Physically Constrained Maximum Likelihood (PCML) Mode Filtering and Its Application as a Pre-Processing Method for Underwater Acoustic Communication

by

Joseph C. Papp

B.S., Worcester Polytechnic Institute, 2006

Submitted to the Department of Electrical Engineering and Computer Science &
Department of Applied Ocean Science and Engineering
in partial fulfillment of the requirements for the degree of
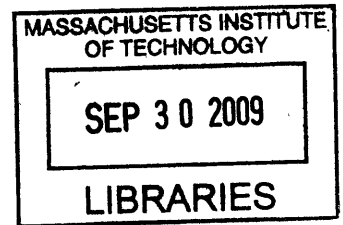
Master of Science

at the

MASSACHUSETTS INSTITUTE OF TECHNOLOGY

and the

WOODS HOLE OCEANOGRAPHIC INSTITUTION

September 2009

© Joseph C. Papp, MMIX. All rights reserved.

ARCHIVES

Author . . . .
Dept. of Electrical Engineering and Computer Science
& Dept. of Applied Ocean Science and Engineering
August 28, 2009

Certified by . . . .
Dr. James C. Preisig
Associate Scientist with Tenure
Thesis Supervisor

Accepted by . . . .
Dr. James C. Preisig
Chair, Joint Committee for Applied Ocean Science and Engineering
Massachusetts Institute of Technology/Woods Hole Oceanographic Institution

Accepted by . . . .
Dr. Terry P. Orlando
Chair, EECS Department Committee on Graduate Students
Massachusetts Institute of Technology

Physically Constrained Maximum Likelihood (PCML) Mode Filtering and Its Application as a Pre-Processing Method for Underwater Acoustic Communication

by

Joseph C. Papp

Submitted to the Department of Electrical Engineering and Computer Science &
Department of Applied Ocean Science and Engineering
on August 28, 2009, in partial fulfillment of the
requirements for the degree of
Master of Science

Abstract

Mode filtering is most commonly implemented using the sampled mode shape or pseudoinverse algorithms. Buck et al [1] placed these techniques in the context of a broader maximum a posteriori (MAP) framework. However, the MAP algorithm requires that the signal and noise statistics be known a priori. Adaptive array processing algorithms are candidates for improving performance without the need for a priori signal and noise statistics. A variant of the physically constrained, maximum likelihood (PCML) algorithm [2] is developed for mode filtering that achieves the same performance as the MAP mode filter yet does not need a priori knowledge of the signal and noise statistics. The central innovation of this adaptive mode filter is that the received signal's sample covariance matrix, as estimated by the algorithm, is constrained to be that which can be physically realized given a modal propagation model and an appropriate noise model.

The first simulation presented in this thesis models the acoustic pressure field as a complex Gaussian random vector and compares the performance of the pseudoinverse, reduced rank pseudoinverse, sampled mode shape, PCML minimum power distortionless response (MPDR), PCML-MAP, and MAP mode filters. The PCML-MAP filter performs as well as the MAP filter without the need for a priori data statistics. The PCML-MPDR filter performs nearly as well as the MAP filter as well, and avoids a sawtooth pattern that occurs with the reduced rank pseudoinverse filter. The second simulation presented models the underwater environment and broadband communication setup of the Shallow Water 2006 (SW06) experiment. Data processing results are presented from the Shallow Water 2006 experiment, showing the reduced sensitivity of the PCML-MPDR filter to white noise compared with the reduced rank pseudoinverse filter. Lastly, a linear, decision-directed, RLS equalizer is used to combine the response of several modes and its performance is compared with an equalizer applied directly to the data received on each hydrophone.

Thesis Supervisor: Dr. James C. Preisig
Title: Associate Scientist with Tenure

.

Acknowledgments

Contents

List of Figures

11

Chapter 1

Introduction

1.1 Motivation

In underwater acoustic communication, the delay spread of the acoustic channel is one of the primary obstacles to reliable, high-rate communication. However, the delay spread of a given mode of propagation is significantly less than the delay spread of the overall acoustic channel. This thesis proposes mode filtering as a pre-processing method to reduce the delay spread of the communication channel and to reduce the number of independent channels that need to be equalized. Having a shorter impulse response reduces the inter-symbol interference encountered between symbols and having fewer independent channels means there are less taps that need to be adjusted in an equalizer. A smaller number of taps means fewer snapshots are required to estimate the data statistics, which allows an equalizer to more quickly adapt to changing channel conditions. The end goal is to communicate faster with lower bit error rates.

There are both adaptive and non-adaptive methods to estimate the signal propagating in a particular acoustic mode. Many of the adaptive methods, such as a filter based on the minimum power distortionless response beamformer, require knowledge of the spatial covariance matrix of the array. The true value of this matrix is not known in practice, and the sample covariance matrix is often used instead. However, a large number of snapshots are required for this matrix to converge close to its true

value. A physically constrained, maximum likelihood method, originally developed by Kraay [4], is proposed to take advantage of the known structure of the vertical pressure field and accurately estimate the covariance matrix with significantly fewer snapshots. Mode filtering reduces the number of independent channels that need to be equalized since the number of energetic modes is usually less than the number of available hydrophones.

The experimental setup analyzed in this thesis is one transmitter and a fixed vertical array of receivers located 19 km away in shallow water (80 meters depth). Low frequency sound signals centered at 203 Hz were used. One application of the processing techniques described in this thesis is autonomous underwater vehicles (AUVs) that want to communicate information back to a central base station as they navigate an area. In this application, the sound source is mounted on the AUV and the receive array would be deployed in a central location. The general application of mode filtering is much more widespread, however, as it can also be used to process signals transmitted across the entire ocean, such as during the Acoustic Thermometry of Ocean Climate experiment [5]. Mode processing is attractive in these applications because the signal energy is trapped in the ocean acoustic waveguide and can travel long distances without bottom or surface interactions.

1.2 Objectives

1. Apply a physically constrained, maximum likelihood method for covariance matrix estimation to adaptive mode filters.

2. Demonstrate broadband mode filtering as a pre-processing method for underwater acoustic communication.

3. Evaluate the performance of an adaptive equalizer to combine the response of several modes after mode filtering.

14

1.3 Organization

Chapter 2 will present the concept of acoustic normal modes and how narrowband mode filters are used to estimate them. Chapter 2 also describes the physically constrained, maximum-likelihood method for covariance estimation developed by Kraay [2]. Chapter 3 presents results from two simulations and results from the Shallow Water 2006 experiment, comparing several mode filtering methods. Chapter 4 evaluates the performance of a linear, decision-directed RLS equalizer both on the outputs of the mode filters and on each hydrophone directly. Chapter 5 describes the conclusions of this work and directions for future research.

Chapter 2

Background

This chapter provides an introduction to the underwater acoustic environment, including representing the vertical pressure field as a sum of normal modes. The chapter describes several narrowband mode filtering techniques. Lastly, a physically constrained, maximum likelihood method for estimating the spatial covariance matrix is presented.

2.1 Normal Modes and Signal Model

Modes are physically motivated, orthogonal basis functions for the vertical sound pressure field. They are derived from solutions to the wave equation and are dependent on frequency and environmental conditions, such as water depth, temperature, salinity, and bottom properties [6]. Modes are only dependent on the local properties of the waveguide, however, and not on the environment between the transmitter and receiver. Equation 2.1 shows the pressure field written as a sum of modes, where $\mathbf{p}(z, f)$ is the complex acoustic pressure at frequency f and depth z, d_m is the mode coefficient of the mth mode, and $\psi_m(z, f)$ is the mth mode shape.

$$p(z, f) = \sum_m d_m(f)\psi_m(z, f) \tag{2.1}$$

Figure 2-1 shows the first four modes of an isovelocity channel that has a pressure release surface and a hard bottom. In this case, the modes are pure sinusoids.

Group velocity is the rate at which signal energy propagates in the horizontal
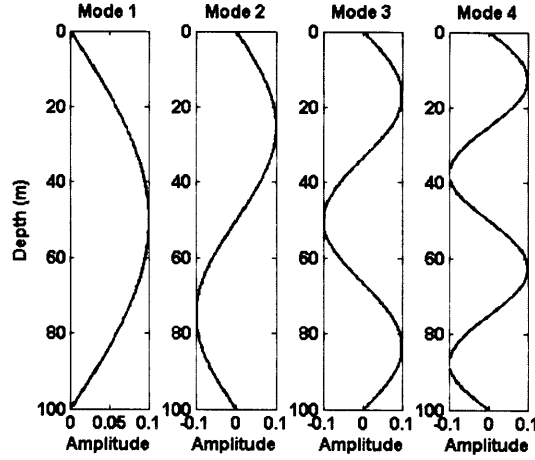
Figure 2-1: First four modes of an isovelocity waveguide

direction, and is defined in general as $c_g = \frac{d\omega}{dk_r}$, where ω is the frequency in radians per second and k_r is the horizontal wavenumber. Each mode has a particular group velocity associated with it at a particular frequency and can be calculated analytically for an isovelocity waveguide. Equations 2.2 and 2.3 express the group velocity of a particular mode in an isovelocity waveguide as function of frequency. The phase velocity is the rate at which the phase of a wave propagates and is defined as $c_p = \frac{\omega}{k_r}$, where k_r is the horizontal wavenumber [6]. In the following equations, c_0 is the medium sound speed, m is the mode number, D is the depth of the waveguide, f is the frequency, and k_r is the horizontal wavenumber.

$$k_r = \sqrt{(2\pi f/c_0)^2 - (m\pi/D)^2} \qquad (2.2)$$

$$c_g = \frac{c_0 k_r}{\sqrt{k_r^2 + (m\pi/D)^2}} = \frac{c_0^2 k_r}{2\pi f} \qquad (2.3)$$

Figure 2-2 shows the group and phase velocities for the first six modes of an isovelocity waveguide, with mode 1 at the center and mode 6 at the top and bottom. The lower order modes have the fastest group velocities, and the higher order modes have

18

slower group velocities. Therefore, the lower order modes will arrive first in this environment. The lower modes have a smaller change in group velocity over a given frequency band, which means they have the shortest dispersion. For each mode, group velocity decreases with decreasing frequency, until it reaches a cutoff frequency where its group velocity becomes zero. For frequencies below this point, the mode is said to be evanescent and its amplitude decays exponentially with range. Figures 2-3 and 2-4 show the effect of the varying group velocity on the modal arrival times for a 20 km waveguide.

Modal Group and Phase Velocity, Modes 1-6
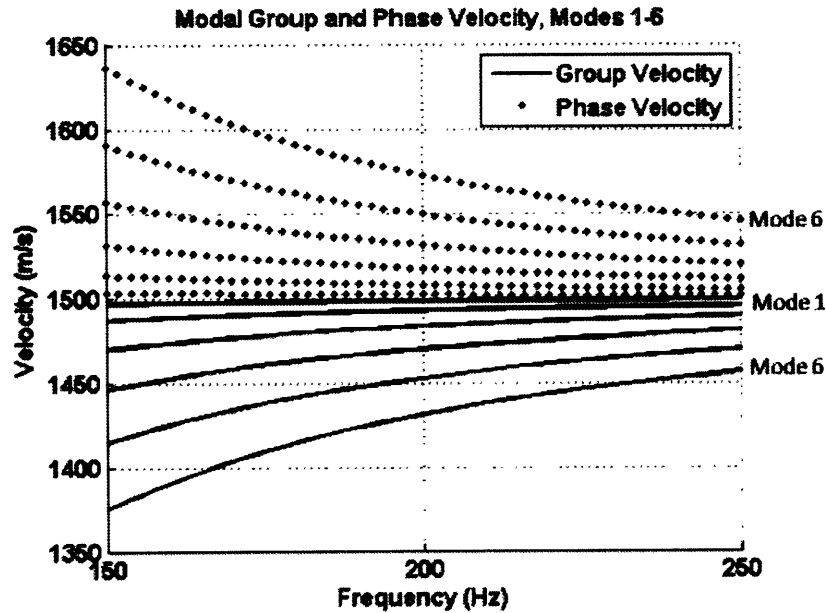
Figure 2-2: Group and phase velocity curves for an isovelocity waveguide showing the frequency dependence. The curves for mode 1 are near the center and mode 6 is on the top and bottom

