

UC San Diego

UC San Diego Electronic Theses and Dissertations

Title

Self-adaptive processes for the mitigation of coherent multipath in ocean acoustics

Permalink

<https://escholarship.org/uc/item/0vz5h4kh>

Author

Higley, William J.

Publication Date

2007

Peer reviewed|Thesis/dissertation

UNIVERSITY OF CALIFORNIA, SAN DIEGO

Self-adaptive Processes for the
Mitigation of Coherent Multipath
in Ocean Acoustics

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

in

Electrical Engineering (Applied Ocean Sciences)

by

William J. Higley

Committee in charge:

Professor William A. Kuperman, Chair
Professor Laurence Milstein, Co-chair
Professor William A. Coles
Professor Bruce Cornuelle
Professor William S. Hodgkiss

2007

Copyright

William J. Higley, 2007

All rights reserved

The dissertation of William J. Higley is approved, and it is acceptable in quality and form for publication on microfilm:

Co-chair

Chair

University of California, San Diego

2007

TABLE OF CONTENTS

Signature Page	iii
Table of Contents	iv
List of Figures	vii
List of Tables	xi
Acknowledgments	xii
Vita and Publications	xiii
Abstract of the Dissertation	xiv
Chapter 1 – Overview	1
1.1 Introduction	1
1.2 Preliminaries	2
1.3 Oceanic Propagation	9
1.4 Time Reversal	13
1.5 Passive Time Reversal	21
1.6 Communications using Time Reversal	24
1.7 Limitations of Time Reversal	29
1.8 Minimum Mean-Squared Error Linear Equalizer	31
1.9 Information Theory	34
1.10 Principal Component Analysis	42
1.11 Summary	45
References	46
Chapter 2 – The relationship between time reversal and linear equalization in digital communications	49
2.1 Introduction	49
2.2 Description of Iterative Time Reversal	50
2.3 Convergence	56

2.4 Conclusion	57
2.5 Acknowledgments.....	57
References.....	58
Chapter 3 – Synthetic aperture time-reversal communications in shallow water:	
Experimental demonstration at sea	59
3.1 Introduction.....	59
3.2 Experimental Details.....	61
3.3 Data Processing.....	64
3.3.A Decorrelation along the track.....	64
3.3.B Extracting bit error rates.....	65
3.3.C Results	73
3.3.D Interleaving communication sequences	74
3.4 Vertical Summation	75
3.5 Conclusion	78
3.6 Acknowledgements.....	78
References.....	81
Chapter 4 – Constellation-constrained capacity of an undersea multiple-input/multiple-	
output (MIMO) channel.....	82
4.1 Introduction.....	82
4.2 Channel Model and Data Acquisition.....	84
4.3 Value of Channel Knowledge.....	86
4.4 Constellation-constrained Capacity	91
4.5 Optimal Transducer Distribution	95
4.6 Conclusion	99
4.7 Acknowledgements.....	100
References.....	101

Chapter 5 – Detection of resonances on a spherical target.....	103
5.1 Introduction.....	103
5.2 Freespace Simulation.....	104
5.3 Waveguide Simulation.....	114
5.4 Experimental Proposal.....	119
5.4 Conclusion.....	119
References.....	121

LIST OF FIGURES

Figure 1.1: a) When a source transmits a signal to a distant receiver in freespace, there is only one path connecting the two points. b) In a multipath environment such as a waveguide, interaction of the sound with the surface and seafloor create multiple paths between the source and the receiver.4

Figure 1.2: a) A pulse (shown) is transmitted from an ultrasonic hydrophone in an underwater waveguide, using a tank set-up described in Appendix A of [6]. b) The received signal, when a pulse is transmitted, is composed of a number of attenuated and time-delayed copies of the original pulse.10

Figure 1.3: a) A pulse is transmitted from an ultrasonic hydrophone in an underwater waveguide, using a tank set-up described in Appendix A of [6]. The received field on a vertical array, arranged a meter from the source, is shown. b) The received field on a horizontal endfire array.12

Figure 1.4: A diagram showing the time reversal process. Initially, a probe source (PS) transmits a pulse, which is recorded on a vertically-aligned source/receiver array (SRA). This signal is time-reversed and back-propagated. The received field on a vertical receiver array (VRA) at the range of the initial probe source is a focus.14

Figure 1.5: The time reversal focus achieved using a single source and receiver. The peak at time zero indicates that the energy has been recompressed into a pulse, although in this case the sidelobes (everything except the peak at time zero) are rather large.15

Figure 1.6: a) The time reversal focus achieved using a single source and a vertical source receiving array (SRA). The received field on the SRA is shown in Fig. 1.3a. b) The time reversal focus achieved using a single source and a horizontal endfire SRA.16

Figure 1.7: a) The time reversal focus, as measured by a vertical receiving array (VRA) at the range of the source, using a vertically-aligned source/receiver array (SRA). b) The time reversal focus, as measured by a VRA at the range of the source, using a horizontally-aligned endfire SRA.18

Figure 1.8: Diagram illustrating passive time reversal.23

Figure 1.9: a) A communications sequence is transmitted from an ultrasonic hydrophone in an underwater waveguide. b) The received signal when the communications sequence of Fig. 1.9a is sent using the impulse response of Fig. 1.2b. ...25

Figure 1.10: Time reversal as used for communications.26

Figure 1.11: a) The received signal after the application of the time reversal process using a single source and single receiver. b) The received signal after the application of the time reversal process using a vertical source/receiver array (SRA).....	28
Figure 1.12: Passive time reversal communications is accomplished by first transmitted the communications sequence preceded by a pulse. Thus, the received signal appears as in the above figure, as the impulse response followed by the communications sequence corrupted by the same impulse response.	30
Figure 2.1: Convergence of iterative time reversal.....	54
Figure 3.1: a) A vertical receiver array (VRA) moored off the coast of Italy received signals from a transducer towed by a moving vessel following the path shown. Each X represents a position in the almost 2-km long synthetic aperture array.....	61
Figure 3.2: Point-to-point time-reversal sidelobe decorrelation. These results are obtained by first normalizing the Green's functions at the distances along the track indicated on the axis, autocorrelating each of them, and then cross-correlating the resulting sidelobes.....	64
Figure 3.3: The results of summing coherently the 65 received waveforms before Doppler correction (in gray) and after Doppler correction (in black). As time progresses, it is shown that the amplitude of the non-corrected sum decreases as the non-corrected signals do not add coherently at later times.....	65
Figure 3.4: The received waveforms before and after synchronization. Synchronization is an issue as time-reversal can't be performed simultaneously from the synthetic aperture array. The signals must be accurately synchronized before time-reversal so they may be summed coherently.....	66
Figure 3.5: a) A typical received signal consisting of an initial pulse and a 10 s communication sequence, after both have propagated through the medium. The selected Green's function is shown in black. b) The same signal, zoomed in to show the detail of the Green's function.	67
Figure 3.6: The Q-function [10] resulting from a single element is shown in black, and from the coherent summation of five widely-spaced elements is shown in gray. It can be seen that the summation results in lower sidelobe levels, which produce lower bit error rates.	69
Figure 3.7: a) The results of incoherently decoding a signal composed of the sum of three widely-spaced synthetic aperture elements. b) The results of decoding the same signal coherently.	70
Figure 3.8: The evolution of the in-phase/quadrature plot as more synthetic elements are summed together before decoding. The elements added are widely-spaced. Binary 0s are in black and binary 1s are in gray.....	71

Figure 3.9: Bit error rate found from 9800 bits, each with a bit length of 1 ms, as a function of the number of transducer positions (or synthetic array elements) for various values of element spacing, δRT .	73
Figure 3.10: Schematic of a synthetic aperture time reversal communication system. For the sake of simplicity, the communication sequences are represented by the symbols X, O, and +.	74
Figure 3.11: The evolution of the in-phase/quadrature plot as more vertical elements are summed together before decoding. Before summation over the vertical elements, a horizontal synthetic aperture of three widely-spaced elements is used.	75
Figure 3.12: Bit error rate found from 9800 bits, each with a bit length of 1 ms, as a function of the number of synthetic array elements, element spacing, δRT , and number of vertical array elements. The bit error rates using only one vertical element can be found in Figure 3.9.	76
Figure 4.1: a) A diagram showing the experimental setup. b) A typical point-to-point impulse response of an underwater acoustic communications experiment performed off the coast of Italy. The transfer function is characterized by a long dispersion time (approximately 70 symbols) and non-Gaussian nature.	85
Figure 4.2: Channel capacity with (solid) and without (dashed) water-filling for various transmitter/receiver sets as a function of input SNR.	87
Figure 4.3: Channel capacity with (solid) and without (dashed) water-filling at a fixed input SNR of 5 dB.	88
Figure 4.4: Upper (solid) and lower (dashed) bounds for the i.i.d. constellation-constrained capacity for various transmitter/receiver sets as a function of input SNR.	94
Figure 4.5: Upper (solid) and lower (dashed) bounds for the i.i.d. constellation-constrained capacity at a fixed input SNR of 5 dB.	95
Figure 4.6: The effect of transducer placement. Ten transmitters are considered in total, with the number of receivers on the abscissa. The number of transmitters is ten less the number of receivers.	96
Figure 5.1: The geometry of the simulation. A vertical array consisting of nine elements is oriented vertically in freespace. The distance between each element is 10 m, and the first element is considered 10 m below some arbitrary origin. The target is located 200 m away from the vertical array.	104
Figure 5.2: The two targets, both located in an aqueous medium.	104

Figure 5.3: The backscattered response on the vertical array from both the bubble (gray) and the shell (black) when excited by a single source on the array.105

Figure 5.4: The received energy across frequency of the backscattered response of the bubble (gray) and the shell (black), averaged across source and receiver. A few of the resonant peaks of the shell are seen at lower frequencies.....107

Figure 5.5: The first principal component from the bubble (gray) and the shell (black). The coherent processing first yields a field, which is shown in these figures. .109

Figure 5.6: The result of the coherent processing on the bubble (gray) and the shell (black). The periodic resonance peaks of the shell are quite pronounced through the frequency range, whereas the response of the bubble is relatively flat, as one would expect for a specular component.110

Figure 5.7: The result of back-propagating the field extracted in Fig 5.5b. The focus is seen only at the location of the target, at a depth of 100 m and a range of 200 m.....111

Figure 5.8: a) In gray is the field previously extracted from the shell and shown in Fig. 5.5b. The windowed portion, shown in black, corresponds to a resonant emission from the target. b) The result of back-propagating the resonant emission.112

Figure 5.9: The back-scattered response from a flush-buried bubble (gray) or shell (black) in a waveguide. The shape of the response is the same in both cases, but each pulse of the response in the case of the bubble is turned into a resonant arrival structure in the case of the shell.....114

Figure 5.10: a) The result of the incoherent (gray) and coherent (black) processing on the non-resonant bubble in a waveguide. b) The result of the incoherent (gray) and coherent (black) processing on the resonant shell in a waveguide.115

Figure 5.11: A comparison of the incoherent (gray) and coherent (black) processing on the flush-buried shell for four different signal-to-noise ratios. Although both methods degrade as noise increases, the coherent method still shows peaks at -10 dB, whereas the incoherent method fails at this SNR.116

Figure 5.12: a) Sample variance of the frequency vectors corresponding to the incoherent processing (gray) and coherent processing (black). b) The sample kurtosis, which is high for peaky vectors, of the frequency vectors corresponding to the incoherent processing (gray) and coherent processing (black).....118

LIST OF TABLES

Table 1.1: A comparison of electromagnetic (EM) and ocean acoustic (OA) propagation, as a way of understanding the source of many unique problems in underwater acoustics.....	3
Table 3.1: Selected values of the bit error rate found from 9800 bits, each with a bit length of 1 ms, as a function of the number of transducer positions (or synthetic array elements) for various values of element spacing, δRT	78
Table 3.2: Selected values of the bit error rate found from 9800 bits, each with a bit length of 1 ms, summed over three vertical elements, as a function of the number of transducer positions (or synthetic array elements) for various values of element spacing, δRT	79

ACKNOWLEDGEMENTS

The text in this dissertation, in part, was originally published in the following papers, of which the dissertation author was the primary researchers and author:

The following manuscript appears in chapter 2:

W.J. Higley, P. Roux, and W.A. Kuperman, "Relationship between time reversal and linear equalization in digital communications," *J. Acous. Soc. Am.*, vol. 120, pp. 35-37, 2006.

The following manuscript appears in chapter 3:

W.J. Higley, P. Roux, W.A. Kuperman, W.S. Hodgkiss, H.C. Song, T. Akal, and M. Stevenson, "Synthetic aperture time-reversal communications in shallow water: Experimental demonstration at sea," *J. Acous. Soc. Am.*, vol. 118, pp. 2365-2372, 2005.

The following manuscript appears in chapter 4:

W.J. Higley, P. Roux, W.A. Kuperman, W.S. Hodgkiss, H.C. Song, T. Akal, and M. Stevenson, "Constellation-constrained capacity of an undersea multiple-input/multiple-output (MIMO) channel," submitted to *IEEE J. Ocean. Eng.*, 2006.

VITA

March 20, 1981	Born, Long Beach, California
2002	B.S. University of California, San Diego
1992-2007	Research Assistant, University of California, San Diego
2007	Ph.D., University of California, San Diego

PUBLICATIONS

W.J. Higley, P. Roux, and W.A. Kuperman, "Relationship between time reversal and linear equalization in digital communications," *J. Acous. Soc. Am.*, vol. 120, pp. 35-37, 2006.

W.J. Higley, P. Roux, W.A. Kuperman, W.S. Hodgkiss, H.C. Song, T. Akal, and M. Stevenson, "Synthetic aperture time-reversal communications in shallow water: Experimental demonstration at sea," *J. Acous. Soc. Am.*, vol. 118, pp. 2365-2372, 2005.

W.J. Higley, P. Roux, W.A. Kuperman, W.S. Hodgkiss, H.C. Song, T. Akal, and M. Stevenson, "Constellation-constrained capacity of an undersea multiple-input/multiple-output (MIMO) channel," submitted to *IEEE J. Ocean. Eng.*, 2006.

ABSTRACT OF THE DISSERTATION

Self-adaptive Processes for the
Mitigations of Coherent Multipath
in Ocean Acoustics

by

William J. Higley

Doctor of Philosophy in Electrical
Engineering (Applied Ocean Sciences)
University of California, San Diego, 2007
Professor William A. Kuperman, Chair
Professor Laurence Milstein, Co-chair

Because electromagnetic waves are strongly attenuated in the ocean, researchers have turned to the use of acoustics for such problems as communications and detection. Acoustic propagation in the ocean is dominated by interactions with the surface and seafloor, leading to complicated arrivals structures. The computing power needed to model this environment makes the use of adaptive processes such as time-reversal an attractive alternative. Further, the use of self-adaptive processes, which use the data itself to influence the processing, have proven to be both robust and powerful. This thesis concerns itself with four such processes. The first derives the canonical minimum mean-squared error linear equalizer (MMSE-LE) as a self-adaptive process related to iterative time reversal. The next chapter analyzes the feasibility of using a horizontally-

aligned synthetic aperture to perform time reversal communications. Also derived in this thesis are bounds on the capacity of a strongly dispersive channel, such as those found in the ocean, with practical constellation-based constraints on the transmitted signal. A final self-adaptive process is described in an attempt to classify a target as resonant or non-resonant based on the measured back-scattered field.

Chapter 1

Overview

1.1 Introduction

Electromagnetic waves such as visible light and the higher-frequency waves used by the typical cell phone are rapidly attenuated in the water. It is estimated, for example, that only 15% of sunlight reaches more than 10m below the surface. On the other hand, acoustic waves, such as the sound of a humpback whale, can propagate great distances, such as that of an entire ocean basin. Because the speed of sound in water is much slower than the speed of light, new problems are introduced not common with electromagnetic waves. This necessitates the introduction of different processing on the data. This thesis concerns itself with four such processes, designed to take advantage of the differences the use of acoustic introduces.

The first of these processes expands the use of a robust and powerful adaptive process known as time-reversal [1]. Chapter 2 examines the self-adaptive processes of iterative time reversal and the canonical minimum mean-squared error linear equalizer (MMSE-LE). Although, neither of these tools is new to this thesis, the chapter shows an unseen equivalence between the two and allows a well-known solution (the MMSE-LE) to be seen in a new light, as a self-adaptive process. The next chapter examines the feasibility of synthetic aperture time reversal as a communications tool. Typically, time reversal requires a large array of transducers, which can be expensive. Synthetic aperture time reversal involves only one transmitter and one receiver. Chapter 4 fills a void in the literature on how to calculate the information theoretic capacity of a strongly-dispersive

multiple-input/multiple-output (MIMO) channel, such as those used in oceanic experiments. The final chapter introduces a self-adaptive process for the detection and classification of targets based on the resonance of a sphere.

1.2 Preliminaries

Although a great deal of productive research has been done in the last twenty-five years in the field of wireless communications, this research is not immediately applicable to the problem of underwater telecommunications because the electromagnetic (EM) waves traditionally used in such wireless communications are quickly attenuated as they propagate through water. Fortunately, as mentioned, acoustic waves are able to propagate long distances in the ocean with minimal attenuation. This motivates the use of acoustic waves, as opposed to electromagnetic waves, as a means of digital communications in the underwater environment.

Acoustic waves are also used instead of electromagnetic waves in the underwater environment with respect to other problems. Detection, for instance, is accomplished using EM waves in the form of RADAR (RADio Detecting And Ranging) in air, and with acoustic waves in the form of SONAR (SOund Navigating And Ranging) underwater.

The underwater utilization of the research performed for in-air applications is not as simple as substituting sound for light. To better understand the unique problems involved in underwater acoustic problems, we compare EM propagation and acoustic propagation in a shallow water oceanic waveguide. A numerical comparison is shown in Table 1.1. Many of the unique problems of underwater acoustic arise from the difference in the speeds of propagation of electromagnetic waves and acoustic waves. Light propagates at

Table 1.1: A comparison of electromagnetic (EM) and ocean acoustic (OA) propagation, as a way of understanding the source of many unique problems in underwater acoustics.

Quantity	Electromagnetic Wave	Ocean Acoustic Wave
Propagating speed (c)	3×10^8 m/s	1.5×10^3 m/s
Carrier frequency (f_c)	10^9 Hz	3×10^3 Hz
Wavelength, $\lambda = c/f_c$	0.3 m	0.5 m
Attenuation in seawater	10 dB/m	2×10^{-4} dB/m
Velocity of mobile/Mach number, $M = v/c$	$v = 65$ mile/hour = 30 m/s $M_{radio} = v/c = 10^{-7}$	$v = 10$ kts = 5 m/s $M_{water} = v/c = 3 \times 10^{-3} > 10^4 M_{radio}$
Doppler shift for a path, $f_d = f_c M$	$f_d = 100$ Hz	$f_d = 10$ Hz Pulse expansion/contraction
Delay spread, T_d Intersymbol interference (ISI), $T_d \times W$	1–2 μs 1–2 symbols	Up to 100 ms 50-100 symbols

3×10^8 m/s, over five orders of magnitude greater than the speed of sound in water, which averages 1.5×10^3 m/s. Because of this difference, the frequencies corresponding to wavelengths of interest change. Typical wireless communications frequencies center around 1 GHz, whereas an underwater communications experiment recently performed [2] centered around 3.5 kHz. In both cases, the wavelength (speed divided by frequency) is between a third of a meter and half a meter. The bandwidth used in wireless communications might be 1 MHz, which is a thousandth of the center frequency. In contrast, the bandwidth used in underwater applications might be 1 kHz, which is a third of the center frequency, making underwater acoustic *ultrawideband* compared to wireless communications. It is interesting to note that the trend in wireless communications research seems to be heading more and more towards being considered ultrawideband, allowing underwater acoustic research the chance to further a field that has been such an influence on it. The speed of propagation also influences Doppler shift and spread

[3],[4]. A person driving with their cell phone can still only move a tiny fraction of the speed of light, say a hundredth of a thousandth of a percent. A submarine can, however, move a greater fraction of the speed of sound, in the range of a few tenth of a percent, which may seem small, but is four orders of magnitude greater than in electromagnetics. At this much larger value, Doppler effects have a large influence on communications requiring additional signal processing to compensate for them. Some problems with Doppler underwater, such as when each path suffers from a different Doppler, are as yet unsolved.

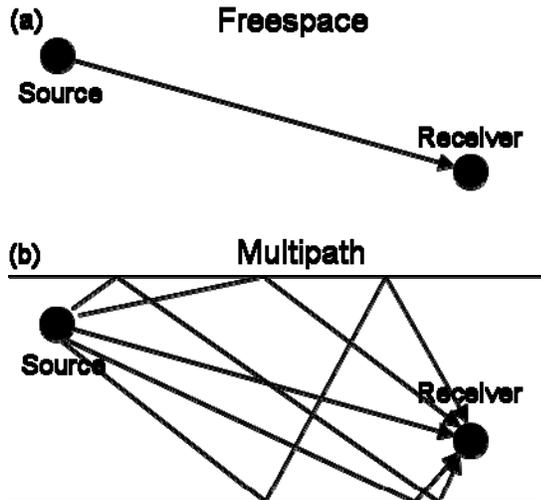


Figure 1.1: a) When a source transmits a signal to a distant receiver in freespace, there is only one path connecting the two points. Thus, the received signal is identical to the transmitted signal, except, perhaps, for an attenuation and/or time-delay. b) In a multipath environment such as a waveguide, interaction of the sound with the surface and seafloor create multiple paths between the source and the receiver. In this case, the received signal receives multiple copies of the transmitted signal, each with its own independent attenuation and time-delay.

With a single source and a single receiver in freespace, there is only one path that connects the two points. Thus, when a signal is transmitted in freespace, the signal at a distant receiver is identical to the transmitted signal, except the received signal may be

attenuated or time-delayed. In contrast, when a signal is transmitted in a shallow water ocean waveguide, there are multiple paths between the source and receiver, as the sound is reflected off the surface and the seafloor. This fundamental difference is diagramed in Fig. 1.1. In an ocean waveguide, multiple copies of a transmitted signal are received at the receiver due to these echoes. Such an environment is called a *multipath* environment, due to the presence of multiple paths between the source and receiver. A pulse, confined to a single time instance, transmitted in a multipath environment, is received at multiple times due to the different paths of propagation. Because the energy in a signal from a single time instance is dispersed to many different times upon reception, we also call such an environment *dispersive*.

Although multipath can occur in both EM and acoustic propagation, the difference in propagation speeds makes multipath a much greater issue in underwater acoustics. In wireless communications, for instance, when using a cell phone from a car traveling the downtown streets of a sprawling metropolis, the multiple paths from source to receiver include the ground and nearby buildings. However, the difference in path lengths may be on the order of a few meters, which is traveled by light in a hundredth of a microsecond. A microsecond-long pulse might be stretched into two microseconds. Such a medium is weakly dispersive. In underwater acoustics, however, the difference in path lengths caused by reflection off the surface and bottom are on the order of hundreds of meters, which sound transverses much more slowly than light. A millisecond-long pulse might be stretched into an arrival structure of a hundred milliseconds or more. The shallow water ocean waveguide is a strongly dispersive medium.

In certain environments, the multipath propagation is rapidly changing and unpredictable. Again, we used the example of using a cell phone from a car, where the multiple paths include the ground and nearby buildings. The paths rapidly change as the source moves, and are unusable. In the ocean, however, when the source and receiver are fixed, the paths change much less rapidly. Because the multiple paths are much the same from use to use, we call this a *coherent* multipath environment.

The presence of multipath in the environment complicates the accomplishment of many tasks that are simple to do in freespace. For example, when attempting communications in the ocean, the desired signal at any given time is corrupted by echoes of the signal at earlier times. Similarly, when attempting to extract resonances from a backscattered field of an object we wish to detect, the resonance is corrupted by multiple overlapping copies of the same resonance from earlier arrivals. In the frequency domain, we have difficulty distinguishing between peaks caused by the resonating object and peaks caused by the interference of different paths.

The processes designed to cope with coherent multipath in an environment can be classified as either *adaptive* or non-adaptive. A non-adaptive process does not utilize the data itself to accomplish its goals. Consider the problem of focusing sound at the location of a pulsating source in a waveguide. The received field is dependant on the location of the source as described by the physics of the waveguide, and these physics are well-known. If we design a model of the waveguide (knowing depth, attenuation properties, etc.) we can guess what the field would look like from a source at every point in the waveguide. Comparing the received field to all of our guesses (so-called replica vectors [5]), and picking the guess that is most similar to the data, we also have a guess as

to the location of the source. We can then do additional processing, knowing the physics of the waveguide, to design a field that focuses well at this point. This is an example of non-adaptive processing. In a different waveguide (i.e. of a different depth), we would have to start over, remodel and redesign our back-propagated field. An example of adaptive processing is time reversal.

To accomplish the same task (focusing sound at the location of a pulsating source) using time reversal, we record the received field, time-reverse it (that is, first send out the sound that was received last), and retransmit it. The result is a focus at the location of the original pulsating source. If we were in a different waveguide, our processing is the same: we record the received field, time-reverse it, and retransmit. In this way, our process adapts itself to the environment. Often, a non-adaptive process requires some sort of modeling, whereas an adaptive process relies only on the data to accomplish its goals. Further, time reversal is an example (albeit a rudimentary one) of a *self-adaptive* process. A self-adaptive process is one that uses the data itself to perform the processing on the data. In the case of time reversal, it is the data itself that is time-reversed and retransmitted.

An example of an adaptive (but not self-adaptive) process is the design of a filter to mitigate multipath using the genetic algorithm. We take the received field and pump it through a bank of random filters, choosing which results we like best and creating a new bank of filters similar to these. We again pump the received field through these filters and again choose those filters with results we like best. By iterating this process, we will eventually obtain a filter with very good results. Interestingly, the same goal can be accomplished self-adaptively as will be described in a later chapter.

The rest of the chapter is organized as follows: oceanic acoustic propagation will be described mathematically in the time-domain, revealing that the introduction of multiple paths can be described as a convolution. The transformation to the frequency-domain is facilitated by the introduction of modal propagation. Time reversal, which is a robust and powerful method of utilizing coherent multipath, is described, with a time-domain (and slightly heuristic) explanation why it works to create of temporal focus. The difference between a vertically-aligned and a horizontally-aligned array are discussed, and the properties of the time reversal focus achieved with both arrays are derived mathematically using modal arguments. The limitations of time reversal are also discussed with the difference between passive and active time reversal introduced as a way of mitigating these limitations. Another method of utilizing multipath, the minimum mean-squared error linear equalizer (MMSE-LE), is introduced. The goal of communications is articulated and the difference between coherent and incoherent communications is made clear.

The mathematical theory of communications, information theory, is briefly described. The concept of entropy is introduced, and it is proven that the variable with the maximum entropy is a Gaussian. Mutual information and capacity are also described. The capacity for a memoryless channel is classically derived. Further, the derivation for the capacity for a dispersive channel, like a ocean waveguide, is also shown. This derivation hinges on the concept of water-filling, which is explained and derived in section 1.9.

1.3 Oceanic Propagation

As stated above, when a signal is transmitted in the ocean, interaction with the surface and seafloor create a multipath channel where multiple copies of the signal are received. We now look at this phenomenon mathematically. We consider a single source hydrophone and a single receiver hydrophone placed at some range, R , away from the source. We can examine propagation in either the frequency domain or the time domain, each leading to a different insight into the nature of propagation.

As the time-domain source signal, $x(t)$, propagates, it interacts with the surface and bottom of the waveguide resulting in a number of copies of the signal at the receiver. Each copy is attenuated by interaction with the bottom and increasing path length, and is time-delayed due to the finite speed of sound. As each path between the source and receiver has different interaction with the bottom and a different distance traveled, the attenuation and time-delay are different for each path. We can write the input-output relationship in the time domain as:

$$y(t) = \sum_p a_p x(t - \tau_p), \quad (1.1)$$

where $x(t)$ and $y(t)$ are the input and output signals at time t , and a_p and τ_p are the attenuation and time-delay of each path, p . For example, if the pulse shown in Fig. 1.2a is transmitted, the received signal takes the shape of the signal shown in Fig. 1.2b.

