



NAVAL POSTGRADUATE SCHOOL

MONTEREY, CALIFORNIA

THESIS

**UNDERWATER MASKED CARRIER ACOUSTIC
COMMUNICATION: MODELING AND ANALYSIS**

by

Ryan R. Ferrao

June 2018

Thesis Advisor:

Justin P. Rohrer

Co-Advisor:

Daphne Kapolka

Approved for public release. Distribution is unlimited.

THIS PAGE INTENTIONALLY LEFT BLANK

REPORT DOCUMENTATION PAGE			<i>Form Approved OMB No. 0704-0188</i>	
Public reporting burden for this collection of information is estimated to average 1 hour per response, including the time for reviewing instruction, searching existing data sources, gathering and maintaining the data needed, and completing and reviewing the collection of information. Send comments regarding this burden estimate or any other aspect of this collection of information, including suggestions for reducing this burden, to Washington headquarters Services, Directorate for Information Operations and Reports, 1215 Jefferson Davis Highway, Suite 1204, Arlington, VA 22202-4302, and to the Office of Management and Budget, Paperwork Reduction Project (0704-0188) Washington, DC 20503.				
1. AGENCY USE ONLY (Leave blank)	2. REPORT DATE June 2018	3. REPORT TYPE AND DATES COVERED Master's thesis		
4. TITLE AND SUBTITLE UNDERWATER MASKED CARRIER ACOUSTIC COMMUNICATION: MODELING AND ANALYSIS			5. FUNDING NUMBERS W7B24	
6. AUTHOR(S) Ryan R. Ferrao				
7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES) Naval Postgraduate School Monterey, CA 93943-5000			8. PERFORMING ORGANIZATION REPORT NUMBER	
9. SPONSORING / MONITORING AGENCY NAME(S) AND ADDRESS(ES) N/A			10. SPONSORING / MONITORING AGENCY REPORT NUMBER	
11. SUPPLEMENTARY NOTES The views expressed in this thesis are those of the author and do not reflect the official policy or position of the Department of Defense or the U.S. Government.				
12a. DISTRIBUTION / AVAILABILITY STATEMENT Approved for public release. Distribution is unlimited.			12b. DISTRIBUTION CODE A	
13. ABSTRACT (maximum 200 words) As naval warfighting capabilities evolve, the need for innovative communication techniques for tactics and command and control increases. Using Navy sonar systems normally reserved for conventional communication in the underwater channel between vessels to transmit a hidden message disguised as ambient ocean noises or biological noise is a possible way to communicate with a low probability of interception or detection. This research applies information hiding via steganography in order to embed and extract bits from an audio file after transmission through a simulated underwater acoustic channel. Specifically, we explore a technique that allows us to communicate such that the transmission appears native to the operating environment. We do this by embedding symbols in an audio source in the frequency domain. We demonstrate the success of our technique via traditional steganography metrics and describe its performance limitations. We find that our scheme can be both imperceptible and robust provided that proper embedding and transmission parameters can be determined.				
14. SUBJECT TERMS audio steganography, covert communications, underwater acoustic communications, biological mimicry, underwater networks			15. NUMBER OF PAGES 89	
			16. PRICE CODE	
17. SECURITY CLASSIFICATION OF REPORT Unclassified	18. SECURITY CLASSIFICATION OF THIS PAGE Unclassified	19. SECURITY CLASSIFICATION OF ABSTRACT Unclassified	20. LIMITATION OF ABSTRACT UU	

THIS PAGE INTENTIONALLY LEFT BLANK

Approved for public release. Distribution is unlimited.

**UNDERWATER MASKED CARRIER ACOUSTIC COMMUNICATION:
MODELING AND ANALYSIS**

Ryan R. Ferrao
Lieutenant, United States Navy
BS, U.S. Naval Academy, 2011

Submitted in partial fulfillment of the
requirements for the degree of

MASTER OF SCIENCE IN COMPUTER SCIENCE

from the

**NAVAL POSTGRADUATE SCHOOL
June 2018**

Approved by: Justin P. Rohrer
Advisor

Daphne Kapolka
Co-Advisor

Peter J. Denning
Chair, Department of Computer Science

THIS PAGE INTENTIONALLY LEFT BLANK

ABSTRACT

As naval warfighting capabilities evolve, the need for innovative communication techniques for tactics and command and control increases. Using Navy sonar systems normally reserved for conventional communication in the underwater channel between vessels to transmit a hidden message disguised as ambient ocean noises or biological noise is a possible way to communicate with a low probability of interception or detection.

This research applies information hiding via steganography in order to embed and extract bits from an audio file after transmission through a simulated underwater acoustic channel. Specifically, we explore a technique that allows us to communicate such that the transmission appears native to the operating environment. We do this by embedding symbols in an audio source in the frequency domain. We demonstrate the success of our technique via traditional steganography metrics and describe its performance limitations. We find that our scheme can be both imperceptible and robust provided that proper embedding and transmission parameters can be determined.

THIS PAGE INTENTIONALLY LEFT BLANK

Table of Contents

1	Introduction	1
1.1	Benefit to the Department of Defense	1
1.2	Research Questions	2
1.3	Scope	2
1.4	Approach and Significant Findings	3
1.5	Organization	3
2	Background	5
2.1	Underwater Acoustic Communication	5
2.2	Steganography Background	11
2.3	Audio Steganography.	16
2.4	Steganography and Covert Underwater Acoustic Communications	23
2.5	Summary	30
3	Methodology	33
3.1	Theory	33
3.2	Implementation	39
3.3	Summary	44
4	Results and Analysis	47
4.1	Steganographic Evaluation	47
4.2	Channel Simulation	52
4.3	Deep Channel Experiment.	59
4.4	Shallow Channel Experiment	62
4.5	Summary	64
5	Conclusions and Future Work	65
5.1	Key Findings	65
5.2	Recommendations for Future Work	65

List of References	67
Initial Distribution List	71

List of Figures

Figure 3.1	Workflow and Module Interaction	44
Figure 4.1	Comparison of Spectra Between Cover and Stego Audio Signals .	48
Figure 4.2	Comparison of Estimated Symbols Constellations.	49
Figure 4.3	Comparison of Average Mean-Square Error (MSE) and Root-Mean-Square-Error (RMSE) for Differing T	51
Figure 4.4	Comparison of Average Peak Signal-to-Noise Ratio (PSNR) and Signal-to-Noise Ratio (SNR) for Differing T	52
Figure 4.5	Travel Time Wrap-Around with 4096 Frequencies at 3km	53
Figure 4.6	Travel Time Calculations for the Two Channels.	54
Figure 4.7	Sound Speed Profile (SSP) for the Deep Channel Model	56
Figure 4.8	Deep Channel Transmission Loss (TL) for Pressure vs. Depth and Range (Decibels (dB) re 1m)	56
Figure 4.9	SSP for the Shallow Channel Model	58
Figure 4.10	Shallow Channel TL for Pressure vs. Depth and Range (dB re 1m)	58
Figure 4.11	Transfer Function Phase Angle at 1km vs. Frequency Bin	59
Figure 4.12	Example of Recovered Symbols vs. Estimated Delay over 100 Bins	60
Figure 4.13	Symbol Recovery Percentage vs. Range for the Deep Channel . .	61
Figure 4.14	Transfer Function Phase Angle at 3km vs. Frequency Bin	62
Figure 4.15	Symbol Recovery Percentage vs. Range for the Shallow Channel	63

THIS PAGE INTENTIONALLY LEFT BLANK

List of Tables

Table 2.1	Audio Steganography Techniques	22
Table 4.1	Deep Channel Monterey-Miami Parabolic Equation (MMPE) Parameters	55
Table 4.2	Shallow Channel MMPE Parameters	57
Table 4.3	Deep Channel Experiment Parameters	60

THIS PAGE INTENTIONALLY LEFT BLANK

List of Acronyms and Abbreviations

AWGN	Additive White Gaussian Noise
BER	Bit Error Rate
bps	Bits Per Second
BPSK	Binary Phase Shift Keying
dB	Decibels
DFT	Discrete Fourier Transform
DI	Directivity Index
DoD	Department of Defense
DSSS	Direct Sequence Spread Spectrum
DT	Detection Threshold
DWT	Discrete Wavelet Transform
FFT	Fast Fourier Transform
FSK	Frequency-Shift Keying
HAS	Human Auditory System
HF	High Frequency
IFT	Inverse Fourier Transform
LF	Low Frequency
LPD	Low Probability of Detection
LPI	Low Probability of Interception

LSB	Least Significant Bit
MMPE	Monterey-Miami Parabolic Equation
MSE	Mean-Square Error
NL	Noise Level
NOAA	National Oceanic and Atmospheric Administration
OFDM	Orthogonal Frequency-Division Multiplexing
PE	Parabolic Equation
PSNR	Peak Signal-to-Noise Ratio
QAM	Quadrature Amplitude Modulation
QIM	Quantization Index Modulation
QPSK	Quadrature Phase Shift Keying
RL	Reverberation Level
RMSE	Root-Mean-Square-Error
SL	Source Level
SS	Spread Spectrum
SNR	Signal-to-Noise Ratio
SSF	split-step Fourier
SSP	Sound Speed Profile
TL	Transmission Loss
TS	Target Strength
UAC	Underwater Acoustic Communications
USN	United States Navy
UUV	Unmanned Underwater Vehicle

Acknowledgments

First and foremost, I would like to thank my advisors, Dr. Daphne Kapolka and Dr. Justin Rohrer. I don't have the words to describe how Dr. Kapolka's patience, encouragement, and wit made this otherwise insurmountable task seem like something I could actually complete. Despite my usual lack of progress, our meetings were always the brightest spot of my day. Dr. Rohrer always treated me as a colleague, and his commitment to quality and excellence helped me to grow as a critical thinker, writer, and student. Thanks to both for never doubting me, even when I doubted myself.

I would also like to thank Dr. Kevin Smith for his guidance concerning modeling, and Charles D. Prince for the constant professional and technical support.

Most importantly, I owe my greatest thanks to my mother and father, who have been and continue to be the most genuine example of love and devotion I can imagine. If I can someday be even half the parent that either of you have been, I will have considered my time a success.

I am truly humbled and grateful to have so many extraordinary people in my life.

THIS PAGE INTENTIONALLY LEFT BLANK

CHAPTER 1:

Introduction

The ability for friendly forces to reliably communicate with one another is a critical component of modern military tactics and operations. The capability to provide command and control as well as situational awareness enables decisive action in meeting our nation's objectives. Conversely, an adversary who seeks to obtain an advantage need only intercept or deny these vital communications in order to significantly disrupt the friendly unit.

Given its existing challenges and limitations, availability of communications in the undersea domain is at increased risk of denial. For this reason, it is advantageous to develop innovative communication methods that have a Low Probability of Detection (LPD) and/or Low Probability of Interception (LPI), thereby allowing us to communicate freely without fear of counter-detection.

This thesis explores the application of steganography to Underwater Acoustic Communications (UAC) such that two nodes could communicate with one another in plain sight unbeknownst to a third party. Jean-Michel Passerieux of Thales Underwater Systems developed an early technique based on this principle which modifies previously recorded noise in order to embed hidden data [1], [2]. We seek to expand upon his work by implementing his concepts in the frequency domain as well as with a different channel simulation technique. Furthermore, we conduct experimentation to better understand the technique's capabilities and limitations.

In this chapter, we state the benefit of our research to the Department of Defense (DoD), state our research objectives, define the scope of research, explain our significant findings, and explain the organization of this thesis.

1.1 Benefit to the Department of Defense

UAC has direct applications to the United States Navy (USN), and are employed by submarines, surface vessels, underwater acoustic modems, Unmanned Underwater Vehicles (UUVs) and underwater sensor nodes. Applying an innovative method for covert UAC

would further enhance the flexibility and capability of tactical communications. UAC is an important topic for military research, as the undersea domain is one of the few remaining natural frontiers.

1.2 Research Questions

The primary research question is whether a steganography based covert UAC technique such as Passerieux's can be effectively demonstrated in the frequency domain, along with how well it performs when applied to a signal whose transmission will be modeled by an accurate underwater acoustic channel model.

Secondary research questions include:

- How well does the algorithm embed data while remaining undetectable (i.e., preserving the original signal's audible and spectral characteristics)?
- What nominal data rates are feasible?
- At what ranges can data be reasonably recovered from the signal?
- How do different inputs to the algorithm affect algorithm performance?
- How well does the algorithm perform in deep vs. shallow water?

1.3 Scope

The primary scope of this research is to provide a proof-of-concept for flexible and robust algorithm for embedding data in existing ambient noise, and to demonstrate data extraction under simulated underwater acoustic conditions. At the conclusion of this study, we provide a framework for follow on research and applications.

Though valuable the following will not be specifically addressed and are beyond the scope of this thesis:

- The effect of Doppler on symbol recovery.
- Receiver and transmitter characteristics, construction, and design.
- Application of different modulation techniques.
- Real-time symbol recovery and receiver synchronization.
- Performance evaluation via at sea testing.

- Use of different frequency bands.
- Compatibility with specific equipment, software, and hardware.
- Resistance to traditional and statistical steganography attacks.

1.4 Approach and Significant Findings

1. This research implemented an audio steganography by cover modification technique based on a scheme first developed by Passerieux, whose goal is to utilize existing background noise in order to appear inconspicuous. Our work builds on his approach by performing data embedding and recovery completely in the frequency domain.
2. We implement our scheme in a 2048 Hz bandwidth centered around 2500 Hz to balance transmission distance with reasonable transmitter power. This band is also representative of many low frequency sonar systems. Our implementation is adaptable to other bandwidths.
3. We show the effects of embedding data on the cover signal, including how these effects can be minimized while balancing the ability to recover data.
4. We model the effects of transmission through two underwater acoustic channels generated by the well-documented Monterey-Miami Parabolic Equation (MMPE). We simulate a generic deep water channel and a challenging shallow water channel.
5. We demonstrate proof of concept for algorithm success for a variety of operating parameters
6. A key challenge to algorithm success is correctly accounting for time delay in order to correctly align symbol periods. Multipath effects contribute to this and can further cause unpredictable results.
7. Background noise significantly degrades performance as range increases. As a result, adequate Source Level (SL) must be chosen in order to balance symbol recovery with detectability.

1.5 Organization

The remainder of this thesis is structured as follows:

Chapter 2: Background

Chapter 2 describes a short history of underwater acoustics research, including applications over time and acoustic modeling. Following that is a description of the important terms and concepts for both UAC and steganography. Next, we describe the existing applications of audio steganography, concluding with present work applying audio steganography techniques to UAC.

Chapter 3: Methodology

In Chapter 3, we begin by outlining the theory behind Passerieux’s technique, as well as the theory of our implementation, including how data is embedded and extracted. The second portion describes the modules which constitute our implementation, as well as a description of the input files used to generate the simulated channel.

Chapter 4: Results and Analysis

Chapter 4 includes the parameters for experimentation, details the environment of our simulated channels, and describes evaluated metrics. It also includes relevant plots and a discussion of the results.

Chapter 5: Conclusions and Future Work

Chapter 5 presents the final conclusions of this research and provides recommendations for future work based on those conclusions.

CHAPTER 2: Background

The study of UAC and steganography are disparate, with few existing applications. To understand the foundation of our research, an overview of concepts and terminology is warranted. In this chapter, we provide pertinent background information and detail the extent of existing work from which to position our research. Section 2.1 describes the history and broad theory of UAC, including the sonar equations and modeling methods. Section 2.2 explains steganography theory, terminology, and critical concepts, while Section 2.3 describes steganography's application to digital audio. Finally, Section 2.4 explains the nature of existing work with regards to audio steganography and UAC.

2.1 Underwater Acoustic Communication

Sound travels through water much more effectively than radio waves and other forms of radiation useful for transmission and reception. Leveraging acoustic wave propagation has enabled humanity to exploit the undersea domain for military, engineering, scientific, commercial, and communications applications.

2.1.1 History

In his widely known work *Principles of Underwater Sound*, Robert Urick presents a concise historical survey of sonar and UAC [3].

First Uses

The submarine bell, used in conjunction with a foghorn for offshore navigation, was the first practical application of underwater sound, and led to the development of echo ranging in the early 20th century. These applications were limited to detecting icebergs and submarine signaling in the years leading up to World War I, after which military applications began to drive innovation. Quartz-steel sandwich projectors using the piezoelectric effect and vacuum-tube amplifiers began to replace older condenser projectors, and constituted the first use of electronics in underwater equipment. By 1918, submarines were receiving echoes at

range and determining target bearings, and surface vessels were towing hydrophone arrays to passively listen while avoiding self-noise.

Wartime

In the years between World War I and World War II, the scientific community gained a greater understanding of underwater sound by exploring the refraction of sound rays produced by temperature, pressure, and salinity changes in the ocean. Better understanding of underwater sound propagation, including how to calculate absorption coefficients were the most significant discoveries of this era. Further advances in electronics improved sonar in terms of power and utility, enabling increased directionality with hydrophones and better performance in listening and echo ranging.

As in World War I, World War II saw increased devotion to the study of underwater sound. Some of the most notable wartime developments were the acoustic homing torpedo, acoustic mines, underwater telephones, and scanning sonar. Further advancements in our understanding of target strength and noise laid the foundation for the sonar equations. Quantitative measures for sound recognition by the human ear, for classifying platforms by frequencies and speeds, and for describing reverberation in the sea laid the framework for the way we understand underwater acoustics today.

Postwar Era

In the postwar years, larger and more powerful equipment has allowed active and passive sonar to operate at lower frequencies and over longer ranges. Arrays placed on the ocean bottom exploit the quiet environment and propagation conditions at these lower frequencies, and the discovery of deep-sea sound channels and the convergence zone expanded the range of UAC applications. This era also saw a much larger focus on non-military applications, some of which include communications, telemetry, beacons, transponders, speedometers, fish finders, wave-height sensors, and fathometers.