19
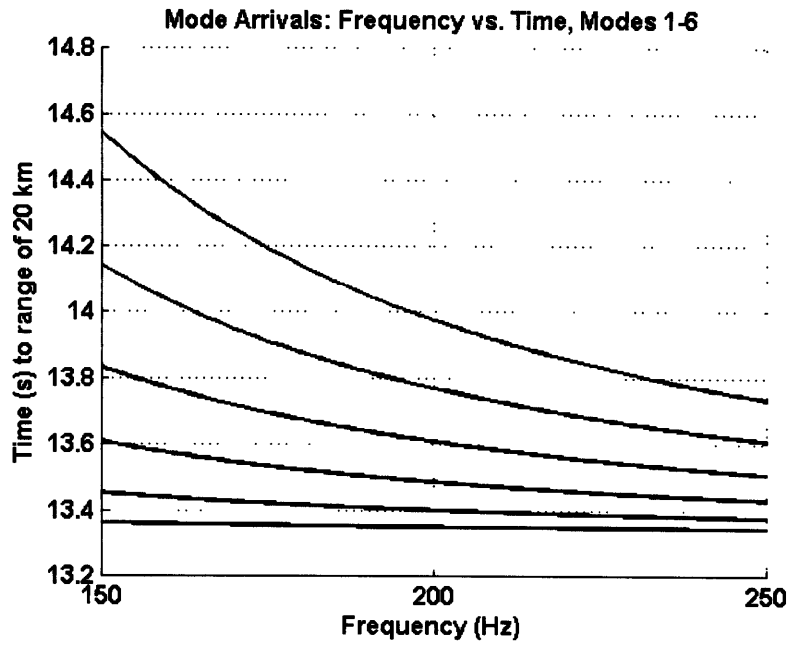
Figure 2-3: Arrival times for a 20 km transmission path for the first six modes of an isovelocity waveguide. Mode 1 is at the bottom and mode 6 is at the top.
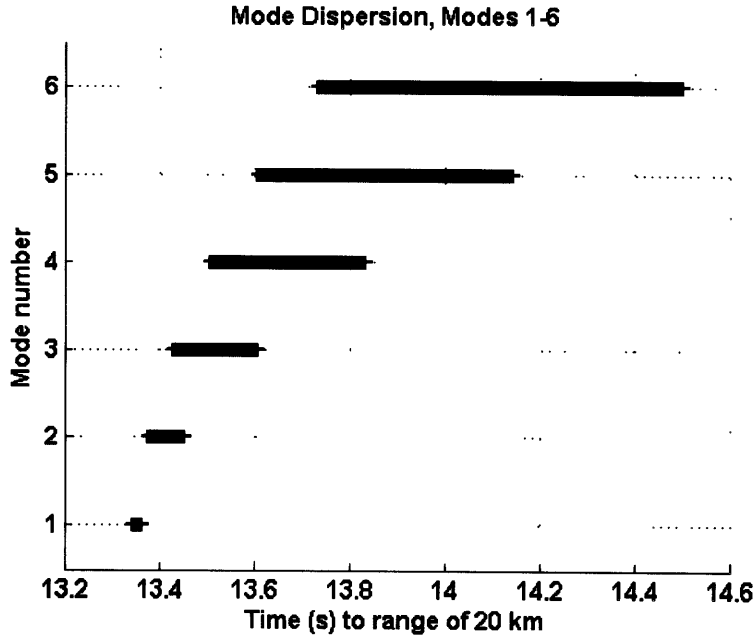
Figure 2-4: Modal dispersion for the first six modes over a 20 km transmission path for an isovelocity waveguide.

20

Equation 2.4 shows a narrowband signal model used to represent the vertical sound pressure field,

$$\mathbf{p} = \boldsymbol{\psi}\mathbf{d} + \mathbf{n} \tag{2.4}$$

where $\mathbf{d}$ is a vector of complex valued mode amplitudes and $\mathbf{n}$ is noise. Both are assumed to be zero-mean. $\boldsymbol{\psi}$ is the sampled mode shapes matrix,

$$\boldsymbol{\psi} = \begin{bmatrix} \psi_1(z_1) & \cdots & \psi_M(z_1) \\ \vdots & \ddots & \vdots \\ \psi_1(z_N) & \cdots & \psi_M(z_N) \end{bmatrix} \tag{2.5}$$

Two noise models are considered in this thesis, spatially white (SW) and Kuperman-Ingenito (KI). Both of these noise models assume the noise is complex and zero mean, and therefore the noise correlation matrix is the same as the noise covariance matrix. The SW noise model assumes the noise at each hydrophone is uncorrelated with the noise at all of the other hydrophones. The covariance matrix for this noise model is $\mathbf{R_n} = \sigma^2\mathbf{I}$. The KI noise model assumes the noise is contained in each mode independently with varying powers, with covariance matrix

$$\mathbf{R_n} = \boldsymbol{\psi} \begin{bmatrix} P_1^2 & 0 & \cdots & 0 \\ 0 & P_2^2 & \ddots & \vdots \\ \vdots & \vdots & \ddots & 0 \\ 0 & \cdots & 0 & P_M^2 \end{bmatrix} \boldsymbol{\psi}^H \tag{2.6}$$

where $P_1^2...P_M^2$ are the powers of each mode in the noise model.

2.2 Narrowband Mode Filters

This section describes several narrowband mode filters and drops the notational dependence of mode shapes on frequency, as this will be implied. Mode filters estimate the complex amplitude of a particular mode at a particular frequency given the pressure field measured with a vertical array of hydrophones. Equation 2.7 expresses the

estimated mode amplitudes as a function of the received vertical pressure field and a linear mode filter, $\mathbf{H}$.

$$\hat{\mathbf{d}} = \mathbf{Hp} \qquad (2.7)$$

The remaining parts of this section will discuss the choice of the filter $\mathbf{H}$ to use in estimating the mode amplitudes.

2.2.1 Sampled Mode Shape (SMS) Mode Filter

The sampled mode shape mode filter is the projection of the vertical pressure field onto each of the sampled mode shapes. As such, it is spatially matched to each mode and is optimal for detecting a single mode in spatially white noise [1, 5]. However, the sampled mode shapes are not orthogonal and the filter suffers from leakage between modes.

$$\mathbf{H}_{SMS} = \begin{bmatrix} \psi_1(z_1) & \cdots & \psi_1(z_N) \\ \vdots & \ddots & \vdots \\ \psi_M(z_1) & \cdots & \psi_M(z_N) \end{bmatrix} = \psi^H \qquad (2.8)$$

2.2.2 Pseudoinverse (PI) Mode Filter

The pseudoinverse mode filter results from taking a pseudoinverse of the sampled mode shape matrix and is the filter that results from choosing $\hat{\mathbf{d}}$ to minimize the squared error between $\psi\hat{\mathbf{d}}$ and $\mathbf{p}$ [1, 7].

$$\mathbf{H}_{PI} = (\psi^H\psi)^{-1}\psi^H \qquad (2.9)$$

The row that filters for a particular mode projects the received array data onto the subspace that is orthogonal to all the other modes. The PI filter is optimal in rejecting interfering modes, but is sensitive to spatially white noise. An alternative that is commonly used is a reduced rank pseudoinversion based on the singular value decomposition. Let

$$\psi = \mathbf{USV}^T \qquad (2.10)$$

22

be the singular value decomposition of ψ, where $\mathbf{S}$ is a diagonal matrix consisting of the singular values of ψ and $\mathbf{U}$ and $\mathbf{V}$ are orthogonal matrices representing the eigenvectors of $\psi\psi^T$ and $\psi^T\psi$ respectively [8].

The pseudoinverse of ψ can be written as

$$\mathbf{H}_{reducedrank} = \psi^+ = \mathbf{V}\mathbf{S}^+\mathbf{U}^T \tag{2.11}$$

where $\mathbf{S}^+$ is the pseudoinverse of the diagonal matrix $\mathbf{S}$, where inverses of the singular values less than some threshold (1/100 of the maximum singular value in our case) are set to zero [8]. This mode filter will be referred to as the reduced rank pseudoinverse filter in the rest of this thesis.

2.2.3 Minimum Power Distortionless Response (MPDR) Mode Filter

The MPDR mode filter minimizes the filter output power subject to the constraint that the desired mode is passed with a gain equal to one. It is based on the MPDR beamformer used in spatial array processing, with a substitution of the mode shape vector in place of the spatial steering vector [9, 10, 11]. The goal is to minimize contributions from interfering signals and noise while preserving the desired mode. The derivation of the filter is as follows. Let the received pressure field be $\mathbf{p}$ and the desired mode be n. $\mathbf{w}_n$ is the weight vector that filters for mode n, with the output of the mode filter being $\mathbf{w}_n^H\mathbf{p}$. The goal of the MPDR filter is to

$$\underset{\mathbf{w}_n}{\text{minimize}} \; \|\mathbf{w}_n^H\mathbf{p}\|^2 = \mathbf{w}_n^H\mathbf{R_p}\mathbf{w}_n \; \text{ such that } \; \mathbf{w}_n^H\psi_n = 1 \tag{2.12}$$

The derivation uses Lagrange multipliers [12, 13]. Let $H(w)$ be the Lagrangian function:

$$H(w) = 1/2 * \mathbf{w}_n^H\mathbf{R_p}\mathbf{w}_n + \lambda(\psi_n^H\mathbf{w}_n - 1) \tag{2.13}$$

$$\nabla_w(H(w)) = \mathbf{0} = \mathbf{R_p}\mathbf{w}_n + \psi_n\lambda \tag{2.14}$$

$$\therefore \mathbf{w}_n = -\mathbf{R_p}^{-1}\psi_n\lambda \tag{2.15}$$

From our constraint,

$$\psi_n^H \mathbf{w}_n = 1 \tag{2.16}$$

$$\psi_n^H(-\mathbf{R_p}^{-1}\psi_n\lambda) = 1 \tag{2.17}$$

$$\therefore \lambda = -(\psi_n^H \mathbf{R_p}^{-1}\psi_n)^{-1} \tag{2.18}$$

$$\therefore \mathbf{w}_n = \frac{\mathbf{R_p}^{-1}\psi_n}{\psi_n^H \mathbf{R_p}^{-1}\psi_n} \tag{2.19}$$

where $\mathbf{R_p}$ is the covariance matrix of the vertical pressure field at a particular frequency. In practice, this matrix is not known a priori and must be estimated from the received data. The sample covariance matrix is one option for estimating $\mathbf{R_p}$, but it requires a large number of snapshots to accurately estimate the true value. Section 2.3 will present an alternative method for estimating this matrix. The linear mode filter $\mathbf{H}$ that estimates all the modes together is formed by combining the weights for each mode:

$$\mathbf{H}_{MPDR} = \begin{bmatrix} \mathbf{w}_1^H \\ \mathbf{w}_2^H \\ \vdots \\ \mathbf{w}_M^H \end{bmatrix} \tag{2.20}$$

2.2.4   Maximum A Posteriori (MAP) Mode Filter

The MAP filter is derived as the linear minimum mean square error filter, and under the assumption the mode coefficients and noise are complex Gaussian random vectors (CGRVs), the filter is also the MAP filter. This mode filter is given by [1, 7]

$$\mathbf{H}_{MAP} = \left(\mathbf{R_d}^{-1} + \psi^H \mathbf{R_n}^{-1}\psi\right)^{-1}\psi^H \mathbf{R_n}^{-1} \tag{2.21}$$

where $\mathbf{R_d}$ is the covariance matrix of the mode coefficients and $\mathbf{R_n}$ is the noise covariance matrix. In practical experiments, neither of these matrices is known exactly and they must be estimated in some way. Chapter 3 discusses this issue further.

24

2.3 PCML Method for Spatial Covariance Matrix Estimation

This section describes a physically constrained, maximum likelihood (PCML) method for estimating the spatial covariance matrix used in adaptive filters. This method was proposed by Kraay in 2003 [4]. For a given linear array, the ensemble covariance matrix is not known in practice and the sample covariance matrix, $\hat{\mathbf{R}}_{data}$, is often used in its place.

$$\hat{\mathbf{R}}_{data} = \frac{1}{L} \sum_{l=1...L} \mathbf{X}_l(\omega)\mathbf{X}_l(\omega)^H \qquad (2.22)$$

where $\mathbf{X}_l(\omega)$ is an array snapshot vector at frequency ω [9]. This matrix is the unconstrained maximum likelihood estimate of the covariance matrix [14]. However, an inaccurate estimate of the true covariance matrix results in poor performance of the adaptive algorithms that make use of it. $\hat{\mathbf{R}}_{data}$ is a good approximation of the true matrix only when a large number of snapshots are used. Various reduced-rank methods have been developed to address this problem, including diagonal loading and dominant mode rejection [4, 9].

The goal of the PCML algorithm is to improve adaptive array processing techniques under snapshot-deficient conditions. The algorithm determines the maximum likelihood estimate of the spatial covariance matrix subject to known physical constraints. The physical constraints come from our knowledge that the received signal is composed of a propagating component plus spatially white sensor noise. As a result of the physical constraints, fewer snapshots are required to obtain an accurate estimate of the spatial covariance matrix.