We can rewrite this equation in discrete time as

$$y[n] = \sum_p h[p] x[n - p], \quad (1.2)$$

where $x[n]$ and $y[n]$ are the input and output signals at discrete time n , and $h[n]$ is the impulse response of the ocean channel. We recognize the form of this equation as a

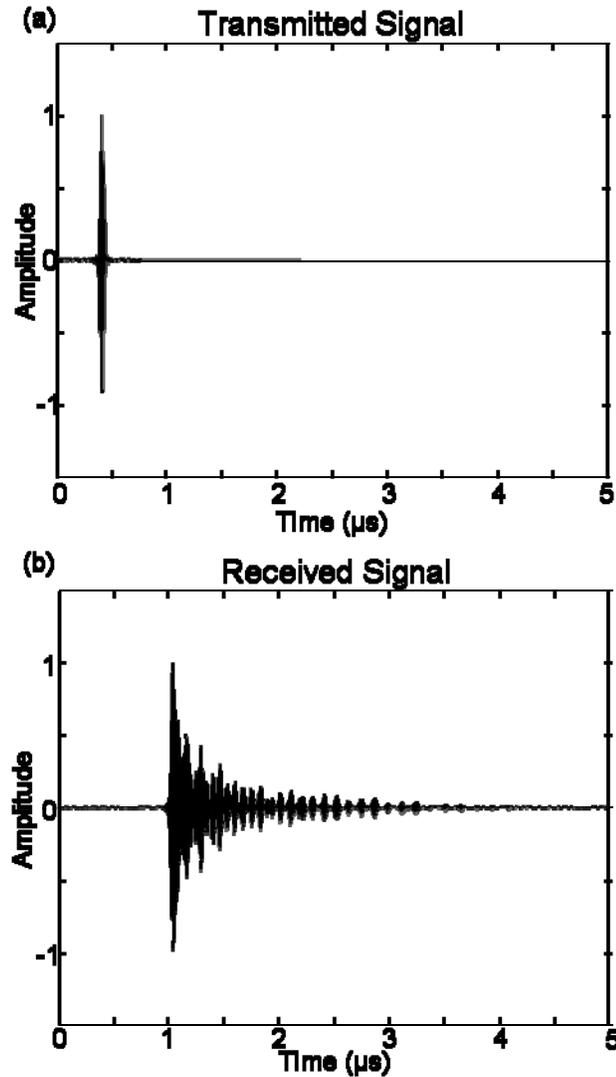


Figure 1.2: a) A pulse (shown) is transmitted from an ultrasonic hydrophone in an underwater waveguide, using a tank set-up described in Appendix A of [6]. This pulse propagates along a number of paths before being received at a receiving hydrophone. b) The received signal, when a pulse is transmitted, is composed of a number of attenuated and time-delayed copies of the original pulse. In this case, the energy of the pulse is dispersed 3 microseconds.

convolution of the input signal and the impulse response. We can identify the signal received in Fig. 1.2b as the impulse response of the ocean channel from a single source to a single receiver. The form of Eq. (1.2) as a convolution motivates a transformation to the frequency domain [7],

$$y(\omega) = h(\omega)x(\omega), \quad (1.3)$$

where $x(\omega)$ and $y(\omega)$ are the input and output at frequency ω , and $h(\omega)$ is the gain at each frequency. Through this transformation, the convolution has thus become a simple multiplication. In a waveguide, the gain at each frequency has the special form [8]:

$$h(\omega) = \frac{ie^{-i\pi/4}}{\rho(z_s)\sqrt{8\pi}} \sum_n \Psi_n(z_s, \omega)\Psi_n(z_r, \omega) \frac{e^{ik_n(\omega)R} e^{-\alpha_n(\omega)R}}{\sqrt{k_n(\omega)R}}, \quad (1.4)$$

where z_s and z_r are the depths of the source and receiver hydrophone, ρ is density, k_n is the horizontal wavenumber of mode n , α_n is the attenuation, and Ψ_n is the mode shape. In effect, at each independent frequency, the source excites each independent mode, which is attenuated but otherwise unperturbed as it propagates to the receiver.

For multiple sources, labeled 1 to N_T , transmitting to multiple receivers, labeled 1 to N_R , superposition holds, yielding the generalized dual equations:

$$\begin{aligned} \mathbf{y}(\omega) &= \mathbf{H}(\omega)\mathbf{x}(\omega) \\ y_i[n] &= \sum_{j=1}^{N_T} \sum_{l=0}^L h_{ij}[l]x_j[n-l] \text{ for } 1 \leq i \leq N_R \end{aligned} \quad (1.5)$$

For example, if a pulse is transmitted from a single source to a vertical array, the field shown in Fig. 1.3a is received. If multiple pulses are transmitted from multiple transmitters, the received field at an array is simply the sum of the fields that would be received from each transmitter. In Fig. 1.3a, we can identify the reflections from the surface and bottom of the waveguide.

Also of interest is the field received on a horizontal array. When the array is oriented along the direction of propagation, such an array is termed an endfire array. Figure 1.3b shows the field received by an array oriented in such a way. At each range, however, the

received signal is shifted by the propagation distance, so that the first arrivals coincide in time.

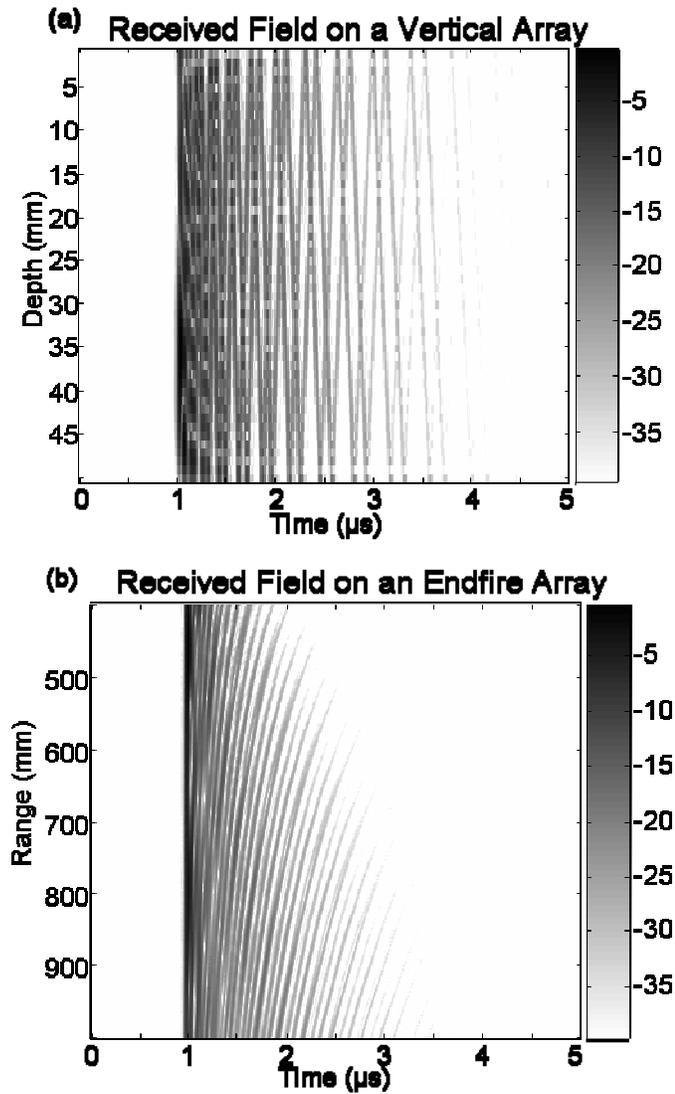


Figure 1.3: a) A pulse is transmitted from an ultrasonic hydrophone in an underwater waveguide, using a tank set-up described in Appendix A of [6]. The received field on a vertical array, arranged a meter from the source, is shown. The reflections of the sound off the surface and bottom of the waveguide are identifiable. b) The received field on a horizontal endfire array, where at each point on the array the received signal is shifted so that the first arrivals align in time. One may notice that as range increases, the amount of dispersion also increases.

1.4 Time Reversal

One powerful and proven self-adaptive technique is time reversal. The technique was first introduced in the ocean using a single transmitter and single receiver [9],[10]. More recently, the idea appeared in acoustic literature in an ultrasonic setting [11],[12]. Following this, time reversal for the use of communications (as well be further analyzed in this thesis) was suggested [13]. The application of time reversal in the ocean was performed [14],[15], with the application of time reversal to underwater communications performed not much later [2],[16]. With time reversal, the multipath dispersion of the ocean can be mitigated. We have already briefly discussed how time reversal can create a focus at the location of a pulsating source. Figure 1.4 diagrams the time reversal process as it was performed in an at-sea experiment [2].

The goal of time reversal is to create a focus in space and time at the location of a pulsating probe source (PS). To accomplish this, the field received as the pulse propagates through the ocean at a source/receiver array (SRA) is time-reversed and back-propagated through the ocean. The resulting field on a vertical receiving array (VRA) at the range of the original source is a spatio-temporal focus. We can look at this process mathematically as follows. For a single source and single receiver, the received signal is, as in Eq. (1.1) and Fig. 1.2b,

$$y(t) = \sum_p a_p x(t - \tau_p). \quad (1.6)$$

The next step in the time reversal process is to time-reverse the received signal, that is replace t with $-t$. This becomes the signal transmitted from the original receiver.

$$y(-t) = \sum_p a_p x(-t + \tau_p) \quad (1.7)$$

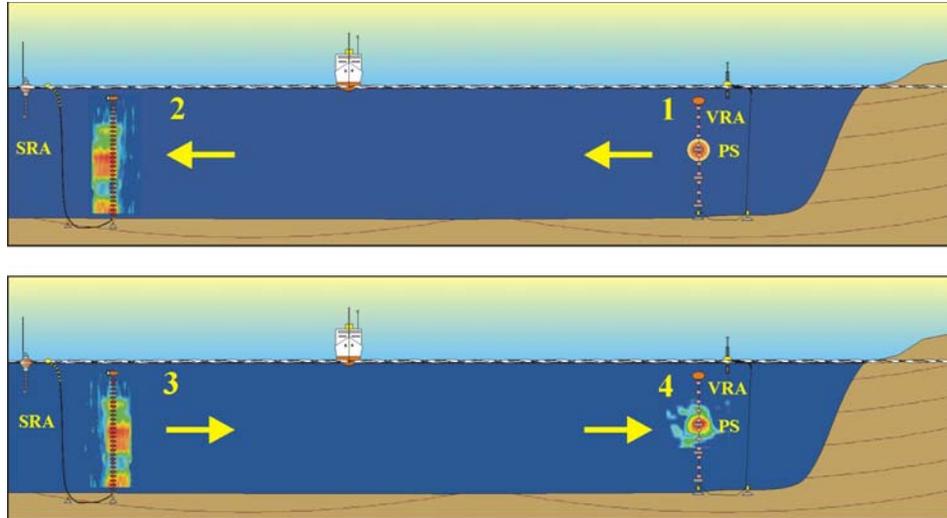


Figure 1.4: A diagram showing the time reversal process. Initially, a probe source (PS) transmits a pulse, which is recorded on a vertically-aligned source/receiver array (SRA). This signal is time-reversed, such that the later arrivals are those first emitted, and back-propagated. The received field on a vertical receiver array (VRA) at the range of the initial probe source is a spatial and temporal focus.

This signal propagates through the same ocean (along the same coherent multipaths) back to the point of the original transmitter. The resulting focus can be written as,

$$f(t) = \sum_q a_q y(-t - \tau_q) = \sum_p \sum_q a_p a_q x(-t + \tau_p - \tau_q). \quad (1.8)$$

Each time τ_p equals τ_q in the double summation, a time-reversed copy of the originally transmitted signal $x(t)$ adds coherently at zero time delay, resulting in a strong copy. All other terms of the double sum add incoherently, resulting in a background of weaker copies. Figure 1.5 shows the focus obtained with a single source/receiver pair, using the same impulse response as Fig. 1.2b.

We can perform the same analysis in the discrete time domain, revealing the matched filter nature of time reversal. The received signal from a single transmitter, in the discrete time domain can be written as, as in Eq. (1.2),

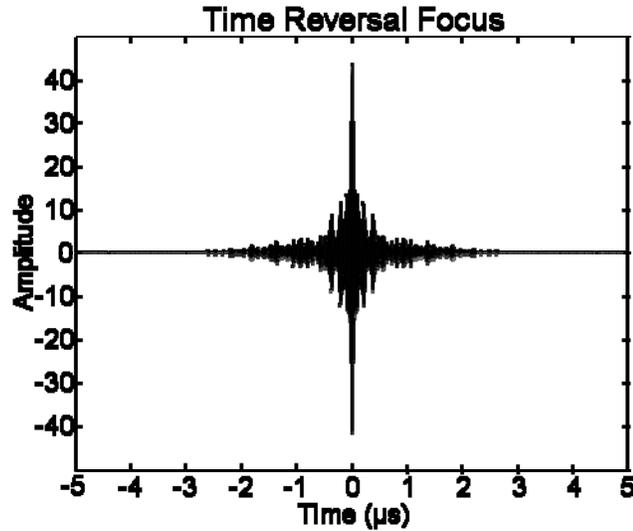


Figure 1.5: The time reversal focus achieved using a single source and receiver. The impulse response from the source to receiver (and from the receiver to source) is shown in Fig. 1.2b. The peak at time zero indicates that the energy has been recompressed into a pulse, although in this case the sidelobes (everything except the peak at time zero) are rather large.

$$y[n] = \sum_p h[p]x[n-p]. \quad (1.9)$$

which we can simplify to,

$$y[n] = h[n] * x[n], \quad (1.10)$$

where $*$ indicates convolution. Time-reversing the received signal and back-propagating it through the same ocean can be written using this notation as,

$$f[n] = h[n] * h[-n] * x[-n]. \quad (1.11)$$

We now turn our attention to an array of receivers. If a signal is transmitted to multiple receivers, to perform time reversal, the entire field is time-reversed and retransmitted. The focus at the location of the original transmitter is then written as,

$$f[n] = \left(\sum_i h_i[n] * h_i[n] \right) * x[-n]. \quad (1.12)$$

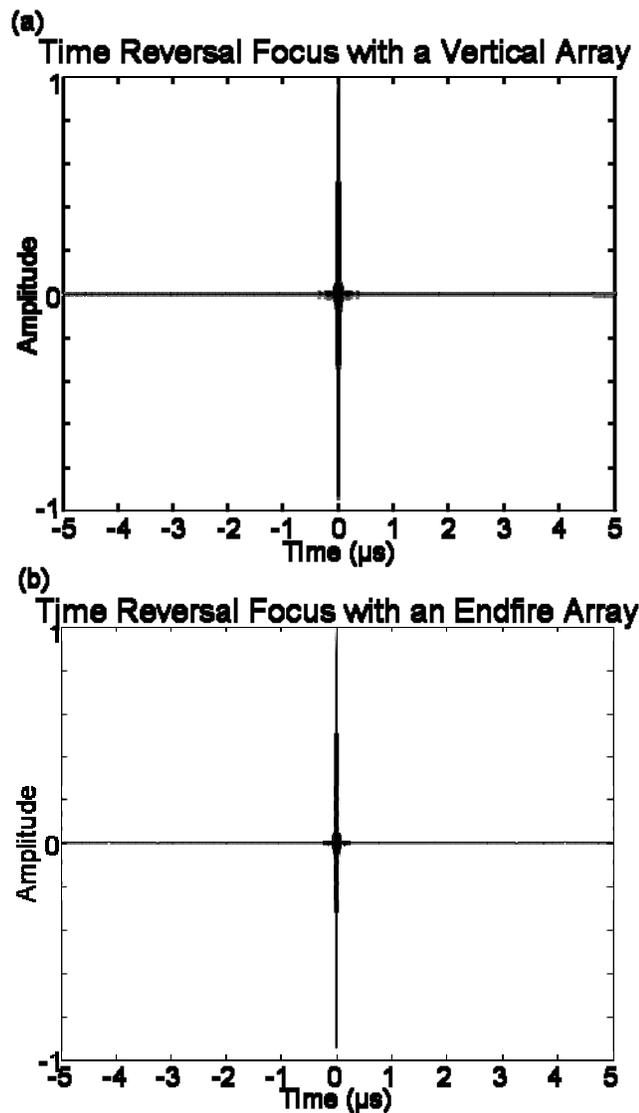


Figure 1.6: a) The time reversal focus achieved using a single source and a vertical source receiving array (SRA). The received field on the SRA is shown in Fig. 1.3a. The peak at time zero is much stronger than the sidelobes in this case, when compared to the focus of Fig. 1.5, demonstrating the value of an array in the time reversal process. b) The time reversal focus achieved using a single source and a horizontal endfire SRA. The received field on the SRA is shown in Fig 1.3b. As in the case of a vertical array, the peak at time zero is much stronger than the sidelobes, when compared to the focus of Fig 1.5.

The parenthetical term of Eq. (1.12) is deemed the Q-function [17], and is used as a measure of the temporal focusing properties of time-reversal. We recognize the Q-function as a sum of the auto-correlations of each impulse response. The results of time reversal for a vertical and endfire array are shown in Fig. 1.6. Comparing the results to Fig. 1.5, we notice that the sidelobes have been significantly reduced by using an array. Turning our attention to the interpretation of Eq. (1.8), we can see that each additional receiver-turned-transmitter adds copies coherently at zero time delay and incoherently elsewhere. This strengthens the copy of the signal at zero time delay in comparison to the sidelobes. So far, we have only looked at the focusing properties of time reversal in the time- domain. We can also examine the spatial focusing properties of time reversal. We consider a vertical array, of which one hydrophone transmits a pulse. The receiving array time-reverses the received field and transmits it back to the vertical array. The resulting time reversal focus can be written as,

$$f_j[n] = \left(\sum_i h_{i,j}[n] * h_{i,src}[-n] \right) * x[-n]. \quad (1.13)$$

The results of time reversal for a vertical and an endfire array are shown in Fig. 1.7. We notice that use of the vertical array results in a sharp focus in the spatial domain, whereas the endfire array provides no such focusing properties [18]. To explain this, we analyze the problem in the frequency domain, using waveguide physics introduced in Eq. (1.4).

We analyze time-reversal in oceanic waveguide environment with a single source and a vertical array of receivers. For this analysis, we will look at each frequency independently to take advantage of the modal solution. We write the received signal

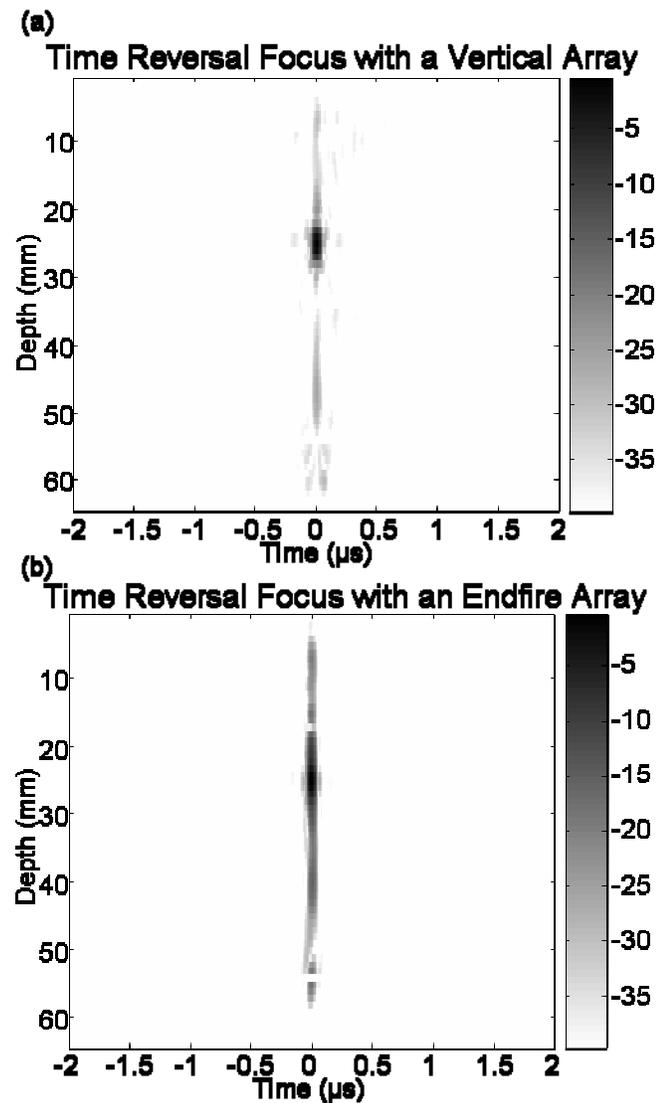


Figure 1.7: a) The time reversal focus, as measured by a vertical receiving array (VRA) at the range of the source, using a vertically-aligned source/receiver array (SRA). The received field on the SRA is shown in Fig. 1.3a. The focus is located at time zero, and at a depth consistent with the location of the probe source (PS). b) The time reversal focus, as measured by a VRA at the range of the source, using a horizontally-aligned endfire SRA. The received field on the SRA is shown in Fig 1.3b. The temporal focusing property is as was shown in Fig. 1.6b, but here, there is weaker spatial focusing than in the vertical SRA case of Fig 1.7a.

from the source at each receiver (suppressing the frequency dependence and attenuation) as,

$$y(z_r) = \frac{ie^{-i\pi/4}}{\rho(z_s)\sqrt{8\pi}} \sum_n \Psi_n(z_s)\Psi_n(z_r) \frac{e^{ik_n R}}{\sqrt{k_n R}} x. \quad (1.14)$$

The next step in the process is to time-reverse the received signal, which is realized in the frequency domain as phase-conjugation, that is replacing i with $-i$. This is retransmitted from the original receivers. The result at the range of the original transmitter can be written as

$$f(z) = \sum_{z_r} \frac{-1}{8\pi\rho(z_s)\rho(z_r)} \sum_m \sum_n \Psi_m(z_s)\Psi_m(z)\Psi_n(z_s)\Psi_n(z_r) \frac{e^{i(k_n-k_m)R}}{\sqrt{k_n k_m R}} x^*. \quad (1.15)$$

We take advantage of two properties of modes. The first is the orthogonality of modes [8]. Specifically,

$$\int_z \frac{\Psi_m(z)\Psi_n(z)}{\rho(z)} dz = \delta_{mn}. \quad (1.16)$$

Assuming our summation over z_r is sufficient enough to satisfy the integral of Eq. (1.16), this simplifies the resulting focus to the form

$$f(z) = \frac{-1}{8\pi\rho(z_r)} \sum_n \frac{\Psi_n(z)\Psi_n(z_r)}{k_n R} x^*. \quad (1.17)$$

The second property of modes we take advantage of is completeness [8]. Specifically,

$$\sum_n \frac{\Psi_n(z_r)\Psi_n(z)}{k_n} = \delta(z-z_r). \quad (1.18)$$

This allows us to finally realize the resulting focus as

$$f(z) = \frac{-1}{8\pi\rho(z_r)R} \delta(z - z_r) x^*. \quad (1.19)$$

The resulting focus at the site of the original transmitter is thus a phase-conjugated and attenuated copy of the transmitted signal. Elsewhere in the water column, at the range of the original transmitter, no signal is received. This is the focus realized in Fig. 1.7a. We now perform the analysis for an endfire array at a single depth z_r . Notice that the received field is now a function of the range, R , not of z_r , as in the vertical case.

$$y(R) = \frac{ie^{-i\pi/4}}{\rho(z_s)\sqrt{8\pi}} \sum_n \Psi_n(z_s)\Psi_n(z_r) \frac{e^{ik_n R}}{\sqrt{k_n R}} x. \quad (1.20)$$

Phase-conjugation and retransmission results in a ‘focus’ at the site of the original transmitter written as,

$$\begin{aligned} f(z) &= \sum_R \frac{-1}{8\pi\rho(z_s)\rho(z_r)} \sum_m \sum_n \Psi_m(z_s)\Psi_m(z)\Psi_n(z_s)\Psi_n(z_r) \frac{e^{i(k_n-k_m)R}}{\sqrt{k_n k_m R}} x^* \\ &= \frac{-1}{8\pi\rho(z_s)\rho(z_r)} \sum_m \sum_n \Psi_m(z_s)\Psi_m(z)\Psi_n(z_s)\Psi_n(z_r) \sum_R \frac{e^{i(k_n-k_m)R}}{\sqrt{k_n k_m R}} x^*. \end{aligned} \quad (1.21)$$

By converting the summation to an integral and normalizing by the length of the aperture, L , we reduce the expression to the following, where R_0 is the range of the first element in the endfire array:

$$f(z) = \frac{-1}{8\pi\rho(z_s)\rho(z_r)} \sum_m \sum_n \Psi_m(z_s)\Psi_m(z)\Psi_n(z_s)\Psi_n(z_r) \frac{1}{L} \int_{R_0}^{R_0+L} \frac{e^{i(k_n-k_m)R}}{\sqrt{k_n k_m R}} dR x^*. \quad (1.22)$$

Because the phase term in the integral varies much more rapidly than the amplitude term, we can hold the range in the denominator constant over the integration, where the constant is chosen to represent an average range of the array. By doing this, the integral reduces to:

$$f(z) = \frac{-1}{8\pi\rho(z_s)\rho(z_r)} \sum_m \sum_n \frac{\Psi_m(z_s)\Psi_m(z)\Psi_n(z_s)\Psi_n(z_r)}{\sqrt{k_n k_m} \left(R_0 + \frac{L}{2}\right) e^{i(k_m - k_n)(R_0 + L/2)}} \text{sinc}\left[\frac{(k_n - k_m)L}{2}\right] x^*. \quad (1.23)$$

The sinc function approaches a delta function in $(k_n - k_m)$ as the length of the aperture is extended to infinity. By replacing the sinc function with a Kronecker delta, δ_{mn} , we are left with:

$$f(z) = \frac{-1}{8\pi\rho(z_s)\rho(z_r)} \sum_n \frac{\Psi_n(z)\Psi_n(z_r)}{k_n \left(R_0 + \frac{L}{2}\right)} \Psi_n^2(z_s) x^*. \quad (1.24)$$

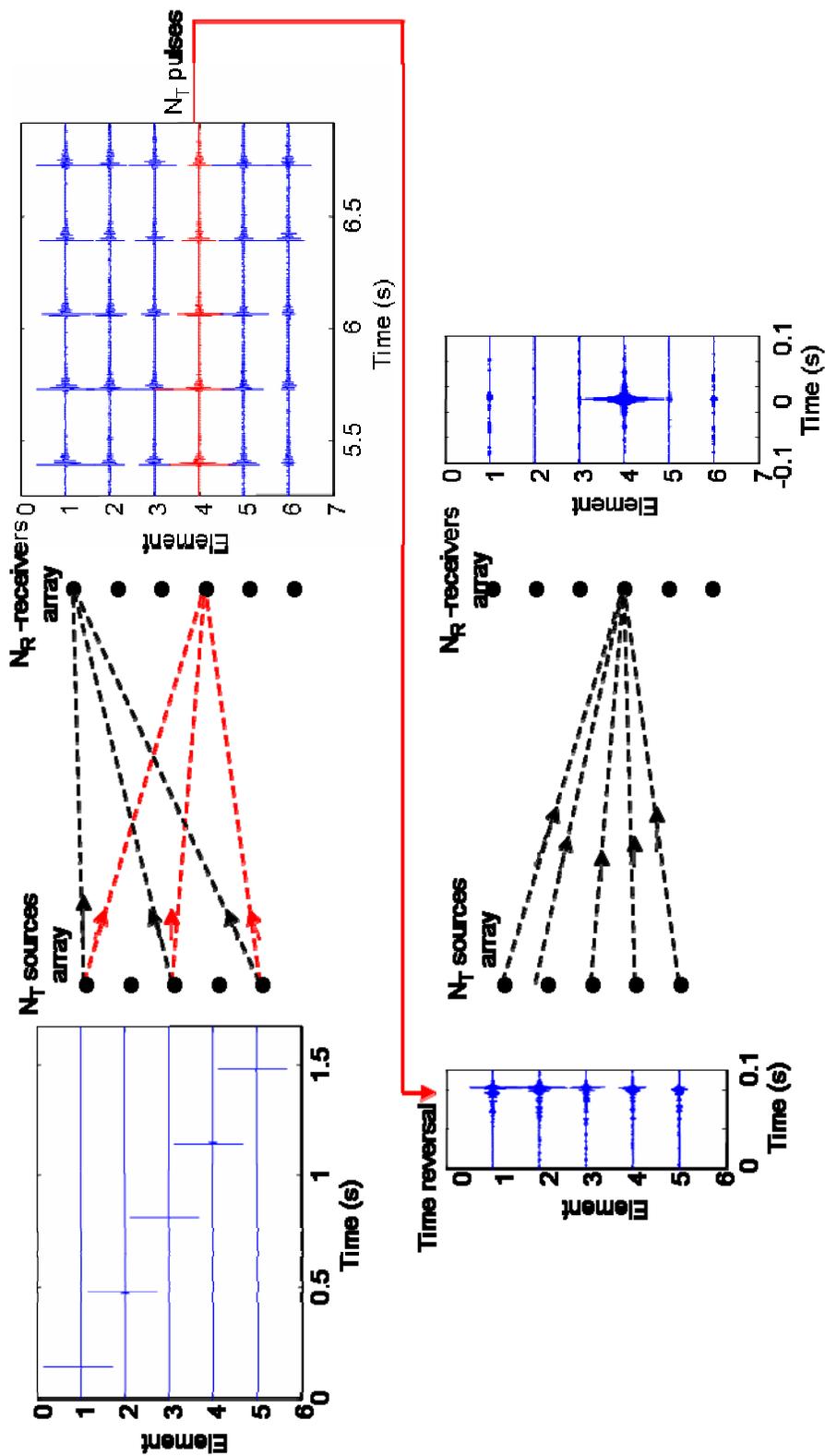
We notice the addition of a mode-squared term in Eq. (1.24) that did not appear in Eq. (1.19). This disallows our use of the completeness of modes to reduce the focus to a delta function in space. In fact, as Fig. 1.7b shows, a spatial focus is not realized with an endfire array. However, a temporal focus, as shown in Fig. 1.6b, is still valuable, and this property can be used to accomplish communications.

1.5 Passive Time Reversal

Typically, the first step of the time reversal process is the emission of a pulse by the probe source. This allows the source/receiver array (SRA) to measure the impulse responses from the probe source to each element of the array. These impulse responses are time-reversed and retransmitted. If the impulse responses were known in advance, say clairvoyantly for argument's sake, the probe source would be unnecessary. The signal transmitted from the SRA would be the time-reversed impulse responses.

When the impulse response from a source to a receiver is identical to the impulse response from the receiver to the source, such a system is called *reciprocal*. This gives

Figure 1.8: Diagram illustrating passive time reversal. The first step of passive time reversal involves measuring the impulse responses between an N_T -element source/receiver array (SRA) and an N_R -element vertical receiver array (VRA). The impulse responses between the SRA and a single element of the VRA are extracted, time-reversed, and retransmitted from the SRA to result in a spatio-temporal focus on an element of the VRA.



us two ways to measure the impulse response between a probe source and an array. Typically, in what is called *active* time reversal, the probe source emits a pulse and the SRA measures the impulse responses that way. However, in what is called *passive* time reversal [19], the each element of the SRA takes turns emitting a pulse, which is recorded on a vertical receiving array (VRA) where a probe source would normally be in active time reversal. By extracting those impulse responses measured by the VRA, time reversing them and retransmitting them from the SRA, a focus is achieved, just as in active time reversal. When the system is reciprocal, active and passive time reversal are mathematically identical. Figure 1.8 illustrates passive time reversal between two arrays with N_T transmitters and N_R receivers.