Today, researchers continue to develop innovative uses of underwater sound. Underwater acoustic modems and sensor networks have expanded our capability to conduct at sea experimentation, area monitoring, and data collection. Maritime industries and the military have each placed a considerable emphasis on developing UUVs, which extend the length

and breadth of research beyond the normal human reach. As digital signal processing and available computing power each continue to improve, so too will the quality and variety of UAC applications.

2.1.2 Performance Factors

Performance factors for underwater communications can best be summarized by the passive and active sonar equations [3], [4]. The passive sonar equation is given by

$$SL - TL \geq NL - DI + DT, \quad (2.1)$$

where SL, Transmission Loss (TL), Noise Level (NL), Directivity Index (DI), and Detection Threshold (DT) are each measured in Decibels (dB). The active sonar equation is

$$SL - 2TL + TS \geq NL - DI + DT \quad (2.2)$$

when noise-limited and

$$SL - 2TL + TS \geq RL + DT \quad (2.3)$$

when reverberation limited. Target Strength (TS) and Reverberation Level (RL) are also measured in terms of dB. Generally, the SL, DI, DT, and TS parameters depend on the characteristics of the equipment and the target, while TL, RL, and NL are depend strongly on the propagation medium [3]. For the purposes of this study, we must elaborate on the parameters determined by the medium, as they will be the focus of simulation.

Transmission Loss

TL can be thought of as having two main components, TL due to geometric spreading and TL due to non-geometric effects such as absorption and scattering [4]. Geometric spreading can be approximated as spherical at short ranges, but it can be dramatically altered by refraction based on the Sound Speed Profile (SSP). The channel's SSP depends on changes in temperature, pressure, and salinity, each of which typically vary more by depth than by horizontal distance. The presence of layers can lead to trapped rays that spread cylindrically past a certain transition range.

Many factors contribute to non-geometric TL. Some sound is absorbed by the seawater itself.

Rays that bounce off the surface experience losses due to scattering, which is determined by sea state roughness, frequency, and angle of incidence. Though rays reflecting off the bottom can be a useful propagation path, bottom bounces experience some loss based on angle of incidence, frequency, as well as the physical properties of the bottom itself such as roughness and bottom type. The combination of multiple arrival paths in a received signal constitutes the multipath effect, which can have a significant and unpredictable effect on the TL and fidelity of an acoustic signal.

Reverberation Level

Back scattering due to heterogeneities like particulates, bubbles, wildlife, the sea surface, and the seafloor can act as an additional noise source in active sonar [4]. The amount of noise produced by this back scatter is called the RL, and depends on the transmit signal as well as the range and ocean properties. In UAC, forward scattering from these inhomogeneities will lead to changes in TL in the received signal.

Noise

Noise comes from a variety of sources and can greatly inhibit performance. NL comes from both ambient noise due to the environment, as well as self-noise of the receiver and electronic noise.

Sources of ambient noise in different frequency regions were summarized by Wenz [5]. At very low frequencies, ocean turbulence and seismic activity are the most prevalent ambient noise sources. Distant shipping and biological noise become significant between 20 Hz to about 500 Hz, beyond which sea state and wind speed dominate. Above 50 kHz, thermal effects on the water molecules themselves contributes more significantly to noise. In shallow water environments, the effects of ambient noise can be magnified due to generally higher levels of human activity, nearby surf, and denser ecosystems.

Self-noise comes from the receiving platform, and can be transferred to the receiver itself through the structure or through the water. Self-noise tends to increase with speed. At low enough speeds, self-noise will be less limiting than ambient noise [4].

2.1.3 Modeling The Underwater Acoustic Channel

As underwater sensor networks, acoustic modems, UUVs, and other nodes which perform UAC become more ubiquitous, mathematical models which can approximate the time and spatial varying channel become a useful tool when at-sea experimentation is not feasible or timely. The overarching framework for most theoretical models of underwater acoustic propagation is the wave equation, whose five canonical solutions make up the basis for the five major model categories [6]. Each model has its own advantages and limitations, and applicability in all cases depends on factors of frequency, range, and the complexity of the ocean environment.

Ray-Theoretical

Ray-theoretical models apply ray tracing theory using the Helmholtz equation in order to calculate TL [6]. These types of models calculate the coherent TL for each eigenray between source and receiver, and then add the pressures coherently or incoherently to calculate the total pressure. Typically ray-theoretical models include a direct path, bottom bounce, and surface bounce eigenray for close ranges and only a direct path at longer ranges, assuming that all others are attenuated away. Ray-theoretical models are useful for deep water, high frequency applications but are not applicable or practical in shallow water environments or at low frequencies [6].

Normal Mode

In simple normal mode theory, it is assumed that the medium is stratified and with cylindrical symmetry, so that the channel changes as a function of depth only. Solving the wave equation with a cylindrical coordinate system and subject to the boundary conditions at the surface and bottom yields a Hankel function for the travelling wave portion and standing waves in the vertical direction [6]. Simple normal mode models assume a single frequency point source, and are limited to frequencies below 500 Hz, as the number of modes needed to accurately estimate TL increase as frequency increases.

Multipath Expansion

Multipath expansion models are similar to normal mode models in that they result from integral expressions of the wave equation, but they only consider certain modes and are

range independent [6]. This approach works well for deep water scenarios at intermediate and high frequencies.

Fast-Field

Fast-field models have their roots in wavenumber integration, and use a similar integral transform as in normal modes but instead replace the first kind Hankel function with the first term of its asymptotic expansion [6]. The infinite integral can be evaluated by Fast Fourier Transform (FFT), which gives values for the scalar potential function of the wave equation at discrete points for particular source and receiver locations. The major downside to fast-field modeling techniques is the inability to handle range dependent calculations.

Parabolic Equation

In Parabolic Equation (PE) models, the elliptical reduced wave equation version of the Helmholtz equation is replaced with a PE, which relies on the assumption that group speed is very close to or equal to compression or shear speeds [6]. The simplified parabolic wave equation depends on range, depth, and azimuth. PE models can be very accurate and exist as both range independent and dependent solutions. Results can be calculated for a wide variety of parameters. The most significant downside to PE acoustic models is the dramatic increase in required computational time at frequencies above 500 Hz. Calculations become increasingly intensive when raising the number of range or depth interval steps, as decreasing the step size means more steps must be calculated in order to achieve the desired prediction range as frequency increases.

2.1.4 The Monterey-Miami Parabolic Equation

For the purposes of this study and its simulations, we opt to implement the MMPE acoustic propagation model [7]. MMPE, like other PE models, uses as its basis the parabolic form of the acoustic wave equation to represent the waveguide's forward propagating acoustic energy, but differs from other implementations in its methods of calculating the solutions to that equation. To do so, MMPE decomposes the acoustic field into a slowly modulating envelope function, with an oscillating phase term at an acoustic frequency, and then makes use of a propagator function to propagate the solution out in range. To compute the PE solution, MMPE uses the split-step Fourier (SSF). The SSF method is advantageous because

it balances computational efficiency (a known problem for most PE models) with generation of range dependent environments, whereas most PE models can only efficiently calculate range-independent solutions.

Though the MMPE model's employment of the SSF method to march the acoustic field solution out in range is relatively efficient, it does suffer from other complexities, for which accuracy suffers. Results from the Shallow Water Acoustic Modeling (SWAM'99) Workshop, which contained several canonical test cases for several different environments and frequencies, are presented for reference in [8]. The flexibility in parameterization, relative efficiency, and granularity trade offs in respect to accuracy available to MMPE make it an adequate model for our research. Additionally, since no modeling will need to be done in real time, the long calculation times inherent to PE models are acceptable.

2.2 Steganography Background

Steganography, a subset of the more general data hiding field, is the security technique of interest to this research, and one whose foundation must be adequately understood.

2.2.1 Steganography Principles

Fridrich informally defines steganography as “the practice of undetectably communicating a message in a cover object” [9]. In this case, the “cover” object refers to the object in which information will be hidden, and usually has little to no intrinsic value other than to contain the valuable hidden message. Applying a steganographic scheme to embed a message into a cover object yields a “stego” object, which a recipient would then use to extract the needed information. The value of steganography, as opposed to more conventional privacy tools such as cryptography, is that it achieves the same goals of private communication between parties while being covert rather than overt. That is, the very existence of the message in question is kept secret from would be attackers and eavesdroppers. In other words, the scheme is LPD. Messages may be passed to and from involved parties without raising suspicion or alerting an unwanted party. Undetectability is therefore the most important aspect of a given steganographic system [9]. The dramatic rise of digitization in multimedia and progress in signal processing capabilities has made steganography and data hiding increasingly easy to implement, and has drawn great interest in steganography as a field of

study for many U.S. agencies.

2.2.2 The Prisoner's Problem

The classical steganography scenario can be best described by Simmons's prisoner's problem [10]. Suppose Alice and Bob are both prisoners in the same prison, but assigned to different cells. The prisoners desire to escape and have resolved to work on a plan to achieve this. They are allowed to communicate with one another under the surveillance of a warden, Eve. If at any point Eve catches wind of their plan, she will disallow further communication and place each prisoner into solitary confinement. The prisoners will use steganography to detail their plans to escape in Eve's plain view.

In the prisoner's problem, we can assume that Eve is aware of the steganographic algorithm but is ignorant of the secret stego key that Alice and Bob would have agreed upon prior to imprisonment. In this way, the security of the system is not reliant upon the secrecy of the system itself, but rather the secrecy of its key [11]. As with encryption, the algorithm in question may fall into the hands of the enemy, and if it does, should not in and of itself compromise the security of the channel. The warden can be passive, and merely observe traffic, or active, and attempt to tamper with or disrupt the steganography or trick the parties into revealing their communication. Both must be considered when designing a secure steganographic system. From a philosophical standpoint, however, it should be noted that all Eve needs to do in order to break the steganographic system in the prisoner's problem is to detect that secret communication is taking place, as opposed to in encryption wherein access to the decrypted content is considered a successful attack. Extracting the secret message is a very difficult problem known as forensic steganalysis, and will not be explored in this research.

2.2.3 The Steganographic Channel

When Alice and Bob begin communicating secretly, they must decide how to define their channel, or comprehensive method of steganography. Fridrich recognizes five basic elements characteristic of a steganographic channel, namely a source of covers, algorithms, keys, messages, and data exchange channel [9].

Cover Source

The cover source is the digital object into which information will be hidden, such as images, audio, and text, and should be an object that can be exchanged overtly. One must be able to make a fundamental assumption about the cover source that makes formal analysis possible, so that its properties can be modeled and analyzed using information theory or complexity theory.

Embedding and Extraction Algorithms

The data-embedding and data-extraction algorithms are the procedures by which the secret bits are placed into and subsequently removed from the cover object. The two functions are normally inverse operations from one another, are not inherently secret, and rely upon a stego key.

Stego Key

The predetermined stego key is a value for determining the protocol for embedding the secret bits. A common method in steganographic schemes is to select the key at random, using a gaussian distribution, from among the set of all possible stego keys. The stego key must be kept secret for the scheme to work.

Message Source

The source of the intended hidden message must be carefully selected as it has strong influence on making the steganographic channel secure, as its size can directly contribute to the detectability of the covert communication. Needing to send consistently large messages and thereby exhausting the storage capacity of the channel can produce large distortions of the cover object, and undermine the goal of hiding information. Typically, the message source is modeled as a random variable covering the sample space of all possible messages.

Primary Communication Channel

The primary communication channel is the medium through which the stego object is transferred from Alice to Bob, or vice versa. As in the prisoner's problem scenario, we must always assume that this channel is being monitored by the warden, Eve, and furthermore, that Eve is behaving in one of three possible manners.

If Eve is a passive warden, she acts merely as an observer, watching all of the traffic but taking no action to interfere with the communication itself. If she is an active warden and suspects hidden communication between Alice and Bob, she can intentionally distort data in the primary communication channel and break the steganographic system. This can be mitigated, however, if Alice and Bob utilize a robust system that can withstand such processing. Finally, the worst case scenario is that Eve can be a malicious warden, guess the steganographic method, and attempt to impersonate either Alice or Bob to confuse or undermine both parties.

The main difference between the active and malicious warden is that an active warden seeks to make steganography between Alice and Bob impossible, whereas the more nuanced malicious warden leverages the stego channel to her advantage, rather than necessarily seek to disrupt or block it. For the purposes of this research, we will assume a passive warden scenario in which an adversary is likely oblivious to the fact that steganography is being employed, or takes no harmful action in the event they are aware.

2.2.4 Embedding and Extraction Methods

Fridrich identifies three fundamental paradigms by which the embedding and extraction algorithms function in relation to their respective covers. These include steganography schemes based on cover selection, cover synthesis, and cover modification, and are elaborated on in the following paragraphs [9].

Steganography By Cover Selection

In a steganography by cover selection scheme, the hidden message is expressed by interpreting a feature inherent to the cover. For this type of scheme to function, Alice must have a selection of static covers, such as image or audio files, from which to choose. A feature specific to that cover will in turn express some information. The embedding algorithm need only be a process that selects from the available covers at random until one suitable for the intended message or portion of the message is found. The stego key is a set of predetermined rules that dictate how Alice and Bob interpret the selected covers.

In practice, many cases of steganography by cover selection involve hash functions. Alice can select a cover from her database, run it through a message-digest function, and compare

the output digest to a desired message bit stream, repeating as necessary with other covers until a match is found, which can then be forwarded to Bob. Bob can obtain the message by extracting the digest.

Steganography by cover selection is advantageous in that the covers are unmodified, therefore there is nothing to introduce artifacts or avenues for statistical exploitation. The biggest disadvantage, however, is that the expected number of tries to find a suitable will increase exponentially as the length of the digest increases, thus precluding the feasibility of longer messages.

Steganography By Cover Synthesis

In a steganography by cover synthesis scheme, Alice generates covers according to the message that needs to be passed, often in conjunction with cover selection to offset the exponential complexity introduced by hashing. Data masking is among the few practical methods of steganography by cover synthesis, and works under the assumption that automated steganalysis tools identify numerical features of a given object, which may or may not be a stego, and compare them with analyzed features obtained from prototypical cover objects. Evading detection is as simple as emulating the statistical features of a cover in the cover space, and not necessarily the observed covers themselves. Radhakrishnan et al. demonstrate a practical method of data masking by using an inverse Wiener filter to make the message look statistically like an audio or image object, giving a covert channel capacity potentially much higher than in traditional steganography [12].

Steganography By Cover Modification

Steganography by cover modification, which is to date the most studied and practiced form of steganography, involves Alice starting with a cover object and making calculated changes to it in order to embed the secret message. Bob receives the stego object, and without requiring the original cover, can extract the message. When communicating, Alice and Bob can select from the set of all possible covers, as well as the sets of keys and messages—some of which depend on the cover itself, to utilize a steganographic scheme that is by definition the pair of embedding and extracting algorithms. Formally, we represent these concepts in terms similar to those used in cryptography: C is the set of cover objects, such that $x \in C$, $\mathcal{K}(x)$ is the set of all stego keys for x , and $\mathcal{M}(x)$ is the set of all messages that

can be communicated in x [9]. A steganographic scheme can therefore be formally defined as an embedding function Emb ,

$$Emb : \mathcal{K}(x) \quad (2.4)$$

an extraction function Ext ,

$$Ext : \mathcal{M}(x) \quad (2.5)$$

such that $\forall x \in C, \forall k \in \mathcal{K}(x)$, and $m \in \mathcal{M}(x)$,

$$Ext(Emb(x, k, m), k) = m. \quad (2.6)$$

The length or number of messages which can be effectively communicated in a given cover x is dependent on the scheme and the cover itself. Fridrich defines several terms to aid in understanding these relationships [9]. Two important concepts which further illustrate this point are the embedding capacity and relative embedding capacity. The maximum quantity of bits or symbols which can be embedded in a given cover element is its embedding capacity, while its relative embedding capacity is the ratio of the embedding capacity to the quantity of individual elements of the cover where a message can be embedded. The concept most critical to the steganographic channel, however, is the steganographic capacity, defined as the maximum quantity of bits or symbols which can be embedded without introducing artifacts susceptible to statistical analyses. The steganographic capacity is typically a great deal smaller than the embedding capacity.

2.3 Audio Steganography

The use of audio as covers, as well as embedding and extracting algorithms tailored to existing formats, has become increasingly relevant as digital data usage spreads. The variety of steganographic criteria and transmission environments has led to a similarly great diversity in techniques. System design and usage is thus both media and situation dependent. The most critical properties associated with audio steganography techniques are the robustness to noise, compression, and signal manipulation, in conjunction with the integrity, security, and volume of the hidden data. Bender et al. describe these properties and explain the typical tradeoffs between capacity and robustness [13]. Another key feature pertinent to audio steganography is to utilize natural limits in the Human Auditory System

(HAS) as a means of undetectability, in addition to the aforementioned susceptibility to statistical analyses. Djebbar et al. compile and discuss current trends in digital audio steganography, explaining their tradeoffs, practicality, and methods [14].

2.3.1 Temporal Domain

Temporal domain audio steganography techniques were among the earliest conceived for the medium, and have the advantage of being relatively simple to understand and implement. As a trade-off, however, robustness and security each suffer, and only a moderate tolerance to noise addition at low levels has been shown. The most robust and secure temporal domain techniques achieve those criteria at the expense of data hiding capacity. Temporal domain techniques often rely on low-bit encoding techniques. To date, few have been developed and employed.

Low Bit Encoding

Low bit encoding, or Least Significant Bit (LSB) encoding, is a traditional steganography method applicable to almost every type of cover element. Its various incarnations each involve deterministically embedding the message bits into the audio cover's LSBs. LSB in its simplest form can support a hiding rate in Bits Per Second (bps) proportional to the audio sampling rate in Hz, but this simplicity comes at the cost of easy message retrieval for a would-be attacker. LSB embedding also suffers from low robustness. Lossy compression, filtration, amplification, and noise addition are each among the inevitable phenomena when processing audio that can cause the message to become significantly corrupted or outright lost. Cvejic et al. demonstrate several variations on this concept showing improved robustness against distortion while preserving perceptual transparency of the stego signal by leveraging multiple LSB layers [15], [16]. Robustness and imperceptibility in most LSB techniques is improved at the cost of hiding rate and capacity.