2.3.1 PCML Covariance Matrix Estimate

The signal snapshots are modeled as series of independent, identically distributed (i.i.d.) zero-mean complex Gaussian random vectors (CGRVs). The joint probability

density function of L such snapshots is [15]

$$p(\mathbf{X}_1, ..., \mathbf{X}_L) = \prod_{l=1...L} \frac{1}{\pi^N |\mathbf{R}|} e^{-\mathbf{X}_l^H \mathbf{R}^{-1} \mathbf{X}_l} \qquad (2.23)$$

where $\mathbf{R}$ is the covariance matrix of those snapshots and N is the number of elements in each vector $\mathbf{X}$. The maximum likelihood (ML) estimate of this covariance matrix given the received snapshots is

$$
\begin{aligned}
\hat{\mathbf{R}}_{ML} &= \arg\max_{\mathbf{R}} p(\mathbf{X}_1, ..., \mathbf{X}_L) \\
&= \arg\max_{\mathbf{R}} \prod_{l=1...L} \frac{1}{\pi^N |\mathbf{R}|} e^{-\mathbf{X}_l^H \mathbf{R}^{-1} \mathbf{X}_l} \\
&= \arg\max_{\mathbf{R}} \pi^{-NL} |\mathbf{R}|^{-L} e^{-\sum_{l=1...L} -\mathbf{X}_l^H \mathbf{R}^{-1} \mathbf{X}_l} \\
&= \arg\max_{\mathbf{R}} \; -log|\mathbf{R}| - \frac{1}{L} \sum_{l=1...L} Tr(\mathbf{X}_l^H \mathbf{R}^{-1} \mathbf{X}_l) \\
&= \arg\max_{\mathbf{R}} \; -log|\mathbf{R}| - Tr(\frac{1}{L} \sum_{l=1...L} \mathbf{R}^{-1} \mathbf{X}_l \mathbf{X}_l^H) \\
&= \arg\max_{\mathbf{R}} \; -log|\mathbf{R}| - Tr(\mathbf{R}^{-1} \hat{\mathbf{R}}_{data}) \qquad (2.24)
\end{aligned}
$$

$$(2.25)$$

where $L(\mathbf{R}, \hat{\mathbf{R}}_{data}) = -log|\mathbf{R}| - Tr(\mathbf{R}^{-1} \hat{\mathbf{R}}_{data})$ is the log-likelihood function. There is no closed form solution for the $\mathbf{R}$ that maximizes this log-likelihood function when physical constraints are imposed. It is, however, possible to use derivatives of the likelihood function in an iterative approach to finding the constrained ML estimate. Kraay derived this method for a spatial beamformer[4]. The covariance matrix is separated into its propagating component plus spatially white sensor noise,

$$[\mathbf{R}]_{ij} = \sigma^2 \delta_{ij} + \frac{1}{(2\pi)^3} \int_{\Omega(\mathbf{k})} P(\omega, \mathbf{k}) [\mathbf{v}(\mathbf{k})]_i [\mathbf{v}^H(\mathbf{k})]_j d\mathbf{k} \qquad (2.26)$$

where $[\mathbf{v}(\mathbf{k})]_n = e^{-j\mathbf{k}^T \mathbf{p}_n}$ is the array manifold vector, $\mathbf{p}_n$ the position vector for the nth array element, $\mathbf{k}$ the spatial wavenumber, and $P(\omega, \mathbf{k})$ the frequency-wavenumber

power spectrum. $\Omega(\mathbf{k})$ is the region of support for the wavenumber field imposed by the wave equation, $\|\mathbf{k}\| = \frac{2\pi}{\lambda}$.

2.3.2 Algorithm Initialization

The PCML algorithm initializes its estimate of the covariance matrix with the sample covariance matrix, $\hat{\mathbf{R}}_{data}$. The initial frequency-wavenumber spectral estimate, $P(\omega, \mathbf{k})$, is estimated using the Capon estimator [11],

$$\hat{P}_0(\omega, \mathbf{k}_n) = \mathbf{w}_{MVDR}^H(\mathbf{k}_n)\hat{\mathbf{R}}_{data}\mathbf{w}_{MVDR}(\mathbf{k}_n) = \frac{1}{\mathbf{v}^H(\mathbf{k}_n)\hat{\mathbf{R}}_{data}^{-1}\mathbf{v}(\mathbf{k}_n)} \qquad (2.27)$$

A good rule of thumb is to initialize the white noise power estimate as one tenth the average diagonal value of $\hat{\mathbf{R}}_{data}$.

$$\hat{\sigma}_0^2 = \frac{1}{10N}Tr(\hat{\mathbf{R}}_{data}) \qquad (2.28)$$

2.3.3 Covariance Matrix Update

The first iterative step is to obtain a new estimate of the covariance matrix. This is done by inverse Fourier transforming the frequency-wavenumber spectrum estimate with respect to the wavenumber $\mathbf{k}$,

$$[\hat{\mathbf{R}}_m]_{ij} = [\mathcal{F}_{Space}^{-1}(\hat{P}_{m-1}(\omega, \mathbf{k}) + \hat{\sigma}_{m-1}^2)]_{i,j} \qquad (2.29)$$

$$[\hat{\mathbf{R}}_m]_{ij} = \hat{\sigma}_{m-1}^2\delta_{ij} + \int_{\Omega(\mathbf{k})} \hat{P}_{m-1}(\omega, \mathbf{k})e^{-j\mathbf{k}^T(\mathbf{p}_i-\mathbf{p}_j)}d\mathbf{k} \qquad (2.30)$$

where $\hat{P}_{m-1}(\omega, \mathbf{k})$ is the estimate of the power at frequency ω coming from the direction specified by the wavenumber $\mathbf{k}$ at iteration $m - 1$. Since $\hat{P}_m$ can only be calculated at discrete points, a covariance matrix taper is used to smooth the estimates of propagating energy around the discrete spatial plane wave samples. Each

sample, $\hat{P}_{m-1}(\omega, \mathbf{k})$, is approximated as a weighted, shifted window in k-space,

$$\hat{P}_{m-1}(\omega, \mathbf{k}) = \sum_n \left[\hat{P}_{m-1}(\omega, \mathbf{k}_n) W(\mathbf{k} - \mathbf{k}_n) \right] \qquad (2.31)$$

The integral in equation 2.30 then becomes a summation,

$$[\hat{\mathbf{R}}_m]_{ij} = \hat{\sigma}^2_{m-1} \delta_{ij} + \left(\sum_n \hat{P}_{m-1}(w, \mathbf{k}_n) e^{-j\mathbf{k}_n^T(\mathbf{p}_i - \mathbf{p}_j)} \right) \mathbf{W}_{ij} \qquad (2.32)$$

where $\mathbf{W}_{ij}$ is the inverse Fourier transform of the covariance matrix taper [16, 17]. The taper primarily used by Kraay was a uniform window given by [4]

$$W(\mathbf{u}) = W(u_x, u_y) = \begin{cases} 1, & \text{for } |u_x| < \frac{\Delta u}{2} \text{ and } |u_y| < \frac{\Delta u}{2} \\ 0, & \text{otherwise} \end{cases} \qquad (2.33)$$

where $\mathbf{u}$ is the wavenumber vector normalized by $\frac{2\pi}{\lambda}$ and Δu is the wavenumber grid spacing. The inverse Fourier transform of this window is

$$\mathbf{W}_{i,j} = \int W(\mathbf{u}) e^{+j\frac{2\pi}{\lambda}\mathbf{u}^T(\mathbf{p}_i - \mathbf{p}_j)} d\mathbf{u} \qquad (2.34)$$

$$= \Delta u \; sinc\left(\frac{2\pi}{\lambda} \frac{\Delta u}{2} (p_{i,x} - p_{j,x}) \right) \Delta u \; sinc\left(\frac{2\pi}{\lambda} \frac{\Delta u}{2} (p_{i,y} - p_{j,y}) \right) \qquad (2.35)$$

2.3.4 Power Spectrum and Noise Power Updates

The frequency-wavenumber power spectral density (PSD) estimates are updated using the gradient of the likelihood function (equation 2.24). The gradient with respect to the power at a given frequency and wavenumber is given by

$$\frac{\partial L(\hat{\mathbf{R}}_m, \hat{\mathbf{R}}_{data})}{\partial \hat{P}_m(w, \mathbf{k}_n)} = -\mathbf{v}^H(\mathbf{k}_n) \hat{\mathbf{R}}_m^{-1} \mathbf{v}(\mathbf{k}_n) + \mathbf{v}^H(\mathbf{k}_n) \hat{\mathbf{R}}_m^{-1} \hat{\mathbf{R}}_{data} \hat{\mathbf{R}}_m^{-1} \mathbf{v}(\mathbf{k}_n) \qquad (2.36)$$

and the gradient with respect to the white noise power is given by

$$\frac{\partial L(\hat{\mathbf{R}}_m, \hat{\mathbf{R}}_{data})}{\partial \hat{\sigma}^2_m} = Tr((\hat{\mathbf{R}}_m^{-1} \hat{\mathbf{R}}_{data} - \mathbf{I}) \hat{\mathbf{R}}_m^{-1}) \qquad (2.37)$$

28

The Hessians are given by

$$\frac{\partial^2 L(\hat{\mathbf{R}}_m, \hat{\mathbf{R}}_{data})}{(\partial \hat{P}_m(w, \mathbf{k}_n))^2} = (\mathbf{v}^H(\mathbf{k}_n)\hat{\mathbf{R}}_m^{-1}\mathbf{v}(\mathbf{k}_n))^2 - 2(\mathbf{v}^H(\mathbf{k}_n)\hat{\mathbf{R}}_m^{-1}\mathbf{v}(\mathbf{k}_n))(\mathbf{v}^H(\mathbf{k}_n)\hat{\mathbf{R}}_m^{-1}\hat{\mathbf{R}}_{data}\hat{\mathbf{R}}_m^{-1}\mathbf{v}(\mathbf{k}_n))$$

$$(2.38)$$

and

$$\frac{\partial^2 L(\hat{\mathbf{R}}_m, \hat{\mathbf{R}}_{data})}{(\partial \hat{\sigma}_m^2)^2} = Tr(\hat{\mathbf{R}}_m^{-1}\hat{\mathbf{R}}_m^{-1}) - 2Tr(\hat{\mathbf{R}}_m^{-1}\hat{\mathbf{R}}_{data}\hat{\mathbf{R}}_m^{-1}\hat{\mathbf{R}}_m^{-1}) \qquad (2.39)$$

which can be used to verify that a maximum of the likelihood function is reached and not a minimum. The multiplicative update proposed by Kraay provides a convenient mapping between the gradients of the likelihood function and a multiplicative scaling factor. The update is chosen such that when the gradient of the likelihood function is zero, the power estimates remain unchanged. The scale factor increases monotonically for a positive gradient and decreases monotonically for a negative gradient.

$$\hat{P}_m(w, \mathbf{k}_n) = \begin{cases} \hat{P}_{m-1}(w, \mathbf{k}_n) * \left[\frac{A-1}{e^{\pi/2}-1} \left(e^{arctan(\alpha\frac{\partial L}{\partial \hat{P}_{m-1}})} - 1 \right) + 1 \right], & \frac{\partial L}{\partial \hat{P}_{m-1}} > 0 \\ \hat{P}_{m-1}(w, \mathbf{k}_n) * \left[\frac{B-1}{e^{-\pi/2}-1} \left(e^{arctan(\beta\frac{\partial L}{\partial \hat{P}_{m-1}})} - 1 \right) + 1 \right], & otherwise \end{cases}$$

$$(2.40)$$

Figure 2-5 shows a plot of this update as a function of A, B, α, and β. A and B are the scale's upper and lower limits, and α and β are parameters that control how quickly the algorithm steps as a function of the gradient. For the white noise power update, Kraay used an additive form since it offered better stability for her environment.

$$\hat{\sigma}_m^2 = \hat{\sigma}_{m-1}^2 + 10^{-4} \frac{\partial L}{\partial \hat{\sigma}_{m-1}^2} \Big/ \left| \frac{\partial^2 L}{(\partial \hat{\sigma}_{m-1}^2)^2} \right| \qquad (2.41)$$