1.6 Communications Using Time Reversal

The goal of communications [20] is to transmit *information* from one place/time to another. In this context, the goal of a receiver is to extract the transmitted information in $x[n]$ from the received signal $y[n]$. This is complicated by the *ambiguities* that arise from the additive noise $w[n]$ and the convolution with $h[n]$.

The simplest way to transmit digital information is on/off keying (OOK) where a pulse is transmitted when a binary 1 is meant, and nothing is transmitted when a binary 0 is meant. This is a special case of amplitude shift keying (ASK) with the amplitudes set to 1 and 0. To transmit information in this way, at each time instant, a pulse is either sent or not sent. The receiver examines how much energy the received signal (which has been corrupted by noise) at each time instant to determine whether or not a pulse was sent at that time.

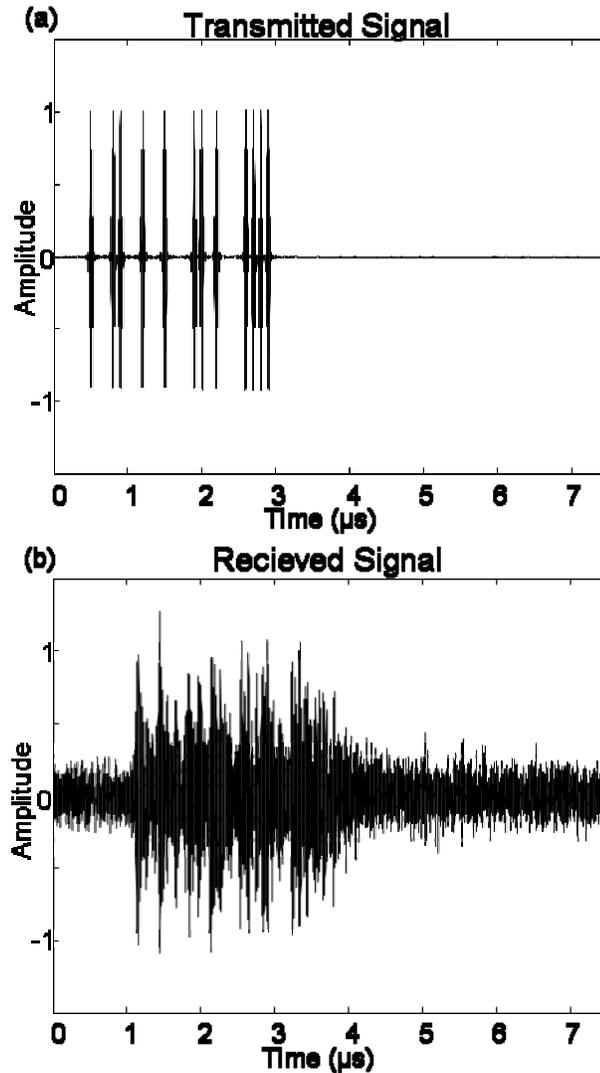


Figure 1.9: a) A communications sequence is transmitted from an ultrasonic hydrophone in an underwater waveguide, using a tank set-up as described in Appendix X. In this case, on/off keying (OOK) is used, where a pulse is sent when a binary 1 is meant and nothing is sent when a binary 0 is meant. b) The received signal when the communications sequence of Fig. 1.9a is sent using the impulse response of Fig. 1.2b. The received signal is corrupted by multiple copies of the original signal, and also additive white Gaussian noise (AWGN). This makes it difficult to interpret when a binary 0 or binary 1 is sent, which was easily apparent from the transmitted sequence.

A sequence of 25 ASK symbols are shown in Fig. 1.9a, and the resulting received signal after propagation and after noise has been added is shown in Fig. 1.9b. As can be seen, it is not longer obvious which time instances are meant to be interpreted as binary

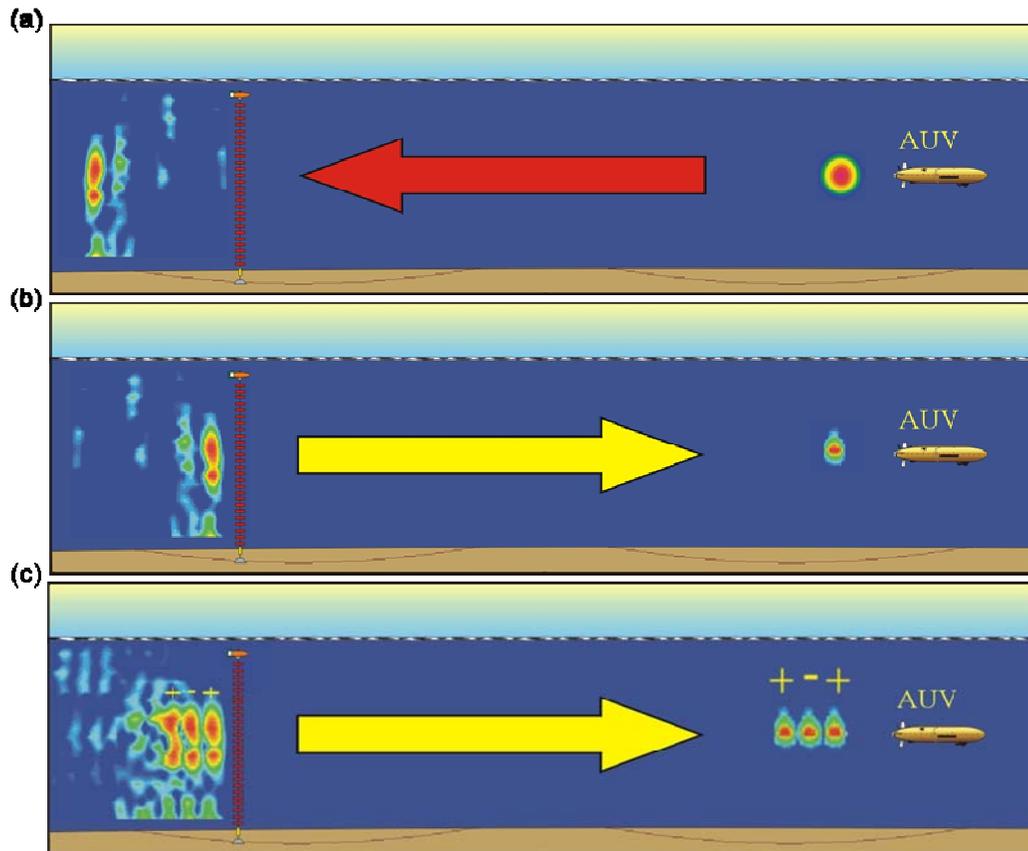


Figure 1.10: Time reversal as used for communications. a) In the first step of the time reversal process, the AUV acts as a probe source, and the pulse propagates to a source/receiver array. b) Normal time reversal, as described in Fig. 1.4, is performed, with the received field being time-reversed and retransmitted resulting in a focus at the location of the original source. c) Time reversal communications is performed by instead repeatedly transmitted the time reversed received field weighted either positively or negatively resulting in similarly weighted foci at the location of the original source.

1s and which are binary 0s. Both the noise and the impulse response have complicated our interpretation.

Another form of amplitude shift keying is called binary phase shift keying (BPSK). In this case, the amplitudes are set to -1 and 1. That is, when a binary 1 is meant, a positive pulse is transmitted and when a binary 0 is meant, a negative pulse is sent. However, because the magnitude of both binary 0s and binary 1s are the same in BPSK,

decoding the bits is no longer a matter of how much energy is in each time bin, but rather the actual amplitude. This is the difference between incoherent and coherent communication.

Time reversal can be used to perform both incoherent [21] and coherent [2] communications, as diagramed in Figure 1.10. To do so, a probe source transmits a pulse, and the received field is measured on a source/receiver array, and just as was described before and as shown in Fig. 1.10a. The normal time reversal process proceeds as is shown in Fig. 1.10b, with the received field being time-reversed and retransmitted resulting in a focus at the location of the probe source. Time reversal communications proceeds as is shown in Fig. 1.10c, with the time-reversed received field being repeatedly transmitted, weighted either positively or negatively. This results in a stream of positive or negative foci at the location of source, which are interpreted as binary digits. Time reversal can also be used to perform incoherent communications by repeatedly transmitting the time-reversed received field, with a weight of either one or zero. This results in a stream of pulses or non-pulses as shown in Fig. 1.11. In Fig. 1.11a, the result of using only a single source and single receiver for digital communications is shown. Although the energy is compressed, interpretation of each bit as either a binary 0 or binary 1 is still difficult. In Fig. 1.11b, a vertical array is used, and performance is greatly enhanced. Decoding of the bits of Fig. 1.11b results in the same bits transmitted in Fig. 1.9a.

It must be noted that time reversal is not the only method used with underwater acoustic communication [22]. Other equalization techniques and retrofocusing techniques have been developed [23],[24].

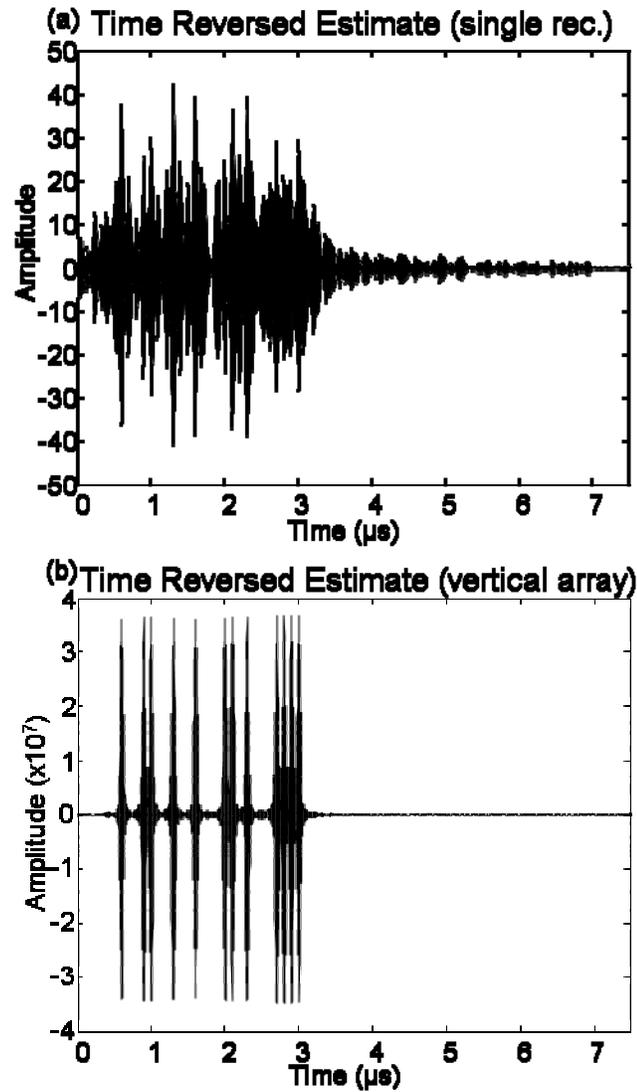


Figure 1.11: a) The received signal after the application of the time reversal process using a single source and single receiver. Although the communication sequence is less dispersed than in the case of Fig 1.9b, it is still difficult to decode (interpret when a binary 0 or binary 1 has been transmitted). b) The received signal after the application of the time reversal process using a vertical source/receiver array (SRA). In this case, the original transmitted signal is almost perfectly reconstructed. At the least, the bits are easily decoded, and the message is perfectly received.

1.7 Limitations of Time Reversal

Time reversal, as has been described, involves a probe source, transmitting to a single receiver or an array of receivers, which back-propagates the signal through the same medium to the original source. We inherently assume that the impulse response is the same from the receiver to the source as it is from the source to the receiver. When this is true, we say the impulse responses are reciprocal. Reciprocity does not always hold, such as when there are currents in the water column. Even when reciprocity does hold, the impulse responses of propagation and back-propagation may be different due to changes in the ocean. As the ocean is in constant flux [25], the impulse responses from one time to another are different. Both of these degrade the performance of the time reversal process, and communications done using the time reversal process.

One technique which we have already discussed that helps alleviate the problems of both of these degradations is passive time reversal. Passive time reversal, as opposed to active time reversal, does not involve the eventual receiver as a source transmitter. To perform passive time reversal communications, the transmitter of a signal precedes the signal with an impulse. The received signal after a passive time reversal communications sequence was transmitted is shown in Fig. 1.12.

The receiver thus records the impulse response (shown in black in Fig. 1.12), and the signal (shown in gray) corrupted by that same impulse response. The receiver extracts the impulse response from the beginning of the signal, time-reversed it, and convolves it with the signal corrupted by the impulse response. If multiple receivers are used, the result is summed across receivers. Mathematically, this is equivalent to Eq. (1.12),

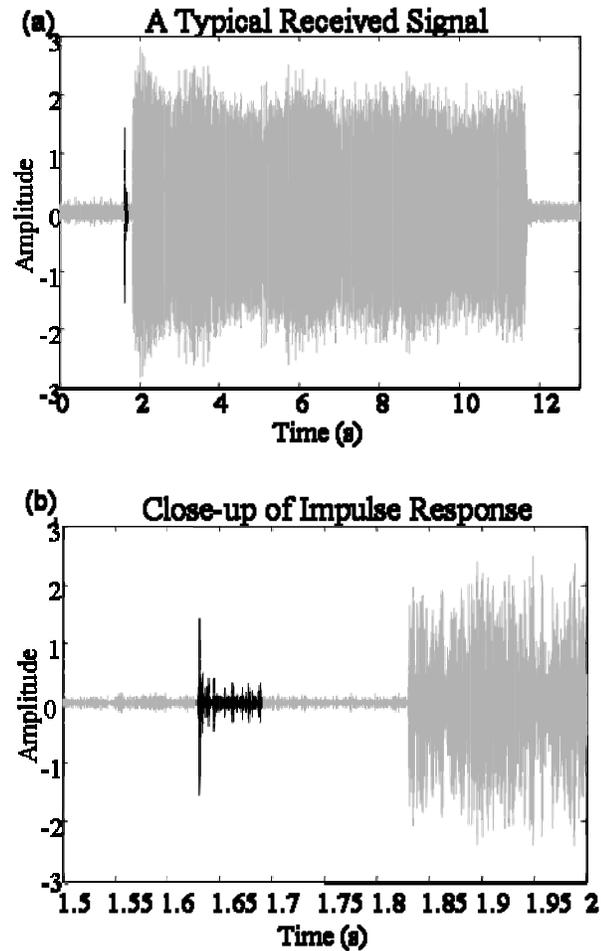


Figure 1.12: Passive time reversal communications is accomplished by first transmitted the communications sequence preceded by a pulse. Thus, the received signal appears as in the above figure, as the impulse response followed by the communications sequence corrupted by the same impulse response. The impulse response (shaded black) is extracted, time-reversed, and convolved with the rest of the communications sequence, thus mitigating the multipath and completing the passive time reversal process.

without the signal itself being time-reversed. Because the propagation of both the pulse and signal are in the same direction, reciprocity need not hold. Also, because the pulse and signal are separated by a much shorter time interval than that used in active time reversal, the effect of a changing medium is reduced.

Comparing Fig. 1.5 and Fig. 1.6, the benefit of using an array in time reversal is evident. Weaker sidelobes are seen when an array is used, as each autocorrelation in the Q-function adds coherently at time delay zero, and incoherently elsewhere. However, for it to be true that they add incoherently at each non-zero time delay, each impulse response of the Q-function needs to be independent. This is not always the case. Consider the limiting case, where a single source transmits to an array of 100 receivers all in the same location. The impulse responses between each source and receiver would be exactly the same, as would each auto-correlation. When summation is performed over these 100 receivers, it is simply a multiplication of the auto-correlation of a single impulse response, and results are identical to that of Fig. 1.5. We examine the independence of the impulses responses from a source to both a vertical and an endfire array. To do so, we measure the decorrelation, that is the maximum of the correlation between two impulse responses at different depths or ranges. The decorrelation matrix of a vertical and endfire array are shown in Fig. 1.13. From these figures, we can see that not each of the 50 depths of the vertical array are independent, with neighboring depths correlating well with each other. In the endfire array, neighboring ranges also share much in common.

1.8 Minimum Mean-Squared Error Linear Equalizer

Although time reversal is a proven and powerful technique for utilizing multipath, we have seen its limitations. By understanding passive time reversal communications as simply correlating the received communication sequence with the impulse response that corrupted it, we may wonder if we could correlate the received impulse response with a

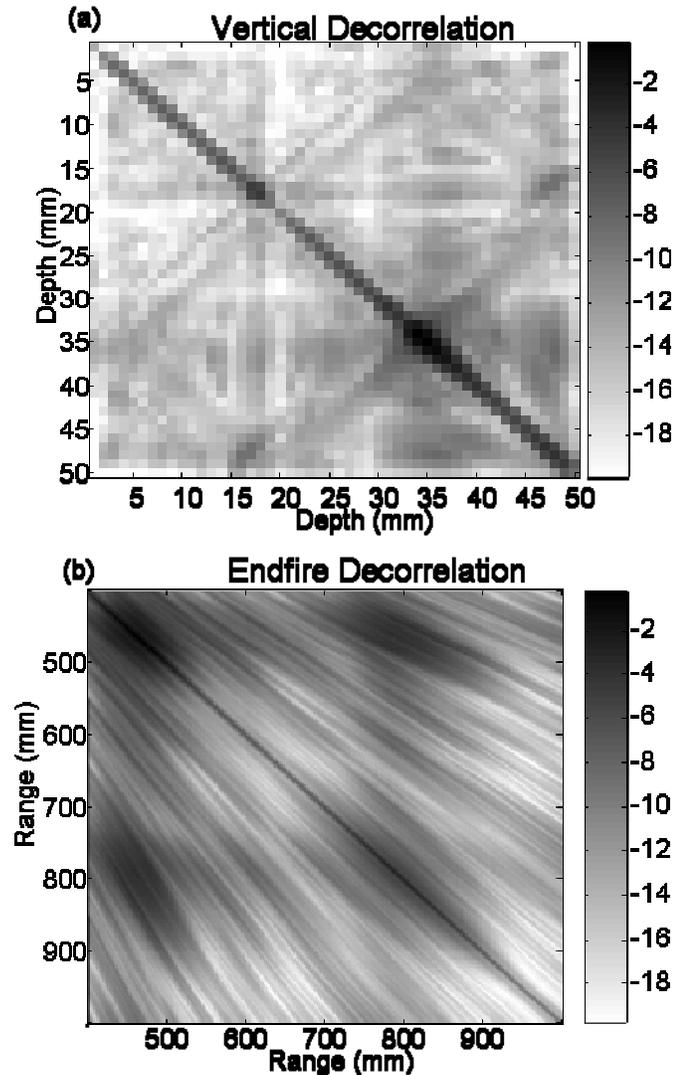


Figure 1.13: (a) A matrix showing the decorrelation of the impulse responses received on a vertical source/receiver array. Two receivers within a millimeter of each other are very similar, and show correlation. If two receivers arranged vertically are more than two millimeters apart, the impulse responses from a single source are independent. This independence is fundamental to the performance of the time reversal process. (b) A matrix showing the decorrelation of the impulse responses received on a horizontal endfire source/receiver array. Two receivers within five millimeters of each other are very similar, and show correlation. If two receivers arranged horizontally are more than ten millimeters apart, the impulse responses from a single source are independent.

something else to achieve better results. Although such additional processing is evidently not necessary when a large array is used, as in Fig. 1.11b, a different filter may be able to

outperform time reversal in the case of a single receiver, as shown in Fig. 1.11a or in the case of small arrays.

If a signal is transmitted from a source to array of receivers, the signal is corrupted by the impulse responses to the array. The goal of receiver is to undo the effect of the propagation by creating a set of filters such that when the outputs of those filters are summed, the result is the original signal. We can write this elegantly in the frequency domain using matrix notation (and suppressing the frequency dependence) as:

$$\mathbf{x}_e = \mathbf{F}\mathbf{y} = \mathbf{F}\mathbf{H}\mathbf{x}, \quad (1.25)$$

where \mathbf{F} is the set of filters we use to recreate our original signal, and \mathbf{x}_e is our estimate of the original signal.

We have discussed passive time reversal as a method of creating such a set of filters. In this case, the filters are time-reversed copies of the impulse response. This is called a matched filter, because the filters and the corrupting impulse responses are the same. We thus write our filters and estimate as:

$$\begin{aligned} \mathbf{F} &= \mathbf{H}^H, \\ \mathbf{x}_e &= \mathbf{H}^H \mathbf{H} \mathbf{x}. \end{aligned} \quad (1.26)$$

We recognize the $\mathbf{H}^H \mathbf{H}$ as the Q-function expressed in the frequency domain.

As a signal propagates through a medium, some energy is lost to attenuation, either in the water or the seafloor. Also, not all of the transmitted energy makes it to the receiver; some is just in the wrong direction. These two phenomena are expressed in Eq. (1.4) by the decaying exponential term $e^{-\alpha_n R}$ and the square root of range in the denominator, respectively. We can compensate for this attenuation through the use of an inverse filter

[26]. An inverse filter amplifies those frequencies that suffer the most attenuation, and can be written as

$$\mathbf{F} = (\mathbf{H}^H \mathbf{H})^{-1} \mathbf{H}^H, \quad (1.27)$$

with the result equal to

$$\mathbf{x}_e = (\mathbf{H}^H \mathbf{H})^{-1} \mathbf{H}^H \mathbf{H} \mathbf{x} = \mathbf{x}. \quad (1.28)$$

Although this may seem like the perfect solution, we haven't considered the effect of noise on the receiver. In reality, every received signal is corrupted, not only by the impulse response, but by noise from a variety of sources which are added to the final signal. We amend our input-output relationship to include noise, \mathbf{w} :

$$\mathbf{y} = \mathbf{H} \mathbf{x} + \mathbf{w}. \quad (1.29)$$

If we apply the inverse filter to a signal corrupted with noise, we run the risk of amplifying the noise to our detriment. Consider an impulse response that is very weak at a specific frequency. The inverse filter attempts to correct this by dramatically amplifying the received signal at this frequency. However, the received signal at this frequency consists of mostly noise, which is also amplified. The solution to this problem is to regularize the inverse, that is, to not let it amplify weak things too much.

Recall that we wish to filter the received vector, \mathbf{y} , with a filter set \mathbf{F} , such that the result is an estimate of \mathbf{x} , which is written as follows:

$$\mathbf{x}_e = \mathbf{F} \mathbf{y} = \mathbf{F} \mathbf{H} \mathbf{x} + \mathbf{F} \mathbf{w}. \quad (1.30)$$

We define the error as the difference between the transmitted signal and our estimate, and the mean-squared error as the average squared value of this error.

$$MSE = E \left[(\mathbf{x} - \mathbf{x}_e)^H (\mathbf{x} - \mathbf{x}_e) \right]. \quad (1.31)$$

We can expand this by substituting Eq. (1.30), and rewrite the mean-squared error as

$$MSE = E \left[(\mathbf{x} - \mathbf{F}\mathbf{H}\mathbf{x} - \mathbf{F}\mathbf{w})^H (\mathbf{x} - \mathbf{F}\mathbf{H}\mathbf{x} - \mathbf{F}\mathbf{w}) \right]. \quad (1.32)$$

Further expansion results in the equation

$$MSE = E \left[\begin{array}{l} \mathbf{x}^H \mathbf{x} - \mathbf{x}^H \mathbf{F}\mathbf{H}\mathbf{x} - \mathbf{x}\mathbf{F}\mathbf{w} - \mathbf{x}^H \mathbf{H}^H \mathbf{F}^H \mathbf{x} + \mathbf{x}^H \mathbf{H}^H \mathbf{n}^H \mathbf{n}\mathbf{F}\mathbf{x} \\ + \mathbf{x}^H \mathbf{H}^H \mathbf{F}^H \mathbf{F}\mathbf{w} - \mathbf{w}^H \mathbf{F}\mathbf{x} + \mathbf{w}^H \mathbf{F}^H \mathbf{F}\mathbf{H}\mathbf{x} + \mathbf{w}^H \mathbf{F}^H \mathbf{F}\mathbf{w} \end{array} \right]. \quad (1.33)$$

We cancel the terms involving a single noise term, as the mean value of additive white Gaussian noise (AWGN) is zero, yielding

$$MSE = E \left[\mathbf{x}^H \mathbf{x} - \mathbf{x}^H \mathbf{F}\mathbf{H}\mathbf{x} - \mathbf{x}^H \mathbf{H}^H \mathbf{F}^H \mathbf{x} + \mathbf{x}^H \mathbf{H}^H \mathbf{F}^H \mathbf{F}\mathbf{H}\mathbf{x} + \mathbf{w}^H \mathbf{F}^H \mathbf{F}\mathbf{w} \right]. \quad (1.34)$$

Assume that signal transmitted from different antennas are independent wide-sense stationary sequences with zero mean and equal power P , and let N_0 be the noise power.

The mean-squared error can then be simplified as

$$MSE = P \cdot \text{tr} \left(\mathbf{I} - \mathbf{F}\mathbf{H} - \mathbf{H}^H \mathbf{F}^H + \mathbf{H}^H \mathbf{F}^H \mathbf{F}\mathbf{H} + \frac{N_0}{P} \mathbf{F}^H \mathbf{F} \right). \quad (1.35)$$

Taking the derivative with respect to \mathbf{F}_{ij} , we may write

$$\frac{\partial MSE}{\partial \mathbf{F}_{ij}} = P \left[\left(\mathbf{F}\mathbf{H}\mathbf{H}^H \right)_{ji} + \left(\mathbf{H}\mathbf{H}^H \mathbf{F} \right)_{ji} - \mathbf{H}_{ji}^H - \mathbf{H}_{ji} + \frac{N_0}{P} \left(\mathbf{F}_{ji} + \mathbf{F}_{ji}^H \right) \right]. \quad (1.36)$$

Setting the derivative to zero for all i and j yields

$$\mathbf{F}\mathbf{H}\mathbf{H}^H - \mathbf{H}^H + \frac{N_0}{P} \mathbf{F} = \mathbf{0}. \quad (1.37)$$

Which can be written more commonly as

$$\mathbf{F} = \mathbf{H}^H \left(\mathbf{H}\mathbf{H}^H + \frac{N_0}{P} \mathbf{I} \right)^{-1}. \quad (1.38)$$

The resulting filter set is called the minimum mean-squared error linear equalizer (MMSE-LE) [27]. Traditionally, when a communications sequence preceded with an impulse, as shown in Fig. 1.12, is received, the impulse response is used to explicitly calculate the MMSE filter set of Eq. (1.38). This filter set is used instead of the matched filter of passive time reversal. In chapter 2, iterative time reversal, a method that converges to the inverse filter of Eq. (1.27) by iteratively reducing sidelobes, is described. In the same chapter, the MMSE-LE is derived by altering iterative time reversal.

1.9 Information Theory

As previously stated, the goal of communications is to retrieve the information in $x[n]$ from the received signal, $y[n]$. This is complicated by the ambiguities that arise through the convolution with $h[n]$ and the additive noise, $w[n]$. Without considering *how* the information is to be retrieved, we can still measure the amount of information that is theoretically retrievable. This measure, termed the channel *capacity* is ultimately a measure of the ambiguity presented by the channel transfer function and noise. Because the calculation of capacity doesn't explain how the information is to be transmitted or retrieved, it serves as an upper bound for communication rates. Although such a rate may be theoretically achievable, it is not necessarily practically achievable.

The amount of information in a random variable, x , is a function of its probability density function (pdf), $p(x)$. This amount of information is called the *entropy* [28] of the random variable, and is written as

$$H(x) = E[-\log p(x)], \quad (1.39)$$

where $E[\cdot]$ is the expected value and apologies are made for the dual use of the letter H . Essentially, the entropy of a random variable is related to its predictability. When the entropy of a random variable is high, the value a particular realization will take is most difficult to predict in advance. In contrast, if the value of a particular realization is known, the entropy is zero.

Of interest is the random variable with the maximum entropy for a given power. Thus with wish to maximize

$$H(x) = -\int p(x) \log p(x) dx \quad (1.40)$$

subject to

$$\int x^2 p(x) dx = \sigma^2 \text{ and } \int p(x) dx = 1. \quad (1.41)$$

This requires, by calculus of variation, maximizing

$$\int \left(-p(x) \log p(x) + \lambda x^2 p(x) + \mu p(x) \right) dx. \quad (1.42)$$

This is maximized when

$$-1 - \log p(x) + \lambda x^2 + \mu = 0. \quad (1.43)$$

Adjusting the values of the constants to match our constraints of Eq. (1.41) results in the following pdf of the variable that maximizes entropy for a given power:

$$p(x) = \frac{1}{\sqrt{2\pi}\sigma} e^{-(x^2/2\sigma^2)}. \quad (1.44)$$

This is recognized as a Gaussian pdf, and indeed a Gaussian random variable has the maximum entropy for a given power [28]. Further, we state the entropy of a complex Gaussian random variable with power σ^2 .

$$H(x) = \log_2 2\pi e\sigma^2 \quad (1.45)$$

The amount of information about the input sequence, x , retrievable from the received sequence, y , is called the *mutual information*, defined as:

$$I(y; x) = H(y) - H(y|x), \quad (1.46)$$

where $H(y|x)$ is the entropy of y given x . The *capacity* of a channel/system, is defined as the maximum mutual information, maximized over the input probability, $p(x)$, subject to relevant constraints. The capacity is thus a function of the channel, noise, and constraints. The value of knowing the capacity is the existence of coding theorems which state that if information is *properly* transmitted at a rate less than the capacity, the probability of error can be made arbitrarily small [28].

