineering*, vol. 30, pp. 338–347, April 2005.

[67] R. K. Andrew, B. M. Howe, J. A. Mercer, and M. A. Dzieciuch, "Ocean ambient sound: comparing the 1960s with the 1990s for a receiver off the california coast," *Acoustics Research Letters Online*, vol. 3, no. 2, pp. 65–70, 2002.

[68] J. A. Hildebrand, "Anthropogenic and natural sources of ambient noise in the ocean," *Marine Ecology Progress Series*, vol. 395, pp. 5–20, 2009.

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

[70] M. Chitre, S. Ong, and J. Potter, "Performance of coded ofdm in very shallow water channels and snapping shrimp noise," in *OCEANS, 2005. Proceedings of MTS/IEEE*, pp. 996–1001, IEEE, 2005.

[71] M. A. Chitre, J. R. Potter, and S.-H. Ong, "Optimal and near-optimal signal detection in snapping shrimp dominated ambient noise," *IEEE Journal of oceanic engineering*, vol. 31, no. 2, pp. 497–503, 2006.

[72] K. Pelekanakis and M. Chitre, "Adaptive sparse channel estimation under symmetric alpha-stable noise," *IEEE Transactions on wireless communications*, vol. 13, no. 6, pp. 3183–3195, 2014.

[73] A. Mahmood, M. Chitre, and M. A. Armand, "Improving psk performance in snapping shrimp noise with rotated constellations," in *Proceedings of the Seventh ACM International Conference on Underwater Networks and Systems*, p. 12, ACM, 2012.

[74] G. M. Wenz, "Acoustic ambient noise in the ocean: Spectra and sources," *The Journal of the Acoustical Society of America*, vol. 34, no. 12, pp. 1936–1956, 1962.

[75] N. R. Council, D. on Earth, L. Studies, O. S. Board, and C. on Potential Impacts of Ambient Noise in the Ocean on Marine Mammals, *Ocean Noise and Marine Mammals*. National Academies Press, 2003.

[76] R. L. Weaver and O. I. Lobkis, "Ultrasonics without a source: Thermal fluctuation correlations at mhz frequencies," *Physical Review Letters*, vol. 87, no. 13, p. 134301, 2001.

[77] O. I. Lobkis and R. L. Weaver, "On the emergence of the greens function in the correlations of a diffuse field," *The Journal of the Acoustical Society of America*, vol. 110, no. 6, pp. 3011–3017, 2001.

[78] B. Froment, M. Campillo, and P. Roux, "Reconstructing the green's function through iteration of correlations," *Comptes Rendus Geoscience*, vol. 343, no. 89, pp. 623 – 632, 2011.

[79] P. Roux, K. G. Sabra, W. A. Kuperman, and A. Roux, "Ambient noise cross correlation in free space: Theoretical approach," *Journal of the Acoustical Society of America*, vol. 117, pp. 79–84, Jan 2005.

[80] K. G. Sabra, P. Roux, and W. A. Kuperman, "Emergence rate of the time-domain green's function from the ambient noise cross-correlation function," *The Journal of the Acoustical Society of America*, vol. 118, no. 6, pp. 3524–3531, 2005.

[81] R. Snieder, "Extracting the green's function from the correlation of coda waves: A derivation based on stationary phase," *Phys. Rev. E*, vol. 69, p. 046610, Apr 2004.

[82] K. G. Sabra, P. Roux, and W. Kuperman, "Arrival-time structure of the time-averaged ambient noise cross-correlation function in an oceanic waveguide," *The Journal of the Acoustical Society of America*, vol. 117, no. 1, pp. 164–174, 2005.

[83] O. A. Godin, "Emergence of deterministic greens functions from noise generated by finite random sources," *Physical Review E*, vol. 80, no. 6, p. 066605, 2009.

[84] S. 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*, vol. 131, no. 5, pp. EL388–EL394, 2012.

[85] L. Stehly, M. Campillo, B. Froment, and R. L. Weaver, "Reconstructing green's function by correlation of the coda of the correlation (c3) of ambient seismic noise," *Journal of Geophysical Research: Solid Earth*, vol. 113, no. B11, 2008.