Echo Hiding

Echo hiding involves embedding the data via small "echoes" of the cover audio signal, in the form of resonance in order to avoid sensing by the HAS. For each data segment, the starting amplitude, offset, and decay rate are adjusted as needed to express the desired symbol, but must be so adjusted such that the added echo effect is imperceptible. Though able to pass

the HAS test, echo hiding techniques can only support low embedding rates and low levels of security. Few recent works present research into echo hiding methods, and the only known applications to date are to audio watermarking.

Silence Intervals

Audio sources that exhibit brief interspersed moments of silence, such as speech signals, may use those properties to embed data in a silence interval scheme. In a typical implementation of this method, the cover's discrete silence intervals and their respective number of samples must be determined, from which the largest integer value less than 2 raised to the power of the number of bits used to represent a given message is subtracted. For example, if the value "8" were to be hidden in a silence interval of length 110 samples, the new interval length would be 95 samples. If the value "6" were to be hidden, the new interval would be 103 samples. Extracting the hidden data makes use of the modulo operator on the new silent interval. Smaller intervals, which in the case of speech usually occur mid-sentence, are left unchanged in order to avoid disturbing the quality of the speech. Though this method produces decent perceptual transparency, it does not survive compression. The changing silence interval lengths result in corrupted data extraction.

2.3.2 Transform Domain

Transform domain audio steganography exploits the inherent properties of both audio signals and the HAS. Processing in the transform domain is dependent on the inherent masking effect wherein weaker frequencies are audibly masked by nearby, stronger resonant frequencies. Stego audio modified in the transform domain achieves its stealthiness by taking advantage of this masking effect. This is achieved by exclusively modifying masked regions, or by slightly modifying the audio signal's samples.

Performing audio steganography in the transform domain as opposed to the time domain shows demonstrably better Signal-to-Noise Ratio (SNR) due to the frequency masking phenomenon [13]. As a result, perceptual models are often used as benchmarks when determining embedding capacities relative to acceptable stego signal distortion. Many techniques have recently evolved which provide security and robustness in the face of common audio signal processes like re-sampling, amplification, and filtration. The chief concern in employing transform domain techniques remains finding a way to preserve

the integrity of embedded data under conditions unlikely to their survival, namely noisy transmission channels, compression, and encoding processes.

Spread Spectrum

Spread Spectrum (SS) is an important communications capability by which a signal can improve its chances of recovery when transmitted through a noisy channel by producing repeated copies of itself throughout the frequency spectrum. SS steganography, in its most basic form, multiplies the message data by an M-sequence code, then hides it in the cover audio [17]. The code is the stego key and must be shared between Alice and Bob. Several variations of the SS concept include Direct Sequence Spread Spectrum (DSSS), hiding the SS data under a frequency mask, combining the technique with phase-shifting, or using the technique in the sub-band domain. Each variation provides a measured increase in robustness and helps in detecting the hidden data. SS techniques are beneficial due to their high robustness, but become vulnerable to time scale.

Discrete Wavelet Transform

Audio Steganography based on the Discrete Wavelet Transform (DWT) is a marriage of LSB embedding and frequency domain calculations. After transforming the cover signal, message data can be hidden in the wavelet coefficients' LSBs. The inverse transform creates the stego signal. The DWT provides both time and frequency relative information, allowing these techniques to be further refined by implementing hearing thresholds when embedding into coefficients, and by simultaneously avoiding embedding in silent portions. The result is a stego signal with minimal perceptible artifacts. Hiding information using the DWT yields very high embedding rates, but the recovery at the receiver side can often be distorted and inaccurate due to the effects of the transform.

Tone Insertion

Steganography by tone insertion leverages the masking property specific to relative powers rather than frequencies. Lower power tones can be inaudible in the presence of higher power tones. Embedding one information bit into a given audio frame involves selecting two tones, each generated at two distinct frequencies, and setting their respective power levels such that a known ratio between the two and the power level of the frame determines whether a "0"

or "1" is inserted. Though embedding rate is limited by the number of audio frames, using appropriate frequencies and sufficiently low power levels balances both robustness and ease of data recovery. At the reception side, each frame is examined, from which the frame's power and the tone frequencies' power are compared to determine the message bit. Audio steganography by tone insertion is resistant to common attacks such as low-pass filtration and bit truncation, but the artifacts associated with inserted tones are easily detected and could possibly lead to the warden extracting the hidden bits.

Phase Spectrum

Embedding with phase coding replaces specific phase information from the cover audio with hidden message data, and works on the fundamental assumption that the HAS is insensitive to the relative phase of a perceived signal's spectral components, provided that the modification to the phase components is kept small [18]. A basic approach is to modulate an independent multi-band phase to insert data into a desired phase component. In their implementation, Gang et al. achieve imperceptibility by using a controlled phase alteration on the cover audio with a Quantization Index Modulation (QIM) method. In QIM, the phase value of a given frequency bin is replaced by the nearest of two possible constellation points for either a "0" or "1." The advantage of phase spectrum techniques is that they are robust to audio compression. Additionally, HAS is not sensitive to phase changes, making it more discreet. As a downside to this fact, however, an active warden can disrupt a phase quantization scheme by quietly inserting imperceptible frequency modulation into the signal.

Magnitude Spectrum

Magnitude spectrum, or amplitude coding, follows the principle that the HAS is more dependent on frequencies as it is more sensitive to signal amplitude values. Djebbar et al. proposed an algorithm which examines a wideband speech spectrum, locates secure spectral embedding areas, and embeds the payload data using a frequency mask 13 dB lower than the cover signal spectrum [19]. The algorithm takes a tolerated distortion level and selects the embedding locations and hiding capacity based on that. For example, frequency components in the 7 kHz to 8 kHz range have negligible contribution to wideband speech, and can largely be replaced by message data. Djebbar's method preserves speech quality and is capable of high embedding capacities.

Cepstral Domain

A cepstrum is the Inverse Fourier Transform (IFT) of the logarithm of a signal's estimated spectrum. Cepstral domain techniques, also known as log spectrum techniques, transform the cover signals into the cepstral domain and embed data into specific cepstrum coefficients selected by statistical mean manipulations. The cepstra of two frequencies in each energetic frame are modified to embed either binary bit. This method is computationally expensive, but has the benefit of high embedding rates and resistance to common signal attacks. Furthermore, performing the cepstrum transform over regions which are perceptually masked in most of the cover audio frames guarantees the same in the stego frames.

Table 2.1. Audio Steganography Techniques

Method	Embedding Tool	Strengths	Weaknesses
Low Bit Encoding	Embed data in least significant bits of cover according to some rule	High bitrate, simple implementation	Data easily extracted or destroyed
Echo Hiding	Introduces echo into cover signal	Robust against lossy compression	Low capacity, low security
Silence Intervals	Hidden data is represented by number of samples in periods of intermittent silence	Robust against lossy compression	Low capacity
Spread Spectrum	Spreads hidden data over all signal frequencies	Robust	Weak to time scale modification
Discrete Wavelet	Embeds data in altered wavelet coefficients	High capacity	Lossy data extraction
Tone Insertion	Embeds data in low power inaudible tones at specific frequencies	Imperceptibility of embedded data	Lack of transparency, not very secure
Phase Spectrum	Modulates the phase of the cover signal	Robust against signal processing changes, data extraction requires knowledge of the original signal	Low capacity
Magnitude Spectrum	Use frequency bands to hide data	Can hide longer messages, less likely to be hurt by transmission errors	Not robust to simple audio manipulation
Cepstral Domain	Embeds data in altered cepstral coefficients	Robust against signal processing changes	High signal distortion, low robustness

A summary of the categories audio steganography techniques. Adapted from [14, Table 4].

2.4 Steganography and Covert Underwater Acoustic Communications

To date, little research exists devoted to steganography applications to UAC, though other forms of covert communication have been proposed. Applying the techniques discussed in 2.3 is much more challenging due to the nature of the channel, as are balancing robustness, imperceptibility, embedding rate, and security. Most importantly, the underwater acoustic channel is dynamic and lossy, with variations in water depth, temperature, salinity, bottom topography/composition, surface roughness, and node distance contributing widely to the stego signal's expected TL. As a result, any feasible covert UAC technique must be resilient to substantial signal degradation and should expect to transmit at very low bitrates. Further, inevitable multipath effects produce interference affecting signal quality. The following subsections detail the extent of existing work particular to UAC, as well as the extent to which each author overcame the challenges associated with the underwater acoustic channel.

2.4.1 Orthogonal Frequency-Division Multiplexing

Leus et al. demonstrate a covert UAC scheme tailored to communications between a UUV and mother platform in [20]. Rather than use a unique data embedding scheme, this application instead takes advantage of multiband Orthogonal Frequency-Division Multiplexing (OFDM) at low SNRs to achieve its covertness. Here, using OFDM as opposed to a Frequency-Shift Keying (FSK) scheme or other chip modulations has the distinct advantage of appearing noise-like. The data transfer at very low SNRs furthers this concept, allowing the messages to be hidden in the presence of ambient noise. The authors chose OFDM as their modulation scheme due to its ability to streamline the equalization process, and multiband over single-band to reduce the receiver side complexity. They also opt to perform the equalization and despreading operations jointly rather than as separate, linear processes.

With a transmitter/receiver pair programmed in this scheme, Leus et al. conducted trials both at sea in two littoral regions as well as in a simulated communication channel using 16 subbands and a total 3.6 kHz bandwidth, with varying time and Doppler spreading. Two user data rates were tested, 78 bps and 4.2 bps. At 78 bps, the performance limits of their scheme are evident, with substantial bit errors occurring at ranges beyond 30 km, and a successful SNR down to -9 and -7 dB, respectively, for each of the two sea channels. In comparison, the 4.2 bps mode transmitted successfully out to 52 km with no bit errors,

with SNRs of -17 and -13 dB. In the first, more benign channel, the 4.2 bps signal limits were extrapolated, not measured, therefore performance is limited by detection and failure of initial synchronization, instead of the OFDM scheme itself. It should also be noted that the second channel was more difficult due to multipath fading and a time-variable Doppler shift induced by the transmitter's reflection from the uneven sea surface. In all cases, covert signals at the lowest achievable SNRs are audibly imperceptible and invisible when viewed in a spectrogram, even if their position in the recorded data is known to a listener.

The disadvantage of a solely modulation based OFDM scheme is the possibility that a trained sonar operator could observe the acoustic features that arise from its signal waveform, which is generated by modulating fixed carrier signals. Furthermore, attenuation is based heavily on distance and the signals are likely to experience frequency selective fading.

2.4.2 Biological Mimicry

Rather than modulate using a scheme whose signal is itself undetectable in the presence of noise due to low SNR, several authors instead opt to exploit the nature of sounds occurring in the underwater acoustic channel in a form of biological mimicry. This more steganographic approach selects sounds based on their beneficial properties, and uses them to obscure the signal carrying the message or convey using a timing sequence. An adversary must think that the noise they observe is occurring naturally.

Masking

Liu et al. present a recent experiment in covert UAC biological mimicry using a DSSS carrier signal masked by relatively louder whale noises in both the time and frequency domain [21]. The communication signal is allowed to be detected, and does not require that its signal level be reduced to the minimum possible. The scheme depends on the warden classifying the signal's observable characteristics as innocuous whale noise. This must hold for both technical displays as well as the HAS.

Liu et al. exploit the HAS masking phenomenon by choosing the relatively low frequency whale noises as the masking sound, as lower frequencies of higher power mask higher frequencies, whereas the opposite is not true for higher power high frequencies. The low frequency signals also aid in longer distance UAC. The authors use a DSSS signal to carry

the actual information in the same frequency range as the whale noise, using a slightly weaker signal amplitude at each successive frequency. To extract the information signal on the receiver's end, they use a frame structure adapted to synchronization, channel estimation, and equalization. The synchronization operation is performed by using a matched filter technique on the original whale signal. Using a high SNR probe signal based on that original whale signal, the channel impulse response can be accurately estimated using a matching pursuit method, which greatly improves the scheme's resistance to multipath effects. Finally, a virtual time reversal mirror equalizer uses the reference whale signal to separate the DSSS signal from the combined signal, thereby obtaining the final virtual received signal and recovered information.

In their simulation, the authors used a smooth humpback whale sound and a DSSS signal in the 600 to 1000 Hz band, with a 10 dB masking strength, in both a BELLHOP simulated shallow water multipath channel and an Additive White Gaussian Noise (AWGN) channel. In this case, the masking strength is the logarithm of the power of the normalized whale sound divided by that of the DSSS signal expressed in dB. The simulated channel was 30m deep, with source and receiver each at 10m and 30km horizontal range. The experiment was repeated over a range of SNRs, as well as masking strengths of 20 and 30 dB to test the effectiveness of the whale sounds interference algorithm and to compare resulting Bit Error Rate (BER). The results indicated that the BERs were predictably higher for the multipath channel than for the gaussian noise channel. With masking strengths lower than 5 dB, the major signal energy is the spread spectrum signal, and below this threshold the masking strength is too low for reliable synchronization. Beyond the 5 dB threshold, BER seemed to be mostly affected by noise. In every case, the BER decreases sharply over the range of -15 to 25 SNR/dB for each masking strength curve.

2.4.3 Environmental Adaptation

In sections 2.4.1 and 2.4.2, the authors focused on LPI and LPD schemes whose reliance on waveform design and undetectability by conventional methods was paramount. Though feasible results have been proven in the respective schemes, the chosen methods pay little concern to the environment in which they are being transmitted and received. To leverage steganography principles in UAC, as is the goal of this study, methods which take more care to conceal the existence of communication, as well as those which pay particular attention to

evaluation by statistical techniques rather than energy intercept detectors, must be explored.

Steganographic Framework

Though intended for covert active sonar tracking rather than UAC, Park et al. present a steganography by cover synthesis method for waveform design which applies relevant concepts [22]. Park et al. employ short-time ambient background modeling using a “time series analysis, phase space representation, and clustering with a time series distance metric [22].” Then, a hidden Markov model captures the time-evolving qualities of the sound. Doing so creates a probability distribution of symbol groups, from which a stochastic matrix and emission distribution function are generated, which together make up the estimated reference background sound model. Park et al. leverage this model by creating a mimic waveform which seeks conform to the statistical properties of the reference. From there, the authors evaluate the security of their algorithm by calculating the Kullback-Leibler divergence and performing hypothesis testing to determine the probability of missed detection, giving them the lower bound of this probability for a potential intercept detector. Kullback-Leibler divergence, also called relative entropy, is a measure of how much a given probability distribution diverges from another, expected distribution, and is a standing metric in steganographic security [23]. Two candidate waveforms were compared: one based on mimicry of the background distribution, and one based on the AWGN distribution, with each being subject to an approximated, simple attenuation factor based on the frequency-selective and range dependent nature of underwater propagation. The mimicked signal outperforms the AWGN waveform across the range of signal strengths, achieving a lower bound probability of missed detection as high as 0.8 at optimal waveform strength as seen from the target.

Recorded Noise

Critical practical limitations of most proposed schemes arise from the human elements involved in detecting covert UAC. Passerieux points out many of these limitations in development of his novel, patent pending stealth communication scheme [2]. A skilled sonar operator armed with a modern suite of audio and time-frequency analysis tools can undermine many of the existing covert UAC techniques. Even if unable to decode messages or conduct an attack, the operator may become a passive warden, aware that communication is taking place. Ubiquitous DSSS techniques, even at lower SNRs, can be susceptible

to detection if the transmission is long enough. OFDM schemes, despite their optimistic low SNR requirements, have not developed to the point where they are resistant to brute force attacks. Schemes based on simulated biological noise, such as whale songs, periodic dolphin clicks, or snapping shrimp, have the distinct challenge of having to successfully emulate nature to high degree and not appear or sound artificial. Several authors propose using recorded biological or weather based noises as covers in their steganography schemes in order to avoid the aforementioned problems, with the added benefit of communication always being suited to the current operating environment. This improves the stego signal's imperceptibility.

Dol et al. are among the few authors to explore a method of UAC using distinct steganographic techniques with application to naval tactics [24]. Their techniques leverage naturally occurring sounds based on the characteristics of their spectrograms, and for cover audio tend to favor whistling sounds with fundamental frequencies and harmonics, as well as biological click sequences. Their work applies three types of audio steganography: tone insertion, phase coding, and silence intervals. Each were detailed in sections 2.3.1 and 2.3.2.

In regard to the time domain techniques, tone insertion and phase coding, robustness is maximized by using frequencies in which the audio energy is concentrated in the carrier signal, as well as by using larger segment sizes, though the latter option comes at the cost of data rate. Both techniques can be combined, but in general modifying the phase is more stealthy than changing the power and amplitude. Implementing a random frequency or phase sequence as a stego key, as well as exploiting harmonics, are additional features which can be implemented to increase security and redundancy, respectively.

For the silence interval technique, Dol replaces the click sounds produced by many cetaceans with an artificial sequence consisting of short, linear frequency modulated pulses centered around a specific frequency, but with random phases [24]. By changing the time sequence between subsequent clicks and providing a tolerance in sequence, selected length data bits can be encoded. The click sequence is further modified to sound more natural by appending extra clicks as a "sweep" at the end of each segment with a rising central frequency, a trait commonly observed in actual marine mammal clicks. The click sequence, silence interval technique is very susceptible to multipath effects, which can add apparent random,

erroneous extra clicks.

To test their scheme, Dol et al. performed sea trials between a frigate-towed transmitting source and a Walrus-class diesel-electric submarine, each on an identical base course and traveling between 6 and 8 knots, and with a gradually increasing range over the course of one hour. The submarine maintained depth between 60–80 meters, while the source remained between 50–55 meters. The sound source consisted of one Low Frequency (LF) transducer ring with a 1–2 kHz range and a High Frequency (HF) one with a 4–8 kHz range. Together, these two transducers cover most of the typical 1–8 kHz range of real marine mammal sounds. A .wav file based on recordings previously obtained by the submarine was prepared for each band demonstrating the three schemes [24].