We are armed, already, to calculate the channel capacity of a discrete Gaussian channel. We imagine each use of the channel to have the input-output relationship:

$$y = hx + w. \quad (1.47)$$

where x is the input symbol, constrained to an average power P , y is the output symbol, h is the attenuation/gain of the channel, and w is a Gaussian random noise variable with an average power N . We can write the capacity of this channel as:

$$C = \max_{p(x)} [I(y; x)] = \max_{p(x)} [H(y) - H(y|x)]. \quad (1.48)$$

The second entropic term can be reduced to $H(w)$, as the only unpredictability of y once x is known derives from the noise variable. It can also be taken out of the maximization since it no longer depends on x .

$$C = \max_{p(x)} [H(hx + w)] - H(w) \quad (1.49)$$

As we derived above, the entropy is maximized when the variable is a Gaussian, which is possible only if and only if x is Gaussian. Since we know the entropy of a Gaussian random variable, we simplify the capacity to:

$$C = \log_2 \left(2\pi e(|h|^2 P + N) \right) - \log_2 \left(2\pi e(N) \right),$$

$$C = \log_2 \left(\frac{|h|^2 P + N}{N} \right) = \log_2 \left(1 + \frac{P}{N} |h|^2 \right). \quad (1.50)$$

More briefly, we have

$$C = \log_2(1 + SNR), \quad (1.51)$$

where SNR is the ratio of *received* signal power to noise power [28].

The above formula is true in the case of a discrete, memoryless Gaussian channel. However, as we learned above, the ocean introduces multipath to any transmission. The received signal is a function of the transmitted signal at multiple times. We say such a channel has memory, and we derive the capacity of such a channel below.

We consider a noisy, dispersive channel similar to that of Eq. (1.2), but with noise added:

$$y[n] = \sum_p h[p]x[n-p] + w[n]. \quad (1.52)$$

We can write this convolution in the frequency domain as a multiplication:

$$y(\omega) = h(\omega)x(\omega) + w(\omega). \quad (1.53)$$

We state that the power transmitted at each frequency is $P(\omega)$, and assume the noise power at each frequency is $N(\omega)$. We constrain our total transmitted power to be P_T , which leads to:

$$\frac{1}{2\pi} \int_{-\pi}^{\pi} P(\omega) d\omega = P_T. \quad (1.54)$$

Looking at each frequency independently, we notice a similarity between Eq. (1.53) equation and Eq. (1.47). We conclude that each frequency behaves as an independent memoryless Gaussian channel, with a capacity (in bits) of $\log_2(1+SNR)$, which is achieved by transmitting a complex Gaussian at each frequency [29]. We can write this as:

$$C = \frac{1}{2\pi} \int_{-\pi}^{\pi} \log_2 \left(1 + \frac{P(\omega)}{N(\omega)} |h(\omega)|^2 \right) d\omega. \quad (1.55)$$

However, we have not yet calculated $P(\omega)$, in terms of the total power P_T . The method of calculating the power in each frequency bin is called *water-filling* [29]. As the capacity is the maximum of the mutual information over input probabilities, we set-up the optimization problem below:

$$C = \max_{P(\omega)} \left[\frac{1}{2\pi} \int_{-\pi}^{\pi} \log_2 \left(1 + \frac{P(\omega)}{N(\omega)} |h(\omega)|^2 \right) d\omega \right]$$

subject to $\frac{1}{2\pi} \int_{-\pi}^{\pi} P(\omega) d\omega = P_T$ and $P(\omega) \geq 0, \forall \omega.$ (1.56)

The constraints come from the total power constraint of Eq. (1.54), and that power must always be positive.

We use the method of Lagrange multipliers to state

$$L(P(\omega)) = \frac{1}{2\pi} \int_{-\pi}^{\pi} \log_2 \left(1 + \frac{P(\omega)}{N(\omega)} |h(\omega)|^2 \right) d\omega$$

$$- \lambda \left(\frac{1}{2\pi} \int_{-\pi}^{\pi} P(\omega) d\omega - P_T \right) - \frac{1}{2\pi} \int_{-\pi}^{\pi} \mu(\omega) (P(\omega) - s^2(\omega)) d\omega. \quad (1.57)$$

Taking the derivative with respect to the Lagrange multipliers and setting each to zero yields the following four equations.

$$\begin{aligned}
\frac{\partial L}{\partial P(\omega)} &= \frac{1}{2\pi \ln(2) \left(P(\omega) + \frac{N(\omega)}{|h(\omega)|^2} \right)} - \frac{1}{2\pi} \lambda - \frac{1}{2\pi} \mu(\omega) = 0, \\
\frac{\partial L}{\partial \lambda} &= - \left(\frac{1}{2\pi} \int_{-\pi}^{\pi} P(\omega) d\omega - P_T \right) = 0, \\
\frac{\partial L}{\partial \mu(\omega)} &= - \frac{1}{2\pi} (P(\omega) - s^2(\omega)) = 0, \\
\frac{\partial L}{\partial s(\omega)} &= \frac{\mu(\omega) s(\omega)}{\pi} = 0.
\end{aligned} \tag{1.58}$$

The second and third equations simply restate our constraints, and the fourth can be interpreted to state that at each frequency, either μ or s must be zero. For the case when μ is zero, we can solve the first equation of Eq. (1.58) for λ .

$$\lambda = \frac{1}{\ln(2) \left(P(\omega) + \frac{N(\omega)}{|h(\omega)|^2} \right)} \tag{1.59}$$

For interpretation purposes, we define a new variable, Θ , which is called the water-fill level.

$$\Theta = \frac{1}{\lambda \ln(2)} \tag{1.60}$$

We can write the transmit power at each frequency as a function of the water-fill level.

$$P(\omega) = \Theta - \frac{N(\omega)}{|h(\omega)|^2} \tag{1.61}$$

For the case when μ is not zero, s must be zero. We examine the third constraint of Eq. (1.58) to reveal that in this case, $P(\omega)$ must also be zero. From these two cases, we see that the transmit power is either zero or as described in Eq. (1.61). Knowing that power

can never be negative allows us to determine the power at any given frequency. We use the notation of a superscript plus sign to indicate the non-negative portion of the argument, and write:

$$P(\omega) = \max\left(0, \Theta - \frac{N(\omega)}{|h(\omega)|^2}\right) = \left(\Theta - \frac{N(\omega)}{|h(\omega)|^2}\right)^+. \quad (1.62)$$

The total power constraint of Eq. (1.54) allows to calculate the water-fill level, Θ , and we may finally write the capacity of a dispersive channel [30] as:

$$C = \frac{1}{2\pi} \int_{-\pi}^{\pi} \log_2 \left(1 + \frac{P(\omega)}{N(\omega)} |h(\omega)|^2 \right) d\omega$$

$$\text{with } P(\omega) = \left(\Theta - \frac{N(\omega)}{|h(\omega)|^2} \right)^+ \text{ and } \frac{1}{2\pi} \int_{-\pi}^{\pi} P(\omega) d\omega = P_T. \quad (1.63)$$

1.10 Principal Component Analysis

The multipath present in a waveguide also complicates the detection and classification of buried targets. In mine countermeasures (MCM), it is preferable to have a large “stand-off” distance. That is, it is desirable to be as far away as possible from the mine and still detect and classify it. As can be seen in Fig. 1.3, dispersion increases with distance, so the further we are from the target, the greater the influence of multipath propagation.

One technique that will be developed to classify a buried target depends on the concept of principal component analysis (PCA). If the target resonates, the resonance is independent of the excitation. We use this fact to attempt to extract the resonance and thus classify the target as a resonant target. The theory of PCA is discussed below.

Imagine we have N observations of the zero-mean random vector \mathbf{x} , which is $M \times 1$. We denote these observations as $\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_N$. We create a sample covariance matrix \mathbf{S} , using the following formula [31]:

$$\mathbf{S} = \frac{1}{N-1} \sum_{i=1}^N \mathbf{x}_i \mathbf{x}_i^H. \quad (1.64)$$

We wish to find a vector that defines the largest portion of the variance in the observations. Stated another way, we wish to find a vector that correlates best with each of the observations. Mathematically speaking, we want a vector, \mathbf{a} , such that the average correlation with the data is maximized. Such a maximization is meaningless without a constraint on the size of \mathbf{a} , so we add the constraint that the magnitude of \mathbf{a} is equal to one. We can write the formula for \mathbf{a} , where the notation “arg max” means “the argument which maximizes”, as follows [31]:

$$\mathbf{a} = \arg \max_{\mathbf{a}} \left(\frac{1}{N-1} \sum_{i=1}^N |\mathbf{a}^H \mathbf{x}_i|^2 \right) \text{ with } \|\mathbf{a}\|^2 = 1. \quad (1.65)$$

We can expand the term in the summation to rewrite Eq. (1.65) as:

$$\mathbf{a} = \arg \max_{\mathbf{a}} \left(\frac{1}{N-1} \sum_{i=1}^N \mathbf{a}^H \mathbf{x}_i \mathbf{x}_i^H \mathbf{a} \right) = \arg \max_{\mathbf{a}} (\mathbf{a}^H \mathbf{S} \mathbf{a}) \text{ with } \|\mathbf{a}\|^2 = 1. \quad (1.66)$$

We can solve this maximization using Lagrange multipliers by defining

$$L(\mathbf{a}) = \mathbf{a}^H \mathbf{S} \mathbf{a} - \lambda (\mathbf{a}^H \mathbf{a} - 1). \quad (1.67)$$

Taking the gradient with respect to \mathbf{a} and equating it to zero yields

$$\frac{\partial L(\mathbf{a})}{\partial \mathbf{a}} = \mathbf{S} \mathbf{a} - \lambda \mathbf{a} = 0. \quad (1.68)$$

Rearranging these terms yields the familiar eigenvalue equation

$$\mathbf{S} \mathbf{a} = \lambda \mathbf{a}. \quad (1.69)$$

Thus, the \mathbf{a} that maximizes equation (1.65) must be an eigenvector of \mathbf{S} . It is in fact, the eigenvector of \mathbf{S} that corresponds to the largest eigenvalue of \mathbf{S} .

We can define a second, third, fourth, etc. principal component as that component orthogonal to the previous components that defines the largest portion of the variance not defined by the previous components. The principal components can be found by taking the eigenvalue decomposition of the sample covariance matrix. The principal components of observed data are simply the eigenvectors of the sample covariance matrix.

We can also define an $M \times N$ matrix, \mathbf{X} , that is the concatenation of each of the observations, such that $\mathbf{X} = [\mathbf{x}_1; \mathbf{x}_2; \dots; \mathbf{x}_N]$. The outer product of a scaled version of this matrix is identical to the sample covariance matrix, \mathbf{S} .

$$\left(\frac{\mathbf{X}}{\sqrt{N-1}} \right) \left(\frac{\mathbf{X}}{\sqrt{N-1}} \right)^H = \mathbf{S} \quad (1.70)$$

Thus if we write the eigenvalue decomposition of \mathbf{S} as

$$\mathbf{S} = \mathbf{U} \mathbf{\Lambda} \mathbf{U}^H \quad (1.71)$$

we can write the singular value decomposition (SVD) of \mathbf{X} as

$$\mathbf{X} = \mathbf{U} \sqrt{(N-1)\mathbf{\Lambda}} \mathbf{V}^H. \quad (1.72)$$

In both cases, the matrix \mathbf{U} is identical. The principal components can also be viewed as the left singular vectors of an appropriately-concatenated matrix of the observations.

1.11 Summary

The rest of the dissertation is structured as follows: Chapter 2 introduces a self-adaptive process called iterative time reversal, which converges to an inverse filter. However, as we have seen in section 1.8, when dealing with noise, the MMSE linear equalizer is a better choice. Chapter 2 derives the MMSE by altering iterative time reversal, thus demonstrating MMSE linear equalization as a self-adaptive process. Typical time reversal experiments achieve results using a large array, the benefits of which are demonstrated in section 1.4. However, this can be expensive and it would be desirable to achieve the same results using only a single source and single receiver. To that end, the feasibility of using a synthetic aperture for time reversal communications is examined in Chapter 3. The capacity of a dispersive channel was derived in section 1.9, but as mentioned, typical time reversal experiments involve large arrays. Bounds on the capacity of the channel between two such arrays in a dispersive medium such as the ocean are derived in Chapter 4. Also in Chapter 4, those bounds are applied to a measured ocean channel. The multipath that complicates communications also complicates the detection and classification of buried targets. Chapter 5 introduces two forms of processing to classify a target as either resonant or non-resonant. These two forms of processing are then applied in a waveguide, showing promising results. The chapter also suggests the set-up for an experiment to test these processes in a practical real-world situation.

REFERENCES

1. P. Roux and M. Tink, "Time reversal in a waveguide: Study of the temporal and spatial focusing," *J. Acous. Soc. Amer.*, vol. 107, pp. 2418-2429, 2000.
2. G.F. Edelmann, T. Akal, W.S. Hodgkiss, S. Kim, W.A. Kuperman, and H.C. Song, "An initial demonstration of underwater acoustic communication using time reversal," *IEEE J. Oceanic Eng.*, vol. 27, pp. 602-609, 2002.
3. T.H. Eggen, A.B. Baggeroer, and J.C. Preisig, "Communications over Doppler spread channels – Part I: channel and receiver presentation," *IEEE J. Oceanic Eng.*, vol. 25, pp. 62-71, 2000.
4. T.H. Eggen, J.C. Preisig, and A.B. Baggeroer, "Communications over Doppler spread channels – II: Receiver characterization and practical results," *IEEE J. Oceanic Eng.*, vol. 28, pp. 612-621, 2001.
5. A.B. Baggeroer, W.A. Kuperman, and P. Mikhalevsky, "An overview of matched field methods in ocean acoustic," *IEEE J. Oceanic Eng.*, vol. 4, pp. 401-424, 1993.
6. S. Walker, "Self-adaptive methods for acoustic focusing and mode extraction in a shallow ocean waveguide," *Ph.D. Thesis, University of California at San Diego*, 2005.
7. A.V. Oppenheim and A.S. Willsky, *Signals and Systems, 2nd Edition*. Prentice Hall, 1997.
8. F.B. Jensen, W.A. Kuperman, M.B. Porter, and H. Schmidt, *Computational Ocean Acoustics*. AIP, 1994.
9. A. Parvelescu and C.S. Clay, "Reproducibility of signal transmissions in the ocean," *Radio Elec. Eng.*, vol. 29, pp. 223-228, 1965.
10. A. Parvelescu, "Matched-signal ('MESS') processing by the ocean," *J. Acous. Soc. Amer.*, vol. 98, pp. 943-960, 1995.
11. M. Fink, "Time reversal mirrors," *Acoustical Imaging*, vol. 25, pp. 1-15.
12. M. Fink, "Time reversed acoustics," *Phys. Today*, vol. 50, pp. 34-40, 1997.
13. P. Roux, "Time-reversal mirrors: Application to waveguide propagation and characterization of fluid flows (French)," *Ph.D. Thesis, University of Paris*, 1997.

14. D.R. Jackson and D.R. Dowling, "Phase conjugation in underwater acoustics," *J. Acous. Soc. Am.*, vol. 89, pp. 171-181, 1991.
15. W.A. Kuperman, W.S. Hodgkiss, H.C. Song, T. Akal, C. Ferla, and D. Jackson, "Phase conjugation in the ocean: Experimental demonstration of an acoustic time-reversal mirror," *J. Acous. Soc. Am.*, vol. 103, pp. 25-40, 1998.
16. D. Rouseff, W.L.J. Fox, D.R. Jackson, and C.D. Jones, "Underwater acoustic communication using passive phase conjugation," *OCEANS, 2002. MTS/IEEE Conf. and Exhib.*, vol. 4, pp. 2227-2230, 2000.
17. T.C. Yang, "Temporal resolution of time-reversal and passive-phase conjugation for underwater acoustic communications," *IEEE J. Oceanic Eng.*, vol. 28, pp. 229-245, 2003.
18. M.R. Dungan and D.R. Dowling, "Orientation effects on linear time reversing array retrofocusing in shallow water," *J. Acous. Soc. Amer.*, vol. 112, pp. 1842-1852, 2002.
19. P. Roux, W.A. Kuperman, W.S. Hodgkiss, H.C. Song, T. Akal, and M. Stevenson, "A nonreciprocal implementation of time reversal in the ocean," *J. Acous. Soc. Amer.*, vol. 116, pp. 1009-1015, 2004.
20. J.G. Proakis, *Digital Communications*. McGraw-Hill, 1995.
21. K.B. Smith, A.A.M. Abrantes, and A. Lazzara, "Examination of time-reversal acoustic in shallow water and applications to noncoherent underwater communications," *J. Acous. Soc. Amer.*, vol. 113, pp. 3095-3110, 2003.
22. D. Kilfoyle and A. Baggeroer, "The state of the art in underwater acoustic telemetry," *IEEE J. Oceanic Eng.*, vol. 25, pp. 4-27, 2000.
23. M. Stojanovic, J.A. Capitovic, and J.G. Proakis, "Adaptive multichannel combining and equalization for underwater acoustic communications," *J. Acous. Soc. Amer.*, vol. 94, pp. 1621-1631, 1993.
24. M. Stojanovic, J.A. Capitovic, and J.G. Proakis, "Phase coherent digital communications for underwater acoustic channels," *IEEE J. Oceanic Eng.*, vol. 19, pp. 100-111, 1994.
25. J.C. Preisig, "Performance analysis of adaptive equalization for coherent acoustic communications in the time-varying ocean environment," *J. Acous. Soc. Amer.*, accepted March 2005.

26. G. Montaldo, M. Tanter, and M. Fink, "Real time inverse filter focusing through iterative time reversal," *J. Acous. Soc. Amer.*, vol. 115, pp. 768-775, 2004.
27. J. Salz, "Digital transmission over cross-coupled linear channels," *AT&T Tech. Jour.*, vol. 64, pp. 1147-1157, 1985.
28. C. Shannon, "A mathematical theory of communications," *The Bell System Tech. Jour.*, vol. 27, pp. 379-423, 623-656, 1948.
29. T. Cover and J. Thomas, *Elements of Information Theory*, John Wiley and Sons, 1991.
30. W. Hirt and J.L. Massey, "Capacity of the discrete-time Gaussian channel with intersymbol interference," *IEEE Trans. on Info. Theory*, vol. 34, pp. 380-388, 1988.
31. T.K. Moon and W.C. Stirling, *Mathematical Methods and Algorithms for Signal Processing*. Prentice Hall, 2000.

Chapter 2

The relationship between time reversal and linear equalization in digital communications

abstract

Iterative time reversal has been suggested as both an efficient method of creating a spatio-temporal focus and for use in telecommunications as a form of equalization. In this paper, the equivalence of a passive, i.e. via computation, iterative time reversal to the Moore-Penrose pseudo-inverse of the propagation matrix is shown. In the context of communications, however, any received signal is corrupted by noise. Therefore, a regularization term is introduced to the iterative equations, causing convergence to the canonical minimum mean-squared error linear equalizer (MMSE-LE). Hence, a relationship between time reversal and equalization is demonstrated.

2.1 Introduction

In digital communications, the performance of a standard time reversal (TR) process has limitations because the reduction in inter-symbol interference is related to a matched-filter process [1]-[3]. On the other hand, the equalization used in coherent communications is more closely related to inverse filtering. Though computationally more intensive than TR, the reduction in inter-symbol interference is potentially much greater, particularly when the ratio of receiving elements to transmission elements is small, and equalization can therefore provide a better approach than standard TR. A recent paper by Montaldo, et al. [4] has described an active iterative form of time reversal to achieve spatio-temporal focusing through a complex medium with greater inter-symbol

interference reduction than standard TR. Additionally, they have suggested that this could be used as an efficient method of equalization [5]. In certain environments, such as underwater communications, repeated propagation through the media as described in the papers is impractical. By performing the iteration passively, via computation, this problem is relieved, at the expense of the additional computation. In this paper, a relationship between TR and equalization is derived by showing that the inclusion of a regularization term in the passive iterative TR process is identical to the minimum mean-square error linear equalizer (MMSE-LE). This allows for a physical insight into the MMSE equalizer.

2.2 Description of Iterative Time Reversal

One goal of communications is to have multiple sources transmit to an array of receivers. Such is the case in underwater acoustic communications with a network of autonomous underwater vehicles (AUVs), where each source corresponds to a different user, or in some forms of array-to-array communications. The goal of passive iterative time reversal is to create a set of filter banks that equalize the received signals, such that the combined impulse response of the channel and each filter bank is a spatio-temporal Kronecker delta function corresponding to each source.

The propagation between each transmitter and receiver element is described by the set of impulse responses $h_{ij}(t)$, $i = 1, 2, \dots, N_R$ and $j = 1, 2, \dots, N_T$. For example, if the signal sent from each transmitter is the time-dependant signal $x_j(t)$, the received signals on the array, in the absence of noise, are

$$y_i(t) = \sum_j h_{ij}(t) * x_j(t), \quad (2.1)$$

where * indicates convolution. Equivalently, in the frequency domain, one may write

$$Y_i(\omega) = \sum_j H_{ij}(\omega) X_j(\omega), \quad (2.2)$$

where capitalization indicates the Fourier transform, and ω is frequency. Writing this in matrix notation yields

$$\mathbf{Y}(\omega) = \mathbf{H}(\omega)\mathbf{X}(\omega). \quad (2.3)$$

When the transmitters wish to send information, they first send a known function followed by the communications sequences, $\mathbf{X}(\omega)$. The receiver array is able to extract a noisy estimate of the unitless channel transfer functions, $\mathbf{H}(\omega)$, from the known part of the signals. The second part of the received signals are the communications sequences convolved with the transfer functions, $\mathbf{H}(\omega)\mathbf{X}(\omega)$. The transfer functions extracted from the first part of the signal are time-reversed to initialize the filter for the first iteration, designated \mathbf{F}_n , where the subscript indicates iteration number,

$$\mathbf{F}_1(\omega) = \mathbf{H}^H(\omega). \quad (2.4)$$

The iterative process begins by passively propagating the filter impulse responses back to the transmitters. The term, “passively propagating”, in this case, is taken to mean replicating, via computation on a computer, the result of physically transmitting the filter impulses and measuring the field back at the original transmitters. This is analytically equivalent to an active time-reversal process when noise is not considered and reciprocity is valid [6]. This results in a combined channel/filter impulse expressed by the following equation, where superscript H indicates conjugate-transpose.

$$\mathbf{R}_1(\omega) = \mathbf{H}(\omega)\mathbf{H}^H(\omega) \quad (2.5)$$

Often, particularly with small arrays, this results in temporal sidelobes that act as inter-symbol interference and degrade the performance of communications systems. The next step in the iterative process is to subtract this result from the objective delta functions, which are constant in frequency, yielding a difference term expressible as

$$\mathbf{D}_1(\omega) = \mathbf{I} - \mathbf{H}(\omega)\mathbf{H}^H(\omega). \quad (2.6)$$

The filter impulse responses are then updated by adding to them the difference term convolved with the time-reversed transfer functions previously obtained. As written in [4], the iterative procedure is the set of equations below, where the frequency dependence has been suppressed for clarity.

$$\begin{aligned} \mathbf{F}_0 &= \mathbf{0} \\ \mathbf{R}_n &= \mathbf{H}\mathbf{F}_n \\ \mathbf{D}_n &= \mathbf{I} - \mathbf{R}_n \\ \mathbf{F}_{n+1} &= \mathbf{F}_n + \mathbf{H}^H\mathbf{D}_n \end{aligned} \quad (2.7)$$

The difference term of the n th iteration can be shown to be

$$\mathbf{D}_n = (\mathbf{I} - \mathbf{H}\mathbf{H}^H)^n, \quad (2.8)$$

which causes the filter responses of the following iteration to be equal to

$$\mathbf{F}_{n+1} = \mathbf{H}^H \sum_{k=0}^n (\mathbf{I} - \mathbf{H}\mathbf{H}^H)^k. \quad (2.9)$$

The summation term in the above equation can be recognized as the Neumann expansion [7] of the matrix inverse, which states

$$\mathbf{A}^{-1} = \sum_{k=0}^{\infty} (\mathbf{I} - \mathbf{A})^k \quad (2.10)$$

given that the norm of $(\mathbf{I} - \mathbf{A})$ is less than one. After many iterations, the filter responses converge to

$$\mathbf{F} = \mathbf{H}^H (\mathbf{H}\mathbf{H}^H)^{-1}. \quad (2.11)$$

This is recognized as the Moore-Penrose pseudo-inverse of the propagation matrix, \mathbf{H} . Finally, this filter set is applied to the received communications sequences, \mathbf{Y} , and the signals, \mathbf{X} , decoded. The problem of the estimate of the transfer function matrix, \mathbf{H} , being noisy is lessened in the case of active iterative time reversal compared to the passive case, as each iteration introduces a different realization of the noise process. However, active iteration is impractical in certain circumstances, such as during an at-sea experiment [2]. Additionally, in the context of communications, additive noise dominates channel estimation error, therefore the transfer functions estimates are usually assumed to be the true transfer functions. A more appropriate goal of passive time reversal would be to create a set of filter impulse responses that minimize the mean-squared error of the received communications sequences. The impulse responses that achieve this goal under a white-noise assumption are governed by the well-know minimum mean-squared error linear equalizer (MMSE-LE) expression

$$\mathbf{F} = \mathbf{H}^H (\mathbf{H}\mathbf{H}^H + \sigma^2 \mathbf{I})^{-1}, \quad (2.12)$$

where σ^2 is the inverse of the signal-to-noise ratio (SNR), calculated as the ratio of the power transmitted from each transmitter, P , and the noise power, N_0 , received at a single receiver.

It is possible, through the addition of a regularization term, to alter the iterative procedure of Eq. (2.7) so that it converges to the MMSE-LE equation stated above in Eq. (2.12). After this modification, the iterative procedure is written as

$$\begin{aligned}
 \mathbf{F}_0 &= \mathbf{0} \\
 \mathbf{R}_n &= \mathbf{H}\mathbf{F}_n \\
 \mathbf{D}_n &= \mathbf{I} - \mathbf{R}_n - \sigma^2 \sum_{k=0}^{n-1} \mathbf{D}_k \\
 \mathbf{F}_{n+1} &= \mathbf{F}_n + \mathbf{H}^H \mathbf{D}_n
 \end{aligned} \tag{2.13}$$

The only difference is the addition of a regularization term in the third iteration equation. This can be recognized as a form of gradient descent solution, similar to the conjugate gradient method [7], to finding the MMSE-LE filter impulse responses. Again, once a number of iterations have been performed, the filter set is applied to the received communications sequences and the signals decoded.

The MMSE-LE has advantages over the inverse filter relating to its performance in a noisy environment. By minimizing the mean-squared error of the communication sequence, it also maximizes the signal-to-interference-plus-noise ratio (SINR). Additionally, convergence is monotonic, as visible in Fig. 2.1, meaning there is no “best number” of iterations, such as was the case in [5]. Recent results also seem to indicate that the MMSE-LE is more robust to channel estimation error [8].

Shown in Fig. 2.1 is a demonstration of the convergence of iterative time-reversal output signal-to-interference-plus-noise to the optimum output SINR of the MMSE-LE. Figure 2.1a shows the output SINR as a function of number of iterations for a sample set of five measured impulse responses [9], which are shown (each shifted in both time and amplitude for visualization purposes) in Fig. 2.1b. The four pairs of curves represent four

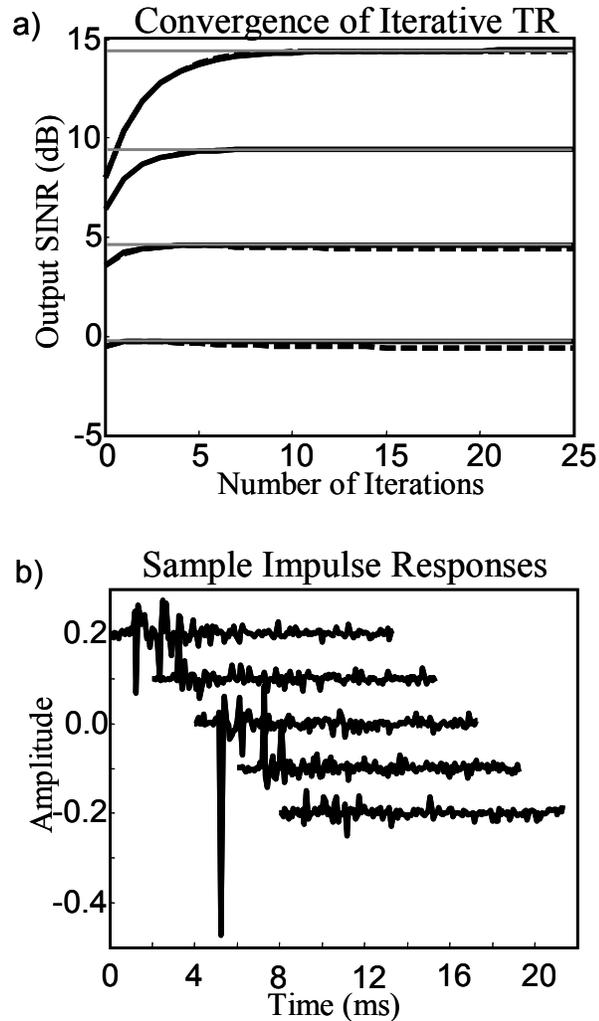


Figure 2.1: Convergence of iterative time reversal. a) The output signal-to-interference-plus-noise ratio (SINR) is shown for four different noise powers (input SNRs), each separated by 5 dB, as a function of number of iterations. The dashed lines correspond to iterative time reversal and the solid lines correspond to the modified iterative time reversal. The optimum SINR calculated explicitly with the MMSE-LE formula is shown in gray. b) The channel used for calculation is a measured at-sea transfer function set with one input, five outputs and 100 taps for each of the five transfer functions.

input SNRs each separated by 5 dB. After each iteration, the combined channel and filter response is converted to the time-domain where SINR calculations are done. The solid lines show the output SINR of modified iterative time-reversal, whereas non-modified iterative time-reversal is shown with dashed curves. Also, the MMSE-LE filter is

calculated explicitly, using the formula of Eq. (2.12), and the combined channel and filter response converted to the time-domain where SINR calculations are done. The results are shown as a thin gray line for each noise power.