[86] M. Campillo and A. Paul, "Long-range correlations in the diffuse seismic coda," *Science*, vol. 299, no. 5606, pp. 547–549, 2003.

[87] C. Sens-Schönfelder, "Synchronizing seismic networks with ambient noise," *Geophysical Journal International*, vol. 174, no. 3, pp. 966–970, 2008.

[88] K. G. Sabra, P. Gerstoft, P. Roux, W. A. Kuperman, and M. C. Fehler, "Extracting timedomain green's function estimates from ambient seismic noise," *Geophysical Research Letters*, vol. 32, no. 3, 2005.

[89] K. G. Sabra, E. S. Winkel, D. A. Bourgoyne, B. R. Elbing, S. L. Ceccio, M. Perlin, and D. R. Dowling, "Using cross correlations of turbulent flow-induced ambient vibrations to estimate the structural impulse response. application to structural health monitoring," *The Journal of the Acoustical Society of America*, vol. 121, no. 4, pp. 1987–1995, 2007.

[90] E. Larose, P. Roux, and M. Campillo, "Reconstruction of rayleigh–lamb dispersion spectrum based on noise obtained from an air-jet forcing," *The Journal of the Acoustical Society of America*, vol. 122, no. 6, pp. 3437–3444, 2007.

[91] A. Duroux, K. G. Sabra, J. Ayers, and M. Ruzzene, "Using cross-correlations of elastic diffuse fields for attenuation tomography of structural damage," *The Journal of the Acoustical Society of America*, vol. 127, no. 6, pp. 3311–3314, 2010.

[92] K. G. Sabra and A. Archer, "Tomographic elastography of contracting skeletal muscles from their natural vibrations," *Applied Physics Letters*, vol. 95, no. 20, p. 203701, 2009.

[93] P. Roux, W. Kuperman, and K. G. Sabra, "Ocean acoustic noise and passive coherent array processing," *Comptes rendus Geoscience*, vol. 343, no. 8, pp. 533–547, 2011.

[94] P. Roux, W. A. Kuperman, and the NPAL Group, "Extracting coherent wave fronts from acoustic ambient noise in the ocean," *The Journal of the Acoustical Society of America*, vol. 116, no. 4, pp. 1995–2003, 2004.

[95] L. A. Brooks and P. Gerstoft, "Greens function approximation from cross-correlation of active sources in the ocean," *The Journal of the Acoustical Society of America*, vol. 126, no. 1, pp. 46–55, 2009.

[96] M. Siderius, C. H. Harrison, and M. B. Porter, "A passive fathometer technique for imaging seabed layering using ambient noise," *The Journal of the Acoustical Society of America*, vol. 120, no. 3, pp. 1315–1323, 2006.

[97] P. Gerstoft, W. S. Hodgkiss, M. Siderius, C.-F. Huang, and C. H. Harrison, "Passive fathometer processing," *The Journal of the Acoustical Society of America*, vol. 123, no. 3, pp. 1297–1305, 2008.

[98] M. Siderius, H. Song, P. Gerstoft, W. S. Hodgkiss, P. Hursky, and C. Harrison, "Adaptive passive fathometer processing," *The Journal of the Acoustical Society of America*, vol. 127, no. 4, pp. 2193–2200, 2010.

[99] L. A. Brooks and P. Gerstoft, "Green's function approximation from cross-correlations of 20–100hz noise during a tropical storm," *The Journal of the Acoustical Society of America*, vol. 125, no. 2, pp. 723–734, 2009.

[100] S. W. Lani, K. G. Sabra, W. S. Hodgkiss, W. Kuperman, and P. Roux, "Coherent processing of shipping noise for ocean monitoring," *The Journal of the Acoustical Society of America*, vol. 133, no. 2, pp. EL108–EL113, 2013.

[101] R. Menon, P. Gerstoft, and W. S. Hodgkiss, "Cross-correlations of diffuse noise in an ocean environment using eigenvalue based statistical inference," *The Journal of the Acoustical Society of America*, vol. 132, no. 5, pp. 3213–3224, 2012.