Dol's results were mixed. Too little data was recovered to enable statistically relevant matching by the tone insertion technique, and its results were consequently left out. Successful decodings were only achieved using the click sequences, yielding a 5% success rate for LF and a 7% success rate for HF. It should be noted that no attempts at equalization or error correction were made. Dol et al. conclude that the silence interval click sequence, despite being noted to sound rather "digital" to the operators, was the best coding technique to prove the concept [24]. This is only provided that the low-energy clicks can propagate to their destination, a fact which was made easier due to the very small amount of multipath propagation present during the experiment. Furthermore, the data capacity is linked to the number of clicks transmitted, thus a necessarily longer sequence will be more likely to arouse suspicion. Though in this brief experiment the overall robustness could not be guaranteed for phase coding, Dol posits that implementing error correction and equalization may make it the most viable overall technique due to its imperceptibility [24].

Passerieux's Technique

Literature to date contains few other examples of background noise recording based techniques. Recognizing the limitations of using an artificially repeated, short recorded click series, Passerieux presented a novel steganography based technique relying only on manipulation of a single continuous audio sample of previously recorded sound [1], [2]. The scheme proposes recording a variable length sound clip for use as a cover, from which a unique auxiliary signal can be computed and inserted back into the original signal, thus producing the stego signal.

On the source side, the algorithm works by dividing the cover signal into a sequence of short, equally sized, non-overlapping intervals, usually several hundred milliseconds long. The number of these segments is proportional to the number of embedded symbols. For each interval, an operator " Ψ " is applied to the complex baseband demodulation of the signal in a frequency band centered on a selected carrier frequency. This allows for optimization with respect to covertness and desired SNR on the receiving end, as well as adaptation to a number of modulation techniques for symbol expression. Though it can be implemented in a variety of ways, the Ψ operator should include some means of incorporating a steganographic key into the embedding scheme, and in Passerieux's implementation provides resistance to Doppler and signal reproduction. The result of the Ψ operation is then multiplied by the desired symbol and a gain value, which must be minimized as much as possible in order to make the transmitted signal look similar to the original signal. This quantity is then added to the initial signal, and each segment is recombined to form the stego signal.

The reception scheme is described in three broad steps. First, since the original signal is not known to the receiver, the auxiliary signal must be approximated from the received stego signal. This will also depend on the gain, as higher values mean that the stego signal will be more distorted. Next, for each interval, the signal must be demodulated into a new signal from which the data will be extracted, which introduces a Doppler compensation term determined at each measurable time interval. This new signal, due to the data embedded in phases, will display periodic peaks at intervals equal to the originally chosen segment size, with amplitudes and phases proportional to the transmitted symbol. From there the symbols can be extracted.

Passerieux evaluated his algorithm in two phases. The data embedding portion was tested on a 10-second sperm whale click signal and 25-seconds of rain noises, each recorded at sea. He inserted bits in the 1500–3000 Hz frequency band, with 0.5-second segments. This equates to 2 bps in his chosen Binary Phase Shift Keying (BPSK) scheme. The gain applied was -10 dB. The spectrograms of each audio cover were compared before and after embedding, revealing no visually detectable artifacts. Moreover, listening to the audio samples claimed no discernible difference.

To test the reception chain, Passerieux simulated the received signal after propagation through an underwater acoustic channel, using a scattering technique described in one

of his previous works [25]. The main parameters include a 2250 Hz carrier frequency, transmitter and receiver range of 5000 meters and at respective depths of 5 and 60 meters, and 100 meters total water depth. The Doppler spread was set to 0.5 Hz, and in order to simulate a non-zero Doppler equivalent to 1.7 m/s relative speed, he resampled the received signal based on sound speed. Remarkably, each of the 16 and 50 estimated symbols were recovered correctly despite low SNR at reception, about 0 dB. This novel concept may be a valuable addition to the steganographic literature with regards to expected robustness.

2.5 Summary

In this chapter, we presented an overview of UAC, including history of applications, the underwater acoustic channel, and methods to approximate the channel. Next, we introduced steganography, including its composite elements, mechanics, and goals. Next, we described existing techniques for applying steganography to audio covers, including the general requirements for performing each technique, the strengths and weaknesses inherent to each, and the critical metrics of robustness, imperceptibility, security, and SNR. In Section 2.4, we discussed UAC, and detailed the numerous challenges necessary to overcome when attempting to use audio steganography techniques as a means of covert UAC. We explained the various works focused particularly on stealth underwater communications for use by submersibles, surface ships, and underwater modems. Though SS and OFDM signal processing-heavy techniques show great promise for future work, we choose to focus on the newer paradigm presented by Dol and Passerieux regarding localized ambient, biological, and weather noises. Passerieux’s technique is of particular interest due to its upfront claim to work within the boundaries of what most in open literature see as practical limitations. We intend to further explore Passerieux’s algorithm as a steganographic technique and analyze its capabilities and limitations. Specifically, we position our study to further investigate the following:

- Performance in terms of traditional steganography metrics, such as relative distortion and quality terms Mean-Square Error (MSE), Root-Mean-Square-Error (RMSE), Peak Signal-to-Noise Ratio (PSNR), and SNR.
- Evaluation of performance via a parabolic equation underwater acoustic channel model.
- Evaluation of algorithm performance at different source and receiver depth and range.

- Evaluation of algorithm performance in a deep water, open ocean environment as well as a constrained, multipath heavy shallow water environment.
- Performance with different length interval segments.
- Evaluation of how much gain increase is tolerable when balancing data recovery with artifact introduction.
- An overall evaluation of robustness, imperceptibility, data rate and the trade-offs between these metrics.

The remainder of this thesis will describe and analyze the approaches taken to implement Passerieux's technique with our chosen deviations, to include the methodology behind channel simulation, and will provide analysis of the results.

THIS PAGE INTENTIONALLY LEFT BLANK

CHAPTER 3: Methodology

In Chapter 2, we discussed the most important concepts of UAC, acoustic channel models, and steganography, then positioned our research in reference to the existing literature. In this chapter, we present a steganography by cover modification, phase spectrum technique in the transform domain based on the principles of Passerieux’s method [1], and simulate a test channel using the MMPE model. Section 3.1 presents the rationale behind our method, including the departure from Passerieux. Section 3.2 details our implementation for calculation, while Section 3.2.1 describes the modules used in the experiment and how they interact with one another. Section 3.2.2 is a description of the input files and post-processing routines used to generate the MMPE channel model.

3.1 Theory

In his original paper and patent, Passerieux proposes to perform his steganographic algorithm in the time domain on the complex baseband signal representation of an original signal obtained after demodulation around a chosen carrier frequency and bandwidth [1], [2]. Passerieux calculates an auxiliary signal, $\Psi_x(t)$ from a noise signal, $x(t)$. $\Psi_x(t)$ is calculated in a manner known to both the transmitter and receiver, with a series of phase terms β_p acting as the stego key. The stego signal to be transmitted, $y(t)$, can then be obtained as

$$y(t) = x(t) + \alpha \cdot a_m \cdot \Psi_x(t), \quad (3.1)$$

where α is a gain term acting as the relative amplitude of the auxiliary signal to the original signal and a_m is the symbol to be embedded (± 1 in a BPSK scheme) in the m^{th} symbol period. With knowledge of the scheme, as well as the pre-shared stego key, the receiver can perform a complementary operation to estimate $\Psi_x(t)$ as $\Psi_y(t)$, and thereafter cross-correlate $\Psi_y(t)$ with the received signal in order to extract the embedded symbols.

The Ψ operator is critical to this scheme. To create the auxiliary signal, $x(t)$ is first divided into small segments of period, T , normally several hundred milliseconds, depending on the desired bitrate. Next, each period, T , is divided into $2p$ sub-intervals of decreasing

and subsequently increasing lengths, indexed from $-p, \dots, -2, -1, -1, -2, \dots, p$. The sub-intervals are separated by guard intervals, both of whose lengths are heuristically chosen. The length of both the sub-intervals and the guard intervals are symmetric around the midpoint of the period segment. The data points in the segment are switched and then time-reversed such that data in the p^{th} sub-interval, $x_p(t)$, would then become $x'_p(t) = x_{-p}(\tau_p - t)$ when switched and time-reversed, where τ_p is the time delay between sub-interval $-p$ and p

3.1.1 Embedding Theory

To leverage the efficiency of the Discrete Fourier Transform (DFT), avoid computational complexity associated with calculating the complex baseband equivalent, and to provide flexibility with frequency selection, we choose to examine an implementation of this concept in the frequency domain. Beginning with a sampled digital signal in the time domain, $x(t)$, which we split into intervals of T seconds as before, we then calculate the DFT of each segment of T seconds. This yields a frequency domain representation of each interval of $x(t)$

$$\hat{X}(k) = \sum_{n=0}^{N-1} x(n) e^{-j \frac{2\pi nk}{N}}, \quad (3.2)$$

where N is the number of sampled audio data points in each T second segment, and k is the frequency of a given bin. This bin frequency is described in terms of the analog frequency as

$$k = \frac{fN}{f_s}. \quad (3.3)$$

Assuming $x(n)$ is a sequence of real numbers, its DFT will have symmetric negative and positive frequency components given by $\hat{X}(k) = \hat{X}^*(N - k)$, where the “*” operator represents the complex conjugate. Next, we take $2p$ sub-intervals and guard intervals in a chosen frequency range, up to the Nyquist frequency, $k = \frac{f_s}{2}$. Time delay and time-reversal in the time domain are analogous to multiplication by a phase term and taking the complex conjugate, respectively, in the frequency domain. For example, if we have chosen two equal length sub-intervals in which to embed our symbols, the data in this period could be

expressed as

$$x_m(n) = \begin{cases} x_1(n) & 0 \leq n \leq n_1 \\ x_2(n) & n_1 + 1 \leq n \leq n_2 \\ x_3(n) & n_2 + 1 \leq n \leq n_3 \\ x_4(n) & n_3 + 1 \leq n \leq N - 1 \end{cases}. \quad (3.4)$$

Combining the sequence given in 3.4 with equation 3.2, the DFT of this original sequence can be expressed in terms of sub-intervals as

$$\begin{aligned} \hat{X}(k) = & \sum_{n=0}^{n_1} x_1(n) e^{-j \frac{2\pi nk}{N}} + \sum_{n=n_1+1}^{n_2} x_2(n) e^{-j \frac{2\pi nk}{N}} \\ & + \sum_{n=n_2+1}^{n_3} x_3(n) e^{-j \frac{2\pi nk}{N}} + \sum_{n=n_3+1}^{N-1} x_4(n) e^{-j \frac{2\pi nk}{N}}. \end{aligned} \quad (3.5)$$

Switching the position of data in sequences 2 and 3 and then time reversing them, the new signals $x'_2(n)$ and $x'_3(n)$ become $x_3(n_1 + n_3 + 1 - n)$ and $x_2(n_1 + n_3 + 1 - n)$, respectively. By letting $n_0 = n_1 + n_3 + 1$, the DFT of the auxiliary signal Ψ_x is given by

$$\begin{aligned} \hat{\Psi}_X(k) = & \sum_{n=0}^{n_1} x_1(n) e^{-j \frac{2\pi nk}{N}} + \sum_{n=n_1+1}^{n_2} x_3(n_0 - n) e^{-j \frac{2\pi nk}{N}} \\ & + \sum_{n=n_2+1}^{n_3} x_2(n_0 - n) e^{-j \frac{2\pi nk}{N}} + \sum_{n=n_3+1}^{N-1} x_4(n) e^{-j \frac{2\pi nk}{N}}. \end{aligned} \quad (3.6)$$

Substituting n' for $n_0 - n$, the DFT of these new sequences can now be expressed as

$$\begin{aligned} \hat{\Psi}_X(k) = & \sum_{n=0}^{n_1} x_1(n) e^{-j \frac{2\pi nk}{N}} + \sum_{n'=n_0-(n_1+1)}^{n_0-n_2} x_3(n') e^{-j \frac{2\pi(n_0-n')k}{N}} \\ & + \sum_{n'=n_0-(n_2+1)}^{n_3} x_2(n') e^{-j \frac{2\pi(n_0-n')k}{N}} + \sum_{n=n_3+1}^{N-1} x_4(n) e^{-j \frac{2\pi nk}{N}}. \end{aligned} \quad (3.7)$$

This can be simplified to

$$\begin{aligned}\hat{\Psi}_x(k) = & \sum_{n=0}^{n_1} x_1(n) e^{-j \frac{2\pi n k}{N}} + e^{-j \frac{2\pi n_0 k}{N}} \sum_{n'=n_0-(n_1+1)}^{n_0-n_2} x_3(n') e^{j \frac{2\pi n' k}{N}} \\ & + e^{-j \frac{2\pi(n_0)k}{N}} \sum_{n'=n_0-(n_2+1)}^{n_0-n_3} x_2(n') e^{j \frac{2\pi n' k}{N}} + \sum_{n=n_3+1}^{N-1} x_4(n) e^{-j \frac{2\pi n k}{N}}.\end{aligned}\quad (3.8)$$

Given that time-reversed sub-intervals are complex conjugates of the original, this manipulated sequence can be expressed as the frequency domain auxiliary signal

$$\hat{\Psi}_X(k) = \hat{X}_1(k) + \hat{X}_2^*(k) e^{-j \frac{2\pi n_0 k}{N}} + \hat{X}_3^*(k) e^{-j \frac{2\pi n_0 k}{N}} + \hat{X}_4(k). \quad (3.9)$$

Next, we can include the additional phase terms in the sub-intervals as Passerieux does in the time-domain, complex baseband version. In our frequency domain implementation, these would be additional complex exponential factors.

The process leading to equation 3.9 can be extended to include additional switched and time-reversed sub-intervals, serving to further obscure the original signal. Each case will add a component to the total DFT consisting of the complex conjugate of the original sub-interval DFT multiplied by a complex exponential of the form $e^{-j \frac{2\pi n_0 k}{N}}$ and by the additional complex exponential from the phase term. If n_0 is equal to N , the value of $e^{-j \frac{2\pi n_0 k}{N}}$ is 1. This corresponds to simply time-reversing the entire sequence.

To demonstrate Passerieux's algorithm for embedding in the frequency domain, assume we have a cover signal in the frequency domain given by

$$\hat{X}(k) = \hat{X}_1(k) + \hat{X}_2(k), \quad (3.10)$$

where $\hat{X}_1(k)$ contains the portion of the signal outside of bands to be used to embed data and $\hat{X}_2(k)$ contains the signal in the frequencies to have data embedded. The auxiliary signal, in the frequency domain, is then

$$\hat{\Psi}_X(k) = \hat{X}_2(k)^* e^{-j \beta_2}, \quad (3.11)$$

where β_2 represents the phase key for the portion of the signal with the embedded data frequencies. In Equation 3.11, the FFT symmetry is preserved such that the time domain representation of the stego signal is real. The process could then be extended to any number of individual frequency bands with different values of the complex exponential for each of them. The frequency domain version of Equation 3.1 to obtain the stego signal using the cover signal and auxiliary signal is therefore

$$\begin{aligned}\hat{Y}(k) &= \hat{X}(k) + \alpha a_m \hat{\Psi}_X(k) \\ &= \hat{X}_1(k) + \hat{X}_2(k) + \alpha \cdot a_m \cdot \hat{X}_2^*(k) e^{-j\beta_2},\end{aligned}\tag{3.12}$$

where α and a_m have the same meaning as they do in equation 3.1.

3.1.2 Extraction Theory

On the receiver end, we start with the output signal $y(n)$. Since the receiver will not know the original auxiliary signal, they must extract the embedded symbols by calculating the auxiliary signal of $\hat{Y}(k)$

$$\hat{\Psi}_Y(k) = \hat{X}_2^*(k) e^{-j\beta_2} + \alpha \cdot a_m \cdot \hat{X}_2(k).\tag{3.13}$$

Next, we take the complex conjugate of the auxiliary signal and multiply it by the received signal, yielding

$$\hat{Z}(k) = \hat{\Psi}_Y^*(k) \cdot \hat{Y}(k) = (\hat{X}_2(k) e^{j\beta_2} + \alpha a_m \hat{X}_2^*(k)) (\hat{X}_1(k) + \hat{X}_2(k) + \alpha \cdot a_m \cdot \hat{X}_2^*(k) e^{-j\beta_2}).\tag{3.14}$$

In Equation 3.13, the $\hat{X}_1(k)$ term can be ignored, since $\hat{X}_1(k)$ and $\hat{X}_2(k)$ have no frequencies in common. The remaining real terms in $\hat{Z}(k)$ together equal

$$2\alpha \cdot a_m \cdot |\hat{X}_2(k)|^2,\tag{3.15}$$

while the other remaining terms are complex. Assuming the amplitudes and phases of the frequency components in $\hat{X}_2(k)$ are random, summing the frequency components of $\hat{Z}(k)$ will make the remaining complex terms tend toward zero. The result of these operations leaves the real terms which will be positive or negative depending on the embedded symbol.

3.1.3 Transfer Function Theory

In this section, we will examine the scheme when the stego signal is transmitted to the receiver through an acoustic channel with transfer function $\hat{H}(k)$.