The PCML algorithm iterates for a number of iterations until the covariance matrix estimate has converged to its most likely value. A diagram of the algorithm from Kraay and Baggeroer's paper [2] is shown in figure 2-6. The PCML frequency wavenumber estimates are the values of $\hat{P}_m(\omega, \mathbf{k})$ at the final iteration of the algorithm. The likelihood function can be calculated at each iteration to provide an indication of whether the algorithm has converged.
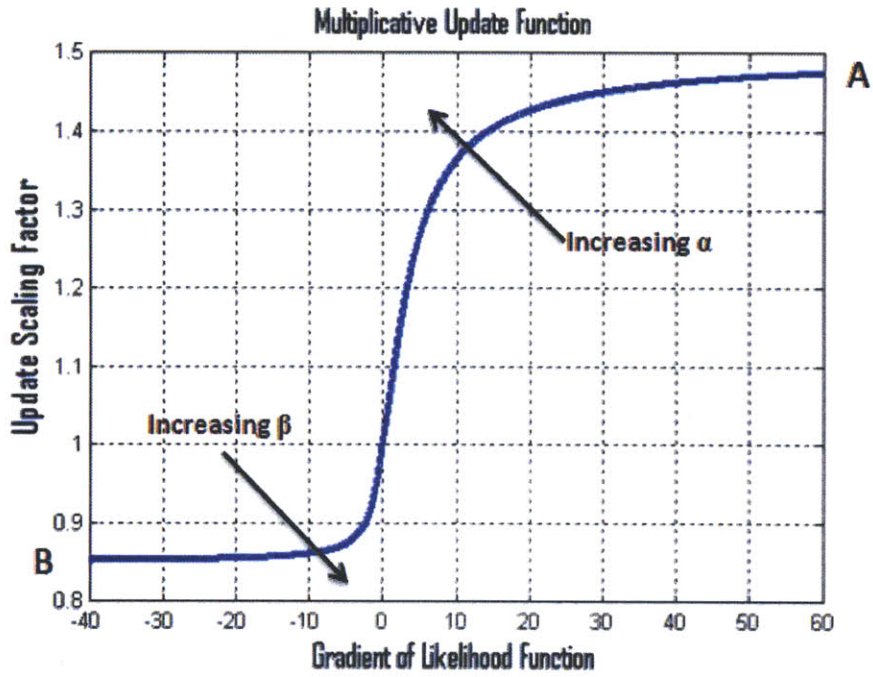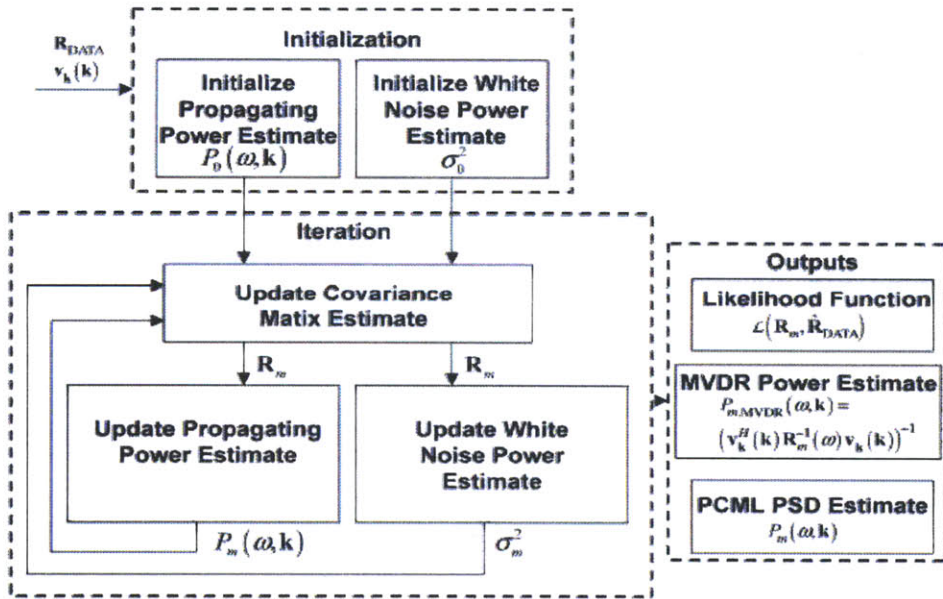
Figure 2-5: Multiplicative Update

Figure 2-6: PCML algorithm by Kraay and Baggeroer [2]

2.4 Broadband Processing

This section describes the broadband processing steps that were used on the simulated data from RAM and on the experimental data from the SW06 experiment. The broadband received data was first bandpass filtered to mitigate out of band noise. A linear-phase bandpass filter with a center frequency of 203 Hz and a one-sided bandwidth of 55 Hz was used. The appropriate narrowband mode filter weights were calculated as described in section 2.2 every 5 Hz between 135 and 265 Hz, and were set to zero outside this region.
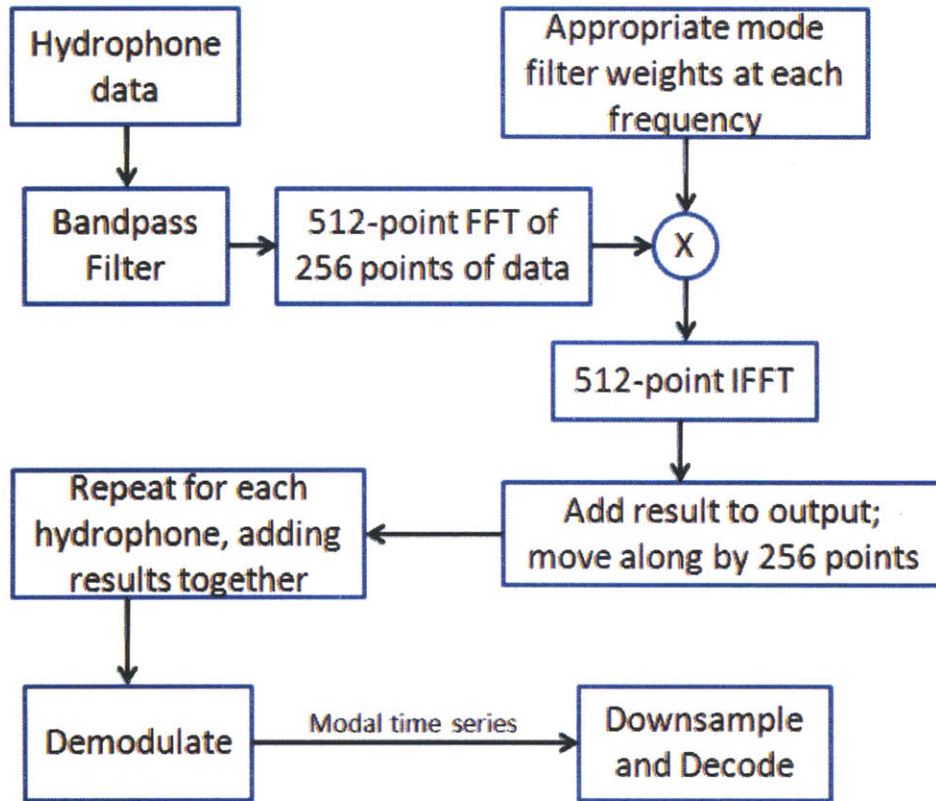

Figure 2-7: Block diagram of the broadband framework used to generate the modal time series

A frequency domain approach was used to implement the filter weights using the overlap-add method [18]. A Fourier transform of length 512 is taken of a data segment of length 256, is multiplied by the desired weights at each frequency, and an inverse Fourier transform is taken. The result is added to the output and the process is

31

repeated starting 256 samples ahead. This process is done at each hydrophone and the outputs are added together and demodulated. Figure 2-7 shows a block diagram of the processing steps. A time domain filter implementation using the Parks-McClellan method was also tested and gave similar results. Only the frequency domain approach is presented in this thesis, however, since the results were nearly the same.

A simple decoder was used to process the modal time series, which has been further downsampled to twice the symbol rate. The decoder was run two times - first taking every other data sample starting with the first, and then taking every other sample starting with the second. The offset which produced the smallest bit error rate was used. The constellation's phase was aligned using a total least squares fit of the symbols through the origin to rotate them to 0/180 degrees. Equations 2.42 and 2.43 show the total least squares algorithm used to determine the angle the data points needed to be rotated by [19]. **a** is a vector of the real parts of the data symbols and **b** is a vector of the imaginary parts.

$$[\mathbf{U} \ \mathbf{S} \ \mathbf{V}] = \text{svd}([\mathbf{a} \ \mathbf{b}]); \tag{2.42}$$

$$\text{angle} = tan^{-1}(-\mathbf{V}(1,2)/\mathbf{V}(2,2)); \tag{2.43}$$

After rotation by this amount, a decision was made for each sample based on which side of the imaginary axis is was on. The bit decisions were compared with the known transmitted data to determine the bit error rate.

Chapter 3

Simulation and Experimental Results

This chapter applies the PCML method for covariance matrix estimation to the underwater environment. The performance of several mode filters is compared in two simulations and experimental data.

3.1 Complex Gaussian Random Vector (CGRV) Simulation Setup

The first simulation analyzed in this thesis models the mode amplitudes and noise as zero-mean complex Gaussian random vectors (CGRVs). Equation 3.1 is a convenient way to express the vertical pressure field as a weighted sum of normal modes plus noise [1].

$$\begin{bmatrix} p(z_1) \\ \vdots \\ p(z_N) \end{bmatrix} = \begin{bmatrix} \psi_1(z_1) & \cdots & \psi_M(z_1) \\ \vdots & \ddots & \vdots \\ \psi_1(z_N) & \cdots & \psi_M(z_N) \end{bmatrix} \begin{bmatrix} d_1 \\ \vdots \\ d_M \end{bmatrix} + \begin{bmatrix} n(z_1) \\ \vdots \\ n(z_N) \end{bmatrix} \qquad (3.1)$$

Written in vector notation, $\mathbf{p} = \psi \mathbf{d} + \mathbf{n}$, where $\mathbf{d}$ is a vector of mode amplitudes, $\mathbf{n}$ is observation noise, and ψ is the matrix of mode shape vectors. Since mode shapes depend on frequency, this is only a narrowband model.

33

In this simulation, the mode coefficient vector, $\mathbf{d}$, is modeled as a zero-mean complex Gaussian random vector. The noise is modeled as a combination of spatially white (SW) and Kuperman-Ingenito (KI) noise [20]. Spatially white noise has a covariance matrix $\mathbf{R}_n = \sigma^2 \mathbf{I}$ and KI noise has a covariance matrix

$$\mathbf{R}_n = \boldsymbol{\psi} \begin{bmatrix} P_1^2 & 0 & \cdots & 0 \\ 0 & P_2^2 & \ddots & \vdots \\ \vdots & \vdots & \ddots & 0 \\ 0 & \cdots & 0 & P_M^2 \end{bmatrix} \boldsymbol{\psi}^H \tag{3.2}$$

where $P_1^2...P_M^2$ are the powers of each mode coefficient in the KI noise model. In this simulation, the noise power in each mode was proportional to 1, 1, 2, 1, 1, 3, 2, 2, and 2 which was chosen as a representative choice. The real and imaginary parts of $\mathbf{d}$ and $\mathbf{n}$ are i.i.d. and Gaussian, and therefore the vectors are proper and their covariance matrices are real [15]. The simulation is similar to the one described by Buck et. al [1] for a shallow water environment. The simulated environment had a typical shallow water sound speed profile and was 80 meters deep. The frequency used for calculating the mode shapes was 200 Hz. A vertical array of 20 equally-spaced hydrophones was used. The location of the bottom hydrophone was fixed at 78 meters down and the depth of the top hydrophone was varied from the water surface to a depth of 40 meters (half the water column). This gradually reduced the fraction of the water column that was spanned by the array. 500 trials were run using independent realizations of the mode coefficients and noise vector. Linear mode filters were applied to the simulated pressure field to obtain an estimate of the mode amplitudes. The error criteria is the mean total squared error, $\|\hat{\mathbf{d}} - \mathbf{d}\|^2$, between the estimate of the complex mode amplitudes and their actual values. The total mode energy is included in plots of the results for reference, and represents the error that would result from estimating each mode to have zero amplitude.