[102] C. Leroy, S. Lani, K. G. Sabra, W. S. Hodgkiss, W. A. Kuperman, and P. 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*, vol. 132, no. 2, pp. 883–893, 2012.

[103] M. G. Brown, O. A. Godin, N. J. Williams, N. A. Zabotin, L. Zabotina, and G. J. Banker, "Acoustic green's function extraction from ambient noise in a coastal ocean environment," *Geophysical Research Letters*, vol. 41, no. 15, pp. 5555–5562, 2014.

[104] S. E. Fried, W. A. Kuperman, K. G. Sabra, and P. Roux, "Extracting the local green's function on a horizontal array from ambient ocean noise," *The Journal of the Acoustical Society of America*, vol. 124, no. 4, pp. EL183–EL188, 2008.

[105] K. F. Woolfe, S. Lani, K. G. Sabra, and W. A. Kuperman, "Monitoring deep-ocean temperatures using acoustic ambient noise," *Geophysical Research Letters*, vol. 42, no. 8, pp. 2878–2884, 2015. 2015GL063438.

[106] K. G. Sabra, B. Cornuelle, and W. Kuperman, "Sensing deep-ocean temperatures," *Phys. Today*, vol. 69, pp. 32–38, feb 2016.

[107] K. G. Sabra, S. Fried, W. Kuperman, and M. Prior, "On the coherent components of low-frequency ambient noise in the indian ocean," *The Journal of the Acoustical Society of America*, vol. 133, no. 1, pp. EL20–EL25, 2013.

[108] K. F. Woolfe, K. G. Sabra, and W. Kuperman, "Optimized extraction of coherent arrivals from ambient noise correlations in a rapidly fluctuating medium," *The Journal of the Acoustical Society of America*, vol. 138, no. 4, pp. EL375–EL381, 2015.

[109] W. Munk, "The sound of climate change," *Tellus A*, vol. 63, no. 2, pp. 190–197, 2011.

[110] W. Munk, P. Worcester, and C. Wunsch, *Ocean acoustic tomography*. Cambridge university press, 2009.

[111] M. J. Buckingham, B. V. Berknout, and S. A. Glegg, "Imaging the ocean with ambient noise," *Nature*, vol. 356, no. 6367, p. 327, 1992.

[112] J. R. Potter and M. Chitre, "Ambient noise imaging in warm shallow seas; second-order moment and model-based imaging algorithms," *The Journal of the Acoustical Society of America*, vol. 106, no. 6, pp. 3201–3210, 1999.

[113] M. Chitre, S. Kuselan, and V. Pallayil, "Ambient noise imaging in warm shallow waters; robust statistical algorithms and range estimation," *The Journal of the Acoustical Society of America*, vol. 132, no. 2, pp. 838–847, 2012.

[114] C. M. A. Verlinden, J. Sarkar, W. S. Hodgkiss, W. A. Kuperman, and K. G. Sabra, "Passive acoustic source localization using sources of opportunity," *The Journal of the Acoustical Society of America*, vol. 138, no. 1, pp. EL54–EL59, 2015.

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

[116] H. Niu, E. Reeves, and P. Gerstoft, "Source localization in an ocean waveguide using supervised machine learning," *The Journal of the Acoustical Society of America*, vol. 142, no. 3, pp. 1176–1188, 2017.

[117] H. Niu, E. Ozanich, and P. Gerstoft, "Ship localization in santa barbara channel using machine learning classifiers," *The Journal of the Acoustical Society of America*, vol. 142, no. 5, pp. EL455–EL460, 2017.

[118] J. Q. D. Tran and W. S. Hodgkiss, "Array surveying using matched field processing," *The Journal of the Acoustical Society of America*, vol. 94, no. 5, pp. 2851–2858, 1993.

[119] W. S. Hodgkiss, D. E. Ensberg, J. J. Murray, G. L. D'Spain, N. O. Booth, and P. W. Schey, "Direct measurement and matched-field inversion approaches to array shape estimation," *IEEE Journal of Oceanic Engineering*, vol. 21, pp. 393–401, Oct 1996.