In the frequency domain, the received signal will be given by

$$\hat{S}(k) = \hat{Y}(k)\hat{H}(k), \quad (3.16)$$

which, expanding $\hat{Y}(k)$, can be expressed as

$$\begin{aligned} \hat{S}(k) &= (\hat{X}_1(k) + \hat{X}_2(k) + \alpha \cdot a_m \cdot \hat{X}_2^*(k)e^{-j\beta_2})\hat{H}(k) \\ &= \hat{X}_1(k)\hat{H}(k) + \hat{X}_2(k)\hat{H}(k) + \alpha \cdot a_m \cdot \hat{X}_2^*(k)e^{-j\beta_2}\hat{H}(k). \end{aligned} \quad (3.17)$$

Obtaining the auxiliary signal of the received signal $\hat{S}(k)$ yields

$$\hat{\Psi}_S(k) = \hat{X}_2^*(k)\hat{H}^*(k)e^{-j\beta_2} + \alpha \cdot a_m \cdot \hat{X}_2(k)\hat{H}^*(k), \quad (3.18)$$

the complex conjugate of which is

$$\hat{\Psi}_S^*(k) = \hat{X}_2(k)\hat{H}(k)e^{+j\beta_2} + \alpha \cdot a_m \cdot \hat{X}_2^*(k)\hat{H}(k). \quad (3.19)$$

Now, computing the product of the received signal with its respective auxiliary signal and dropping the terms containing $\hat{X}_1(k)$ gives

$$\begin{aligned} \hat{Z}(k) &= \hat{\Psi}_S^* \cdot \hat{S}(k) \\ &= (\hat{X}_2(k)\hat{H}(k)e^{+j\beta_2} + \alpha \cdot a_m \cdot \hat{X}_2^*(k)\hat{H}(k))(\hat{X}_1(k)\hat{H}(k) + \hat{X}_2(k)\hat{H}(k) \\ &\quad + \alpha \cdot a_m \cdot \hat{X}_2^*(k)e^{-j\beta_2}\hat{H}(k)) \\ &= (\hat{X}_2(k)e^{+j\beta_2} + \alpha \cdot a_m \cdot \hat{X}_2^*(k))(\hat{X}_2(k) + \alpha \cdot a_m \cdot \hat{X}_2^*(k)e^{-j\beta_2})\hat{H}(k). \end{aligned} \quad (3.20)$$

Unlike in Equation 3.14, however, none of the frequency components of this new $\hat{Z}(k)$ can be guaranteed to be purely real, since none of the terms contain the complex conjugate of the transfer function $\hat{H}(k)$. Additionally, due to the fact that the phase of the transfer function depends on frequency, the components which were previously real in the case of the transmit signal, $2\alpha \cdot a_m \cdot |\hat{X}_2(k)|^2$, will tend to cancel each other out. If the transfer

function is known, this problem can be overcome if it is divided out of the $\hat{Z}(k)$ term, though the transfer function is usually not known exactly.

As a compromise, we can attempt to undo the phase shift imposed by applying the transfer function in the time domain. We do this by continually adjusting the starting point of the received signal to find the optimal symbol recovery performance. This is equivalent to estimating the transfer function as a simple time advance, or $\hat{H}_e(k) = e^{j\frac{2\pi(k-1)f_s r}{\text{nfft} \cdot c}}$, where nfft is the number of FFTs, r is range in meters, and c is 1500 m/s. In a dispersion-free channel, this completely cancels the effect of the channel, since the transfer function for simple spherical spreading without dispersion is $\hat{H}(k) = \frac{1}{r}e^{-j\frac{2\pi(k-1)f_s r}{\text{nfft} \cdot c}}$. However, algorithm performance in both the time domain and the frequency domain will suffer as dispersion increases.

In Section 3.2, we present our implementation and method for experimentation, to include code structure, module interaction, and simulated acoustic channel setup.

3.2 Implementation

To demonstrate performance of Passerieux’s method in the frequency domain, we programmed a series of modules for embedding symbols, extracting symbols, and applying the transfer function generated from MMPE, as well as a main module to manage parameter selection and data collection. Each module was written in MATLAB R2017b due to its ease of matrix operations and useful built-in functions for signal processing [26]. MMPE generates a `.mat` file containing the transfer function to be used in conjunction with the other modules.

3.2.1 Modules

The three main modules perform the embedding, extraction, and transfer function operations, and work with one another to complete the simulated data exchange.

Embedding

The function `passerieuxembedfd.m` implements the operations described in Section 3.1.1. `passerieuxembedfd.m` takes `coverAudio`, `T`, `alpha` as inputs. `coverAudio` is assumed to be an audio file in uncompressed `.wav` format, from which only one channel

will be used for calculation. The T is the symbol period and α is the relative amplitude of the auxiliary signal. `passerieuxembedfd.m` reads the `coverAudio` sample frequency for use in the algorithm.

`passerieuxembedfd.m` outputs eight variables, x , y , M , N , $nfft$, a , `start_index`, and `stop_index`. x is the original sampled audio, with truncated extra points after division into even symbol periods. y is the stego audio with embedded symbols, where the number of embedded symbols depends on the chosen T . Though the algorithm is computed in the frequency domain, both x and y are time domain signals. M , N , and $nfft$ are matrix dimensions which must be saved for use in subsequent operations. a is a pseudo-random stream of symbols, which are always $+1$ or -1 in our BPSK scheme. `start_index`, and `stop_index` are index markers for the sub-intervals and guard bands. We heuristically choose $p = 4$, such that the total number of sub-intervals as described in Section 3.1 are eight, the sub-intervals are symmetric around the midpoint, are each one-third as long as the previous sub-interval, and with guard interval one-quarter length of the associated sub-interval. `passerieuxembedfd.m` also calculates additional indices to ensure no information is embedded in the guard bands. The phase key, β_p , is a simple, symmetric series of terms: $\frac{\pi}{4}, \frac{3\pi}{4}, -\frac{3\pi}{4}, -\frac{\pi}{4}$. β_p is hard-coded into `passerieuxembedfd.m`.

`passerieuxembedfd.m` always embeds symbols in a 2048 Hz band centered around a 2500 Hz center frequency. This places the embedded symbols in a band most frequently associated with low frequency sonar systems, which provide an adequate trade-off between transmission ranges and transmitter size and required power. In Passerieux's original paper, he uses a 1500 Hz band centered around 2250 Hz [1]. The 2048 band simplifies applying the MMPE transfer function, whose number of frequencies is always an integer power of two.

Extraction

`passerieuxextractfd.m` is the partner function to `passerieuxembedfd.m`, which implements the operations described in Section 3.1.2, and takes y , M , N , $nfft$, `start_index`, and `stop_index` as arguments. y is a time domain stego signal. M , N , and $nfft$ are matrix dimensions based on the length of y and the T second symbol period used in a given call to `passerieuxembedfd.m`. `start_index` and `stop_index` mark the relative sub-interval and guard band starting points based on the embedding process. For these

reasons, calls to `passerieuxextractfd.m` must follow calls to `passerieuxembedfd.m`. `passerieuxextractfd.m` outputs `z_sum`, which is an array representing the terms from equations 3.14 and 3.15. The sign of each element `z_sum`'s real term represent the recovered symbols.

For optimization, and to account for the phenomenon described in the final paragraph of Section 3.1.3, `passerieuxextractfd.m` iterates through possible values of estimated delay in order to find the highest portion of recovered symbols, corresponding to proper symbol period cross-correlation. In practice, one would have to find this by maximizing the absolute value of `z_sum`'s real portion.

Transfer Function

The function `mmpettransfer.m` applies the channel transfer function from MMPE to a stego signal in order to approximate a received signal, from which data extraction can be performed to test algorithm effectiveness after transmission through a simulated underwater channel. `mmpettransfer.m` takes `y`, `M`, `N`, `nfft`, `range`, `pressd`, and `SNR` as inputs. `range` and `pressd` are an array of range or depth values and the associated transfer function for a given frequency band, respectively. Both parameters are generated from MMPE post-processing, whose operation will be described in Section 3.2.2. `SNR` is a dB value that will correspond to noise that the transmitter must overcome in order to successfully recover symbols at the receiver. `mmpettransfer.m` outputs an expected received signal, `ys`, which represents a time-domain representation of the value in Equation 3.16.

To account for dispersion, `mmpettransfer.m` pads the symbol period matrix representation of `y` with zeroes, according to the value of `nfft`. After multiplying by the transfer function, which is the frequency-domain equivalent of convolution, overlapping segments of each successive are added coherently such that the reshaped signal represents the received signal with added dispersion length.

Since the MMPE transfer function only covers a 2048 Hz frequency band, the transfer function for other frequencies in the signal are set to approximate spherical spreading, with zero phase change. This provides a more intuitive received signal when listening to the audio or viewing the frequencies on a spectrogram.

3.2.2 MMPE

Accurate representation of an acoustic channel transfer function is critical to estimating algorithm performance. MMPE generates three binary files, `press.bin`, `avr.bin`, and `avz.bin`, each containing a small header detailing the important calculation information, and, more importantly the PE field function particle velocity component [27]. The post-processing routines `peout1.m` and `peout2.m` use these binary files and provide a variety of useful processing routines. Calculating these all-important binaries depends on the execution of a single executable, `mmpe2dbbv2_64`, which contains options for prioritizing accuracy or efficiency, and whose output depends on parameters specified in seven `.inp` files, which define the environment, source characteristics, and source position in 2D space.

Main Input

`pefiles.inp` is the main input file for MMPE. It specifies what the other inputs are, as well as the default output name for the three binary files. The line containing the parameters “`nzout`, `depmin[m]`, `depmax[m]`” specify the requested number of points, and the minimum and maximum depth to output, in meters. Line two has the “`nrout`, `rngmin[km]`, and `rngmax[km]`” fields which are self explanatory. The final line, containing “`nz`, `dr[km]`, `depcalc[m]`, and `c0[m/s]`” represents the vertical FFT size, range step in kilometers, the maximum calculation depth in meters, and reference sound speed in meters per second. `nz` must be an integer power of two. If any values are set to zero, their value will default to the value specified in their respective input files [27].

Environmental Data

`pessp.inp` defines the vertical SSP for the water column. The boolean value at the start of the file specifies whether or not water volume attenuation should be used, and should normally be set to 1. The number of SSPs and the individual points defining the profile itself constitute the majority of the file [27].

`pebath.inp` defines the ocean bottom lengths and locations. The user can specify multiple depths to simulate shelves, sloping bottoms, or a simple flat bottom [27].

`pebotprop.inp` defines the properties of the ocean floor itself. These properties are range dependent, but not depth dependent. The user must specify the number of bathymetry points,

as well as commensurate floating point numbers for range , sound speed, sound speed gradient, density, compression attenuation, shear speed, and shear attenuation. `pebotprop.inp` essentially defines the bottom type and its physical characteristics [27].

`pedbath.inp` defines the deep layer bathymetry beneath that defined by `pebotprop.inp`, in absolute depth from the surface. An excessively deep bathymetry value will not be used if it exceeds maximum computational depth. Shallower depths will replace the upper layer and act as a outcropping for its length value. This file is formatted in the same manner as `pebath.inp` [27].

`pedbotprop.inp` is the partner file to `pedbath.inp`, containing the deep layer's acoustic properties. It has identical format to `pebotprop.inp` [27].

Source Data

`pesrc.inp` defines all the the source node's characteristics, and can be configured for a point source or a vertical hydrophone source. This is based on the value set for array length. To set the source as a point source, the user set must set length to zero. If a point source is selected, the value for "D/E angle" also defaults to zero. `pesrc.inp` also has fields for center frequency in Hz, frequency bandwidth in Hz, and number of frequencies, which must be a power of two. `pesrc.inp` also defines the source depth, in meters. Each field in `pesrc.inp` is a floating point number, with the exception of number of frequencies. Great care must be taken to select a sufficient number of frequencies, as too few can result in travel time wrap around, and the frequency domain equivalent of incorrect phase terms.

Post-Processing

After calculating the MMPE binary files, one can run the MMPE post-processing files `peout1.m` and `peout2.m` to obtaining meaningful data. `peout1.m` prompts the user for the name of the binary input file and initializes the routine by reading in the header information, and giving the file ID for subsequent processing. `peout2.m` gives several options for processing the data. Running `peout.m` will prompt the user to select from one of six options:

1. Output starting field data
2. Compute data for single radial

3. Compute data for single range
4. Compute data for single depth
5. Compute data for single interface
6. Compute travel time data

Selecting any of these options will prompt the user to enter the value at which to calculate the output, such as the frequency, range or depth. `peout2.m` will output dB TL plots for the particle velocity components, as well as pressure, and will prompt the user to save the data after review. The “save” option will retain the complex quantities not explicitly displayed by the TL plots. When saving the particle velocity files, the amplitudes are scaled appropriately relative to pressure.

Figure 3.1 details the overall scheme, including major inputs, outputs, and modules.

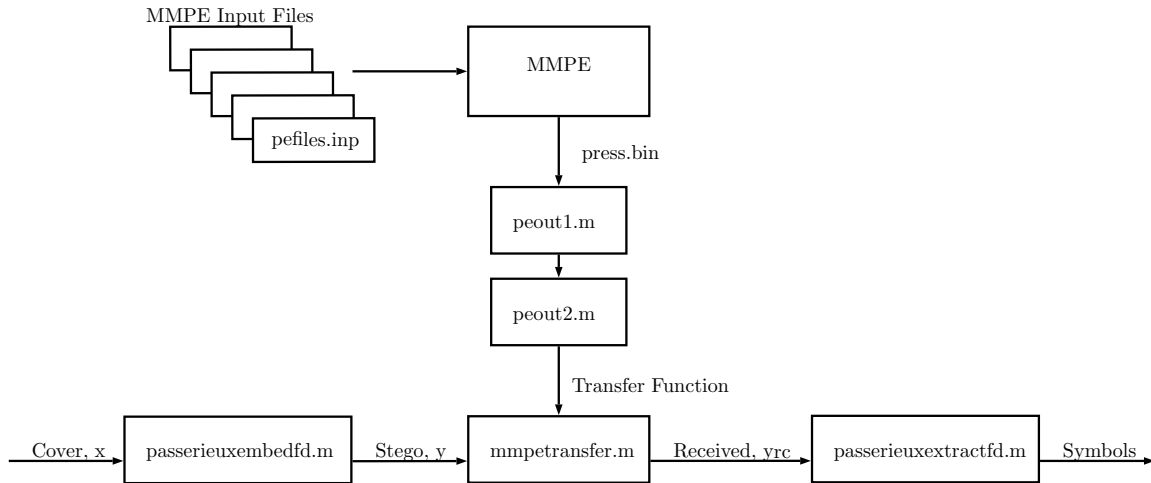


Figure 3.1. Workflow and Module Interaction

3.3 Summary

In this chapter, we explained the theory of our frequency domain steganography implementation, including how data is embedded and how it can be extracted. We also explained how a transfer function approximating the frequency response of an acoustic channel would theoretically impact our stego signal, as well as how to mitigate negative effects of applying such a transfer function. Next, we explained the modules we developed to implement these

approaches in MATLAB, using digital audio as our cover signal and an MMPE generated transfer function to modify our stego signal. We concluded with a description of how MMPE inputs define the environment and source, as well as how the output can be manipulated to display a variety of useful acoustic data. In Chapter 4, we define parameters for testing, construct MMPE channel models, and present the results of experimentation. Further, we evaluate the effectiveness and performance of our scheme.

THIS PAGE INTENTIONALLY LEFT BLANK

CHAPTER 4:

Results and Analysis

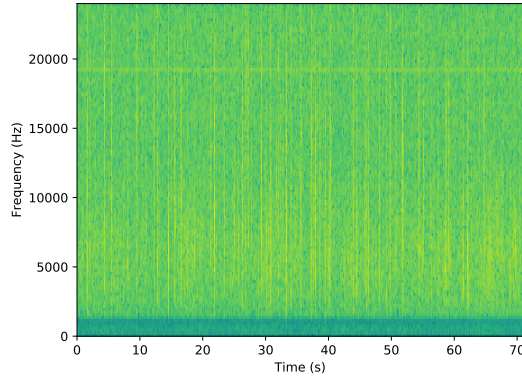
In Chapter 3, we discussed the theory supporting our frequency domain implementation of Passerieux’s steganography by cover modification technique, as well as the tools by which we will measure the theoretical performance of that technique. In this section, we present the results of our simulations and experiments.

4.1 Steganographic Evaluation

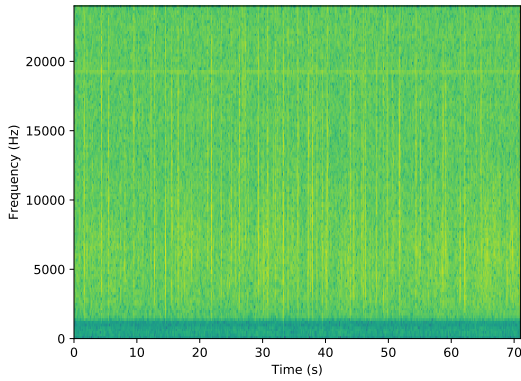
We begin by performing the embedding and extraction algorithms on a given cover audio signal and show theoretical perfect bit recovery absent losses from the transmission medium or noise. For this embedding procedure, we use one channel of a 48 kHz sampled, uncompressed .wav file, `SpermWhaleNormalClicks.wav` obtained from the National Oceanic and Atmospheric Administration (NOAA) Southwest Fisheries Science Center catalog of cetacean sounds [28]. We use this cover audio file as the basis for all further experimentation as well. The embedding frequencies are a 2048 Hz bandwidth centered on 2500 Hz. Additionally, we use nominal parameters from Passerieux’s original paper, namely $T=0.50$ seconds for a 2 bps bitrate and gain $\alpha=0.3$ corresponding to -10 dB [1]. Next, we measure the effectiveness of the embedding scheme by traditional steganography metrics MSE, RMSE, PSNR, and SNR.

4.1.1 Spectrogram and Constellation

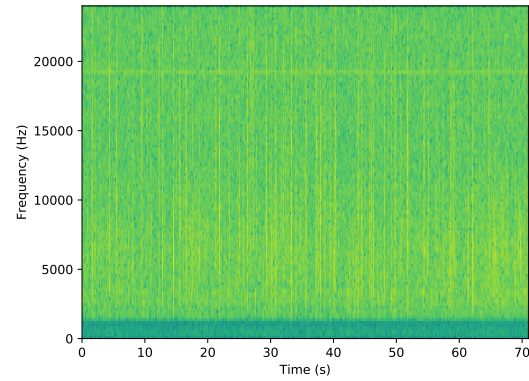
We first examine the impact of embedding on the original signal, with the goal being to preserve its characteristics. Viewing the spectrogram of both the cover and stego audio reveals little difference in the plots, even for higher values of α . This would be difficult for an operator to distinguish on a time-frequency display, let alone after transmission through an underwater acoustic channel. Figure 4.1 shows a comparison of the spectrum for the full 71-second length of the audio signal, as well as two example stego signals, each with 142 embedded bits.