3.2 Application of the PCML Algorithm to the Underwater Environment

Before presenting the results of the CGRV simulation, this section first presents the application of Kraay's PCML algorithm to the problem of estimating complex mode amplitudes. For the underwater environment model of equation 2.1, the covariance matrix of the acoustic pressure field can be decomposed into a propagating modal component plus spatially white sensor noise. In this case, there are a discrete number of propagating modes to sum instead of an infinite number of spatial plane waves to integrate over, so no covariance matrix taper is necessary. Thus equations 2.30 and 2.32 become

$$\hat{\mathbf{R}} = \hat{\sigma}^2 \mathbf{I} + \sum_{i=1\ldots M} \hat{P}(w, \boldsymbol{\psi}_i) \boldsymbol{\psi}_i \boldsymbol{\psi}_i^H \qquad (3.3)$$

The second change is that instead of steering the beamformer to a spatial direction, it is steered to a particular mode. The steering vector, $\mathbf{v(k)}$, becomes

$$\mathbf{v(k}_n) = \boldsymbol{\psi}_n \qquad (3.4)$$

With these modifications, the PCML algorithm developed by Kraay can be applied to determine the maximum likelihood estimate of the covariance matrix given the physical constraint. In Kraay's simulation, the window had a factor of $1/(\Delta u)^2$ in it, where Δu is the normalized wavenumber grid spacing of 0.1. This caused the final PSD estimates to be larger than the initial power estimates since $\hat{\mathbf{R}}$ is scaled by a window in equation 2.32. The P_m estimates converge to 100 times the values they would have converged to had the factor of $1/(\Delta u)^2 = 1/100$ not been included in the equation. For the underwater environment, there is no window and the initial $\hat{P}_0$ values are already very close to the final PSD estimates. Therefore, the upper limit of the multiplicative update used in the PCML algorithm can be set around 1.5 instead of 125. The PCML algorithm is less sensitive to the choices of α and β, however these parameters still need to be small enough to ensure minimal ripple in the steady state and large enough for the solution to converge in a reasonable number of iterations.

Figures 3-1 through 3-4 show the effect of these parameters on the log-likelihood as a function of iteration number.
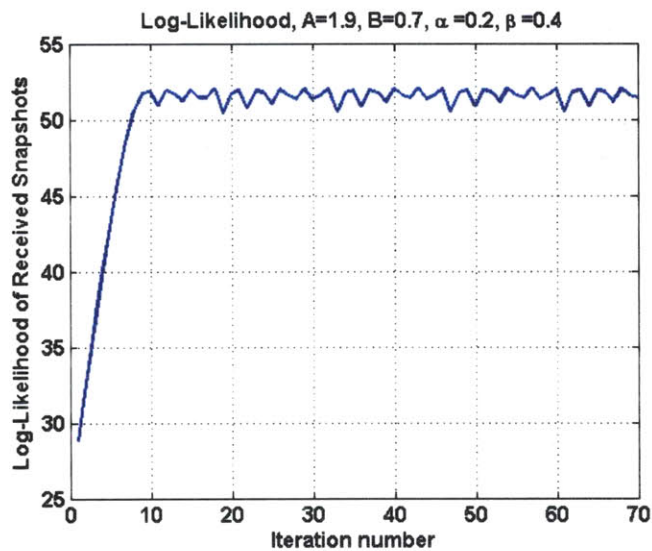

Figure 3-1: Large values of A and small values of B cause ripple in the steady state.
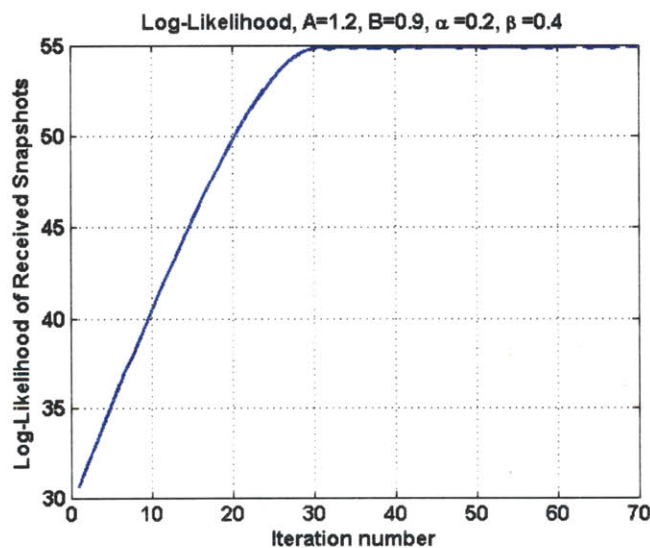

Figure 3-2: Small values of A and higher values of B require a larger number of iterations for the algorithm to converge.

In the spatial simulation done by Kraay, the likelihood function was believed to be highly multimodal with respect to $\hat{\sigma}^2$ [4]. The simulated annealing method was proposed by Kraay as an alternative to the gradient update, and this technique is
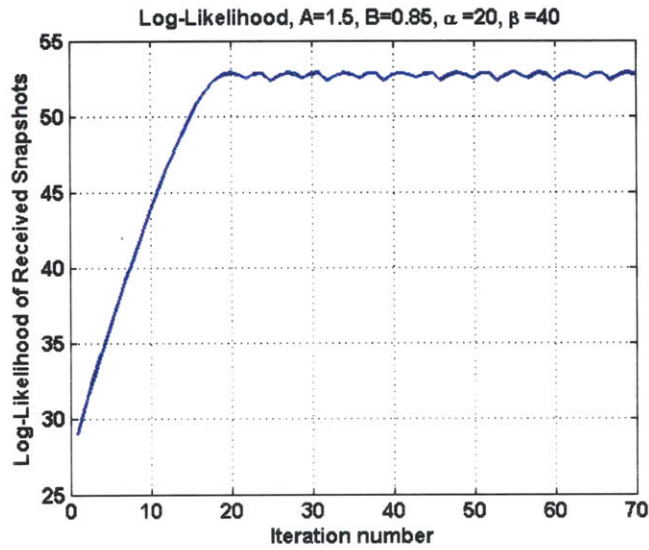
Figure 3-3: Large values of α and β cause ripple in the steady state.
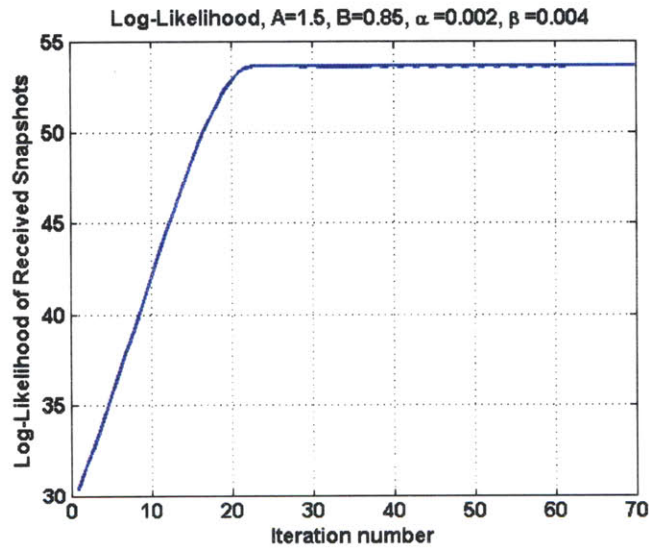
Figure 3-4: Small values of α and β require a larger number of iterations for the algorithm to converge.

able to overcome such local maximums in the likelihood function. Though more computationally intensive, this method provides greater stability and robustness in finding the global maximum since it has the ability to escape local maxima [21]. However, figure 3-5 shows a plot of the likelihood as a function of the white noise power estimate for the mode simulation being investigated. The likelihood function

in this case is not multimodal, and the gradient update proposed by Kraay can be used without concern for getting caught in a local maximum. The curve decreases monotonically with increasing $\hat{\sigma}^2$ beyond the range that is shown. The spike in the likelihood function near zero occurs because of poor matrix conditioning and can be avoided with a non-zero initialization of the estimate. In this thesis, only the gradient technique is used since it is less computationally intensive and provides excellent results.
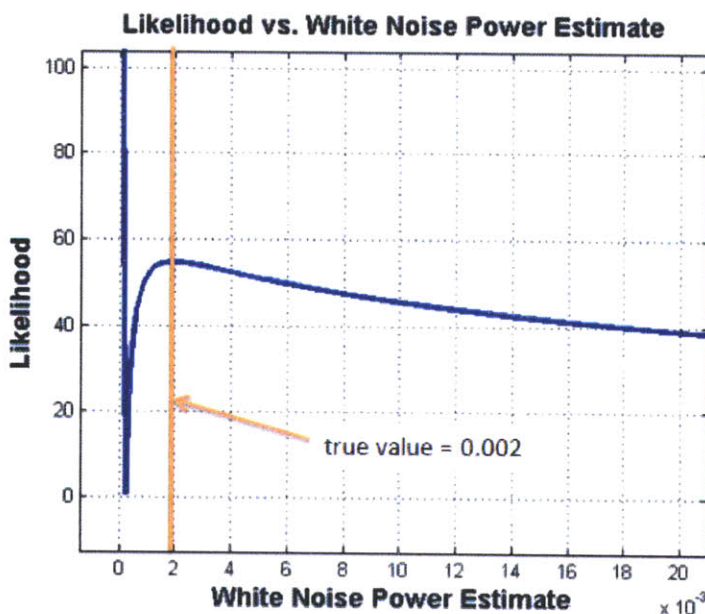

Figure 3-5: Likelihood of the 50 received data snapshots as a function of the white noise power estimate, $\hat{\sigma}^2$.

3.3   CGRV Simulation Results

This section presents the results of the simulation described in the section 3.1. The filters compared were the pseudoinverse, sampled mode shape, reduced rank pseudoinverse, MAP, MPDR, PCML-MAP, MPDR-unweighted, and PCML-MPDR filters. The total mode energy is indicated with the black '+' line. The MPDR filter uses the sample covariance matrix formed from simulated data snapshots. The

MPDR-unweighted filter is the MPDR filter that uses an unweighted sum of the outer products of the mode shape vectors as its covariance matrix. That is,

$$\mathbf{R}_{unweighted} = \sum_{i=1...9} \psi_i \psi_i^H \tag{3.5}$$

The PCML-MAP filter uses the mode power estimates from the final iteration of the PCML algorithm in the variances of $\mathbf{d}$ and assumes that the modes are independent. The filter also assumes there is no KI noise present. That is, the filter uses

$$\mathbf{R_d} = \begin{bmatrix} \hat{P}_1^2 & 0 & \cdots & 0 \\ 0 & \hat{P}_2^2 & \ddots & \vdots \\ \vdots & \vdots & \ddots & 0 \\ 0 & \cdots & 0 & \hat{P}_M^2 \end{bmatrix} \tag{3.6}$$

and

$$\mathbf{R_n} = \hat{\sigma}^2 \mathbf{I} \tag{3.7}$$

in the MAP filter equation (equation 2.21). The PCML-MPDR filter uses the estimate of the covariance matrix from the final iteration of the PCML algorithm in the MPDR mode filter (equation 2.19).

In the PCML algorithm, the multiplicative update was used to update both the mode power estimates and the white noise power estimate. The update amount was calculated with A=1.5, B=0.85, α=0.2, and β=0.4. The algorithm consistently converged within about 100-150 iterations using these parameters. 50 data snapshots were used to initialize the adaptive algorithms used in this simulation, except where stated otherwise. The signal powers varied in each mode, and were proportional to 1, 2, 5, 6, 3, 1, 0.5, 0.5, and 0.1 respectively in each of the first 9 modes.

The sawtooth pattern in the reduced rank PI and in the unweighted MPDR curves is a result of the changing number of singular values used in the pseudoinversion as the condition number of either ψ or $\mathbf{R}_{unweighted}$ changes. This effect will be explored later in the section. The PCML-MAP filter performs within a fraction of a dB of

the true MAP filter. The MPDR filter using the sample covariance matrix requires tens of thousands of snapshots to perform similarly to the PCML-MPDR filter with only 50 snapshots. Therefore, there is a significant advantage to using the PCML algorithm to estimate the spatial covariance matrix used in the MPDR filter. Figure 3-6 shows the effect of the number of snapshots on the MPDR mode filter, with and without using the PCML method for the covariance matrix estimation. It takes more than 10,000 snapshots for the unconstrained sample covariance matrix to converge to its true value. When the PCML algorithm is used, the matrix converges within 20 snapshots.
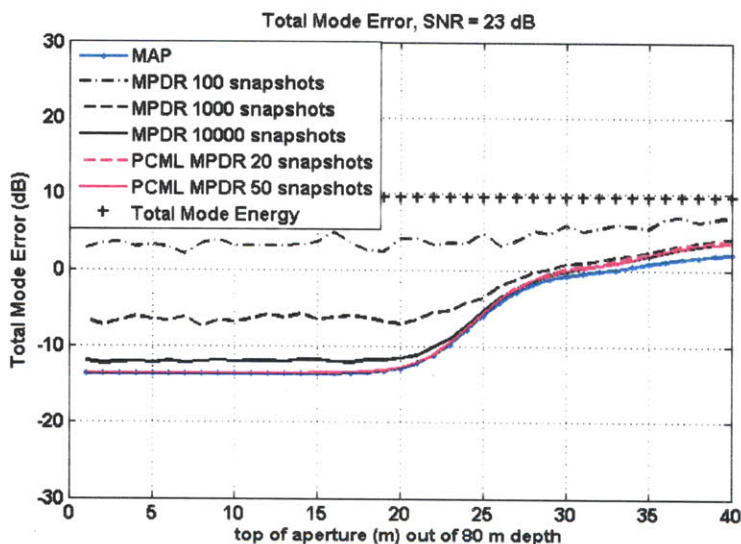
Figure 3-6: Varying the number of snapshots used in the MPDR filter, with and without the PCML algorithm.