[120] M. G. Morley, S. E. Dosso, and N. R. Chapman, "Array element localization using ship noise," *The Journal of the Acoustical Society of America*, vol. 125, no. 3, pp. 1403–1409, 2009.

[121] G. J. Heard, M. McDonald, N. R. Chapman, and L. Jaschke, "Underwater light bulb implosions: a useful acoustic source," in *OCEANS '97. MTS/IEEE Conference Proceedings*, vol. 2, pp. 755–762 vol.2, 1997.

[122] S. E. Dosso and N. E. B. Collison, "Acoustic tracking of a freely drifting sonobuoy field," *The Journal of the Acoustical Society of America*, vol. 111, no. 5, pp. 2166–2177, 2002.

[123] N. E. B. Collison and S. E. Dosso, "Acoustic tracking of a freely drifting sonobuoy field: experimental results," *IEEE Journal of Oceanic Engineering*, vol. 28, pp. 544–551, July 2003.

[124] A. Plinge, F. Jacob, R. Haeb-Umbach, and G. A. Fink, "Acoustic microphone geometry calibration: An overview and experimental evaluation of state-of-the-art algorithms," *IEEE Signal Processing Magazine*, vol. 33, pp. 14–29, July 2016.

[125] I. McCowan, M. Lincoln, and I. Himawan, "Microphone array shape calibration in diffuse noise fields," *IEEE Transactions on Audio, Speech, and Language Processing*, vol. 16, pp. 666–670, March 2008.

[126] R. Biswas and S. Thrun, "A distributed approach to passive localization for sensor networks," in *In Proceedings of the National Conference on Artificial Intelligence (Vol. 20, No. 3, p. 1248).*, 2005.

[127] S. Thrun, "Affine structure from sound," in *Advances in Neural Information Processing Systems 18 Neural Information Processing Systems*, pp. 1353–1360, 2005.

[128] M. Pollefeys and D. Nister, "Direct computation of sound and microphone locations from time-difference-of-arrival data," in *2008 IEEE International Conference on Acoustics, Speech and Signal Processing*, pp. 2445–2448, March 2008.

[129] J. Wendeberg, T. Janson, and C. Schindelhauer, "Self-localization based on ambient signals," *Theoretical Computer Science*, vol. 453, pp. 98–109, 2012.

[130] S. Burgess, Y. Kuang, J. Wendeberg, K. Åström, and C. Schindelhauer, "Minimal solvers for unsynchronized tdoa sensor network calibration," in *International Symposium on Algorithms and Experiments for Sensor Systems, Wireless Networks and Distributed Robotics*, pp. 95–110, Springer, 2013.

[131] S. Zhayida, S. S. Rex, Y. Kuang, F. Andersson, and K. Åström, "An automatic system for acoustic microphone geometry calibration based on minimal solvers," *arXiv preprint arXiv:1610.02392*, 2016.

[132] R. Pinkel, M. Goldin, J. Smith, O. Sun, A. Aja, M. Bui, and T. Hughen, "The wirewalker: A vertically profiling instrument carrier powered by ocean waves," *Journal of Atmospheric and Oceanic Technology*, vol. 28, no. 3, pp. 426–435, 2011.

[133] F. B. Jensen, W. A. Kuperman, M. B. Porter, and H. Schmidt in *Computational Ocean Acoustics*, ch. 6, Springer, 2011.

[134] A. Viterbi, "Error bounds for convolutional codes and an asymptotically optimum decoding algorithm," *IEEE Transactions on Information Theory*, vol. 13, pp. 260–269, April 1967.

[135] K. S. Arun, T. S. Huang, and S. D. Blostein, "Least-squares fitting of two 3-d point sets," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. PAMI-9, pp. 698–700, Sept 1987.

[136] M. A. Fischler and R. C. Bolles, "Random sample consensus: A paradigm for model fitting with applications to image analysis and automated cartography," *Commun. ACM*, vol. 24, pp. 381–395, June 1981.

[137] F. A. Everest, R. W. Young, and M. W. Johnson, "Acoustical characteristics of noise produced by snapping shrimp," *The Journal of the Acoustical Society of America*, vol. 20, no. 2, pp. 137–142, 1948.