(a) Spectrum of Cover Audio



(b) Spectrum with Embedded Bits, $\alpha = 0.32$



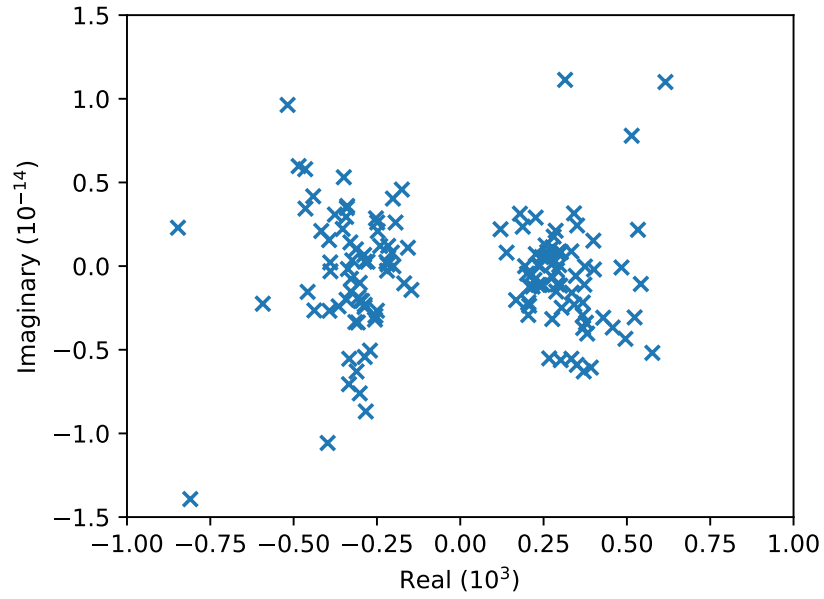
(c) Spectrum with Embedded Bits, $\alpha = 0.89$

Figure 4.1. Comparison of Spectra Between Cover and Stego Audio Signals

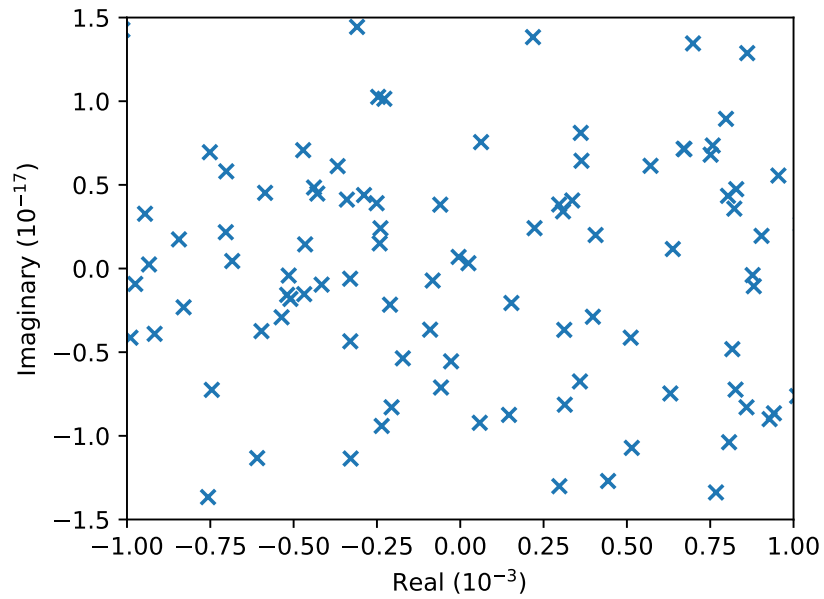
The user can also listen to the stego audio to qualitatively assess whether there are obvious artifacts or tones. Increasing the gain introduces greater signal distortions due to the additional information added by the auxiliary signal term. Likewise, a gain of zero reproduces the cover signal as the stego signal.

A visual representation of symbol recovery can be obtained by plotting the complex values which make up the demodulated values in the `z_sum` array resulting from the extraction algorithm. This is similar in concept to a typical constellation diagram for a BPSK signal transmission. Here, the sign of the real component corresponds to the estimated symbol for a symbol period, with 1 or -1 corresponding to 1 or 0, respectively. A clear separation of points on a scatter plot represents ideal symbol recovery. Figure 4.2 displays two

polar scatter plots representing the constellation for an extremely successful and extremely unsuccessful symbol recovery.



(a) Perfect Symbol Recovery with $T = 0.5$ and $\alpha = 0.3$



(b) Ambiguous Symbol Recovery with $T = 0.5$, $\alpha = 0.3$, and No Estimated Delay

Figure 4.2. Comparison of Estimated Symbols Constellations.

Noting the difference in scale, Figure 4.2 demonstrates the effect of maximizing this separation. Higher gain values for α increase the magnitude of these terms. Not properly accounting for the estimated delay as described in Section 3.2.1 will obscure these points. As performance degrades, the constellation converges increasingly toward the center and giving an expected 50% symbol recovery.

4.1.2 Steganography Measures

The impact of embedding algorithm on the signal can be further evaluated by calculating several of the well known quantitative steganography measures. To calculate these results we perform the embedding procedure using $\text{Alpha} = 0.11, 0.16, 0.22, 0.32, 0.45, 0.63, \text{ and } 0.89$. These amplitude ratios correspond roughly with -19 to -1 dB in 3 dB increments. We perform embedding using these values for $T = 0.50, 0.25, \text{ and } 0.10$ corresponding to 2, 4, and 10 bps respectively. Values of T any lower than 0.10 begin to incur MATLAB errors due to insufficient precision and calculation resolution. Each point represents the average for that value over five runs, after performing the embedding procedure after reseeding the random number generator for the random embedded bits.

MSE and RMSE

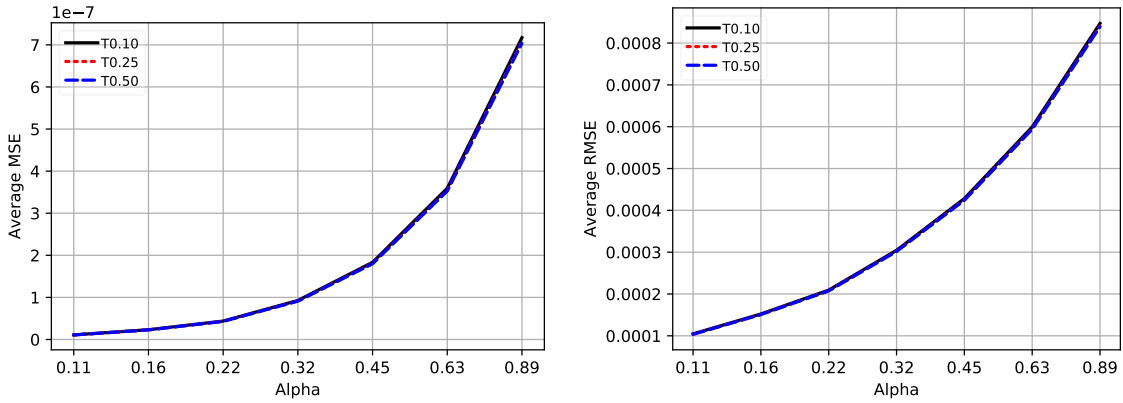
The MSE is an average measure of the energy of embedding changes [9]. The MSE is given by

$$\text{MSE} = \frac{1}{n} \sum_{i=1}^n |x[i] - y[i]|^2, \quad (4.1)$$

where x and y are the cover and stego audio in sampled points and n is the total number of sampled points. The RMSE is given by

$$\text{RMSE} = \sqrt{\text{MSE}}. \quad (4.2)$$

and represents the standard deviation of the stego signal from the cover signal. Figure 4.3 displays the results of the MSE and RMSE calculations.



(a) Average MSE vs α

(b) Average RMSE vs α

Figure 4.3. Comparison of Average MSE and RMSE for Differing T

In Figure 4.3 we see that MSE and RMSE are very close for each value of T, but grow exponentially for successive values of α . Overall, these measures are extremely small even for larger values of α .

PSNR and SNR

The PSNR and SNR are two other measures of relative distortion [9]. PSNR is given by

$$\text{PSNR} = 10 \log_{10} \frac{x_{\max}^2}{\text{MSE}}, \quad (4.3)$$

where $x_{\max}^2$ is the max value that an item in x can achieve. Since the MATLAB `audioread` function returns normalized values between 1.0 and -1.0 of type double, $x_{\max}^2$ is simply 1 [26].

The SNR in this case is in terms of cover audio to stego audio, with the “noise” representing the modifications brought about during the embedding process. It is given by

$$\text{SNR} = 10 \log_{10} \frac{\sum_{i=1}^n |x[i]|^2}{\text{MSE}}. \quad (4.4)$$

In general, SNR above 30 dB means that the audio quality in a stego signal has been perfectly preserved, while those with SNR below 20 dB are considered “noisy” signals [14].

We calculate PSNR and SNR using the same embedding parameters and averaging described at the beginning of Section 4.1.2. Figure 4.4 displays the results of the PSNR and SNR calculations.

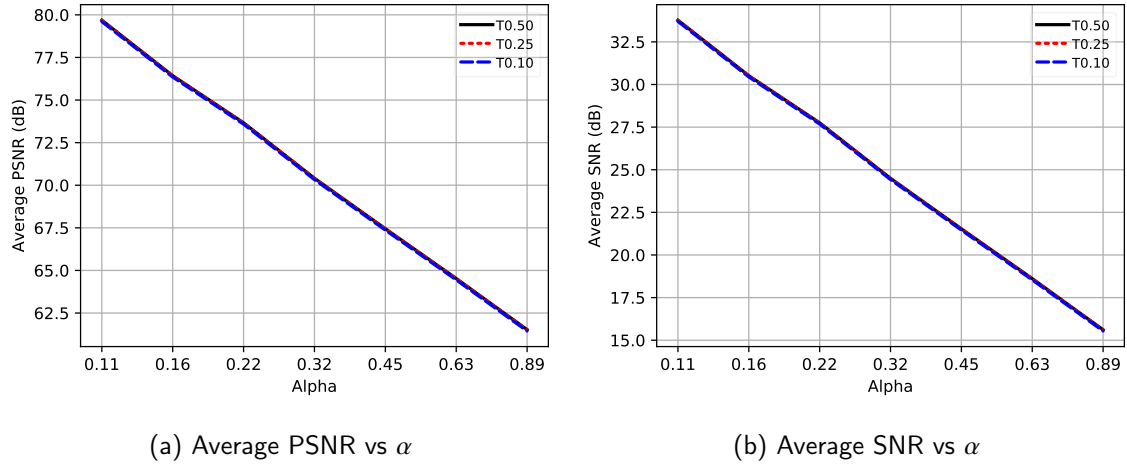


Figure 4.4. Comparison of Average PSNR and SNR for Differing T

As with MSE and RMSE, adjusting the symbol period has little effect on the resulting PSNR and SNR, which both decrease linearly with α . The audio quality is technically only preserved for α values less than 0.16, or -16 dB or less. However, the average SNR remains above that which would be considered noisy until α is approximately 0.50, including the $\alpha = 0.3$ value used in the original paper [1].

4.2 Channel Simulation

To simulate the impact of transmission through an underwater acoustic channel on our ability to extract the embedded symbols, we generate two models via the MMPE. The first represents a simple, open-ocean channel with a deep bottom to minimize possible multipath effects and demonstrate the effect of spherical spreading. The second corresponds to a fairly challenging shallow water channel used for at-sea experimentation by Song and Hodgkiss [29]. Both channels were generated using the “efficiency” option in the MMPE executable.

It is important to note that selection of simulation parameters may result in exceeding the

limits of the model, which can limit the angles of elevation and depression, and, in extreme cases, cause errors during executable runtime. Without always having a clear warning of this, simulation parameters must be chosen by trial-and-error to ensure sufficient resolution for ranges, depths, and frequencies. To ensure adequate frequency resolution, we examine the the travel time output at max possible range for each channel. After running the MATLAB post-processing routines on the calculated binaries, we select `peout2.m` option ‘6’ and output over the entire water column, to ensure no wrap-around. An example of travel time wrap-around is given in figure 4.5.

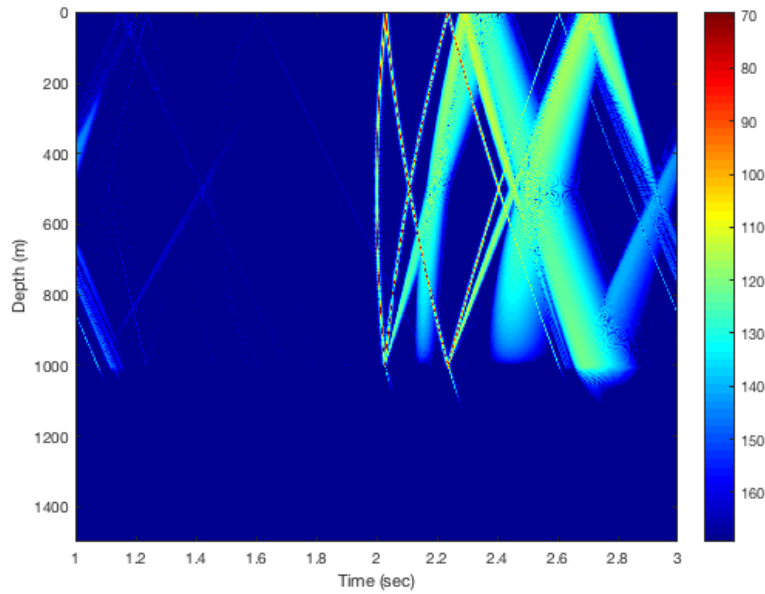
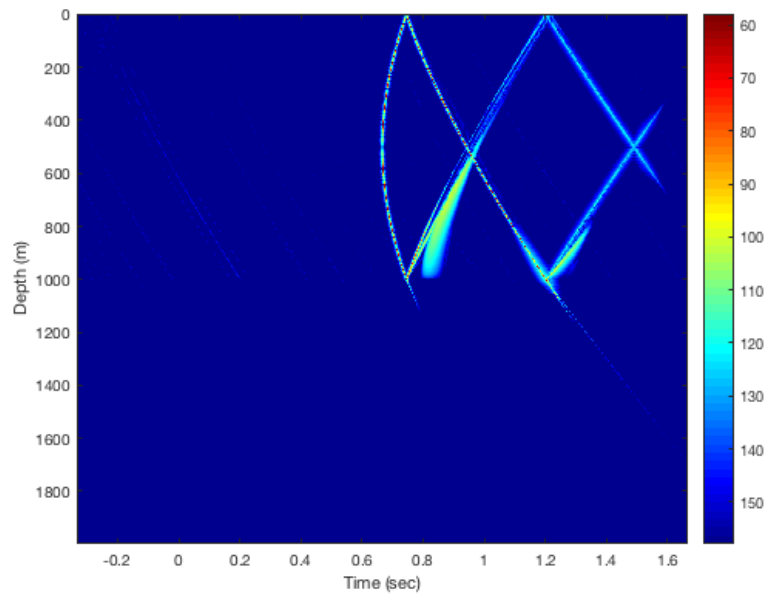
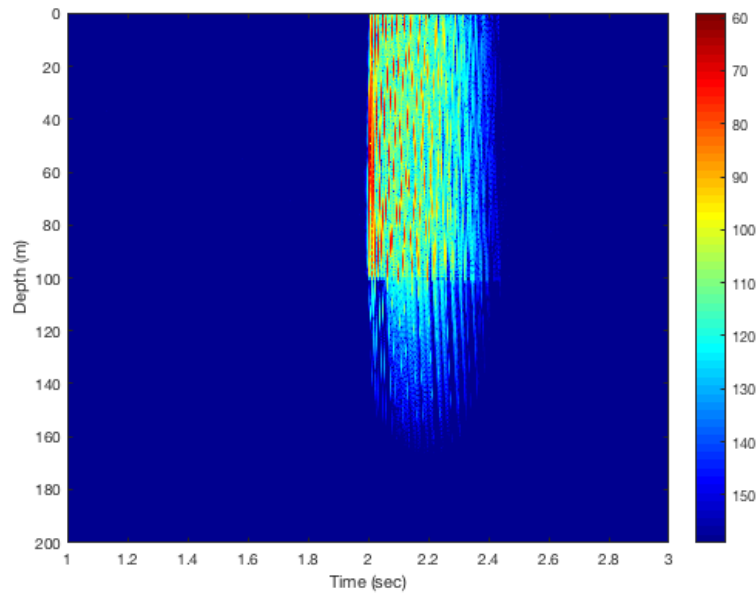


Figure 4.5. Travel Time Wrap-Around with 4096 Frequencies at 3km

Increasing the number of frequencies calculated by the next higher power of two may eliminate this phenomenon. This will result in drastically longer calculation times and much larger binary output files. For the purpose of these experiments, the number of frequencies was limited to 4096. Figure 4.6 displays the travel time calculations for the two simulated channels.



(a) Travel Time for the Deep Channel, 4096 Frequencies, 1km.



(b) Travel Time for the Shallow Channel, 4096 Frequencies, 3km.

Figure 4.6. Travel Time Calculations for the Two Channels.

4.2.1 Deep Channel Setup

The deep channel simulates a generic, open-ocean channel. Its parameters are described in Table 4.1.