Figures 3-7 through 3-12 show the results of this simulation for varying SNR and noise types. The PCML-MAP filter consistently performs the same as the true MAP filter, and the curves lie on top of each other. In the presence of KI noise, the PCML-MPDR filter also performs the same as the MAP filter, however it performs slightly worse in SW noise. The unweighted MPDR filter uses a pseudoinversion to compute the inverse of its covariance matrix, and which often results in a similar sawtooth pattern to that observed in the reduced rank PI filter. Even when the optimal number of singular values are used with the reduced rank PI filter, it still

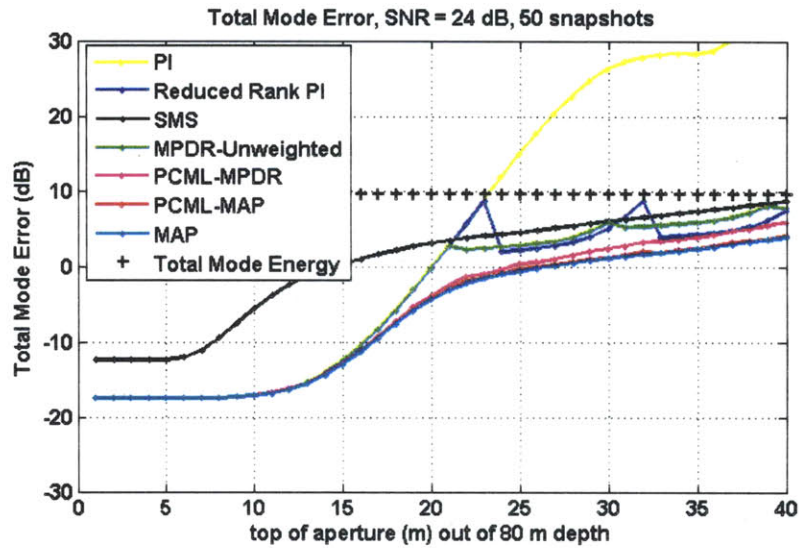does not perform as well as the adaptive methods do when SW noise is present.

Figure 3-7: Comparison of various mode filtering methods, SW noise.
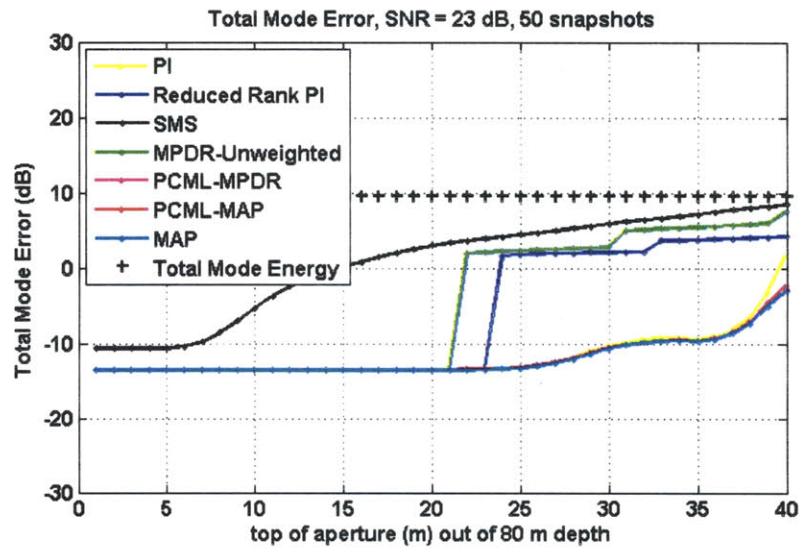

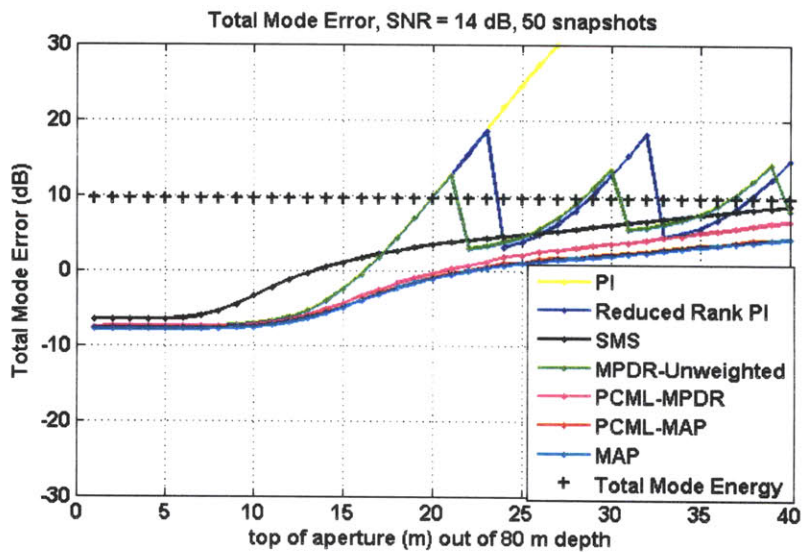Figure 3-8: Comparison of various mode filtering methods, KI noise.
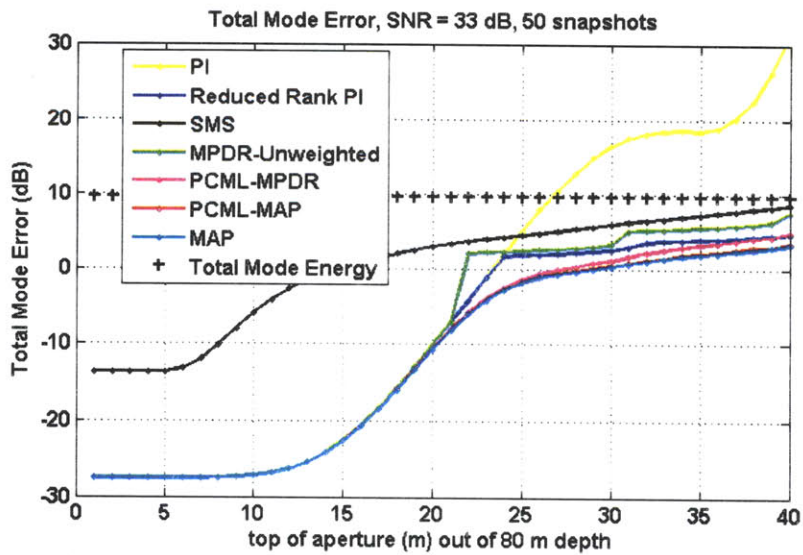
Figure 3-9: SW noise, low SNR.
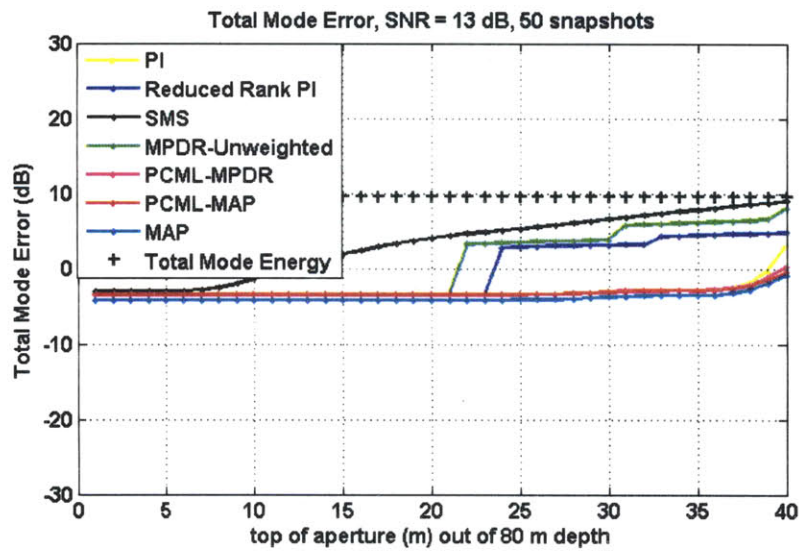

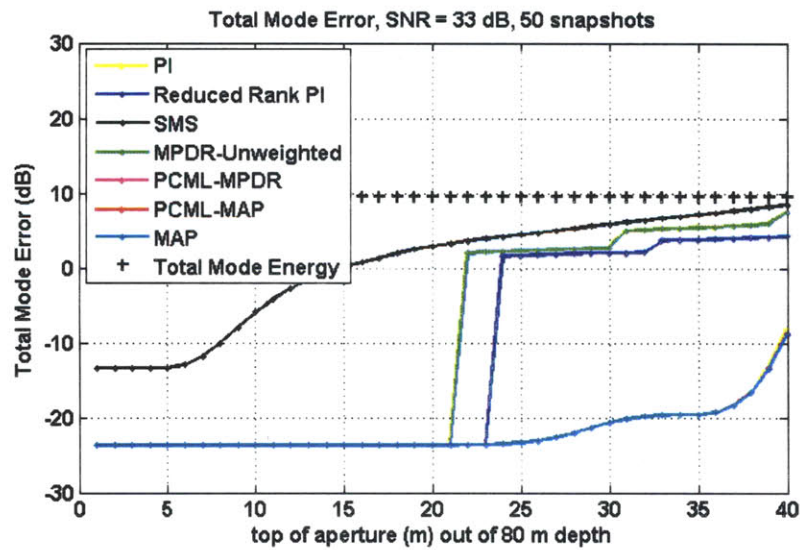Figure 3-10: SW noise, high SNR.

Figure 3-11: KI noise, low SNR.

Figure 3-12: KI noise, high SNR.

43

The last part of this section presents the effect of the number of singular values included in the reduced rank PI filter. In the figures below, the 'Reduced rank PI 0.5' filter, for example, indicates that singular values within 0.5 times the maximum singular value were included in the pseudoinversion. With KI noise, it is better to use more singular values to better null the interfering nodes. With SW noise, it is better to use fewer singular values so the filter is less sensitive to the white noise. This can been seen with equation 3.8. If there is no SW noise, then it is best to include all singular values in the pseudoinversion so $\boldsymbol{\psi}^+\boldsymbol{\psi}$ will fully cancel. In the presence of SW noise, it is best to use fewer singular values so the spatially white noise, $\mathbf{n}_{SW}$, is not magnified from the inversion of the small singular values of $\boldsymbol{\psi}$.

$$\hat{\mathbf{d}} = \boldsymbol{\psi}^+\mathbf{p} = \boldsymbol{\psi}^+\boldsymbol{\psi}\mathbf{d} + \boldsymbol{\psi}^+\boldsymbol{\psi}\mathbf{d}_{KI} + \boldsymbol{\psi}^+\mathbf{n}_{SW} \tag{3.8}$$

The PI filter still does not do as well as the PCML-MPDR or MAP filters when SW noise is present even when the correct number of singular values is used.
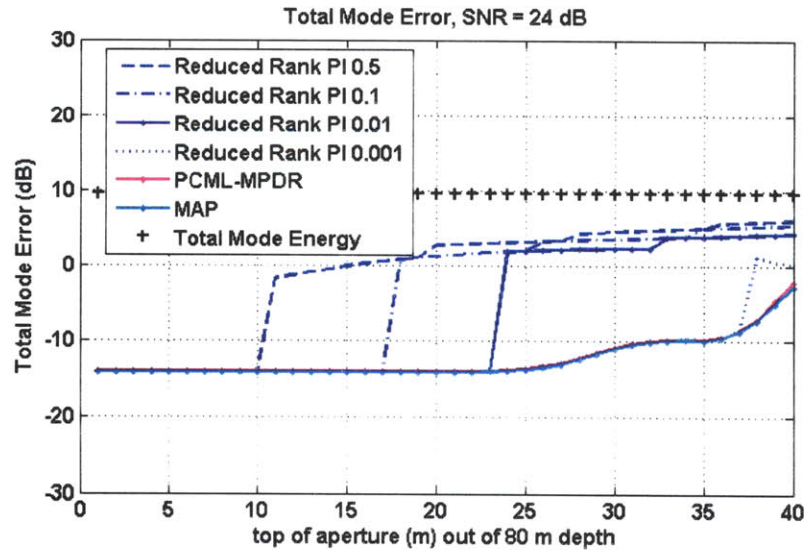

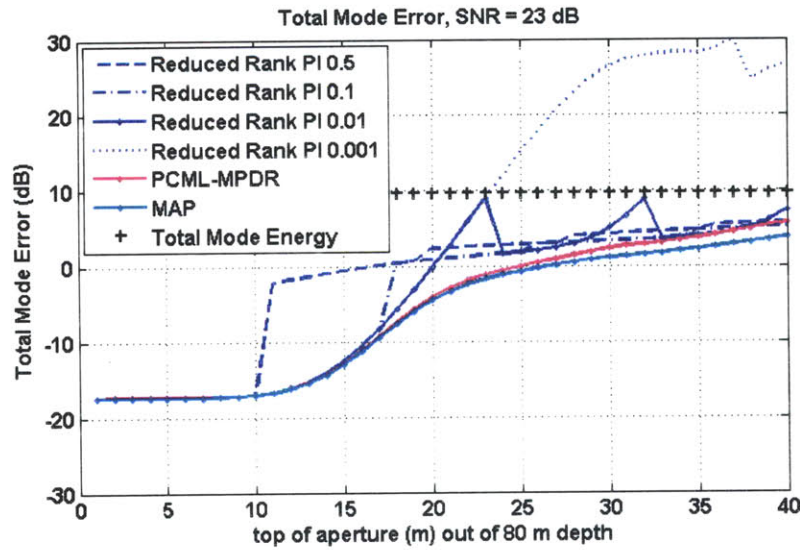Figure 3-13: PI filter performance in the presence of KI noise.