Through analysis of the iterative process, one can see the MMSE-LE as a filter set that attempts to cancel its own sidelobes in the time domain, but is regularized so that the gain in the frequency domain is not too large, taking into account the fact that the communications sequence is noisy.

2.3 Convergence

The modified iterative time reversal procedure converges so long as the Neumann expansion of the matrix inverse is valid, that the norm of $(\mathbf{I} - \mathbf{H}\mathbf{H}^H - \sigma^2\mathbf{I})$ is less than one. This isn't a restrictive constraint, as the received signal, \mathbf{Y} , can be multiplied by a constant to ensure this condition is met without loss of optimality, in the sense of maximizing SINR. Scaling the received signal, \mathbf{Y} , is equivalent to scaling both the transfer function matrix, \mathbf{H} , and the noise, thus leaving the SINR, unchanged.

The speed of convergence is determined, as in many gradient methods, by the eigenvalue spread, in frequency, of the matrix $\mathbf{H}\mathbf{H}^H + \sigma^2\mathbf{I}$. The larger the spread, the longer the iterative algorithm takes to converge. As can be seen in Fig. 2.1, convergence occurs with less iteration at lower signal-to-noise ratios because the eigenvalue spread is smaller, as they are dominated by the constant noise components.

2.4 Conclusion

It has been shown that iterative time reversal can be performed passively, resulting in a procedure that converges to the Moore-Penrose pseudo-inverse of the propagation matrix. More importantly, it has been shown that a minor modification, the inclusion of a regularization term, alters the procedure so that it converges to the MMSE linear equalizer. Thus, one can view the MMSE equalizer equivalently as a regularized iterative time reversal process. The iteration systematically reduces temporal sidelobes and the regularization limits the amplification in the frequency domain preventing noisy channels being included in the signal estimate.

2.5 Acknowledgements

This chapter, in full, is a reprint of the material as it appears in the Journal of the Acoustical Society of America, July 2006 with authors, W.J. Higley, P. Roux, and W.A. Kuperman. The dissertation author was the primary investigator and author of this paper. The authors wish to acknowledge helpful and insightful discussions with Bruce Cornuelle. This research was supported by the Office of Naval Research.

REFERENCES

1. T.C. Yang, "Temporal resolution of time-reversal and passive-phase conjugation for underwater acoustic communications," *IEEE J. Oceanic Eng.*, vol 28, pp 229-245, 2003.
2. W.A. Kuperman, W.S. Hodgkiss, H.C. Song, T. Akal, C. Ferla, and D. Jackson, "Phase conjugation in the ocean: Experimental demonstration of an acoustic time-reversal mirror," *J. Acous. Soc. Amer.*, vol. 98, pp. 25-40, 1998.
3. G.F. Edelmann, T. Akal, W.S. Hodgkiss, S. Kim, W.A. Kuperman, and H.C. Song, "An initial demonstration of underwater acoustic communication using time reversal," *IEEE J. Oceanic Eng.*, vol. 27, pp. 602-609, 2002.
4. G. Montaldo, M. Tanter, and M. Fink, "Real time inverse filter focusing through iterative time reversal," *J. Acous. Soc. Amer.*, vol. 115 (2), pp. 768-775, 2004.
5. G. Montaldo, G. Lerosey, A. Derode, A. Tourin, J. de Rosny, and M. Fink, "Telecommunication in a disordered environment with iterative time reversal," *Waves Random Media*, vol. 14, pp. 287-302, 2004.
6. P. Roux, W.A. Kuperman, W.S. Hodgkiss, H.C. Song, T. Akal, and M. Stevenson, "A nonreciprocal implementation of time reversal in the ocean," *J. Acous. Soc. Amer.*, vol. 116, pp. 1009-15, 2004.
7. T.K. Moon and W.C. Stirling, *Mathematical Methods and Algorithms for Signal Processing*, Prentice Hall, 2000.
8. M. Ding, B.L. Evans, and I. Wong, "Effect of channel estimation error on bit rate performance of time domain equalizers," *Signals, Systems and Computers, Conference Record of the Thirty-Eighth Asilomar Conference on*, vol. 2, pp. 2056-60, 2004.
9. H.C. Song, W.S. Hodgkiss, W.A. Kuperman, M. Stevenson, and T. Akal, "Improvement of time reversal communications using adaptive channel equalizers," *IEEE J. Oceanic Eng.* (in press), 2006.

Chapter 3

Synthetic aperture time-reversal communications in shallow water: Experimental demonstration at sea

abstract

Time reversal has been shown as an effective way to focus in both time and space. The temporal focusing properties have been used extensively in underwater acoustics communications. Typical time-reversal communication experiments use vertical transducer arrays both to increase the signal-to-noise ratio and decrease the temporal sidelobes created in the time reversal process. Comparable temporal focusing is achieved using a horizontal array. In this paper, synthetic aperture time reversal communications are accomplished, requiring only two transducers (one transmitter and one receiver). Deriving results from an at-sea experiment, this work confirms the viability of synthetic aperture time reversal communications.

3.1 Introduction

Time reversal has been shown to be an effective way to combat the temporal spreading classically observed in a waveguide [1]. Indeed, after pulse transmission from a point source, if the waveforms recorded at a set of receivers are time-reversed and retransmitted simultaneously into the medium, the result is spatial focusing and temporal compression. Spatial focusing means that the time-reversed field is strong at the location of the original source and relatively weaker elsewhere in the medium. Specifically, spatial focusing implies that the ratio of the intensities between any point not at the focus and that of the focus is low. Temporal compression means that the time-reversed signal

at the source has the same pulse width to the signal previously emitted by the source. Because time reversal refocuses the desired signal in time, it has been applied to underwater communications [2-4].

Recent time reversal communications experiments [1-3] have used a vertical array of transducers, moored to the ocean bottom or hung over the side of a ship. Indeed a vertical array provides spatial diversity that is turned into an advantage when time-reversal is performed [5]. Proposed in this work is a communication system using a synthetic aperture [6] time reversal array that would require only two transducers, a transmitter and a receiver, of which at least one is moving. The advantage of a synthetic aperture array is that it provides spatial diversity with the use of only one source. However, the drawback of a synthetic aperture array is that the waveforms are received at different times, as the transducer is only at one place at any given time. This makes synchronization an issue. Steps must be taken to coherently sum the waveforms received on a synthetic aperture array. The motion of the transducer also induces a Doppler shift which may be different for each waveform, or even time varying. This must also be corrected before the received waveforms can be summed coherently.

Our proposed synthetic aperture communication system is based on the idea that a horizontal time reversal array focuses well in the temporal domain [7]. Time reversal, as a simple and computationally inexpensive form of pulse compression, allows for a faster bit rate than systems without such pulse compression because inter-symbol interference is mitigated. To affirm the feasibility of the proposed communication system to an oceanic environment, an experiment has been performed off the coast of Italy. The goal is to study, experimentally, the acoustical physics of a one-channel time-reversal

communication system in a multipath environment. At this stage, no any encoder/decoder (like turbo codes [8], for example) or equalization [9] to potentially further reduce the error rate at the output of the communication system has been employed.

During the at-sea experiment, passive time reversal [4] rather than active time reversal is used. As long as the communications sequence sent from each synthetic aperture element is not too long, the Green's function used in passive time reversal can be viewed as time-invariant. This is in contrast to the concept that as the transducer moves from synthetic aperture element to element, the Green's functions vary.

The purpose of the experiment is to demonstrate the low bit error rates achievable at sea using a two-element system and passive synthetic aperture time reversal. Section 3.2 describes the experiment performed at sea and section 3.3 details the processing done on the acquired data to extract the bit error rates. In section 3.4, the possibility of using a few vertically oriented hydrophones, which spans only a small portion of the water column, to further lower the bit error rates of the system is examined.

3.2 Experimental Details

A vertical receiver array (VRA) was moored off the coast of Italy, approximately 30 km north of the Island of Elba, where the mean depth, D , was approximately 120 m. In this environment, the sound speed was relatively constant, ranging from 1506 to 1508 m/s. From various geoacoustic experiments in the same area, the bottom properties were obtained through inversion yielding a sound speed of 1600 m/s and an attenuation of 0.3 dB/ λ . This vertical array consisted of 32 elements each separated by 2 m, spanning from

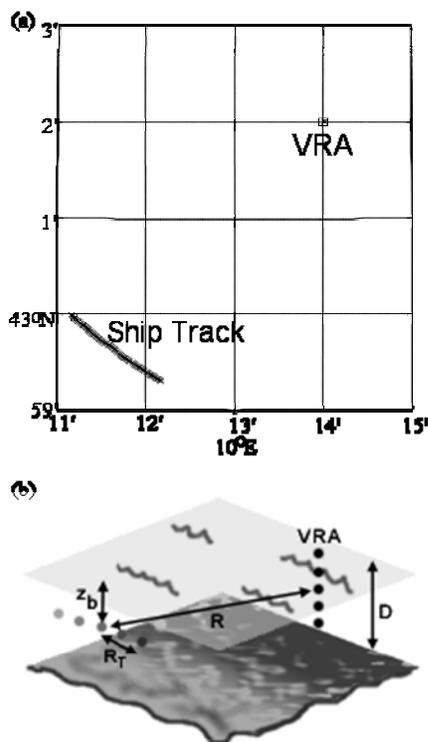


Figure 3.1: a) A vertical receiver array (VRA) moored off the coast of Italy received signals from a transducer towed by a moving vessel following the path shown. Each X represents a position in the almost 2-km long synthetic aperture array. It is of note that the synthetic aperture array is oriented in such a manner that the VRA is approximately broadside to the track. The average distance between the moving source and the VRA is 5.5 km. b) The same set-up from another angle, showing the depth of the water, D , the range of the synthetic aperture array, R , the depth of the array, z_b , and the distance along the track of a single element, R_T .

30 to 92 m in the water column. A transducer was towed (at a 35-m depth) by a moving vessel at a speed of 2 knots on a track shown in Fig. 3.1. It is noticeable that the track is oriented such that the VRA is broadside to the synthetic aperture array (as viewed from above), as opposed to endfire. In a range-independent environment, the Green's function should not *a priori* vary over such a track path, thus limiting the gain in spatial diversity. However, at the frequency of 3 kHz, the wavelength (0.5 m) is very small compared to the propagation range ($R \sim 5.5$ km) so that a minimal bathymetry change leads to a

noticeable modification of the Green's function. Also, the path is such that the VRA is not perfectly broadside to the track, and the small amount of range difference between elements may be enough to induce changes in the Green's function. A measure of the Green's functions change is discussed quantitatively below. The moving transducer transmitted a 1 ms pulse at 3 kHz, followed by a 10-s communication sequence coded using binary Amplitude Shift Keying (ASK), also at 3 kHz, consisting of 9800 bits with a bit length of 1 ms. The initial choice of an incoherent communication scheme (versus a coherent scheme like Binary Phase Shift Keying) was driven by its robustness to Doppler or synchronization issues [10]. The pulse and the communication sequence were separated by 200 ms so that they do not overlap after propagation through the dispersive channel. The purpose of the pulse is to probe the medium and to provide the Green's function that will be used to deconvolve the communication sequence [4]. The incident field was then received by all 32 channels on the vertical array, producing a matrix of Green's functions. In section 3.3, passive time-reversal is performed on one element of the vertical array only, whereas a 4-m vertical aperture made of three contiguous elements is used in section 3.4. The pulse and communication sequence were transmitted every thirty seconds, each time producing a new Green's function matrix. During a 33 minute long track, 65 such matrices of Green's functions were acquired separated by a mean distance of 30 m. The synthetic aperture array thus consisted of up to 65 elements (though, typically no more than four at a time are examined in this paper) spanning just less than 2 km. The signal-to-noise ratio of the communication sequence varied between 25 and 30 dB over the ship track (see Fig. 3.5).

3.3 Data Processing

3.3.A. Decorrelation along the track

As can be seen in Fig. 3.1, the track of the moving transducer creates a synthetic aperture array, such that the VRA (as viewed from above) is oriented broadside to this array. Considering, for now, only one channel on the VRA (depth ~ 38 m), the source depth, receiver depth, and range between the source and the receiver can all be considered approximately constant. To distinguish between different transmissions of the moving transducer, the variable, R_T , which is the distance along the track at the time of transmission, is introduced. Thus, the Green's function between a point on the track and the VRA can be written as $G(t, R_T)$. The fact that the synthetic aperture is oriented broadside to the VRA rather than endfire does not present a problem for the time-reversal process since the Green's function changes azimuthally, thereby providing the required spatial diversity. The change in Green's function is calculated as follows: The Green's functions are normalized to have unit power and autocorrelated. Each autocorrelated Green's function corresponds to the point-to-point time reversal between one VRA element and the moving source (see Eq. (3.1)). Spatial diversity will be used efficiently if the time-reversal sidelobes cancel each other along the source track, i.e. the temporal sidelobes of the autocorrelated Green's function decorrelates as the source moves. The point-to-point time-reversal sidelobes decorrelation is shown in Fig. 3.2. It can be noticed that the Green's function changes enough to affect the sidelobes of the autocorrelation over a distance that varies between 100-300 m along the track.

The change in Green's functions along the track can be attributed to either the 30 m azimuthal or the 3 m radial distance (with respect to the VRA) between each source

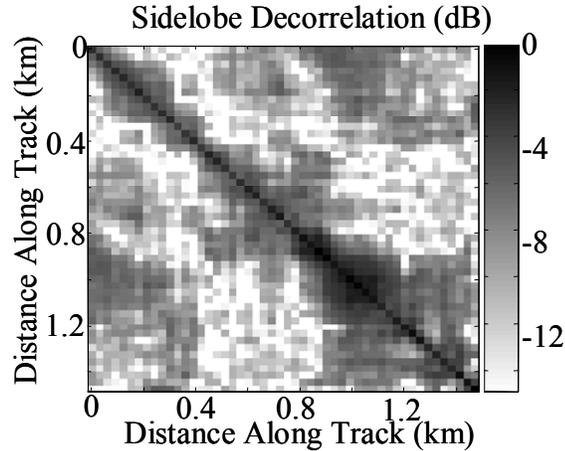


Figure 3.2: Point-to-point time-reversal sidelobe decorrelation. These results are obtained by first normalizing the Green's functions at the distances along the track indicated on the axis, autocorrelating each of them, and then cross-correlating the resulting sidelobes. The results are a measure of how much the Green's function changes as a function of the distance along the track. The distance between Green's functions that results in a total decorrelation of the point-to-point time reversal sidelobes varies between 100-300 m.

position. In the extreme case, where 300 m along the track is necessary to decorrelate the autocorrelation sidelobes, this corresponds to less than 300 m azimuthal distance and 30 m radial distance. There is an ambiguity as to which of these distances causes the decorrelation. However, this radial distance corresponds to 60 wavelengths, which is enough to cause a decorrelation of the field in the waveguide. In the other extreme case, where only 100 m is necessary to decorrelate the autocorrelation sidelobes, it is likely that a combination of azimuthal and radial distance changes are responsible for the decorrelation.

3.3.B. *Extracting bit error rates*

After propagation through the channel, the received signals are not matched in frequency due to time-dependant Doppler shift. The Doppler shift may be either due to relative motion between the two transducers, or a mismatch between the sampling

frequencies of the transducers. Also, as the signals are not received simultaneously, they are not synchronized in time. Both of these effects must be compensated before the signals can be added coherently.

To correct for the Doppler shift, an FFT is taken of the zero-padded communications sequence. As the Green's function of this data set have a strong ballistic path and the communications are encoded with ASK, in the absence of Doppler shift, one would expect a strong peak at 3 kHz, the carrier frequency. Due to Doppler effects, this peak is shifted a small amount, ranging from .4-.9 Hz. If the motion of the transducer were endfire, one would expect the Doppler shift to be greater. The signal is then dilated or compressed in time so that the value of the peak is shifted back to 3 kHz. Fig. 3.3 shows the coherent sum over the synthetic aperture of all received time-domain signals (after synchronization and time reversal described below) with no Doppler correction in gray, and with Doppler correction in black. It is seen that the two sums are similar in

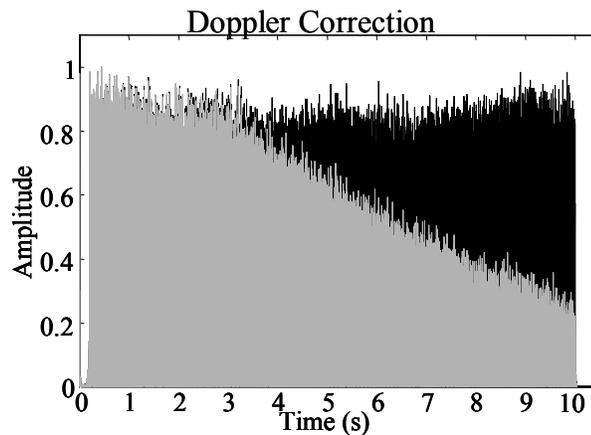


Figure 3.3: The results of summing coherently the 65 received waveforms before Doppler correction (in gray) and after Doppler correction (in black). As time progresses, it is shown that the amplitude of the non-corrected sum decreases as the non-corrected signals do not add coherently at later times.

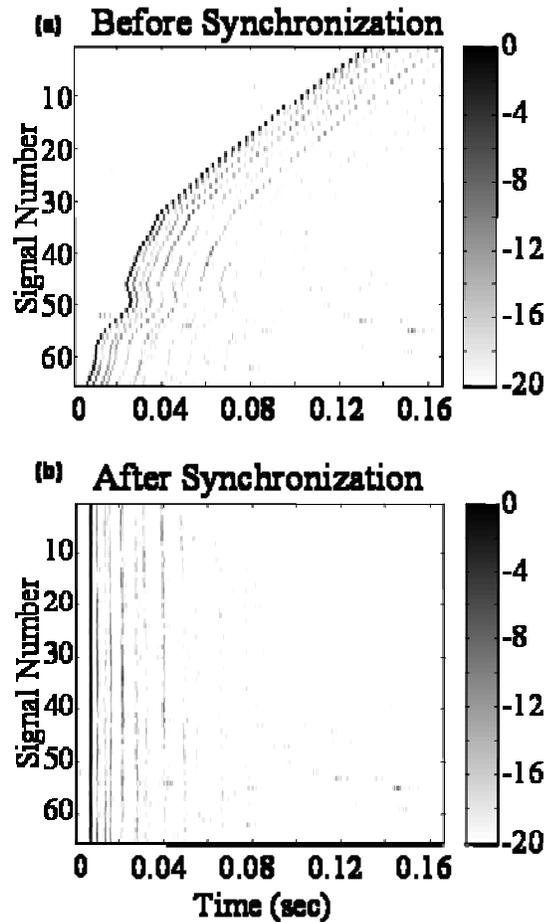


Figure 3.4: The received waveforms before and after synchronization. Synchronization is an issue as time-reversal can't be performed simultaneously from the synthetic aperture array. The signals must be accurately synchronized before time-reversal so they may be summed coherently. Shown are the received waveforms in dB, normalized to the maximum.

amplitude at the beginning, but that as time progresses, the amplitude of the non-corrected sum decreases. This is due to the fact that the non-corrected signals do not add coherently at later times.

Once the signal has been Doppler corrected, the next step is to synchronize the signals in time. In order to do that, the signal with the highest signal-to-noise ratio is used as a reference. As mentioned earlier, the received signal consists of two parts, the

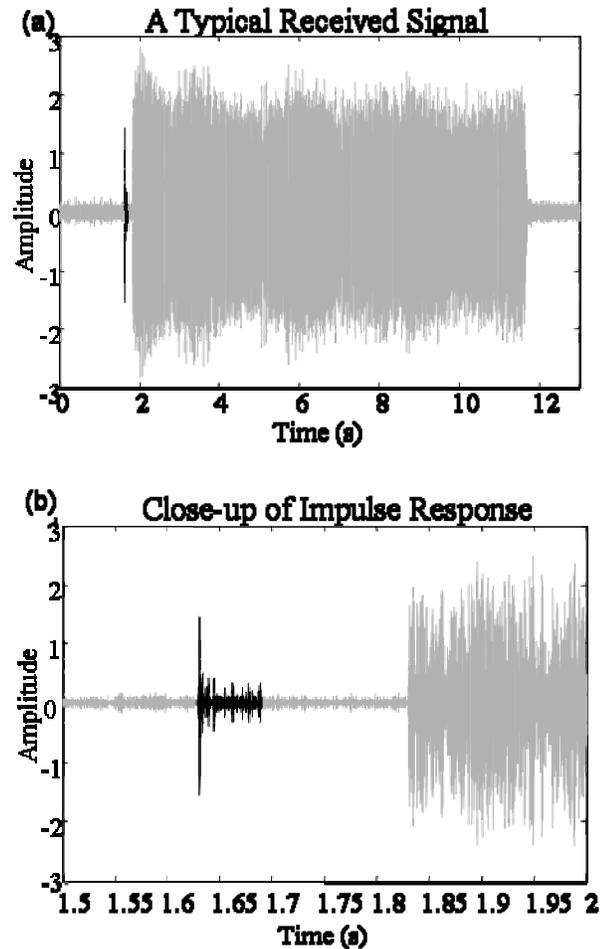


Figure 3.5: a) A typical received signal consisting of an initial pulse and a 10 s communication sequence, after both have propagated through the medium. The selected Green's function is shown in black. b) The same signal, zoomed in to show the detail of the Green's function, which has spread from an original 1 ms pulse, to a 70 ms long dispersed signal. Signal-to-noise ratio is about 25 dB.

dispersed pulse, which has become the Green's function, and the dispersed communication sequence. In order to synchronize the waveforms, one may focus on the Green's function. Knowing the time between the beginning of the Green's function and the communication sequence (200 ms), the synchronization of the Green's functions along the track yields the synchronization of all the communication sequences. Fig. 3.4

shows the 65 Green's functions received before and after synchronization. For each of the 65 received waveforms the following steps are taken. The Green's functions are time-gated to select the first arrival of the impulse response. The first arrivals are then synchronized along the aperture using cross-correlation. After synchronization of the first arrival a larger time window, approximately 70 ms, is taken and designated the Green's function. Figure 3.5 shows a typical signal received on the VRA, with the selected Green's function darkened. In the case of a set of Green's functions lacking a dominant first arrival, the method of synchronization would have to be modified [6]. For example, a particularly strong and stable path of the Green's function could be chosen in place of the first arrival. In the case of low signal-to-noise ratios, one could consider that the only thing in common between any two received waveforms is the communication sequence, $m(t)$ (see Eq. (3.1)). Cross-correlation between received waveforms will exhibit a peak when the embedded sequence in each waveform overlaps. The time at which these peaks occur correspond to the relative delay in each received signal. By removing this delay, the signals are synchronized.

After synchronization, each Green's function is extracted, time reversed and convolved with the communication sequence to cancel the multi-path propagation as classically done in passive time-reversal. The result of these convolutions is sixty-five waveforms that can be summed coherently to produce the message estimate, written as

$$\hat{m}(t) = m(t) * \sum_j [G(-t, R_{Tj}) * G(t, R_{Tj})], \quad (3.1)$$

where $m(t)$ is the coded communication sequence. The summation part of the equation has been referred to as the Q-function [10]. A typical Q-function for one element (which

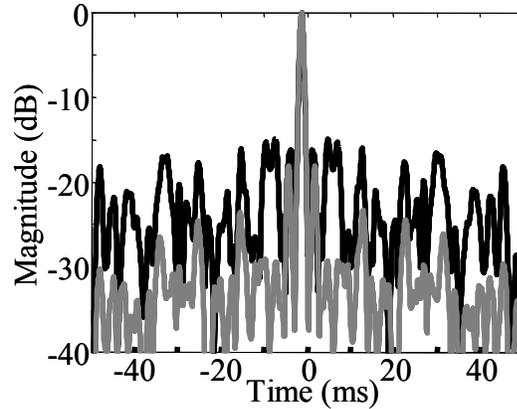


Figure 3.6: The Q-function [10] resulting from a single element is shown in black, and from the coherent summation of five widely-spaced elements is shown in gray. It can be seen that the summation results in lower sidelobe levels, which produce lower bit error rates.

is the point-to-point time-reversal described in Section 3.3.A) and for five elements is shown in Fig. 3.6. As expected, the time-reversal sidelobes are reduced as more elements are added coherently, suggesting that better communication can be achieved with a time reversal synthetic aperture.

Although the modulation used (ASK) is an incoherent modulation, it can be demodulated either incoherently or coherently. In either case, the message estimate is base-banded by multiplication with a sinusoid at the carrier frequency and filtered. The resulting waveform is then resampled at the symbol rate producing N complex numbers, where N is the number of symbols. To demodulate incoherently, the absolute value of each of these numbers is compared to a threshold to determine whether the bit is a binary 1 or a binary 0. The threshold is chosen as half the mean of the N positive numbers. Fig. 3.7a shows the resulting values from incoherent modulation before the hard decision is made. The figure was made using only three widely-spaced synthetic aperture elements

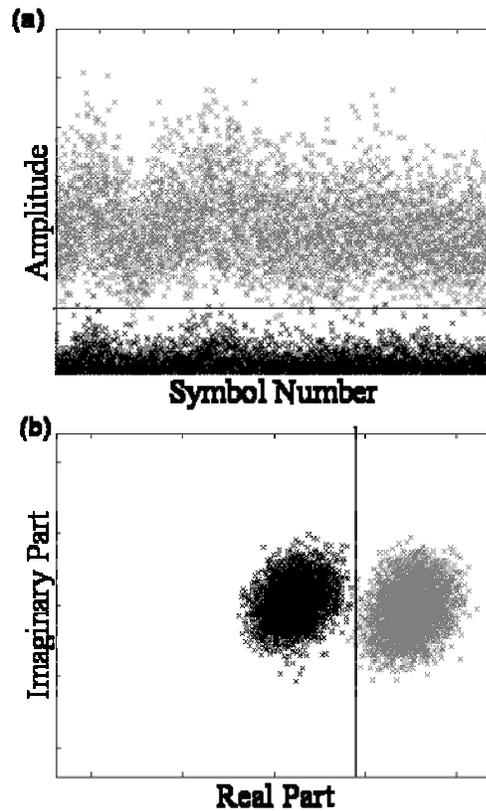


Figure 3.7: a) The results of incoherently decoding a signal composed of the sum of three widely-spaced synthetic aperture elements, that is synthetic aperture elements that have decorrelated Q-function sidelobes. Binary 0s are shown in black, binary 1s are shown in gray, and a threshold is indicated by a dark line. The number of errors in the case shown was 31, giving a bit error rate of .32%. b) The results of decoding the same signal coherently. Again, binary 0s are shown in black, binary 1s are shown in gray, and a threshold is indicated by a dark line. The number of errors in this case, decoded coherently was 7, giving a bit error rate of .071%.

with decorrelated Q-function sidelobes. In Fig. 3.7a, binary 0s are shown in black and binary 1s are shown in gray. The threshold is shown by a dark line. To demodulate coherently, the real part of each of the N complex numbers is compared to a threshold to determine whether the bit is a binary 1 or a binary 0. The threshold is chosen as the mean of the real part of the N complex numbers. Fig. 3.7b shows the resulting values from the coherent modulation before the hard decision is made. This figure was made using the

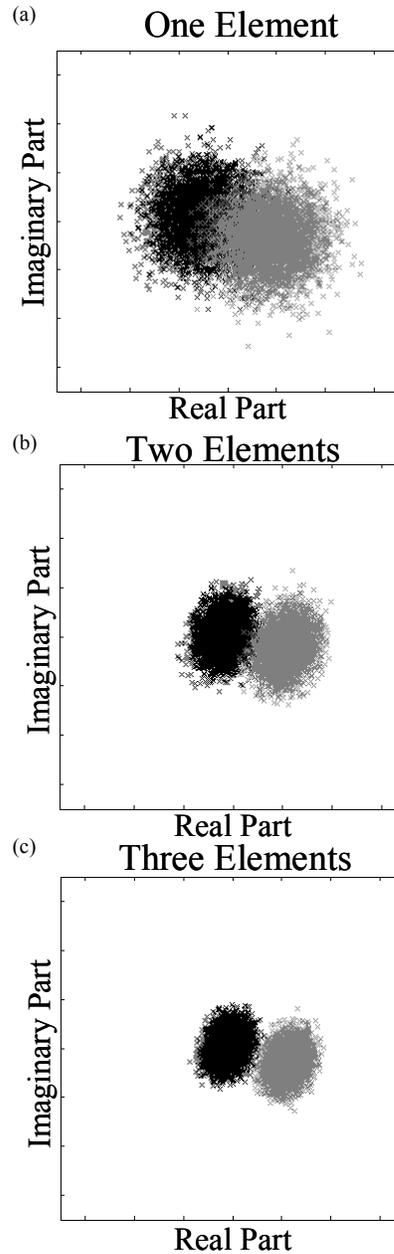


Figure 3.8: The evolution of the in-phase/quadrature plot as more synthetic elements are summed together before decoding. The elements added are widely-spaced. Binary 0s are in black and binary 1s are in gray.

same three widely-spaced synthetic aperture elements. Again, binary 0s are shown in black, binary 1s are shown in gray, and threshold is shown by a dark line. The

comparison between the estimate sequence and the sequence sent produces a bit error rate. From Fig. 3.7, it can be deduced that coherent demodulation is not only possible, but gives lower bit error rates than the incoherent demodulation. From this, one may anticipate that it is possible to use synthetic aperture time reversal with more complicated, coherent modulation techniques such as BPSK or QPSK. As coherent demodulation gives lower bit error rates, the rest of the paper presents results for coherent modulation. Lower bit error rates are expected as more synthetic aperture elements are added together, which is seen in Fig. 3.8. This figure shows the evolution of the in-phase/quadrature plot as synthetic aperture elements are summed. As seen in Fig. 3.8, the result of using more synthetic aperture elements are tighter clouds of points and ultimately, lower bit error rates.