Table 4.1. Deep Channel MMPE Parameters

Input File	Parameter	Values
pefiles.inp	nzout, depmin[m], depmax[m]	500, 0, 2000
	nrout, rngmin[km], rngmax[km]	500, 0, 1.0
	nz, dr[km], depcalc[m], c0[m/s]	0, 0.0, 0.0, 1500.0
pessp.inp	Water volume attenuation	1
	Number of SSPs in first radial	1
	range[m], Number SSPs in depth	0.0, 501
pebath.inp	Number of bathymetry points	1
	range[km] and depth[m]	0.0, 1000
pebotprop.inp	Number of range points	1
	range[km], sound speed[m/s], gradient[1/s], density [g/cm ³], compression attenuation[dB/m/kHz], shear speed[m/s], shear attenuation[dB/m/kHz]	0.0, 1700.0, 0.0, 1.8, 0.25 0.0, 0.0
pesrc.inp	Source depth[m]	500
	Array length[m]	0.0
	D/E angle [deg]	0.0
	Center frequency [Hz]	2500
	Frequency bandwidth [Hz]	2048
	Number of frequencies	4096

The deep bathymetry files `pebath.inp` and `pebotprop.inp` are left at their default values. Figure 4.7 shows the SSP for the deep channel, as defined in `pessp.inp`. Figure 4.8 displays the TL through depth and range of the entire channel for the 2500 Hz center frequency.

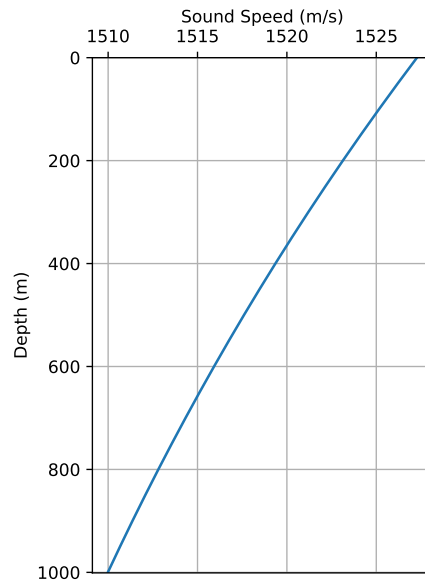


Figure 4.7. SSP for the Deep Channel Model

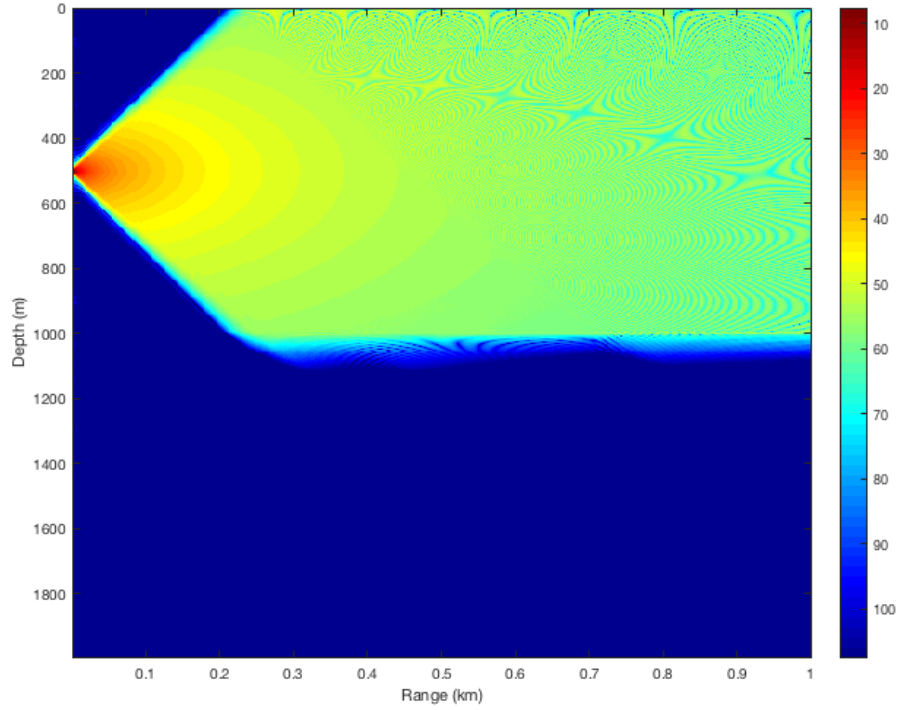


Figure 4.8. Deep Channel TL for Pressure vs. Depth and Range (dB re 1m)

4.2.2 Shallow Channel Setup

The shallow channel simulates Song and Hodgkiss's challenging shallow water environment [29]. The shallow water channel is meant to demonstrate performance under stronger multipath conditions, but has the added benefit of being calculated out to three kilometers in range. The parameters for generating the shallow channel are described in Table 4.2.

Table 4.2. Shallow Channel MMPE Parameters

Input File	Parameter	Values
pefiles.inp	nzout, depmin[m], depmax[m]	500, 0, 200
	nrout, rngmin[km], rngmax[km]	500, 0, 3.0
	nz, dr[km], depcalc[m], c0[m/s]	0, 0.0, 0.0, 1500.0
pessp.inp	Water volume attenuation	1
	Number of SSPs in first radial	1
	range[m], Number SSPs in depth	0.0, 9
pebath.inp	Number of bathymetry points	1
	range[km] and depth[m]	0.0, 100.0
pebotprop.inp	Number of range points	1
	range[km], sound speed[m/s], gradient[1/s], density [g/cm ³], compression attenuation[dB/m/kHz], shear speed[m/s], shear attenuation[dB/m/kHz]	0.0, 1700.0, 0.0, 1.8, 0.25 0.0, 0.0
pessp.inp	Source depth[m]	60
	Array length[m]	0.0
	D/E angle [deg]	0.0
	Center frequency [Hz]	2500
	Frequency bandwidth [Hz]	2048
	Number of frequencies	4096

As with the deep channel, the deep bathymetry files `pebath.inp` and `pebotprop.inp` are left at their default values. Figure 4.9 shows the SSP for the shallow channel as defined in `pessp.inp`. This less generic SSP is representative of a profile measured in the Monterey Bay. Figure 4.10 displays the TL through depth and range of the entire channel for the 2500 Hz center frequency.

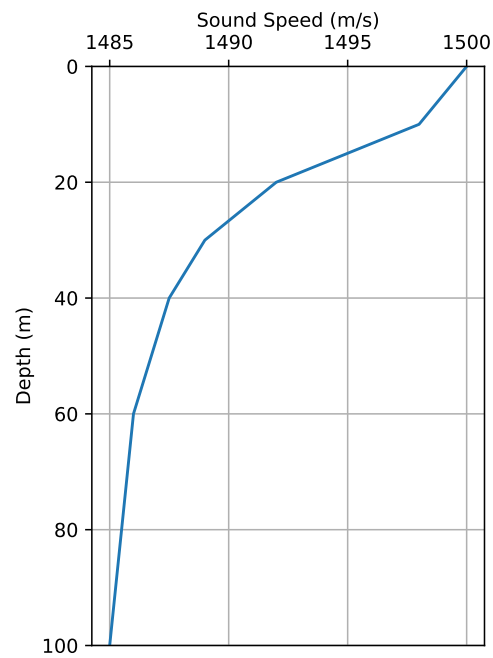


Figure 4.9. SSP for the Shallow Channel Model

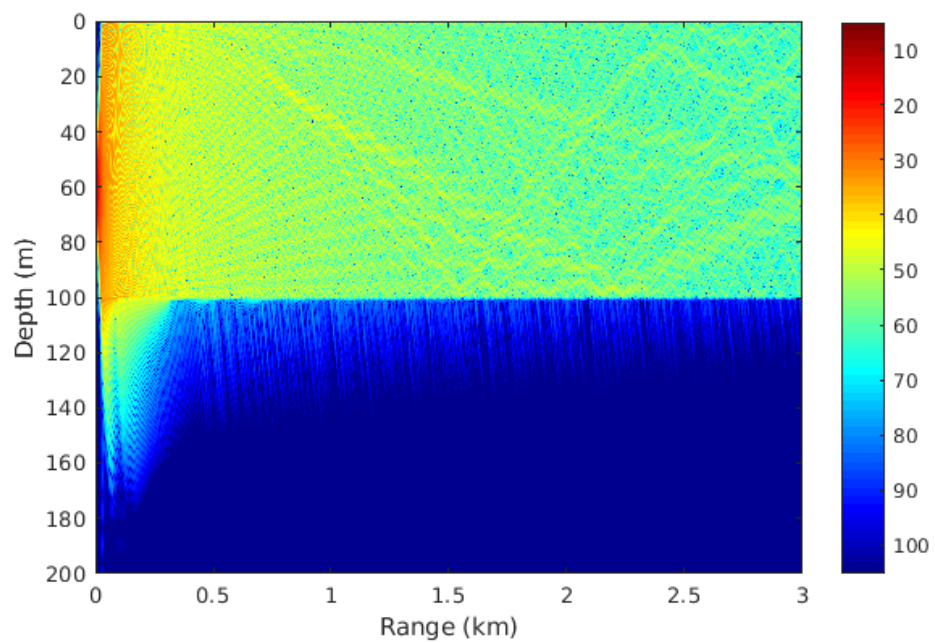


Figure 4.10. Shallow Channel TL for Pressure vs. Depth and Range (dB re 1m)

4.3 Deep Channel Experiment

To measure algorithm performance in the deep channel, we first obtain the transfer function for our 2048 Hz frequency band. We initialize the environment by running the post-processing routine `peout1.m`, and when prompted, specify the deep channel binary generated by MMPE. Next, we run `peout2.m`, selecting “Compute data for single depth,” choosing 500 meters, such that the receiver depth matches the source depth. This outputs a `.mat` data file containing several matrices, one of which is `pressd`. `pressd` is a 500 by 4096 matrix of complex values representing the transfer function for every 0.5 Hz value in our frequency band, across every range bin from zero to our maximum range. This transfer function is applied to the transmit signal by the `mmpetransfer` module. Figure 4.11 displays the calculated transfer function’s phase angle at maximum range over each of the frequency bins.

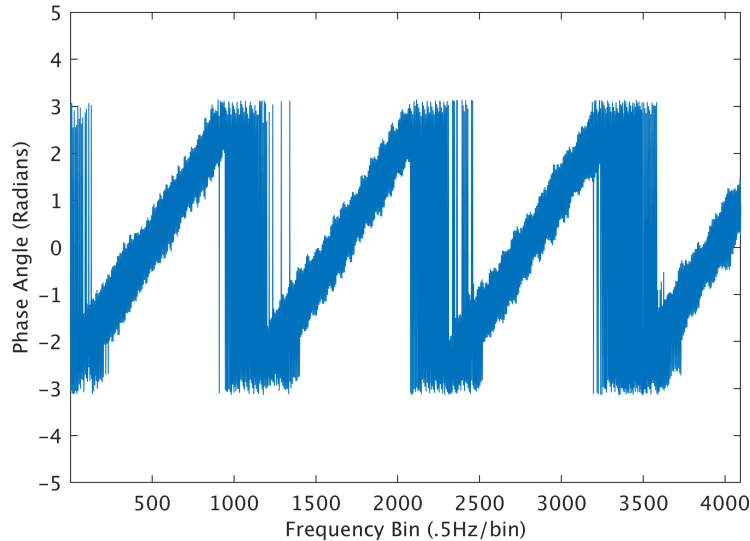


Figure 4.11. Transfer Function Phase Angle at 1km vs. Frequency Bin

To demonstrate the impact of the channel on our ability to recover symbols, we perform the embedding procedure, simulate the received signal based on the channel characteristics, and then apply extraction with varying T and α values identical to those described in Section 4.1. Furthermore, we add white Gaussian noise with standard deviation equal to the transmit signal’s to simulate noise at the receiver side, accounting for the effect over distance such that the receiver sees lower SNR. In reality, this noise would be spectrum dependent. To

achieve a receiver side SNR of 0 dB, we initially set the SNR to 60 dB, such that the nominal TL loss for spherical spreading offsets this value. We also set the SNR to 40 dB to compare the results for more noise and an overall lower SNR at the receiver. We perform each simulation in 20 meter increments from zero to 1000 meters. Table 4.3 summarizes the experiment's simulation parameters.

Table 4.3. Deep Channel Experiment Parameters

Parameter	Values
Symbol Period (T)	0.5s, 0.25s, 0.10s
Auxiliary Signal Gain (α)	0.11, 0.16, 0.22, 0.32, 0.45, 0.63, 0.89
SNR	60 dB, 40 dB
Range	20m to 1000m, in 20m Increments

Bit percentages at each range are calculated by finding the estimated delay value over the first 100 bins which yielded the highest bit percentage. Beyond that value, bit percentages tend to oscillate around 50%. Figure 4.12 demonstrates the importance of correctly finding the estimated delay to symbol recovery. Maximizing the spread of the real terms in z_sum is functionally equivalent to finding the correct estimated delay. Figure 4.13 displays the results of each experiment.

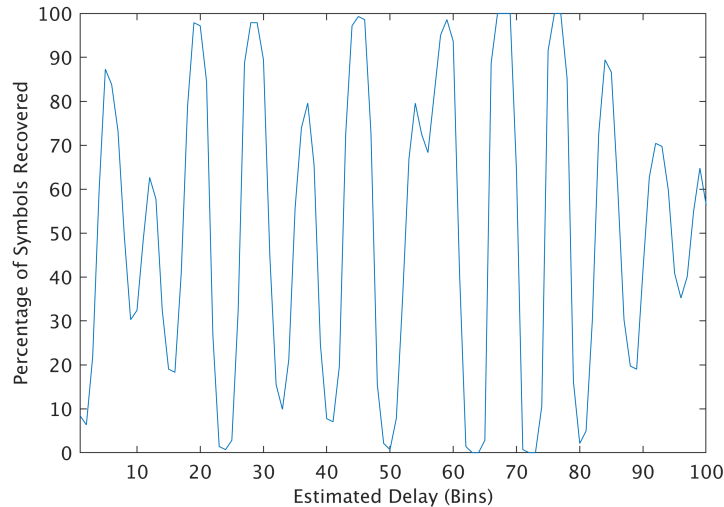
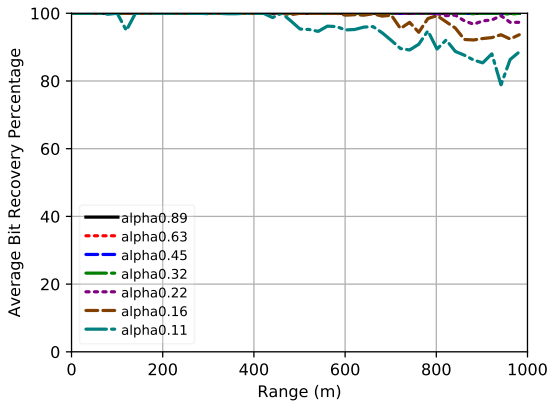
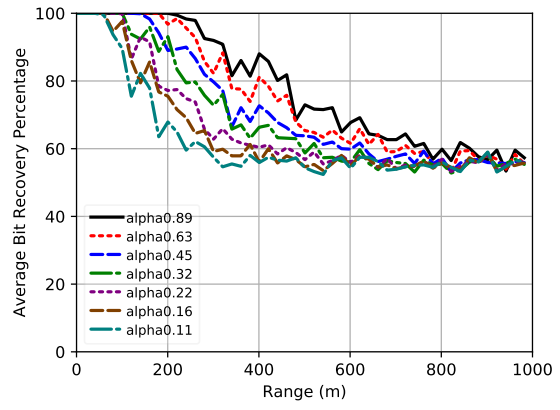


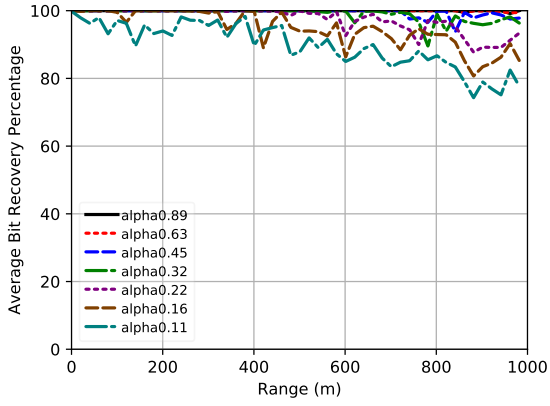
Figure 4.12. Example of Recovered Symbols vs. Estimated Delay over 100 Bins



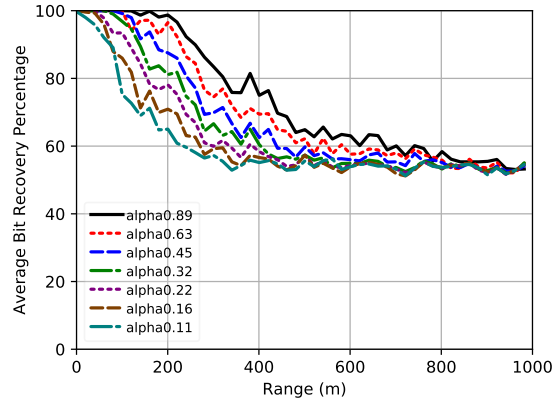
(a) $T = 0.5s$, $SNR = 60dB$



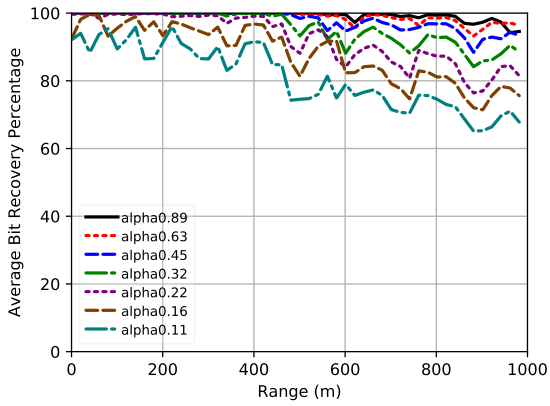
(b) $T = 0.5s$, $SNR = 40dB$



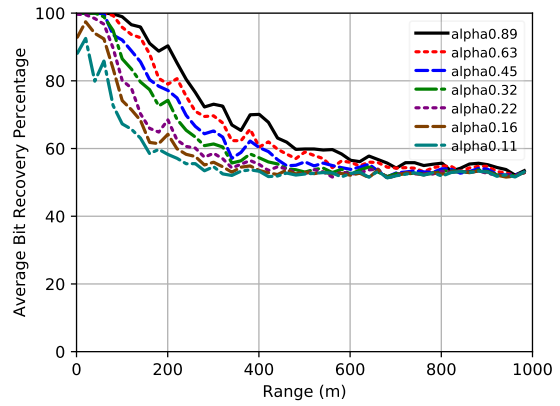
(c) $T = 0.25s$, $SNR = 60dB$



(d) $T = 0.25s$, $SNR = 40dB$



(e) $T = 0.10s$, $SNR = 60dB$



(f) $T = 0.10s$, $SNR = 40dB$

Figure 4.13. Symbol Recovery Percentage vs. Range for the Deep Channel

Predictably, overall performance suffers for lower α values, representing the trade-off between imperceptibility and robustness. Lowering the symbol period, though effectively increasing the bitrate, also hinders performance, likely due to the data points being more profoundly impacted by dispersion. Noise represents the largest contribution to performance degradation, with values approaching the worst-case 50% mark in each of the 40dB scenarios. The erratic percentages from point to point may be best explained by imprecision in finding estimated delay, as well as the effects of multipath arrival, each of which will impact the other. In every scenario, performance suffered over range, especially for low α values.