Total Mode Error, SNR = 23 dB

Figure 3-14: PI filter performance in the presence of SW noise.

3.4 RAM Simulation Setup

Simulations of an underwater environment similar to that observed during Shallow Water 2006 were run with a Matlab version of RAM (Range-dependent Acoustic Model). RAM is a parabolic equation based code for simulating underwater sound propagation [22]. The sound speed profile was range varying and was determined from CTD casts taken along the transmission path of the SW06 experiment. Figure 3-15 shows a plot of the sound speed profile at the receive location and the corresponding mode shapes. The environment is assumed to have cylindrical symmetry. The values of the impulse response, H(z,f), at frequencies ranging from 100-300 Hz and depths ranging from the surface to the bottom were calculated, and H(z,f) was assumed to be zero outside this region. The inverse discrete Fourier transform was taken at each depth to form h(z,t), the impulse response as a function of depth and time.

Figure 3-16 shows a plot of the simulated impulse response of the channel. The second mode arrives first in this environment followed by the first mode. There was minimal energy in modes higher than mode 2, which was different from the experimental results of the SW06 experiment. This is likely because the simulated environment was only minimally range-varying, and there is only a small amount of

mode coupling. The simulated environment's sound speed profile is specified at three distinct locations, whereas in the SW06 experiment it continually varied along the transmission path.
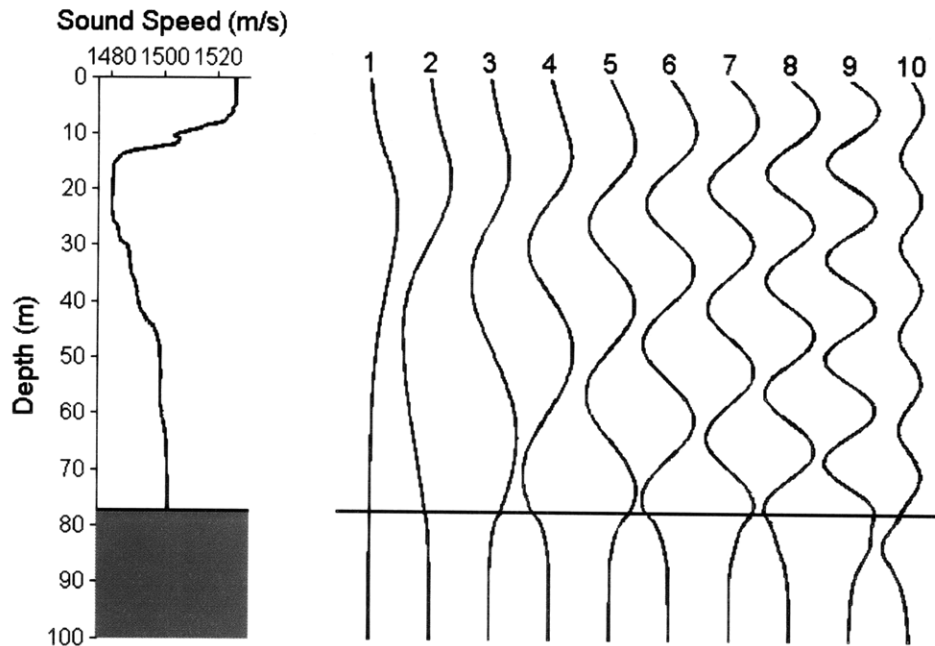

Figure 3-15: Sound speed profile and 200 Hz mode shapes for the shallow water environment used in the CGRV simulation and at the receiver location of the RAM simulation.

A binary phase shift keyed (BPSK) signal with a center frequency of 203 Hz and a one-sided bandwidth of 51 Hz was generated and convolved with the impulse response determined by RAM for each simulated hydrophone depth. A combination of Kuperman-Ingenito (KI) and spatially white (SW) broadband noise was added to the signals. The simulated receive array had nine equally spaced hydrophones at depths of 17.25, 24.75, 32.25, 39.75, 47.25, 54.75, 62.25, 69.75, and 77.25 meters, which coincide with hydrophone locations used during the SW06 experiment. Maximum-length sequences were transmitted and it is possible to generate estimates of the effective impulse response of the channel by computing the circular cross-correlation of the transmitted signal with the output of the mode filter [23].

3.5 RAM Simulation Results

This section presents the results of the RAM simulations described in the section 3.4. One performance metric used to compare the types of mode filters is the estimated channel impulse response. Since m-sequences were transmitted, this was done by circularly cross-correlating the output of the mode filter with the transmitted data. The height and narrowness of the impulse response peak relative to the noise floor is an indicator of the signal's strength. A second performance metric is the mean square error between the estimated response of a particular mode and its true response, as determined by filtering with 80 equally spaced hydrophones in the water column and bottom. The last performance metric considered is the bit error rate for the simulated communication channel when a mode filter is used to pre-process the data before decoding without an equalizer.

The PCML-MAP filter uses an estimate of the noise covariance matrix generated from a recording of the noise only, which in experimental work would be recorded immediately before or after a given transmission. As in the CGRV simulation, the
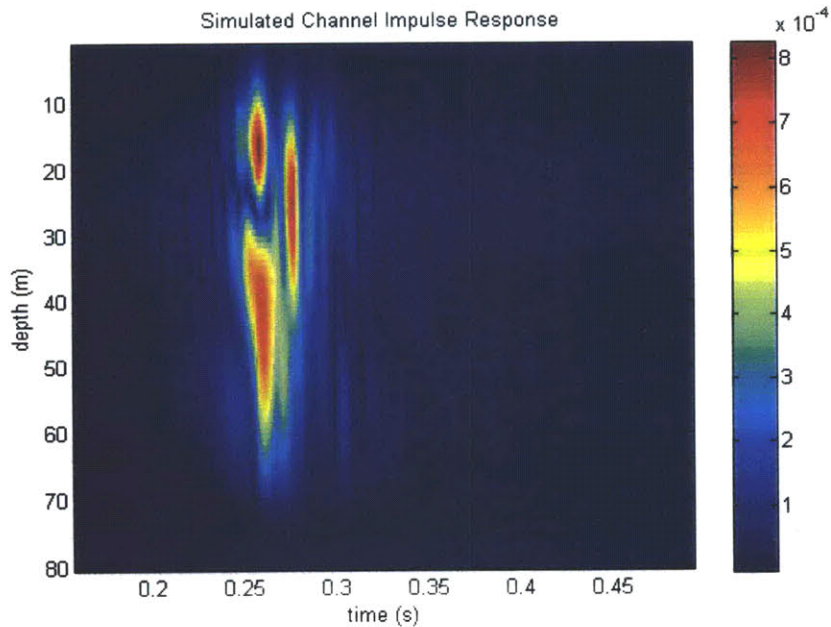
Figure 3-16: Channel impulse response as generated by RAM.

47

covariance matrix of the signal's mode coefficients is estimated via the PCML algorithm for this mode filter. The MAP filter uses the true mode covariance matrix of the signal alone and the true noise covariance matrix, as determined using the 80-hydrophone array. This filter is included for comparison purposes, but in practice these matrices would not be known precisely. Each of these mode filters was computed every 5 Hz between 135 and 265 Hz and the filters were applied using the overlap-add method [18]. A detailed description of the short-time Fourier techniques used to apply the narrowband mode filters to broadband signals is found in [5] and [24], and is summarized in section 2.4.

Figures 3-17, 3-18, and 3-19 show the simulated impulse responses for modes 1-3 using the full 80-element array. The plots from all the mode filters except MAP fall directly on top of each other. Mode 2 has the strongest response, as evidenced by the height of its peak relative to its sidelobes and by the width of its mainlobe. The impulse response plots are normalized by dividing by the sum of the squared values of the noise floor, so the height of the impulse response estimate is an indication of signal strength.
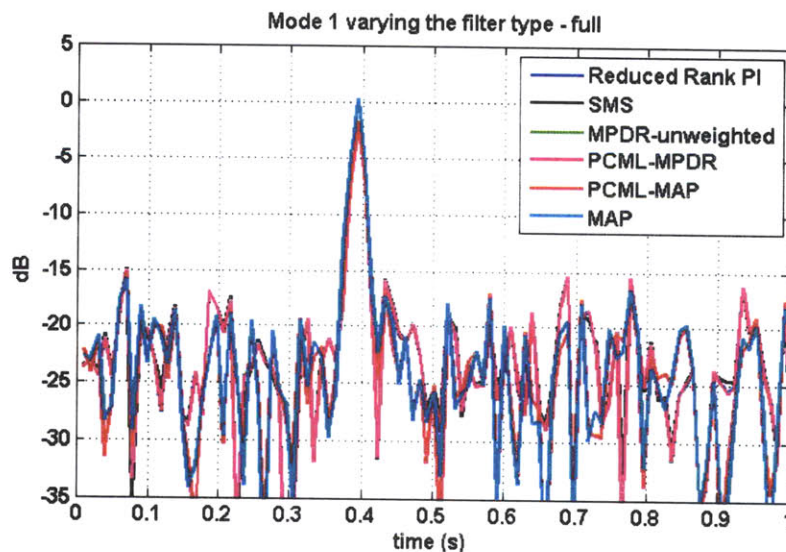
Figure 3-17: Estimated channel impulse response using 80 hydrophones.
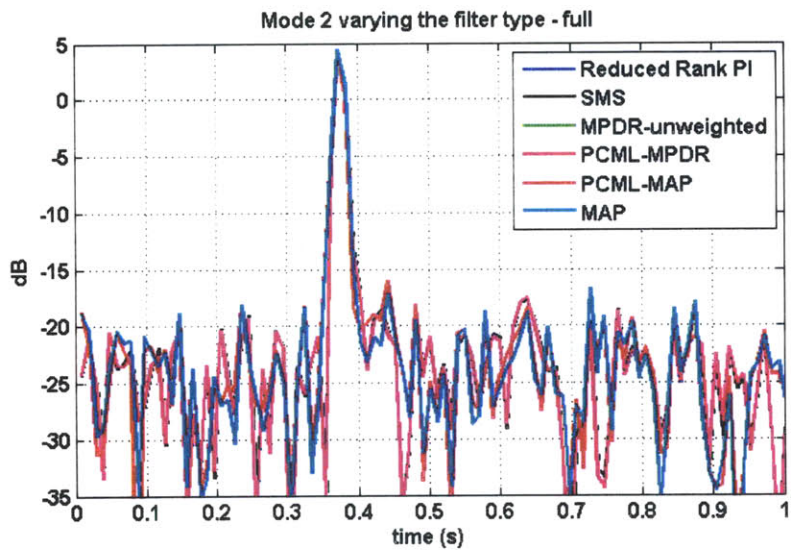
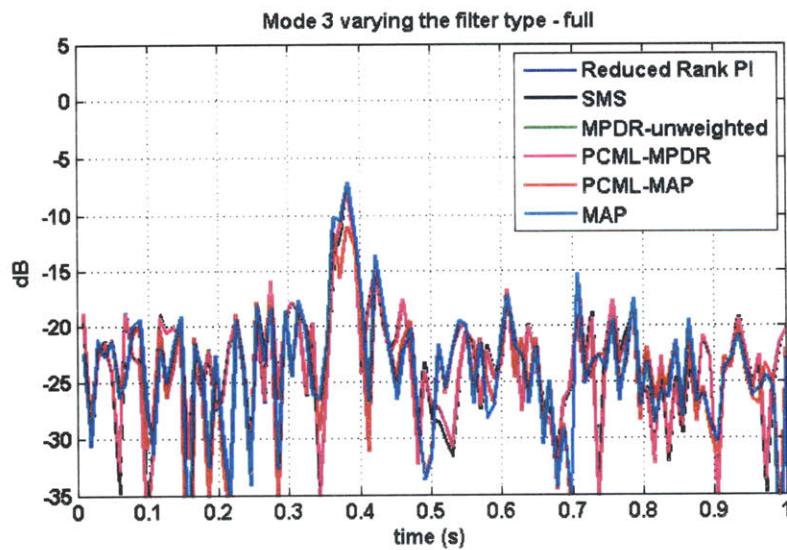Figure 3-18: Estimated channel impulse response using 80 hydrophones.