3.3.C. Results

We present the bit error rates as a function of two parameters of the synthetic aperture array: the number of elements used in summation and the element spacing. Despite the orientation of the track, such that the VRA is broadside to it, it is shown that the Q-function sidelobes decorrelate as the distance along the track increases, thus element spacing plays an important role in determining bit error rates. Results are shown as a grayscale plot in Fig. 3.9. The bit error rate decreases as the number of elements is increased and also decreases as the element spacing is increased. However, by increasing the number of elements used, as this is a synthetic aperture, the effective data rate is

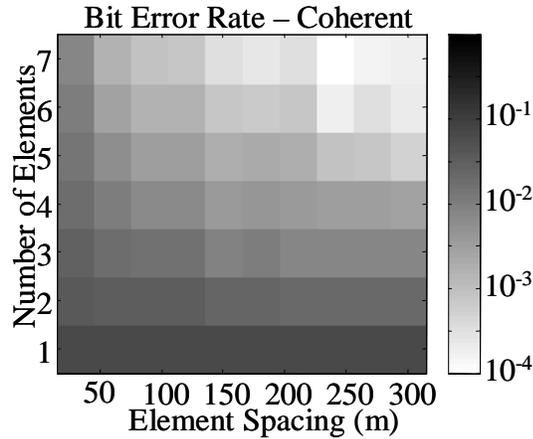


Figure 3.9: Bit error rate found from 9800 bits, each with a bit length of 1 ms, as a function of the number of transducer positions (or synthetic array elements) for various values of element spacing, δR_T .

diminished. Using only one element in the system, corresponding to a data rate of 1000 bps, yields a bit error rate of 7%. By using two elements spaced 300 m apart, corresponding to a data rate of 500 bps, the bit error rate is dramatically lowered to 2.1%. An increase to three elements with a mean spacing of 300 m, corresponding to a data rate of 333 bps, yields a bit error rate below 1%.

3.3.D. Interleaving communication sequences

If one were to transmit as quickly as allowed by the equipment, the element spacing, δR_T , would be small, requiring a large number of transducer positions to achieve a low bit error rate. For instance, in Table 3.1, it is shown that it takes seven elements to achieve a bit error rate of .75% when the element spacing is only 30 m, but when the element spacing is 300 m, the same bit error rate is achieved with only three elements. The spacing of 300 m is chosen from the decorrelation of the Q-function sidelobes shown in Fig. 3.2. If one were to simply increase the aperture by increasing the element spacing,

δR_T , there would be times when the transducers were unused since it is necessary to wait for one of the transducers to move and the aperture to further increase. In this spare time between transmissions of the first communication sequence, other sequences can be sent. By interleaving communication sequences, the data rate is the same as the fastest allowed by the equipment, but the bit error rate will be significantly lower. Fig. 3.10 demonstrates this concept graphically.

(a)XXXXOOOO++++
 (b)X X X X O O O O + + + +
 (c)XO+XO+XO+XO+

Figure 3.10: Schematic of a synthetic aperture time reversal communication system. For the sake of simplicity, the communication sequences are represented by the symbols X, O, and +. a) Multiple sequences transmitted as quickly as allowed by the equipment. b) The same sequences, with the spacing between transmissions increased in order to increase the length of the synthetic aperture endfire aperture. This creates a better focus and decreases the bit error rate. c) By interleaving sequences, one may achieve the same bit error rate as in step b, but maintains the same data rate as in step a.

3.4 Vertical Summation

The experiment at sea was done using a vertical array consisting of thirty-two elements, only one of which being used so far in decoding the data. By using up to three contiguous elements, one may examine the decrease in bit error rate that would occur if one were to use a 4-m long vertical array (centered at 38 m) that spanned only a small portion of the 120-m deep water column. Such an array would still be relatively cheap, one of the primary advantages of the proposed system. The in-phase/quadrature plots

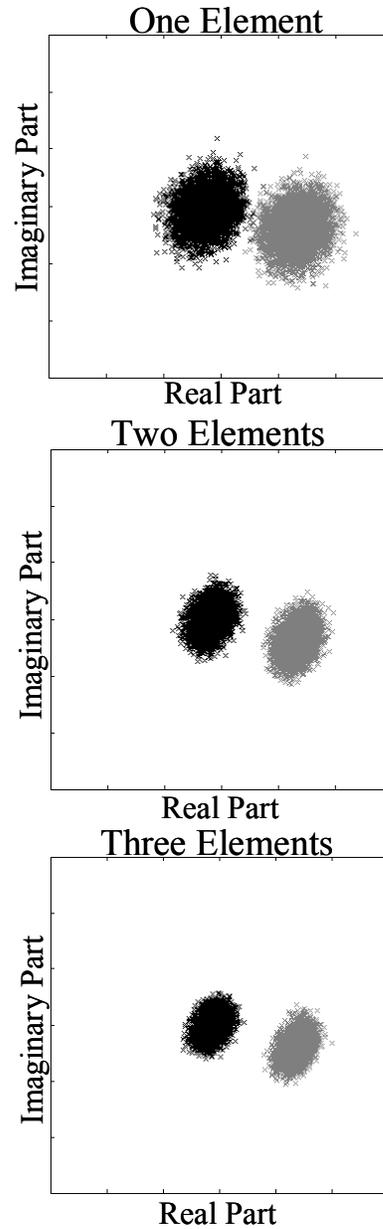


Figure 3.11: The evolution of the in-phase/quadrature plot as more vertical elements are summed together before decoding. Before summation over the vertical elements, a horizontal synthetic aperture of three widely-spaced elements is used.

resulting from the summation over a small number of vertically oriented elements are shown in Fig. 3.11. The summation is done vertically after the horizontal summation had been done over three widely-spaced synthetic aperture elements. The bit error rates using

a small number of vertically oriented elements are presented in Fig. 3.12. The addition of a few vertical elements dramatically lowers the bit error rates of the system. For example, with a synthetic aperture of four elements, a vertical line array of three elements yields 12 waveforms that can be coherently summed. The results demonstrate that with a vertical line array of three elements, at three locations forming a synthetic horizontal line array, bit error rates lower than 10^{-5} can be achieved with a 1 ms bit length.

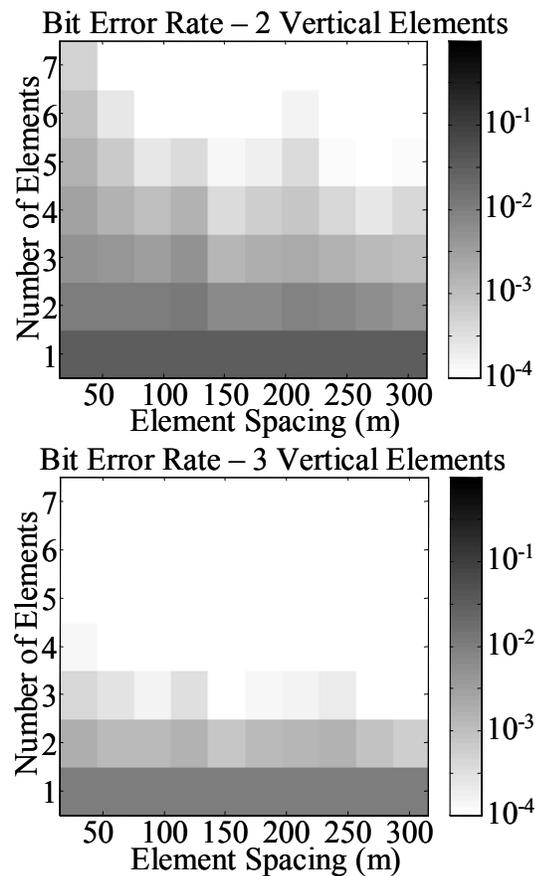


Figure 3.12: Bit error rate found from 9800 bits, each with a bit length of 1 ms, as a function of the number of synthetic array elements, element spacing, δR_T , and number of vertical array elements. The bit error rates using only one vertical element can be found in Figure 3.9.

3.5 Conclusion

This paper serves to confirm the feasibility of an inexpensive communication system using a synthetic aperture time reversal array at sea. The main benefit of this system is that it requires only two transducers (one transmitter and one receiver), significantly decreasing the technological cost compared to typical time reversal communication schemes. The disadvantage of this system is the repeated transmission of the communication sequence tends to lower to data rate. As in many communication systems, a tradeoff must be made between the data rate and the desired bit error rate. It is also shown that the addition of a few transducers oriented vertically in addition to the horizontal aperture significantly decreases the bit error rate without a significant increase in cost or complexity of the system. Finally, it is reiterated that this study is limited to how synthetic aperture time reversal can utilize shallow water propagation complexity; additional communication algorithms can always be appended to increase performance.

3.6 Acknowledgements

This chapter, in full, is a reprint of the material as it appears in the Journal of the Acoustical Society of America, October 2002 with authors, W.J. Higley, P. Roux, W.A. Kuperman, W.S. Hodgkiss, H.C. Song, and T. Akal. The dissertation author was the primary investigator and author of this paper.

Table 3.1. Selected values of the bit error rate found from 9800 bits, each with a bit length of 1 ms, as a function of the number of transducer positions (or synthetic array elements) for various values of element spacing, δR_T .

δR_T	Data Rate in bits per second (Number of Elements in the Synthetic Aperture)						
	1000 (1)	500 (2)	333 (3)	250 (4)	200 (5)	167 (6)	143 (7)
30 m	6.9e-2	4.2e-2	2.8e-2	2.0e-2	1.5e-2	1.0e-2	7.5e-3
60 m	-	3.6e-2	2.0e-2	1.1e-2	6.1e-3	3.1e-3	1.6e-3
150 m	-	2.6e-2	9.5e-3	4.1e-3	1.9e-3	8.5e-4	3.2e-4
300 m	-	2.1e-2	7.6e-3	2.9e-3	5.3e-4	2.3e-4	2.0e-4

Table 3.2. Selected values of the bit error rate found from 9800 bits, each with a bit length of 1 ms, summed over three vertical elements, as a function of the number of transducer positions (or synthetic array elements) for various values of element spacing, δR_T .

TABLE 3.2
BIT ERROR RATE FOR VARIOUS ELEMENT SPACINGS OVER THREE VERTICAL ELEMENTS

δR_T	Data Rate in bits per second (Number of Elements in the Synthetic Aperture)						
	1000 (1)	500 (2)	333 (3)	250 (4)	200 (5)	167 (6)	143 (7)
30 m	1.0e-2	1.9e-3	4.4e-4	1.4e-4	7.3e-5	2.5e-5	1.4e-5
60 m	-	1.3e-3	2.8e-4	5.3e-5	1.8e-5	3.7e-6	0
150 m	-	7.6e-4	7.1e-5	4.1e-6	2.2e-6	0	0
300 m	-	6.1e-4	7.7e-5	2.3e-5	4.1e-6	0	0

REFERENCES

1. W.A. Kuperman, W.S. Hodgkiss, H.C. Song, T. Akal, C. Ferla, and D. Jackson, "Phase conjugation in the ocean: Experimental demonstration of an acoustic time-reversal mirror," *J. Acous. Soc. Amer.*, vol. 98, pp. 25-40, 1998.
2. G.F. Edelmann, T. Akal, W.S. Hodgkiss, S. Kim, W.A. Kuperman, and H.C. Song, "An initial demonstration of underwater acoustic communication using time reversal," *IEEE J. Oceanic. Eng.*, vol. 27, pp. 602-609, 2002.
3. K.B. Smith, A.A.M. Abrantes, and A. Larraza, "Examination of time-reversal acoustics in shallow water and applications to noncoherent underwater acoustic communications," *J. Acous. Soc. Amer.*, vol. 113, pp. 3095-3110, 2003.
4. D. Rouseff, D. R. Jackson, W. L. J. Fox, C. D. Joes, J. A. Ritcey, and D. R. Dowling, "Underwater acoustic communication by passive phase conjugation: Theory and experimental results," *IEEE J. Oceanic Eng.*, vol. 26, pp. 821-831, 2001.
5. P. Roux and M. Fink, "Time reversal in a waveguide: Study of the temporal and spatial focusing," *J. Acous. Soc. Amer.*, vol. 107, pp. 2418-2429, 2000.
6. S. Stergiopoulos, "Implementation of adaptive and synthetic-aperture processing schemes in integrated active-passive sonar systems" *Proceedings of the IEEE*, vol. 86, pp. 358-398, 1998.
7. M.R. Dungan and D.R. Dowling, "Orientation effects on linear time reversing array retrofocusing in shallow water," *J. Acous. Soc. Amer.*, vol. 112, pp. 1842-1852, 2002.
8. C. Berrou and A. Glavieux, "Near optimum error correcting coding and decoding: turbo-codes", *IEEE Transactions on communications*, 44 (10), pp. 1261-1271, 1996.
9. G.F. Edelmann, S. Kim, W.S. Hodgkiss, W.A. Kuperman, H.C. Song, and T. Akal, "Combining and comparing time reversal processing and adaptive channel equalization for communications sequences," *J. Acous. Soc. Amer.*, vol 112, p. 2447, 2002.
10. T.C. Yang, "Temporal resolution of time-reversal and passive-phase conjugation for underwater acoustic communications," *IEEE J. Oceanic Eng.*, vol 28, pp 229-245, 2003.

Chapter 4

Constellation-constrained capacity of an undersea multiple-input/multiple-output (MIMO) channel

abstract

The use of multiple receivers and/or multiple transmitters in communications has recently been shown to increase the effective capacity of a channel without any additional bandwidth or power. Closed-forms expressions for the capacity of such a channel with an additional constraint that all transmitted symbols be independent and identically distributed from a given constellation are not currently known and techniques involving simulation are computationally prohibitive for many transmitters or long dispersion times. Previously introduced bounds on a single-input/single-output channel so constrained are generalized to the multiple-input/multiple-output case. This allows the question of optimal distribution of a fixed number of transducers to be addressed.

4.1 Introduction

As communications systems strive for higher and higher data rates, they are limited by constraints on power, bandwidth, and complexity. Recent information theoretic results suggest that an additional domain can be used to achieve greater spectral efficiencies, that of multiple transmitters and receivers. Through the use of multiple transmitters and receivers, the effective channel has a greater capacity with no additional power or bandwidth.

To achieve these predicted spectral efficiencies, however, requires a channel such that the matrix of channel gains at each frequency has full rank and independent entries and

that perfect estimates of these gains are available at the receiver. Additional gains are to be had if perfect estimates of these gains are available at the transmitter.

Although no additional power or bandwidth is used to achieve greater capacities with a multiple-input/multiple-output (MIMO) system, complexity is an issue. The capacity of a system that constrains each transmitter to independent and identically distributed (i.i.d.) symbols drawn from a given constellation is of interest, as such a system is relatively easy to implement. Such MIMO systems arise from the use of arrays that have been recently demonstrated in underwater acoustic communication experiments [1]. Equalization techniques in underwater acoustic communications experiments have recently involved time reversal [2]-[5], optimum retrofocusing techniques [6], and more advanced methods [7],[8]. Although, the capacity for the MIMO underwater acoustic channel with a simple power constraint can be examined in simulation [9], closed form expressions are unknown for the capacity of such channels with a constellation constraint at each transmitter. The capacity is calculable for memoryless channels through numerical integration and for dispersive channels through simulation. Unfortunately, such simulation becomes computationally prohibitive for systems with many transmitters/receivers or long dispersion times. Various bounds on the capacity, previous applied to single-input/single-output (SISO) channels, can be extended to the MIMO channel.

An underwater acoustics communications experiment was performed off the coast of Italy to acquire a matrix of transfer functions for analysis. The goal of this paper is two-fold. First, calculation of the capacity of a measured MIMO ISI channel with and without channel knowledge at the transmitter is done and compared. Second, and more restrictive,

an i.i.d. constellation constraint is imposed at the transmitter. Upper and lower bounds for the capacity of the channel under this constraint are derived and applied to the data set. The generalization of these bounds allows the question of optimum placement of a fixed number of transducers to be answered.

Section 4.2 defines our channel model and describes the acquisition of a data set that adheres to this model. In Section 4.3, the capacity of the channel with and without channel knowledge at the transmitter is calculated and compared. In Section 4.4, bounds for the channel with an i.i.d. constellation constraint imposed are derived and applied to the data set. Section 4.5 uses the expressions for and bounds on capacity to determine the optimum placement of a fixed number of transducers.

4.2 Channel Model and Data Acquisition

The linear time-invariant discrete-time multiple-input/multiple-output intersymbol interference (MIMO ISI) channel with N_T transmitters and N_R receivers is considered. The channel length (in symbols) is denoted L . Thus, one may write the received signals as

$$y_i[n] = \sum_{j=1}^{N_T} \sum_{l=0}^{L-1} h_{ij}[l] x_j[n-l] + w_i[n] \quad \text{for } 1 \leq i \leq N_R, \quad (4.1)$$

where $x_j[n]$ is the transmitted symbol from transmitter j at time n , $h_{ij}[n]$ is the transfer function between transmitter j and receiver i , and $w_i[n]$ is additive Gaussian noise. In this paper, the noise is assumed to be both temporally and spatially white for convenience. This may be written more simply in vector notation as

$$\mathbf{y}[n] = \sum_{l=0}^{L-1} \mathbf{h}[l]\mathbf{x}[n-l] + \mathbf{w}[n]. \quad (4.2)$$

Additionally, in the frequency domain, this equation is written as

$$\mathbf{Y}(\omega) = \mathbf{H}(\omega)\mathbf{X}(\omega) + \mathbf{W}(\omega), \quad (4.3)$$

where ω is frequency and capitalization indicates the Fourier transform of a quantity. This channel model arises in a number of settings, such as magnetic recording, wireless communications, and underwater acoustic communications.

To measure an underwater acoustic communications transfer function matrix, an experiment was performed off the coast of Italy. The experiment (FAF05) represents a joint effort between NATO Underwater Research Centre (NURC) and the Marine Physical Laboratory (MPL) at Scripps Institution of Oceanography. Two vertical arrays were moored 4 km from each other, approximately 15 km north of the Island of Elba, where the mean depth was approximately 120 m. In this environment, the sound speed was slightly downward refracting, with a mean sound speed of 1508 m/s. From various geoacoustic experiments in the same area, the bottom properties were obtained through inversion yielding a sound speed of 1600 m/s and an attenuation of 0.3 dB/ λ . The two vertical arrays consisted of a 32 element receive array with 2 m spacing and a 29 element source/receive with 2.8 m spacing. The transfer function matrix between these two arrays was measured between 3 and 4 kHz, appropriately filtered, basebanded, and downsampled in time. Further, the matrix was downsampled in space to yield the transfer function matrix between two 10-element arrays, with vertically spacings of 6 m and 8.4 m. The matrix was scaled by a normalizing constant such that amount of energy in any point-to-point impulse response was less than one. Fig. 4.1a shows the geometry of the

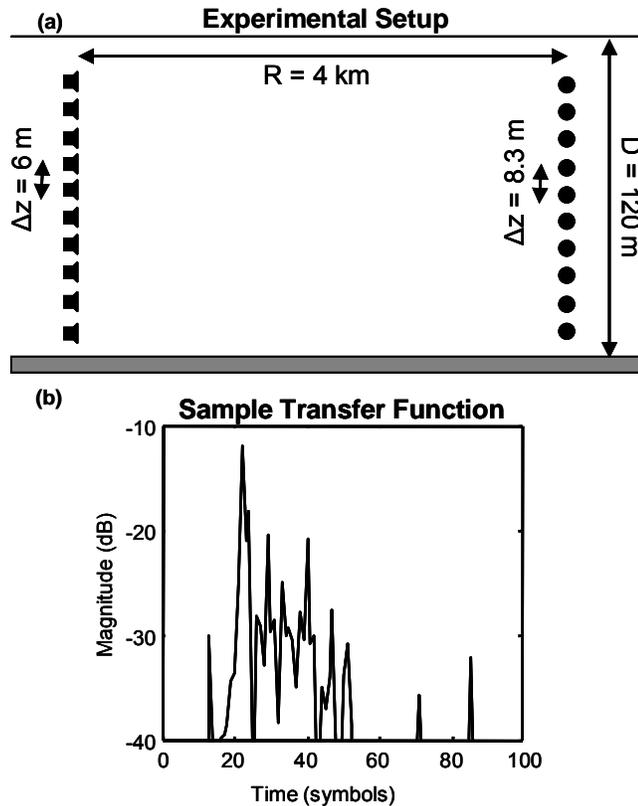


Figure 4.1: a) A diagram showing the experimental setup. The impulse responses used represent two arrays of 10 elements, the sources separated by 6 m and the receivers by 8.4 m. The arrays are situated in 120 m deep water approximately 4 km apart. b) A typical point-to-point impulse response of an underwater acoustic communications experiment performed off the coast of Italy. The transfer function is characterized by a long dispersion time (approximately 70 symbols) and non-Gaussian nature.

setup, and Fig. 4.1b shows a typical point-to-point impulse response, characterized by a long dispersion time (approximately 70 symbols) and the presence of sparse multipath arrivals. For the analysis presented in this paper, these impulse responses will be considered the true, noise-free impulse responses.

4.3 Value of Channel Knowledge

The information theoretic capacity of this channel [10] when channel knowledge is

not available at the transmitting side can be calculated [11] as

$$C_{UG} = \frac{1}{2\pi} \int_{-\pi}^{\pi} \log \left(\det \left(\mathbf{I}_{N_r} + \frac{P_t}{\sigma^2 N_r} \mathbf{H}^H(\omega) \mathbf{H}(\omega) \right) \right) d\omega, \quad (4.4)$$

where σ^2 is the noise variance, P_t is the total transmitted power, $\det(\bullet)$ is the determinant of a matrix, and \mathbf{I}_N is the $N \times N$ identity matrix. When channel knowledge is available at the transmitting end, power can be optimally allocated to maximize capacity in a process known as water-filling. Water-filling is the solution to the optimization problem of maximizing capacity with the constraints that total power is fixed and power must be positive. The capacity when water-filling is employable is calculated [12],[13] as

$$C_{WF} = \frac{1}{2\pi} \int_{-\pi}^{\pi} \sum_n \log \left(1 + \frac{P_n(\omega) \lambda_n(\omega)}{\sigma^2} \right) d\omega$$

$$\text{with } P_n(\omega) = \left(\Theta - \frac{\sigma^2}{\lambda_n(\omega)} \right)^+ \text{ and } \frac{1}{2\pi} \int_{-\pi}^{\pi} \sum_n P_n(\omega) d\omega = P_t. \quad (4.5)$$

where $\lambda_n(\omega)$ is the n th eigenvalue of $\mathbf{H}^H(\omega) \mathbf{H}(\omega)$, $P_n(\omega)$ is the power allocated to the n th channel, Θ is the water-fill level, and $(a)^+$ is defined as $\max(a, 0)$. Input signal-to-noise ratio (SNR) is defined as the total transmitted power divided by the noise power, i.e. P_t/σ^2 .

Figure 4.2 compares the two expressions for capacity. It can be seen in Fig. 4.2a, 4.2b, and 4.2c that the gains achievable through water-filling decrease at higher signal-to-noise ratio (SNR) and tend towards zero when the number of receivers is equal to or greater than the number of transmitters. This is true as long as the matrices $\mathbf{H}^H(\omega) \mathbf{H}(\omega)$ are of full rank. In contrast, when the number of transmitters is greater than the number of receivers, the difference tends towards a positive non-zero constant, as seen in Fig. 4.2d.

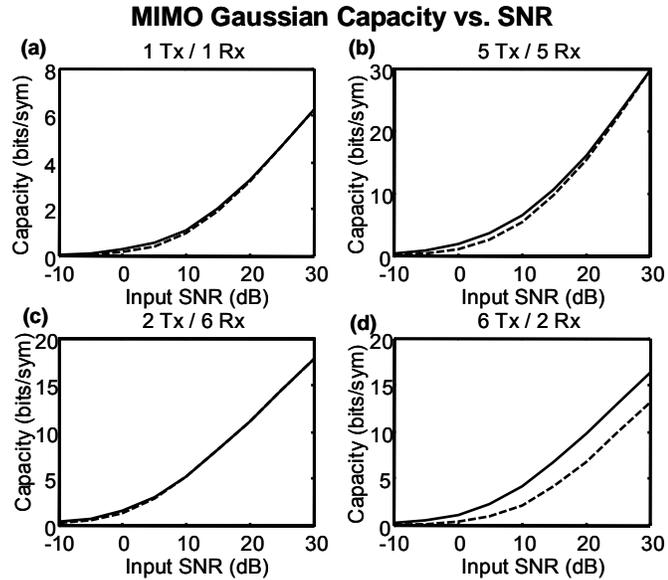


Figure 4.2: Channel capacity with (solid) and without (dashed) water-filling for various transmitter/receiver sets as a function of input SNR. When the number of receivers is greater than or equal to the number of transmitters, the value of channel knowledge is quite small, that is, water-filling shows little benefit. When the number of transmitters is greater than the number of receivers, water-filling can increase the information rate.

Thus, when the number of transmitters is greater than the number of receivers, having channel knowledge at the numerous transmitters significantly increases capacity. For communications in such cases, this implies that there may be some benefit in using some sort of retrofocusing technique [6] or another technique that uses channel knowledge. Also, it may be noted trivially that at very low SNR, the gains achievable through water-filling go to zero, as both measurements of capacity also go to zero.

Figure 4.3 examines the differences between the two expressions when the number of transducers is varying and SNR is held constant at 5 dB. Although it has been noted above that in the case of equal number of transmitters and receivers, the gains achievable through water-filling disappear at high SNR, one can see in Fig. 4.3a that at SNRs where

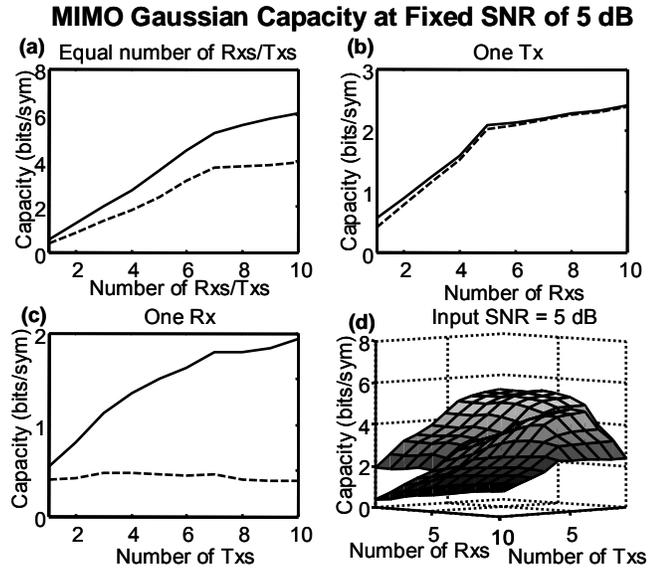


Figure 4.3: Channel capacity with (solid) and without (dashed) water-filling at a fixed input SNR of 5 dB. a) An equal number of receivers and transmitters is considered showing a linear gain up to seven transducers on each end. b) Using one transmitter and a varying number of receivers. c) Using one receiver and a varying number of transmitters. d) A surface plot of capacity as a function of the number of transmitters and receivers at an input SNR of 5 dB with (upper surface) and without (lower surface) water-filling.

there are gains to be had, and the number of transmitters and receivers are equal, increasing this number results in an increase in the gains achievable through water-filling. It is also seen in Fig. 4.3a that both measurements scale linearly with $\min(N_R, N_T)$, up to seven transducers. With more transducers, the increases in capacity are small, indicating there are approximately seven fundamental sub-channels in the water column for this experimental configuration. That is, that the average rank of $\mathbf{H}^H(\omega)\mathbf{H}(\omega)$ is approximately seven. It was seen in Fig 4.2c, that that gains achievable through water-filling tend towards zero at high SNRs when the number of receivers is equal to or greater than the number of transmitters. Similarly, when the SNR and number of transmitters are

held fixed, these gains tend toward zero as the number of receivers is increased, as seen in Fig. 4.3b. The opposite case is displayed in Fig. 4.3c which shows the case of one receiver and a varying number of transmitters with the SNR held fixed. The gains achievable through water-filling increase as the number of transmitters is increased, as adding more transmitters does not change the capacity of the system in the non-waterfilled case. Thus, to efficiently use multiple transmitters in such a situation, they must utilize knowledge of the channel. A comprehensive comparison of the two capacity measurements at an input SNR of 5 dB is shown in Fig. 4.3d.

To achieve the capacity calculable by the above formulas, one must transmit Gaussian symbols. Most systems are constrained to transmit symbols from a finite set, such as phase-shift keying (PSK) or quadrature-amplitude modulation (QAM), etc. A more informative measure may be the capacity of such a system, a constellation-constrained capacity. No closed expressions for such a capacity are known to exist. The constellation-constrained capacity of a memoryless channel [14] is calculable through numerical integration.

For a dispersive channel, simulations can be done to estimate this system capacity [15],[16]. This technique has also been extended to MIMO ISI channels [17]. The technique, based on the product-sum algorithm, becomes computationally prohibitive as the channel impulse responses become long or as the number of transmitters becomes large. In such cases, one may resort to bounding the capacity of the system, but previously introduced bounds [18],[19] are applicable only to single-input/single-output (SISO) channels. The extensions to MIMO ISI channels are discussed below.

4.4 Constellation-constrained Capacity

First, it is noted that the capacity formula of Eq. (4.4) is an upper bound for the system capacity, as the Gaussian is the random variable with the highest entropy for a given average power. Second, as an extension of the matched filter upper bound of Shamai, Ozarow, and Wyner [18], one can write

$$C_{UM} = \sum_{j=1}^{N_T} I \left(\left(\sum_{i=1}^{N_R} \sum_{n=0}^{L-1} h_{ij}^2[n] \right) x_j + v; x_j \right), \quad (4.6)$$

where x_j is random variable with the same distribution as the input symbols at transmitter j , v is a Gaussian random variable with the same distribution as that of w , and $I(a;b)$ is the mutual information between a and b .