4.4 Shallow Channel Experiment

The shallow water experiment examines algorithm performance under more challenging conditions. The experiment is conducted using the same methods and parameters described in Section 4.3, but with a 60 meter source and receiver depth. In this experiment, multipath effects are more prominent. Figure 4.14 displays the transfer function phase angle at 3000 meters for each of the 4096 frequency bins.

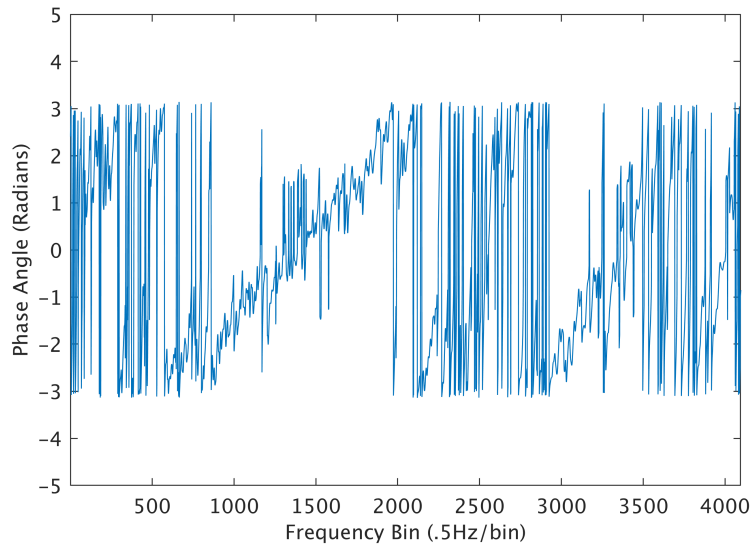
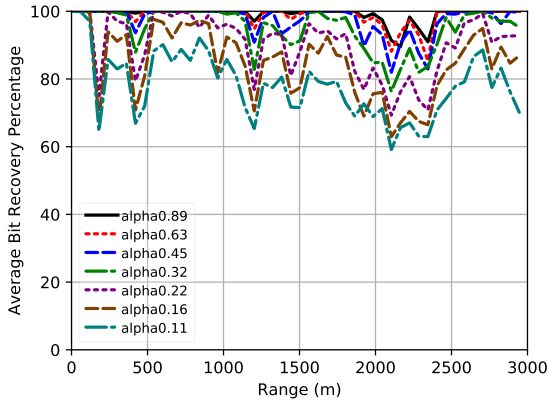
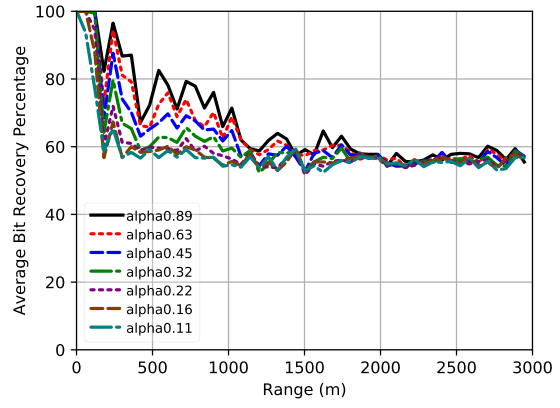


Figure 4.14. Transfer Function Phase Angle at 3km vs. Frequency Bin

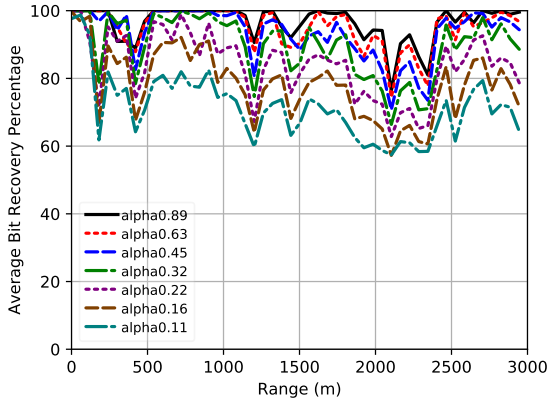
The results of the shallow water experiment are presented in Figure 4.15



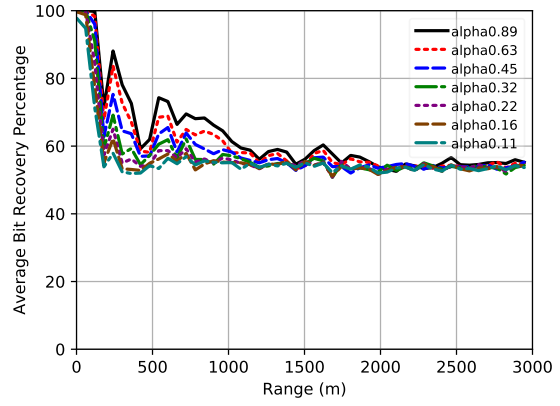
(a) $T = 0.5s$, $SNR = 60dB$



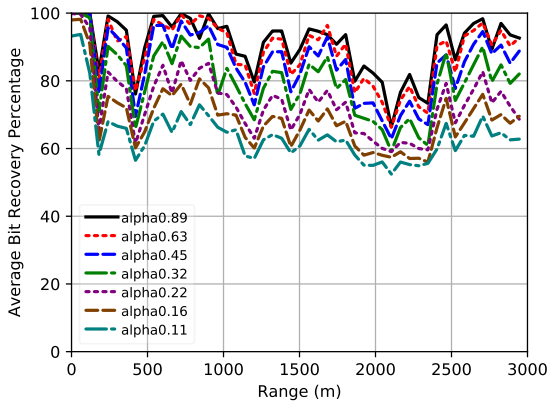
(b) $T = 0.5s$, $SNR = 40dB$



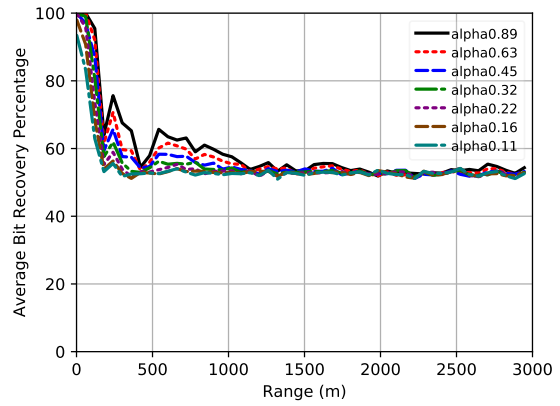
(c) $T = 0.25s$, $SNR = 60dB$



(d) $T = 0.25s$, $SNR = 40dB$



(e) $T = 0.10s$, $SNR = 60dB$



(f) $T = 0.10s$, $SNR = 40dB$

Figure 4.15. Symbol Recovery Percentage vs. Range for the Shallow Channel

The results of the shallow channel experiment are similar to the deep channel experiment with some notable exceptions. Overall performance is worse than in the deep water scenario, even for high α values and longer symbol periods. For each plot, bit recovery percentage shows several notable localized dips and spikes, shown across each T and α value. This may be explained by the effects of constructive or destructive interference due to the surface reflections, bottom bounces, and direct paths expected in such an environment. Noise is still the most significant factor in worsening performance, with symbol recovery percentages for the 40 dB scenarios dropping off rapidly after several hundred meters.

4.5 Summary

In this Chapter, we explained the factors used to evaluate our scheme, including steganographic distortion and audio quality measurements, demonstrating the extent to which our embedding algorithm preserves the characteristics of the cover audio signal. Next, we described how we set up our MMPE defined channels and their associated transfer functions, as well as how we mitigated potential simulation problems. We concluded the chapter with a discussion of our experimental results and their possible explanations. In Chapter 5, we will summarize our work, present our findings, and detail areas for further research.

CHAPTER 5:

Conclusions and Future Work

In this thesis, we have discussed UAC and steganography at length, to include how these concepts may be linked in order to provide an innovative tactical capability to military platforms, UUVs, and underwater acoustic modems. We have examined the extent of present research using steganography for covert UAC and further explored Passerieux's steganography by cover modification technique using recorded ambient noise.

This thesis presented new work based on Passerieux's method and takes a modified approach to embed and extract hidden symbols in the frequency domain. We explained the theory supporting our implementation and demonstrate symbol recovery. Then, we evaluated the impact of embedding via several common steganography metrics, and examined the impact of simulated ocean channels and noise upon our ability to recover information.

5.1 Key Findings

It is possible to implement Passerieux's method in the frequency domain and recover symbols hidden in the phase, and recover those symbols at short ranges. The α values best for minimizing distortion, which are anything less than about 0.5, also show high success rates in both channel scenarios. Higher bit rates, though negligible in terms of stego signal distortion, negatively impact symbol recovery. There is an observable trade-off between robustness and imperceptibility.

Multipath arrival and background noise are the most likely contributors to performance degradation. If this scheme were to be employed in a practical scenario, determining the appropriate SL to transmit the stego signal which balances the goal of inconspicuous steganographic communication with reliable message recovery would be paramount.

5.2 Recommendations for Future Work

An important goal of this research was to lay the foundation for future work, and identify areas that can be improved or expanded upon. At present, we recommend the following as

focus areas for future work:

- Implementing spectrum-dependent noise to better understand its effects on performance.
- Developing guidance for choosing appropriate transmission SL and algorithm parameters as a function of channel characteristics, ambient noise, desired robustness, and desired covertness.
- Improving understanding of limits to performance by simulating the effects of transmission through an acoustic channel with greater accuracy, such as the MMPE “accuracy” option.
- Developing a more realistic method for estimating delay/symbol period alignment.
- An evaluation of security via steganalysis, information theory, and general resistance to active and passive warden scenarios.
- Overall algorithm and scheme improvement, focusing on robustness, imperceptibility, and capacity.

Other areas for recommended future research include:

- Performing at-sea testing in real time.
- Experimenting among a greater variety of simulated channels, to include those that are representative of specific locations.
- Evaluating performance using other existing models or modeling techniques.
- Developing a method to resist Doppler effects, or deal with a moving transmitter and/or receiver.
- Implementing and measuring the effects of different modulation techniques, such as Quadrature Phase Shift Keying (QPSK) and Quadrature Amplitude Modulation (QAM).
- Experimenting among a variety of cover signals with different properties.
- Experimenting with different bandwidths and center frequencies.
- Determining optimal sub-interval and guard band sizes.
- Developing a method for receiver synchronization.
- Developing higher layer protocols to allow interface with existing equipment.
- Developing software modules for operator usability.
- Implementing and testing a complete, receiver/transmitter stegosystem.

List of References

- [1] J. M. Passerieux, “Stealth underwater acoustic communications based upon steganography techniques,” in *2nd International Conference and Exhibition on Underwater Acoustics*, 2014.
- [2] J. M. Passerieux, “Method and system for acoustic communication,” May 25 2017, US Patent App. 15/319,703.
- [3] R. J. Urick, *Principles of Underwater Sound*, 3rd ed. Los Altos, CA: McGraw-Hill, 1983.
- [4] L. E. Kinsler, A. R. Frey, A. B. Coppens, and J. V. Sanders, *Fundamentals of Acoustics Fourth Edition*. Hoboken, NJ: John Wiley & Sons, Inc., 1999.
- [5] G. M. Wenz, “Acoustic ambient noise in the ocean: Spectra and sources,” *The Journal of the Acoustical Society of America*, vol. 34, no. 12, pp. 1936–1956, 1962. Available: <https://doi.org/10.1121/1.1909155>
- [6] R. Su, R. Venkatesan, and C. Li, “A review of channel modeling techniques for underwater acoustic communications,” *Eighteenth Annual Newfoundland Electrical and Computer Engineering Conference*, 2010.
- [7] K. B. Smith, “Convergence, stability, and variability of shallow water acoustic predictions using a split-step Fourier parabolic equation model,” *Journal of Computational Acoustics*, vol. 9, no. 01, pp. 243–285, 2001.
- [8] A. Tolstoy, K. Smith, and N. Maltsev, “The SWAM’99 workshop - An overview,” *Journal of Computational Acoustics*, vol. 09, no. 01, pp. 1–16, mar 2001.
- [9] J. Fridrich, *Steganography in Digital Media: Principles, Algorithms, and Applications*. Cambridge, UK: Cambridge University Press, 2010.
- [10] G. J. Simmons, *The Prisoners’ Problem and the Subliminal Channel*. Boston, MA: Springer US, 1984, pp. 51–67. Available: https://doi.org/10.1007/978-1-4684-4730-9_5
- [11] A. Kerckhoffs, “La cryptographie militaire,” *Journal des Sciences Militaires*, 1883.
- [12] R. Radhakrishnan, M. Kharrazi, and N. Memon, “Data masking: A new approach for steganography?” *Journal of VLSI Signal Processing Systems for Signal, Image and Video Technology*, vol. 41, no. 3, pp. 293–303, Nov 2005. Available: <https://doi.org/10.1007/s11265-005-4153-1>

- [13] W. Bender, D. Gruhl, N. Morimoto, and A. Lu, "Techniques for data hiding," *IBM Systems Journal*, vol. 35, no. 3.4, pp. 313–336, 1996. Available: <https://ieeexplore.ieee.org/document/5387237/>
- [14] F. Djebbar, B. Ayad, K. A. Meraim, and H. Hamam, "Comparative study of digital audio steganography techniques," *EURASIP Journal on Audio, Speech, and Music Processing*, vol. 2012, no. 1, p. 25, Oct 2012. Available: <https://doi.org/10.1186/1687-4722-2012-25>
- [15] N. Cvejic and T. Seppanen, "Increasing the capacity of lsb-based audio steganography," in *Multimedia Signal Processing, 2002 IEEE Workshop on*. IEEE, 2002, pp. 336–338.
- [16] N. Cvejic and T. Seppanen, "Increasing robustness of LSB audio steganography using a novel embedding method," in *International Conference on Information Technology: Coding and Computing, 2004. Proceedings. ITCC 2004.*, April 2004, vol. 2, pp. 533–537 Vol.2.
- [17] D. Kahn, "Cryptology and the origins of spread spectrum: Engineers during World War II developed an unbreakable scrambler to guarantee secure communications between Allied leaders; actress Hedy Lamarr played a role in the technology," *IEEE Spectrum*, vol. 21, no. 9, pp. 70–80, Sept 1984.
- [18] L. Gang, A. N. Akansu, and M. Ramkumar, "MP3 resistant oblivious steganography," in *2001 IEEE International Conference on Acoustics, Speech, and Signal Processing. Proceedings (Cat. No.01CH37221)*, 2001, vol. 3, pp. 1365–1368 vol.3.
- [19] F. Djebbar, B. Ayad, K. Abed-Meraim, and H. Hamam, "Unified phase and magnitude speech spectra data hiding algorithm," *Security and Communication Networks*, vol. 6, no. 8, pp. 961–971, 2013. Available: <http://dx.doi.org/10.1002/sec.644>
- [20] G. Leus and P. A. van Walree, "Multiband OFDM for covert acoustic communications," *IEEE Journal on Selected Areas in Communications*, vol. 26, no. 9, pp. 1662–1673, December 2008.
- [21] S. Liu, G. Qiao, A. Ismail, B. Liu, and L. Zhang, "Covert underwater acoustic communication using whale noise masking on DSSS signal," in *OCEANS-Bergen, 2013 MTS/IEEE*. IEEE, 2013, pp. 1–6.
- [22] J. D. Park and J. F. Doherty, "A steganographic approach for covert waveform design," in *MILCOM 2016 - 2016 IEEE Military Communications Conference*, Nov 2016, pp. 180–185.
- [23] C. Cachin, "An information-theoretic model for steganography," in *Information Hiding*. Springer, 1998, pp. 306–318.

- [24] H. Dol, B. Quesson, and F. Benders, “Covert underwater communication with marine mammal sounds,” *Undersea Defence Technology-UDT Europe 2008, 10-12 June 2008, Glasgow, UK*, 2008.
- [25] F. X. Socheleau, C. Laot, and J. M. Passerieux, “Concise derivation of scattering function from channel entropy maximization,” *IEEE Transactions on Communications*, vol. 58, no. 11, pp. 3098–3103, November 2010.
- [26] The MathWorks Inc., *MATLAB R2017b*. Available: <https://www.mathworks.com/>
- [27] *MMPE Introduction*, U.S. Office of Naval Research Ocean Acoustics Library, 2006. Available: <http://oalib.hlsresearch.com/PE/MMPE/mmpeintro.html>
- [28] NOAA, “Cetacean sounds,” 2014. Available: <https://swfsc.noaa.gov/textblock.aspx?Division=PRD&ParentMenuId=148&id=5776>
- [29] H. Song and W. Hodgkiss, “Efficient use of bandwidth for underwater acoustic communication,” *The Journal of the Acoustical Society of America*, vol. 134, no. 2, pp. 905–908, 2013.

THIS PAGE INTENTIONALLY LEFT BLANK

Initial Distribution List

1. Defense Technical Information Center
Ft. Belvoir, Virginia
2. Dudley Knox Library
Naval Postgraduate School
Monterey, California