The benefit of this bound is that it is expressed as a sum of the mutual informations achievable through memoryless channels and is thus easily calculable through numerical integration. As each term in the summation corresponds to one of the transmitters it is possible to consider a different constellation for each transmitter. The matched filter upper bound corresponds to the sum of the capacities of each transmitter sending only one symbol at different times. It can also be seen as the result of hypothetically capturing all the energy associated with each transmitted symbol, or the result of having perfect past and future decisions.

A lower bound on the capacity of a SISO ISI channel is to filter the symbols with the minimum mean-squared error decision-feedback equalizer (MMSE-DFE) forward filter and to assume perfect feedback [18],[19]. Shamai and Laroia [19] conjecture that replacing the resulting interference-plus-noise variable by a Gaussian variable of the same power does not increase the mutual information. The concept can be extended by

considering a multi-channel DFE [20],[21], with each transmitter sending a symbol from a given set at symbol rate.

In order to derive the infinite-length multi-channel DFE, the combined response of the channel impulse response, a matched filter, and appropriate sampling is considered, resulting in the hermitian matrix polynomial, $\mathbf{G}(z)$. This combined channel transfer matrix is used in the following decomposition [20]:

$$\mathbf{P}^T \cdot \left(\mathbf{G}(z) + \frac{\sigma^2}{P_t} \mathbf{I} \right) \cdot \mathbf{P} = \mathbf{S}^H(1/z^*) \cdot \Sigma \cdot \mathbf{S}(z), \quad (4.7)$$

where \mathbf{P} is a permutation matrix, $\mathbf{S}(z)$ is causal and minimum phase, and Σ is diagonal consisting of the elements ζ_j^2 for $1 \leq j \leq N_T$. Through the application of the appropriate forward and feedback filters, the signal-to-noise ratios at the input of the decision devices read

$$SNR_j = \frac{P_t}{\sigma^2} \zeta_j^2. \quad (4.8)$$

The elements ζ_j^2 can be influenced by the selection of the permutation matrix \mathbf{P} , and can hence be optimized according to some criterion. An appropriate optimization, in the context of this paper, would be to maximize the capacity. Thus one may write a MIMO extension of the Shamai-Laroia conjectured lower bound as

$$C_{LCa} = \max_{\mathbf{P}} \left[\sum_{j=1}^{N_T} I(\zeta_j x_j + v; x_j) \right], \quad (4.9)$$

where, again, x_j is a random variable with the same distribution as the input symbols at transmitter j , and v is a Gaussian random variable with the same distribution as w . The factors ζ_j^2 can be viewed as power degradation factors that arise due to the memory of the

MIMO channel. Another way of viewing the Shamai-Laroia conjecture is to consider these power degradation factors independent of the input symbol distribution, as using Gaussian symbols in Eq. (4.9) results in the capacity of the channel with unknown channel transfer functions in Eq. (4.4). This means that for a MIMO ISI channel with a Gaussian input symbol distribution, the MMSE-DFE achieves capacity [22].

Unfortunately, for systems with many transmitters, the maximization over the permutation matrix, \mathbf{P} , becomes computationally prohibitive. An even lower bound can be more easily calculated by neglecting the maximization over \mathbf{P} , and replacing the permutation matrix in Eq. (4.7) by the identity matrix. It has been suggested [20] that this maximization is only significant when the channel impulse responses are short or consist of only a few dominant taps. Also, the permutation matrix has little effect as the number of transmitters becomes large with respect to the channel lengths or at high SNR. One may write this more calculable lower bound as

$$C_{LCb} = \sum_{j=1}^{N_T} I(\varsigma_j x_j + \nu; x_j). \quad (4.10)$$

A still lower bound of great practical interest is that achievable with a finite-length MMSE-DFE [21]. In this case, the SNR degradation factors ς_j are replaced by their finite-length counterparts, denoted ξ_j . This generates a class of non-decreasing lower bounds as the length of the equalizer filters increases. One may write this new bound as

$$C_{LCc} = \sum_{j=1}^{N_T} I(\xi_j x_j + \nu; x_j). \quad (4.11)$$

Although the three lower bounds can be ordered, the differences between them should be very small for the data set to be examined. In this case, the channel impulse

responses and corresponding finite-length decision-feedback filters are sufficiently long (approximately 70 and 300 taps, respectively). Additionally, because the lengths of the channel impulse responses are so long, the interference-plus-noise variable would be nearly Gaussian, and the Shamai-Laroia conjectured lower bound is more likely to be true.

The true capacity of the constellation-constrained system is bounded above by the matched filter and Gaussian i.i.d. upper bounds in Eqs. (4.6) and (4.4) and below by the generalized Shamai-Laroia conjectured lower bound in Eq. (4.11). The tightness of these bounds is of importance, as the true capacity lies between the least upper bound and the greatest lower bound. Of the three bounds, the generalized Shamai-Laroia conjectured lower bound is of the greatest practical interest, as it most nearly approximates the information rates achievable with a practical system, specifically, a MIMO finite-length MMSE-DFE. In further discussion, for any transducer set and SNR, the minimum of Eqs. (4.4) and (4.6) will simply be referred to as, “the upper bound,” and Eq. (4.11) will simply be referred to as “the lower bound.”

The tightness of the bounds follows a similar pattern to the gains achievable through water-filling discussed previously. The bounds converge at low SNRs as they both tend towards zero, and at high SNRs as they both tend to the entropy of the input symbols. Between these two extremes, the tightness of the bounds is dependent on the number of transmitters and receivers. When the number of receivers is equal to or greater than the number of transmitters (see Fig. 4.4a, 4.4b, and 4.4c), the bounds are relatively tight compared to the case when the number of transmitters is greater than the number of receivers (see Fig. 4.4d). The bounds in both Figs. 4.4 and 4.5 are calculated using a

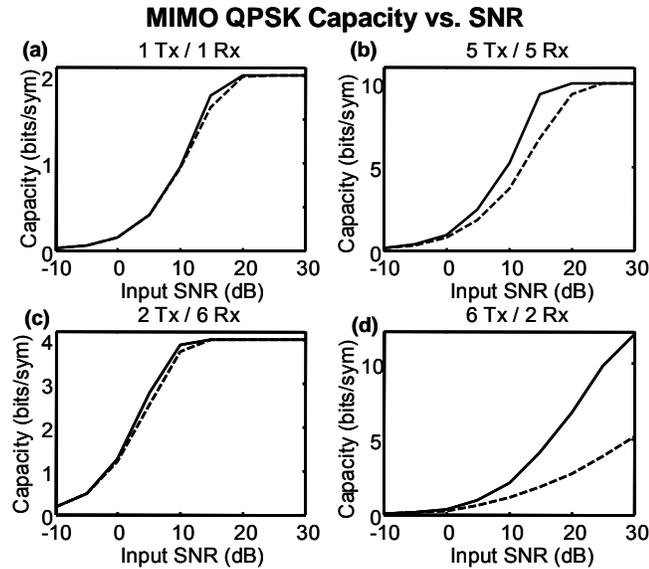


Figure 4.4: Upper (solid) and lower (dashed) bounds for the i.i.d. constellation-constrained capacity for various transmitter/receiver sets as a function of input SNR. The bounds are shown to be tight when the number of receivers is greater than or equal to the number of transmitters. The bounds are not tight when the opposite is true, and the number of transmitters is greater than the number of receivers.

QPSK constellation at all transmitters. Furthermore, Fig. 4.5 is calculated at an input SNR of 5 dB.

When input SNR is held constant and number of transducers is varied, the difference between the two bounds again looks similar to the comparison of the two Gaussian capacities of Eqs. (4.4) and (4.5). With an equal number of transmitters and receivers, the bounds are less tight when this number is increased, as seen in Fig. 4.5a. It is of great importance to note that the bounds (and presumably the true constellation-constrained capacity) grow linearly with $\min(N_R, N_T)$, up to the number of independent sub-channels. As mentioned in Section 4.3, for this experiment, the number is about seven. This shows that the gains from multiple transducers are limited by the number of fundamental sub-channels dictated by the ocean environment and array geometries. The bounds are tight

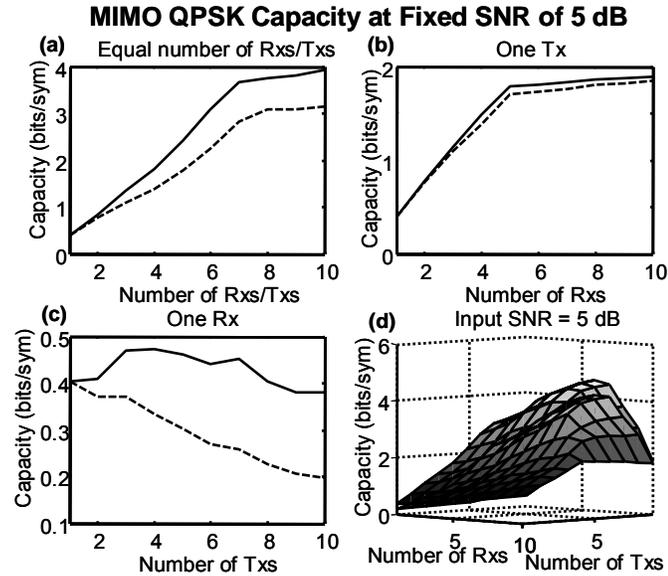


Figure 4.5: Upper (solid) and lower (dashed) bounds for the i.i.d. constellation-constrained capacity at a fixed input SNR of 5 dB. a) An equal number of receiver and transmitters is considered showing a linear gain up to seven transducers on each end. b) Using one transmitter and a varying number of receivers. c) Using one receiver and a varying number of transmitters. d) A surface plot of the upper and lower bound for the i.i.d. constellation-constrained capacity as a function of the number of transmitters and receivers at an input SNR of 5 dB with (upper surface) and without (lower surface) water-filling.

with only one transmitter and a varying number of receivers, with the difference between the two nearly constant past four receivers (Fig. 4.5b). With only one receiver and a varying number of transmitters, however, the bounds are the least tight. The lower bound (but not necessarily the true constellation-constrained capacity) strangely decreases as more transmitters are added (Fig. 4.5c). In such “overloaded” instances, it is better to leave some transmitters inactive, than to allocate some power to have them send symbols from the given constellation. This corresponds to setting x_j to zero for some transmitter(s) j , rather than QPSK, as shown. A comprehensive display of both bounds at 5 dB is shown in Fig. 4.5d.

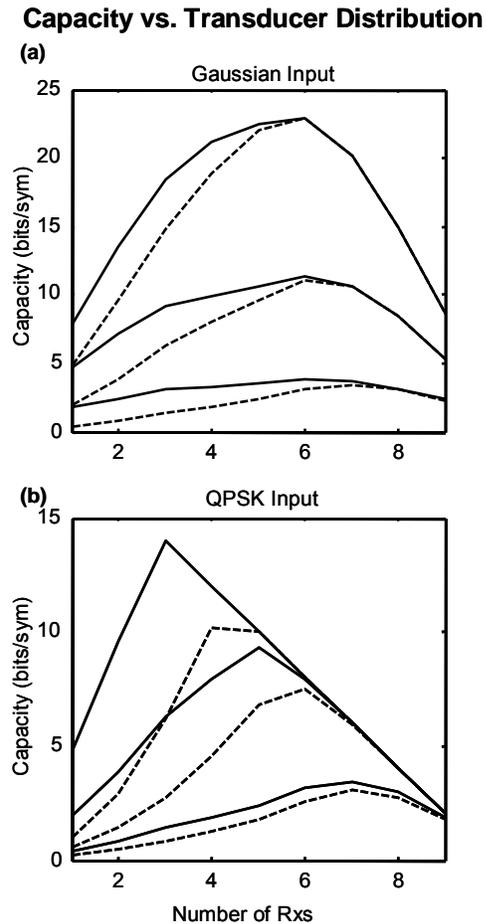


Figure 4.6: The effect of transducer placement. Ten transmitters are considered in total, with the number of receivers on the abscissa. The number of transmitters is ten less the number of receivers. a) Both Gaussian capacities with (solid) and without (dashed) channel knowledge at the transmitter at input SNRs of 5, 15, and 25 dB. b) Upper (solid) and lower (dashed) bounds on the QPSK-constrained capacities at input SNRs of 5, 15, and 25 dB.

4.5 Optimal Transducer Distribution

The effect of array size at both the transmitter and receiver on the channel capacity has been seen. A slightly different problem is now examined where a fixed number of total transducers is considered. Fig. 4.6 shows the results of this analysis. If one has ten total transducers to distribute as either receivers or transmitters, the question of how to

optimally distribute them is of interest. For example, would it be better to have one transmitter and nine receivers, nine transmitters and one receiver, or something in between these two extremes. The Gaussian capacity with (solid) and without (dashed) channel knowledge at the transmitting end is shown in Fig. 4.6a for input SNRs of 5, 15, and 25 dB. The previously discussed phenomenon of water-filling gains with more transmitters than receivers is evident. Also, the position of the capacity maximum without water-filling, which occurs at six receivers (and four transmitters) at 25 dB, shifts as a function of input SNR. At 5 dB, this maximum is moved to seven receivers (and three transmitters). At even lower input SNRs (not shown), the maximum shifts further towards more receivers. Thus, at lower input SNRs, the array gain at the receiver outweighs the gain in capacity from spatial diversity. Interestingly, the shift in maximum is not present when channel knowledge is available at the transmitter.

When the transmitting end is i.i.d. constellation-constrained, a similar shift in maximum is seen. Fig. 4.6b shows the upper (solid) and lower (dashed) bounds of a system constrained in this manner with ten total transmitters for input SNRs of 5, 15, and 25 dB. Again, the previously discussed phenomenon of looser bounds with more transmitters than receivers is evident. The shift in maximum is more prevalent in this case, where the optimum number of receivers is greater for lower input SNR, where the array gain at the receiver outweighs both the gains in capacity from spatial diversity and increased input entropy.

4.6 Conclusion

In this paper, the capacity of a measured at-sea data set has been analyzed in number of ways. First, the Gaussian capacity with and without channel knowledge at the transmitter was calculated and analyzed. The importance of channel knowledge at the transmitter in a number of cases was scrutinized showing that it is most important when the number of transmitters is greater than the number of receivers. This indicates that, in such cases, a retro-focusing scheme that uses channel knowledge should be beneficial. Secondly, SISO ISI bounds on the i.i.d. constellation-constrained capacity were generalized to the MIMO case and studied. These bounds were least tight when the number of transmitters was greater than the number of receivers. The generalization of these bounds allowed the question of optimum transducer placement to be answered. For a fixed number of transducers, the optimum number of transmitters and receivers can be found. At lower input SNRs, the array gain at the receiver outweighs the gain in capacity due to spatial diversity. For an i.i.d. constellation-constrained system, the array gain outweighs the gain in capacity due to both spatial diversity and increased input entropy. The generalization of these bounds allows other similar questions to be answered. Cost functions can be created and maximized when the cost of receivers and transmitters differs. Also, the optimum distribution of transducers when uplink and downlink requirements differ can be calculated. Such cases arise in the context of underwater acoustics with autonomous underwater vehicles (AUVs), which transmit far more information than they receive.

4.7 Acknowledgments

This chapter, in full, is a reprint of the material as it was submitted to the IEEE Journal of Oceanic Engineering with authors, W.J. Higley, P. Roux, W.A. Kuperman, W.S. Hodgkiss, H.C. Song, M. Stevenson, and T. Akal. The dissertation author was the primary investigator and author of this paper. The authors wish to thank Dr. John Proakis for his insight into this work. This work was supported by the Office of Naval Research.

REFERENCES

1. D. Kilfoyle and A. Baggeroer, "The state of the art in underwater acoustic telemetry," *IEEE J. Ocean. Eng.*, vol 25, pp. 4-27, 2000.
2. G. Edelmann, W.S. Hodgkiss, S. Kim, W.A. Kuperman, H.C. Song, and T. Akal, "An initial demonstration of underwater acoustic communication using a time reversal mirror," *IEEE J. Ocean. Eng.*, vol. 27, pp. 602-609 (2002).
3. D. Rouseff, D. Jackson, W. Fox, C. Jones, J. Ritcey, and D. Dowling, "Underwater acoustic communications by passive phase-conjugation," *IEEE J. Ocean. Eng.*, vol. 26, pp. 821-831, 2001.
4. T. Yang, "Temporal resolution of time-reversal and phase phase-conjugation for underwater acoustic communications," *IEEE J. Ocean. Eng.*, vol 28, pp. 229-245, 2003.
5. H. Song, P. Roux, W.S. Hodgkiss, W.A. Kuperman, T. Akal, and M. Stevenson, "Multiple-input/multiple-output coherent time reversal communications in a shallow water acoustic channel, *IEEE J. Ocean. Eng.*, to be published.
6. M. Stojanovic, "Retrofocusing techniques for high rate acoustic communications," *J. Acou. Soc. Am.*, vol. 117, pp. 1173-1185, 2005.
7. M. Stojanovic, J.A. Capitovic, and J.G. Proakis, "Adaptive multi-channel combining and equalization for underwater acoustic communications," *J. Acous. Soc. Am.*, vol. 94, pp. 1621-1631, 1993.
8. D. Kilfoyle, J. Preisig, and A. Baggeroer, "Spatial modulation experiments in the underwater acoustic channel," *IEEE J. Ocean. Eng.*, vol 30, no. 2, pp. 406-415, 2005.
9. T.J. Hayward and T.C. Yang, "Single and multi-channel underwater acoustic communication channel capacity: A computational study," submitted to *J. Acou. Soc. Am.*, 2006.
10. A. Goldsmith, S.A. Jafar, N. Jindal, and S. Vishwanath, "Capacity limits of MIMO channels," *IEEE J. Select. Areas on Comm.*, vol. 21, pp. 684-702, 2003.
11. G.J. Foschini and M.J. Gans, "On limits of wireless communication in a fading environment when using multiple antennas," *Wireless Pers. Commun.*, vol. 6, no. 3, pp. 311-335, Mar. 1998.
12. I.E. Telatar, "Capacity of multiantenna Gaussian channels," *Eur. Trans. Telecommun.*, vol. 10, pp. 585-595, Nov./Dec. 1999.

13. G. Raleigh and J.M. Cioffi, "Spatio-temporal coding for wireless communication," *IEEE Trans. Comm.*, vol. 46, pp.357–366, Mar.1998.
14. G. Ungerboeck, "Channel coding with multilevel/phase signals," *IEEE Trans. Inform. Theory*, vol. 28, pp. 55-67, Jan. 1982.
15. D. Arnold and H.-A. Loeliger, "On the information rate of binary-input channels with memory," *Proc. IEEE Int. Conf. Comm.*, vol. 9, pp. 2692–2695, June 2001.
16. H.D. Pfister, J.B. Soriaga, and P.H. Siegel, "On the achievable information rates of finite state ISI channels," *Proc. IEEE Global Comm. Conf.*, vol. 5, pp. 2992–2996, Nov. 2001.
17. Z. Zhang, T.M. Duman, and E.M. Kurtas, "Achievable information rates and coding for MIMO systems over ISI channels and frequency-selective fading channels," *IEEE Trans. Comm.*, vol. 52, no. 10, pp. 1698-1710, Oct. 2004.
18. S. Shamai, L.H. Ozarow, and A.D. Wyner, "Information rates for a discrete-time Gaussian channel with intersymbol interference and stationary inputs," *IEEE Trans. Inform. Theory*, vol. 37, pp. 1527–1539, Nov. 1991.
19. S. Shamai and R. Laroia, "The intersymbol interference channel: Lower bounds on capacity and channel precoding loss," *IEEE Trans. Inform. Theory*, vol. 42, pp. 1388–1404, Sept. 1996.
20. R. Fischer, "Sorted spectral factorization of matrix polynomials in MIMO communications," *IEEE Trans. Comm.*, pp. 945-951, Jun. 2005.
21. N. Al-Dhahir, A. H. Sayed, "The Finite-Length Multi-Input Multi-Output MMSE–DFE," *IEEE Trans. Sig. Proc.*, pp. 2921–2936, Oct. 2000.
22. T. Guess and M.K. Varanasi, "An information-theoretic framework for deriving canonical decision-feedback receivers in Gaussian channels," *IEEE Trans. Inform. Theory*, vol. 51, no. 1, pp. 173-187, Jan. 2005.

Chapter 5

Detection of Resonances on a Spherical Target

abstract

The use of low frequency (2-15 kHz) acoustic signals for mine countermeasures have been motivated by the chance to use scattering features of buried targets (e.g. the radius, wave thickness, and elastic properties) in the classification of the target. At frequencies, lower than traditionally used, sound penetrates into the sediment better and excites less signal-generated noise off the seabed. The structural waves of typical mine-sized targets are enhanced at these frequencies. This paper introduces and compares two forms of processing designed to classify a target as resonant or non-resonant based on the backscattered field. The two forms of processing are distinguished as an incoherent, energy-based method, and a coherent method based on the concept of principal component analysis (PCA).

5.1 Introduction

When performing mine countermeasures (MCM), it is preferable to have a large standoff distance. That is, it is desirable to detect and classify a mine while as far away as possible from it. This desire leads to the use of lower frequencies, which propagate further horizontally. This decreases the resolution of traditional acoustic imaging techniques. Also, the increase in standoff distance leads to an increase in the influence of the waveguide in introducing multipath.

Recent experiments [1],[2] have attempted to use low frequency (2-15 kHz) signals as part of the target detection and classification problem of MCM. At this low frequency, although the resolution of traditional imaging is limited, other properties such as the radius, wall thickness, and elastic properties of the target can be extracted.

Lower frequencies also benefit from greater seabed penetration. Attenuation above the grazing angle decreases approximately linearly with frequency, and below the grazing angle the evanescent lateral wave field in the bottom penetrates deeper [3]. The signal-generated noise from reverberation off surface features of the seafloor (such as sand ripples) will also decrease with frequency [4].

The problem of using target resonances to classify a target is complicated in a waveguide by the multipath in the environment. This paper compares two forms of processing on two targets in both freespace and in a waveguide. The two forms of processing are classified as an incoherent, energy-based processing method, and a coherent method based on the concept of principal component analysis (PCA). Our simulations show the benefit of the coherent method over the incoherent method in classifying a target as resonant or non-resonant. Finally, this paper proposes an experiment using a horizontally-aligned array towed from an autonomous underwater vehicle (AUV).

5.2 Freespace Simulation

A freespace simulation is done to test the feasibility of our processing. In an aqueous medium with a density of 1 g/cm^3 and a sound speed of 1507 m/s, there is an array of nine hydrophones arranged vertically with a separation distance of 10 m. The target is

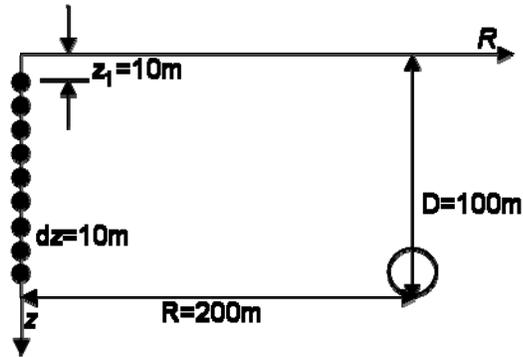


Figure 5.1: The geometry of the simulation. A vertical array consisting of nine elements is oriented vertically in freespace. The distance between each element is 10 m, and the first element is considered 10 m below some arbitrary origin. The target is located such that the bottom of the target is 10 m below the lowest element, and 200 m away from the vertical array.

situated such that the bottom is located 10 m below the lowest source/receiver, and the center is located 200 m away horizontally. The geometry of the set-up is shown in Fig. 5.1. Two targets are analyzed: a bubble and a spherical shell. The bubble is a spherical vacuum with a radius of .53 m. The spherical shell has an inner radius of .5 m and an outer radius of .53 m. The interior of the shell is a vacuum, and the shell itself is composed of steel with a density of 7.7 g/cm^3 , a shear sound speed of 2940 m/s and a compressional sound speed of 5880 m/s. The two targets are shown in Fig. 5.2.

Each hydrophone transmits and pulse which is back-scattered by the target and received on all nine hydrophones. This creates a three-dimensional matrix for each target as a function of transmitter depth, source depth, and frequency. The matrices for the bubble and the sphere are denoted as $\mathbf{K}_b(z_t, z_s, \omega)$ and $\mathbf{K}_s(z_t, z_s, \omega)$, respectively. By taking the inverse Fourier transform along the frequency domain, one can examine the received field for any given transmitter. In Fig. 5.3, the freespace responses from the first transmitter for both the bubble and the shell are shown. The response from the bubble is

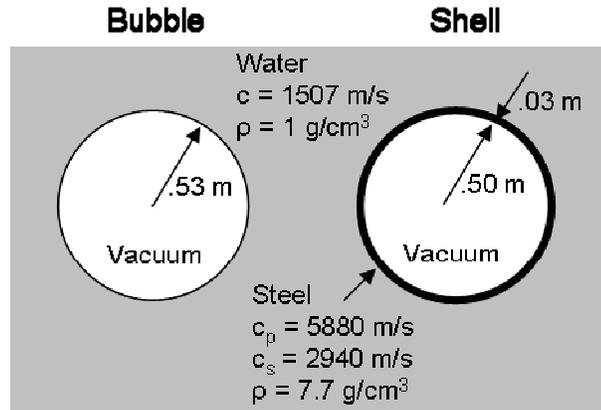


Figure 5.2: The two targets, both located in an aqueous medium. The first target is a bubble, whose interior is a vacuum. The radius of .53 m is matched to the outer radius of the second target, a resonant shell. The wall thickness of the shell is 3 cm, and it is also filled with a vacuum. The speed of sound and density of the water are 1507 m/s and 1 g/cm³ respectively. The shell is composed of steel with a density of 7.7 g/cm³. The sound speeds of compression and shear waves in the steel are 5880 m/s and 2940 m/s respectively.

composed of only a specular component, that field that reflects off the surface of the bubble. However, the response from the shell is composed of the same specular component (possibly phase-shifted) and an additional resonant component that arrives later.

Given the matrices $\mathbf{K}_b(z_t, z_s, \omega)$ and $\mathbf{K}_s(z_t, z_s, \omega)$, it is desirable to find some processing that will distinguish between the two. As can be seen in Fig. 5.3, in the freespace case, one could simply time-gate out the specular and be left with the resonance, if it existed. However, this method will not work in a waveguide, as will be seen later, so other processing methods need be examined.

The first method considered is an incoherent, energy-based method. The matrix \mathbf{K} can be reduced to a one-dimensional vector by summing incoherently across transmitters

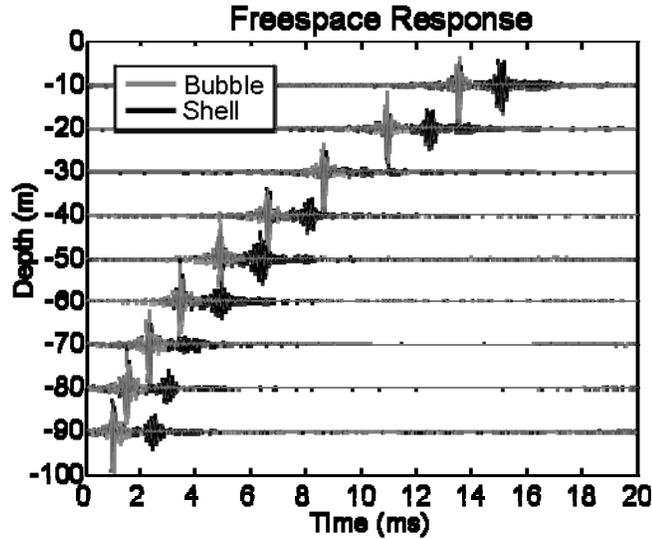


Figure 5.3: The backscattered response on the vertical array from both the bubble (gray) and the shell (black) when excited by a single source on the array. The response from the bubble consists of a strong, single, specular echo, while the response from the shell is composed of both that same specular response, and a resonant component that arrives later.

and receivers. The processing can be expressed mathematically, where N is the number of hydrophones, as

$$\mathbf{e}(\omega) = \sum_{z_t=1}^N \sum_{z_r=1}^N |\mathbf{K}(z_t, z_r, \omega)|^2. \quad (5.1)$$

This term can also be expressed in terms of the singular value decomposition (SVD) [5]. Each frequency of \mathbf{K} represents an $N \times N$ matrix. Taking the singular value decomposition of this matrix yields N singular values. Thus, there are N singular values at each frequency. The sum of these singular values squared also yields the vector $\mathbf{e}(\omega)$. This is shown mathematically below as [6]:

$$\begin{aligned} \text{svd}(\mathbf{K}(z_t, z_r, \omega)) &\Rightarrow \sigma_i(\omega) \text{ for } i=1,2,\dots,N \\ \sum_{i=1}^N \sigma_i^2(\omega) &= \sum_{z_t=1}^N \sum_{z_r=1}^N |\mathbf{K}(z_t, z_r, \omega)|^2 = \mathbf{e}(\omega). \end{aligned} \quad (5.2)$$

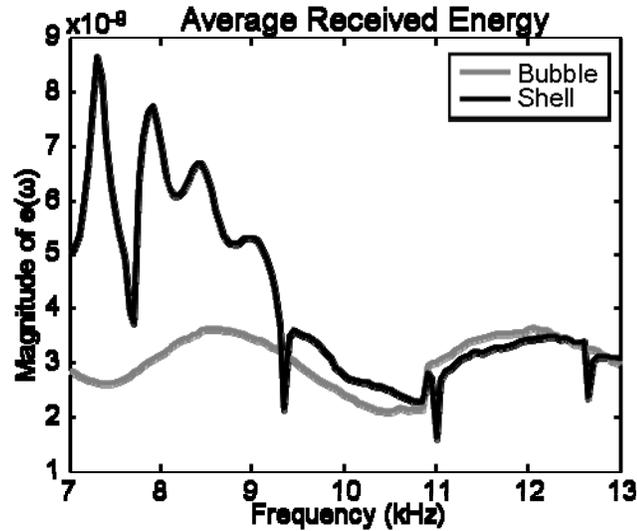


Figure 5.4: The received energy across frequency of the backscattered response of the bubble (gray) and the shell (black), averaged across source and receiver. A few of the resonant peaks of the shell are seen at lower frequencies.

Shown in Fig. 5.4 are the results of this energy-based processing. The resonances of the shell are seen as peaks in the average energy at specific frequencies. The bubble shows no such peaks in the frequency domain, and it rather flat across frequency, as one would expect from a specular echo.

The second method considered is a coherent one related to principal component analysis. We reduce the three-dimensional matrix $\mathbf{K}(z_t, z_r, \omega)$ into a two-dimensional matrix, labeled $\hat{\mathbf{K}}(z_t, z_r \cdot \omega)$, via concatenation. The received field in the frequency domain from each source can be thought of as a separate observation of a random vector. From these observations, a sample covariance matrix can be created as follows:

$$\mathbf{S}(z_r \cdot \omega) = \frac{1}{N-1} \sum_{z_t=1}^N \hat{\mathbf{K}}(z_t, z_r \cdot \omega) \hat{\mathbf{K}}^H(z_t, z_r \cdot \omega). \quad (\text{S.3})$$

Taking the first eigenvector of the matrix \mathbf{S} yields the first principal component of the data [6]. The first principal component is the energy-constrained vector which maximizes the average correlation with the data. It can be thought of as the part of the data that is the most similar across all observations, or in the case examined here, across all sources. We write the first principal component mathematically as:

$$\mathbf{p}(z_r \cdot \omega) = \arg \max_{\mathbf{p}} \sum_{z_t=1}^N \left| \mathbf{p}^H(z_r \cdot \omega) \hat{\mathbf{K}}(z_t, z_r \cdot \omega) \right|^2. \quad (\text{S.4})$$

As stated above, \mathbf{p} is also the first eigenvector of the sample covariance matrix, \mathbf{S} . The vector \mathbf{p} can be found another way, using singular value decomposition. If we consider \mathbf{S} to have the following eigenvalue decomposition,

$$\mathbf{S} = \mathbf{U} \mathbf{\Lambda} \mathbf{U}^H, \quad (\text{S.5})$$

we can write the singular value decomposition of $\hat{\mathbf{K}}$ as

$$\hat{\mathbf{K}} = \mathbf{U} \sqrt{(N-1)\mathbf{\Lambda}} \mathbf{V}^H. \quad (\text{S.6})$$

In both cases, the matrix \mathbf{U} is identical. The principal component, \mathbf{p} , can be thought of as either the first eigenvector of \mathbf{S} , or the first left singular vector of $\hat{\mathbf{K}}$.

Deconcatenation of the vector $\mathbf{p}(z_r \cdot \omega)$ results in a matrix $\mathbf{P}(z_r, \omega)$ that is a function of receiver and frequency. Taking the inverse Fourier transform along the frequency domain yields a single time-domain field for each matrix, \mathbf{K}_b and \mathbf{K}_s . These fields are plotted in Fig. 5.5. The fields shown below can be thought of as the parts of the received fields that are most independent of source depth. In the case of the bubble, there is a strong first arrival as there is similarity between the specular echoes for each source. In the case of the shell, the resonance is nearly separated from the specular, as the resonance

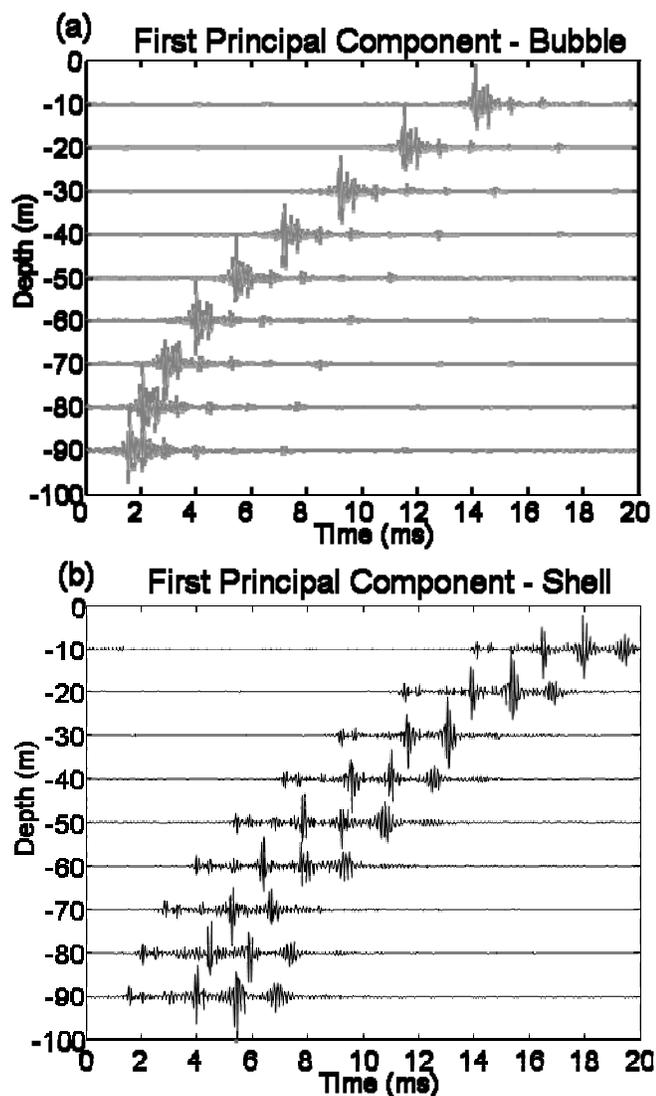


Figure 5.5: The first principal component from the bubble (gray) and the shell (black). The coherent processing first yields a field, which is shown in these figures. The field extracted from the bubble is composed of a strong first arrival and weaker later arrivals which may be attributed to weak resonances of the bubble. On the other hand, the extracted field from the shell's response is composed of periodic pulses corresponding to the resonances propagating around the shell.

is entirely source independent. The resonant component of the received field is independent of excitation, and shows up prominently in this processing.

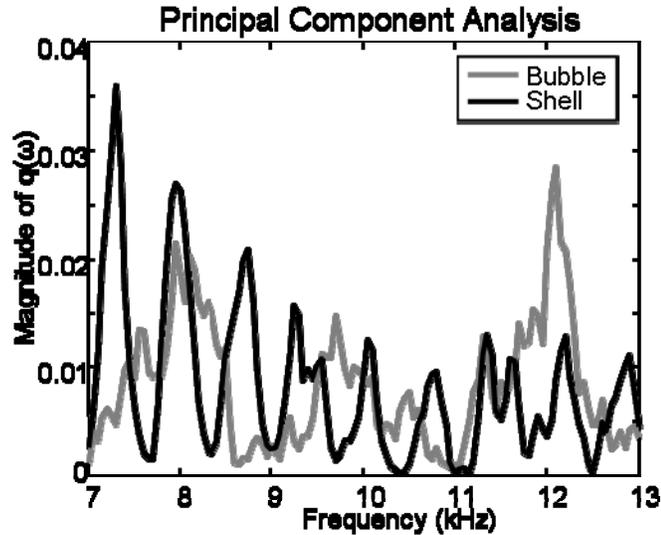


Figure 5.6: The result of the coherent processing on the bubble (gray) and the shell (black). The periodic resonance peaks of the shell are quite pronounced through the frequency range, whereas the response of the bubble is relatively flat, as one would expect for a specular component.

We can reduce the matrix $\mathbf{P}(z_r, \omega)$ into a frequency vector by summing incoherent across receiver. We can compare this new vector, $\mathbf{q}(\omega)$, with the previously calculated incoherent vector, $\mathbf{e}(\omega)$. The new vector is written mathematically as:

$$\mathbf{q}(\omega) = \sum_{z_r=1}^N |\mathbf{P}(z_r, \omega)|^2. \quad (\text{S.7})$$

The vector $\mathbf{q}(\omega)$ is calculated for both the bubble and the sphere, with the results shown in Fig. 5.6. The processing requires the results for both the bubble and the sphere to have the same magnitude, when summed across frequency, as the matrix \mathbf{U} is unitary, regardless of the relative energy in the matrix \mathbf{K} . The result of the principal component analysis processing done on the bubble is relatively flat, compared to the results obtained when done on the shell. The resonances of the shell appear again as peaks in the frequency domain, with a spacing of approximately 675 Hz. This corresponds to a

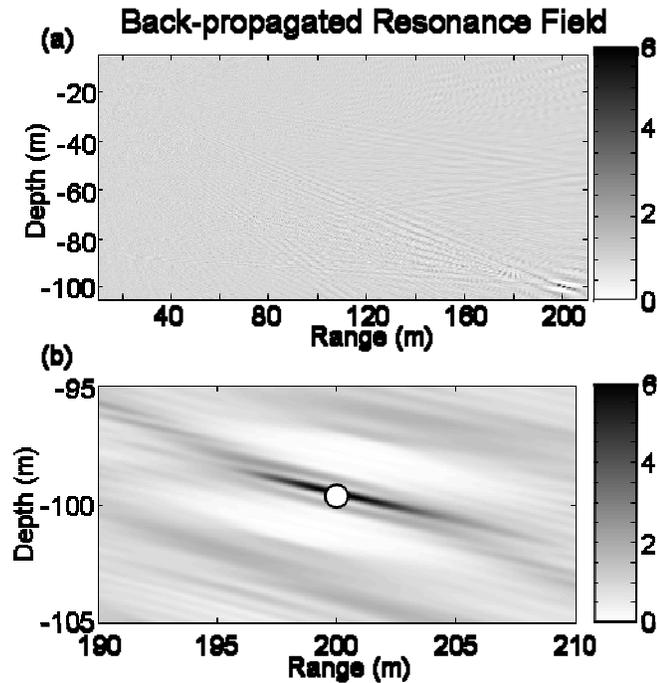


Figure 5.7: The result of back-propagating the field extracted in Fig 5.5b. The focus is seen only at the location of the target, at a depth of 100 m and a range of 200 m.

repeated time-delay of about 1.5 ms, a spacing that is seen in Fig 5b. This is the time it takes the resonance to travel around the sphere before emitting again. From this, we can calculate the group speed of the A_0 wave as approximately 2120 m/s.

Comparing Fig. 5.4 and Fig. 5.6, we notice that the principal component analysis is able to distinguish resonant peaks at 10 kHz and above, whereas the energy-based analysis cannot. Also, those peaks at frequencies less than 10 kHz are more pronounced than in the energy-based analysis.

Returning our attention to the time-domain resonance field of Fig. 5.5b, expressed by the matrix $\mathbf{P}_s(z_r, \omega)$, an incoherent back-propagation of this field can be calculated. The result of this calculation is shown in Fig. 5.7. We see that the back-propagated resonance field focuses at a location just above the sphere. The relative tightness of the focus is

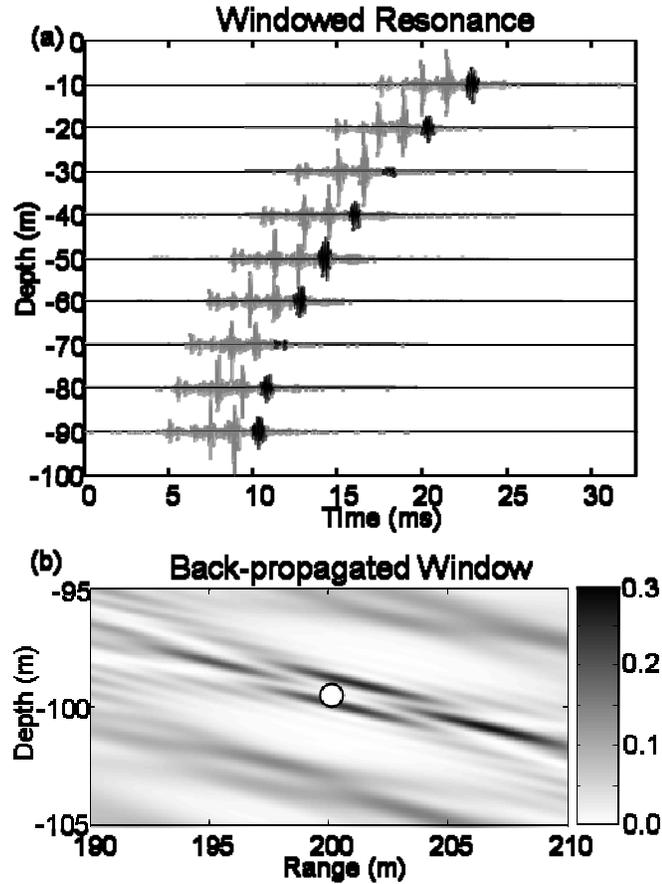


Figure 5.8: a) In gray is the field previously extracted from the shell and shown in Fig. 5.5b. The windowed portion, shown in black, corresponds to a resonant emission from the target. b) The result of back-propagating the resonant emission shows a pair of foci both above and below the location of the target, corresponding to the two locations on the target from which resonances are emitted. The other two foci, in front of and behind the target, are assumed sidelobes, artifacts of the processing.

related to the vertical aperture of the hydrophone array used. Back-propagation of the matrix $\mathbf{P}_b(z_r, \omega)$ results in a similar focus near the location of the sphere. A difference between the two back-propagated fields can be found by windowing out a single resonance of the matrix $\mathbf{P}_s(z_r, \omega)$ in the time domain (as shown in Fig. 5.8a) and back-propagating only that piece. As seen in Fig. 5.8b, this results in a group of foci in the area around the sphere. The two middle foci are interpreted as the two patches on the

sphere from which resonances are emitted to the array. The other two, in front and behind the sphere, are interpreted as sidelobes, artifacts of the processing.

1.3 Waveguide Simulation

Having established two forms of processing, an incoherent energy-based method yielding $\mathbf{e}(\omega)$, and a coherent method based on principal component analysis which yields $\mathbf{q}(\omega)$, the case of a buried sphere in a waveguide is now examined. The array is situated vertically as before, in a 100-m deep underwater waveguide. The target is situated 200 m away horizontally, as before, but instead of the bottom of the target being at a depth of 100 m, the top is located at that same position. The target is therefore buried, such that the top of the target is tangent to the waveguide bottom. The noise-free matrices $\mathbf{K}_b(z_t, z_s, \omega)$ and $\mathbf{K}_s(z_t, z_s, \omega)$ are acquired as before, and an additional matrix $\mathbf{K}_r(z_t, z_s, \omega)$ is calculated from a seafloor reverb model, to allow us to examine our processing in the presence of signal-generated noise. The backscattered response of the flush-buried sphere in a waveguide is shown in Fig. 5.9 for the two targets: the bubble and the sphere. The shape of the back-scattered response is the same, but each pulse of the bubble's response is convolved with the resonance in the case of a shell.

The two above-mentioned method of processing are applied now to the case of a flush-buried target. For comparison purposes, each of the vectors $\mathbf{e}(\omega)$ and $\mathbf{q}(\omega)$ are normalized such that the energy in each vector is one, which we denote $\hat{\mathbf{e}}(\omega)$ and $\hat{\mathbf{q}}(\omega)$. Figure 5.10 shows the result of both forms of processing, without noise, on both flush-buried targets. In the case of the bubble (Fig. 5.10a), there are no noticeable peaks found in either form of processing, as we would expect from specular echoes. Both the

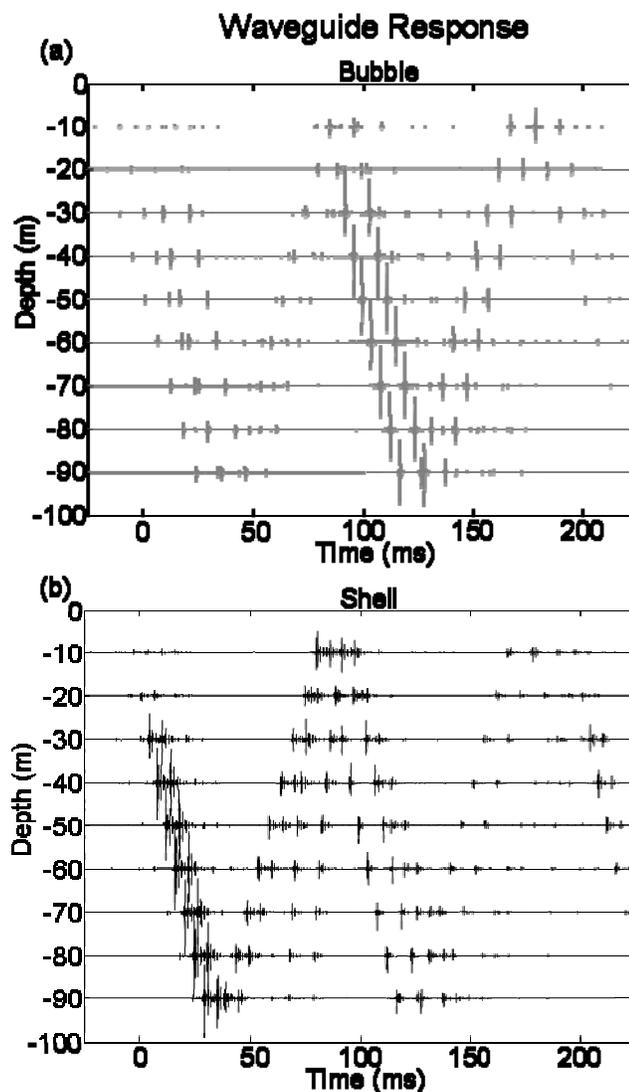


Figure 5.9: The back-scattered response from a flush-buried bubble (gray) or shell (black) in a waveguide. The shape of the response in the same in both cases, but each pulse of the response in the case of the bubble is turned into a resonant arrival structure in the case of the shell.

incoherent and coherent forms of processing yield resonant peaks when applied to the flush-buried shell (Fig. 5.10b), but the method based on principal component analysis causes these peaks to be more pronounced.

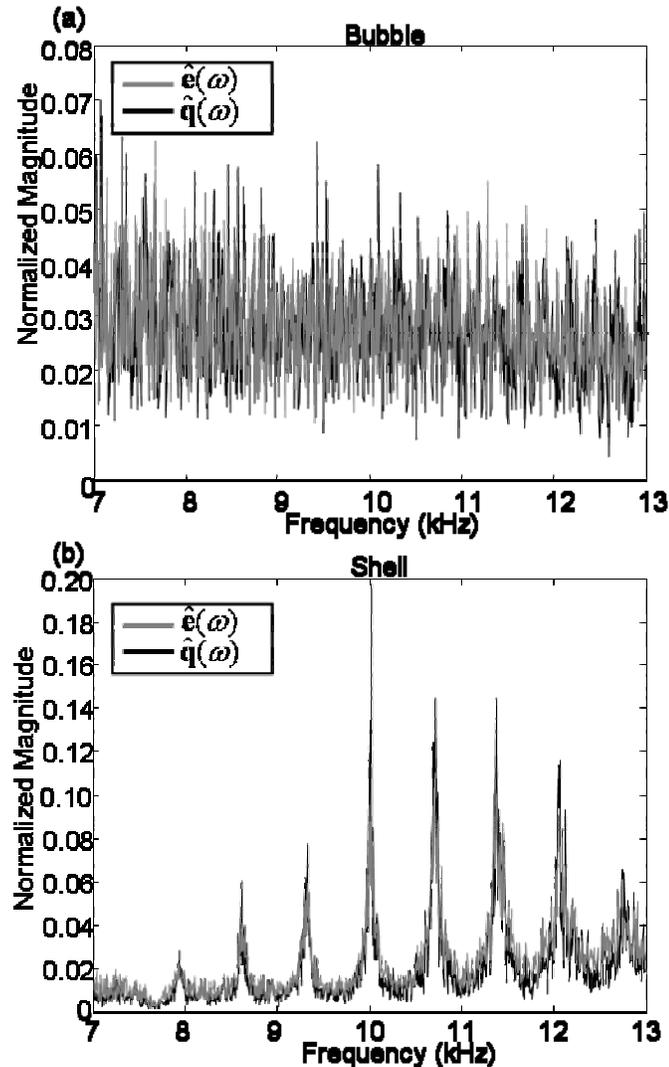


Figure 5.10: a) The result of the incoherent (gray) and coherent (black) processing on the non-resonant bubble in a waveguide. The frequency response is relatively flat, as we would expect from a purely specular echo. b) The result of the incoherent (gray) and coherent (black) processing on the resonant shell in a waveguide. The periodic resonant peaks are clearly identified, and correspond well to those found in freespace (Fig. 5.6).

By examining the results of the processing under various levels of signal-generated noise, a difference is again seen in the results of the incoherent and coherent processing. Figure 5.11 shows the results of both forms of processing on the flush-buried shell for four different noise levels. In Fig. 5.11a, the results are very similar to Fig. 10, with both

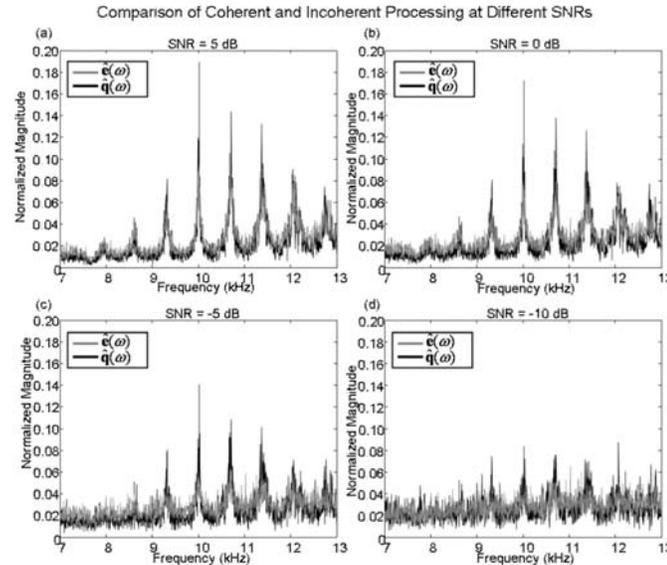


Figure 5.11: A comparison of the incoherent (gray) and coherent (black) processing on the flush-buried shell for four different signal-to-noise ratios. Although both methods degrade as noise increases, the coherent method still shows peaks at -10 dB, whereas the incoherent method fails at this SNR.

forms of processing showing strong resonant peaks, and those of the coherent PCA-based method a little stronger. In Fig. 5.11b and 5.11c, a degradation of performance for both methods is witnessed, with the coherent method still performing better than the incoherent energy-based method. In Fig. 5.11d, the incoherent method fails, whereas small peaks at the resonant frequencies are still seen using the coherent method.

The main conclusion drawn from Fig. 5.10 and 5.11 is the better, more peaked performance of the coherent method, $\hat{\mathbf{q}}(\omega)$, over the incoherent method, $\hat{\mathbf{e}}(\omega)$. These results can be distilled into a classification metric by calculating the sample variance or sample kurtosis of the above vectors and comparing these values at various signal-to-noise ratios to the variance or kurtosis of the same processing performed on a bubble.

Figure 5.12 shows the result of just such a processing. Although the sample variance the coherent processing (black in Fig. 5.12a) is, in general, higher than that of the incoherent

processing (gray in Fig. 5.12a) the separation of the coherent method when applied to the shell (solid line) and the bubble (dashed line) is greater than the separation of the incoherent method applied to the shell and the bubble. The coherent method recognizes a difference between the two down to an SNR of -10 dB, whereas the incoherent method fails at this reverb level. This result was previously seen in Fig. 5.11d. Since the coherent PCA-based method yields a more peaked result, the sample kurtosis of $\hat{\mathbf{e}}(\omega)$ and $\hat{\mathbf{q}}(\omega)$, shown in Fig. 5.12b, yields another classification metric. In this case, the coherent method recognizes a difference between the shell and the bubble at SNRs as low as -20 dB, whereas this metric is rather poor for the incoherent method, failing at an SNR of -5 dB.

5.4 Experimental Proposal

Although both forms of processing perform well when a vertical array is used at high signal-to-noise levels, such a configuration would be impractical to realize at-sea for real-world application. Of great interest would be the results of this processing performed on a dataset realized from a horizontal array towed from an AUV.

The strength of the processing, in both the incoherent and coherent case, is based on “averaging” away the frequency-domain peaks caused by waveguide interference, leaving only those peaks emitted from a resonant target. Therefore, such a towed source would have to sample the wavenumber space efficiently, which is done much more easily with a vertical array than a horizontal one [7]. The best solution to this would seem to be

a broadside array as close as feasible to the target. The point of acoustic detection and classification of a target in MCM, of course, is to not be close to the target, so a compromise in distance must be made.

5.5 Conclusion

This paper has introduced and compared two forms of processing to classify a target as either resonant or non-resonant by using the back-scattered field from that target. The two processing methods are an incoherent, energy-based method, and a coherent PCA-based method. Comparison of the two methods in simulation shows the benefit of the coherent method over the incoherent method in detecting resonances in the presence of noise. Lastly, an experiment to test the practicality of the processing for real-world MCM is proposed.

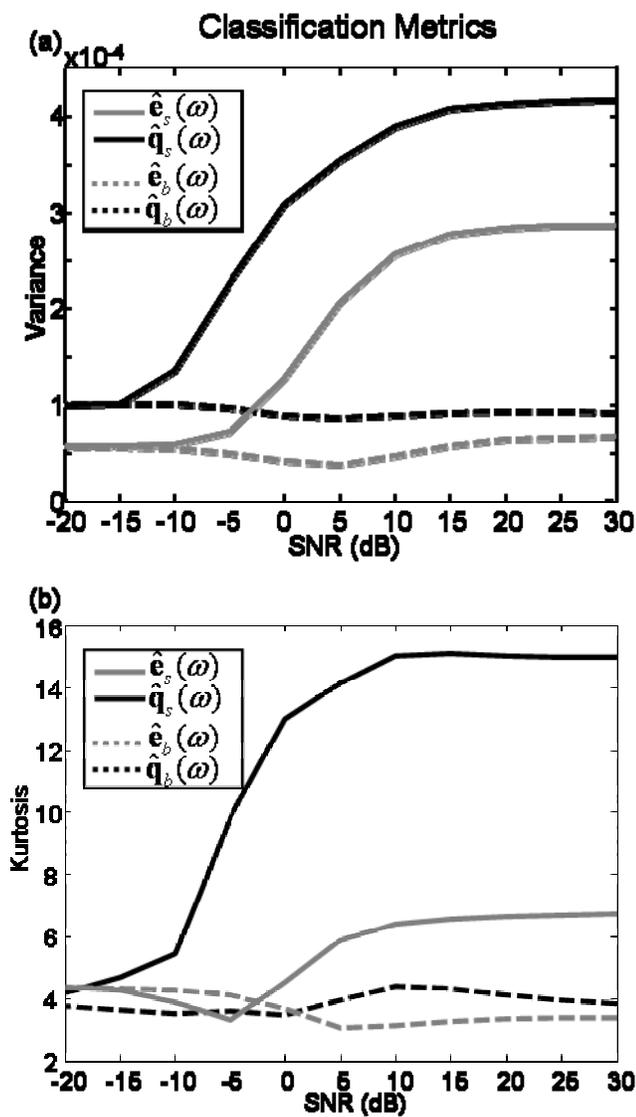


Figure 5.12: a) Sample variance of the frequency vectors corresponding to the incoherent processing (gray) and coherent processing (black). The solid lines correspond to the processing performed on the flush-buried shell, and the dashed lines correspond to the processing performed on the bubble for comparison. The coherent processing outperforms the incoherent processing by approximately 5 dB, using this metric. b) The sample kurtosis, which is high for peaky vectors, of the frequency vectors corresponding to the incoherent processing (gray) and coherent processing (black). The solid lines correspond to the processing performed on the flush-buried shell, and the dashed lines correspond to the processing performed on the bubble for comparison. Using this metric, classification using the coherent PCA-based method is possible as low as -20 dB.

REFERENCES

1. A. Tesei, A. Maguer, W.L.J. Fox, R. Lim, and H. Schmidt, "Measurements and modeling of acoustic scattering from partially and completely buried spherical shells," *J. Acous. Soc. Am.*, vol. 112, pp. 1817-1830, 2002.
2. A. Tesei, W.L.J. Fox, A. Maguer, and A. Løvik, "Target parameter estimation using resonance scattering analysis applied to air-filled, cylindrical shells in water," *J. Acous. Soc. Am.*, vol. 108, pp. 2891-2900, 2000.
3. A. Maguer, W.L.J. Fox, H. Schmidt, E. Pouliquen, and E. Bovio, "Mechanisms for subcritical penetration into a sandy bottom: Experimental and modeling results," *J. Acous. Soc. Am.*, vol. 107, pp. 1215-1225, 2000.
4. H. Schmidt and J. Lee, "Physics of 3-d scattering from rippled seabeds and buried targets in shallow water," *J. Acous. Soc. Am.*, vol. 105, pp. 1605-1617, 1999.
5. C. Prada, S. Manneville, D. Spoliansky, and M. Fink, "Decomposition of the time reversal operator: Detection and selective focusing on two targets," *J. Acous. Soc. Am.*, vol. 99, pp. 2067-2076, 1996.
6. T.K. Moon and W.C. Stirling, *Mathematical Methods and Algorithms for Signal Processing*, Prentice Hall, 2000.
7. M.R. Dungan and D.R. Dowling, "Orientation effects on linear time reversing array retrofocusing in shallow water," *J. Acous. Soc. Amer.*, vol. 112, pp. 1842-1852, 2002."