Figure 3-19: Estimated channel impulse response using 80 hydrophones.

49

Figures 3-20, 3-21, 3-22, and 3-23 show the simulated impulse responses using 9 equally spaced hydrophones. The plot for mode 4 is included for reference, however almost no signal is resolvable in that mode or any higher order modes in this simulation.
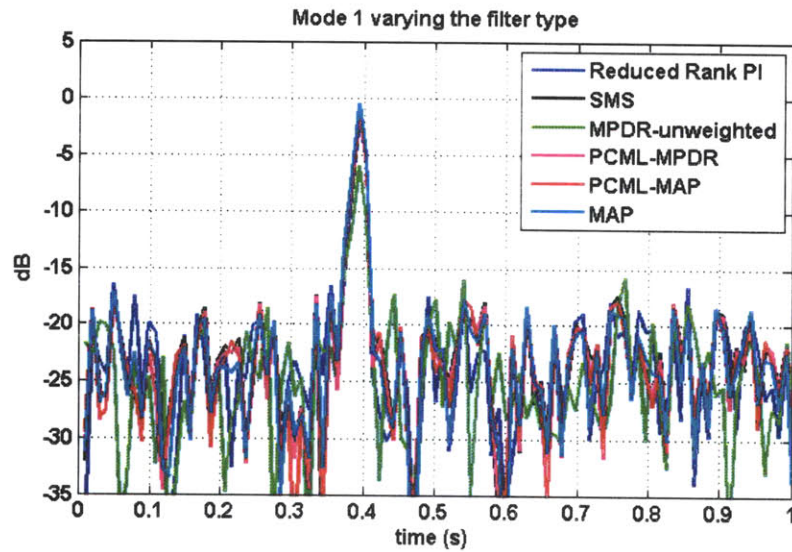

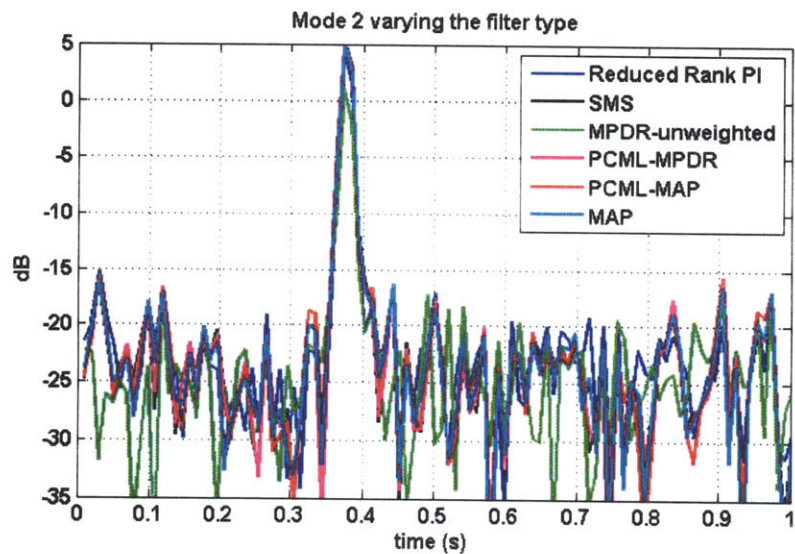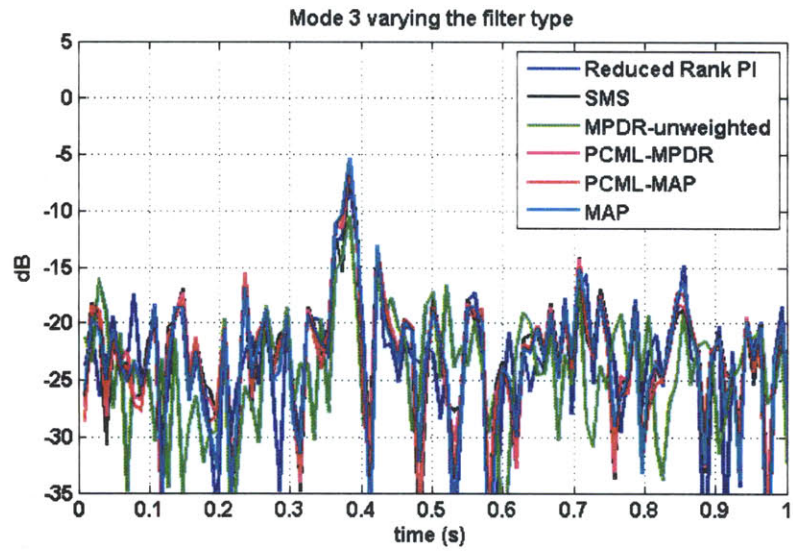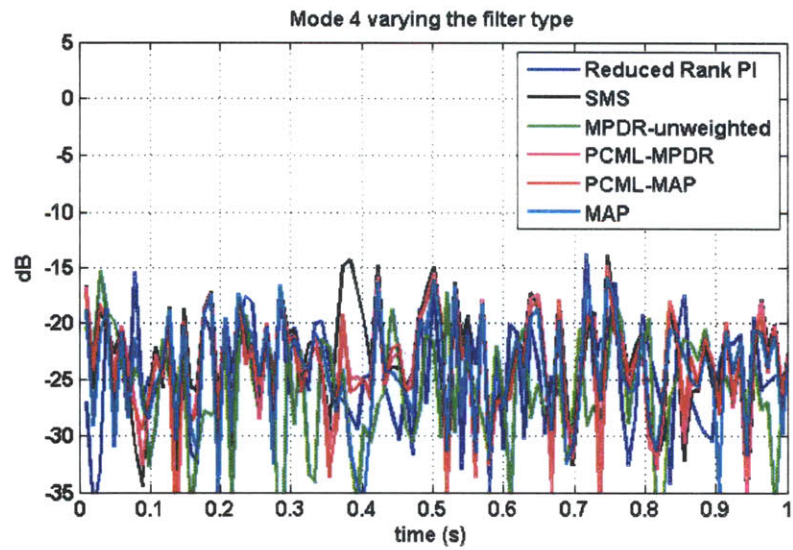Figure 3-20: Estimated channel impulse response using 9 hydrophones.

Figure 3-21: Estimated channel impulse response using 9 hydrophones.

Figure 3-22: Estimated channel impulse response using 9 hydrophones.

Figure 3-23: Estimated channel impulse response using 9 hydrophones.

The following figures show the simulated impulse response using 4 hydrophones. In this case, the mode filters have a difficult time differentiating between the modes, and some of the signal power from mode 2 is observed in the other modes.
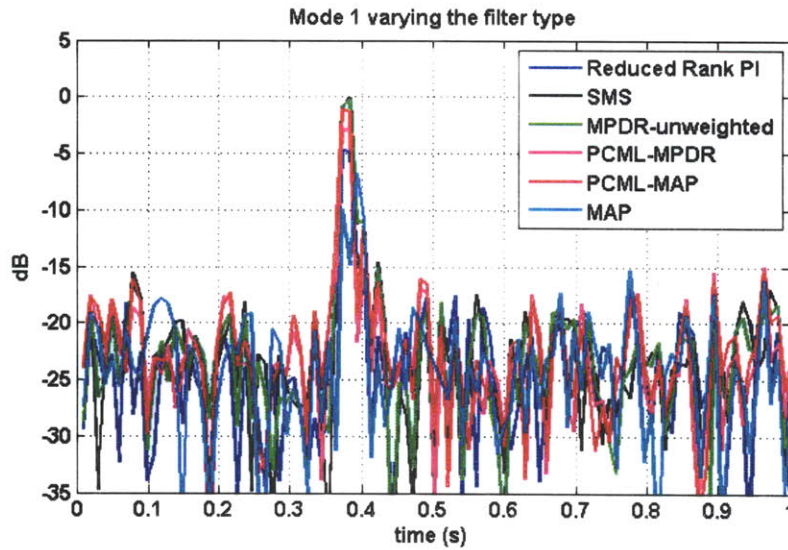

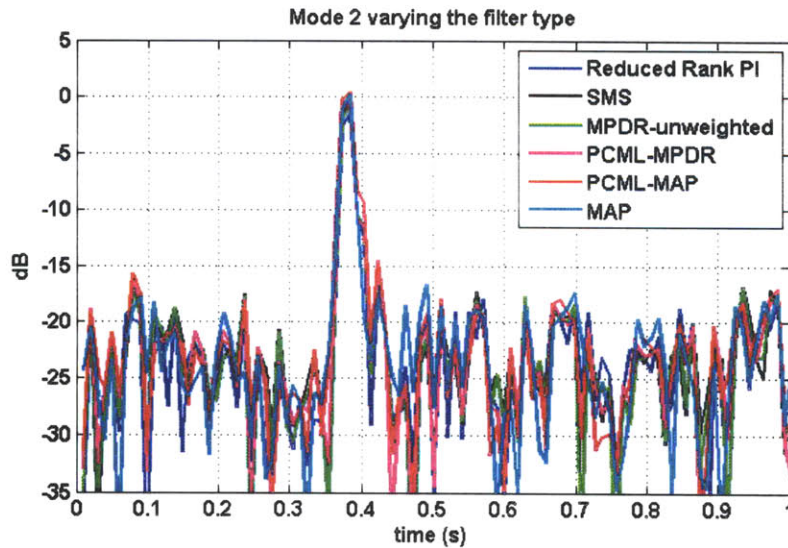Figure 3-24: Estimated channel impulse response using 4 hydrophones.

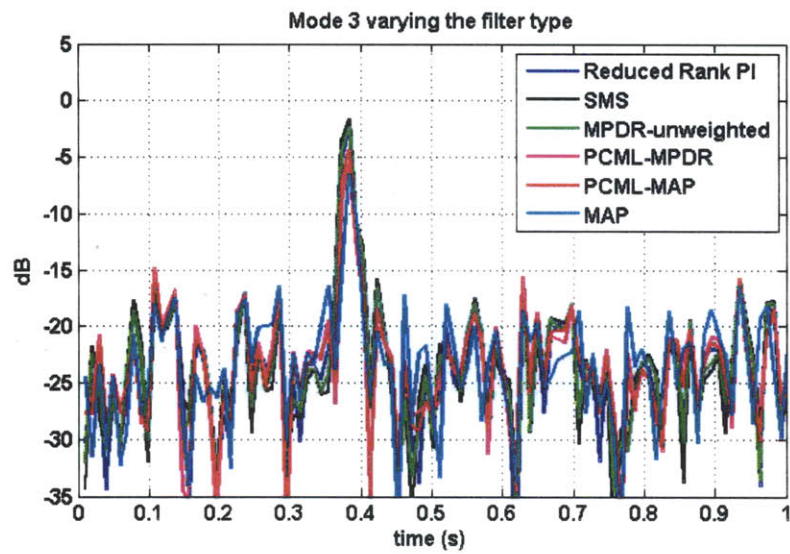Figure 3-25: Estimated channel impulse response using 4 hydrophones.

Figure 3-26: Estimated channel impulse response using 4 hydrophones.

Figures 3-27 and 3-28 show the simulated impulse response of the first four modes using 9 and 4 hydrophones respectively. As the number of hydrophones is reduced, it becomes more difficult to differentiate between each mode and all the modes begin to appear similarly.
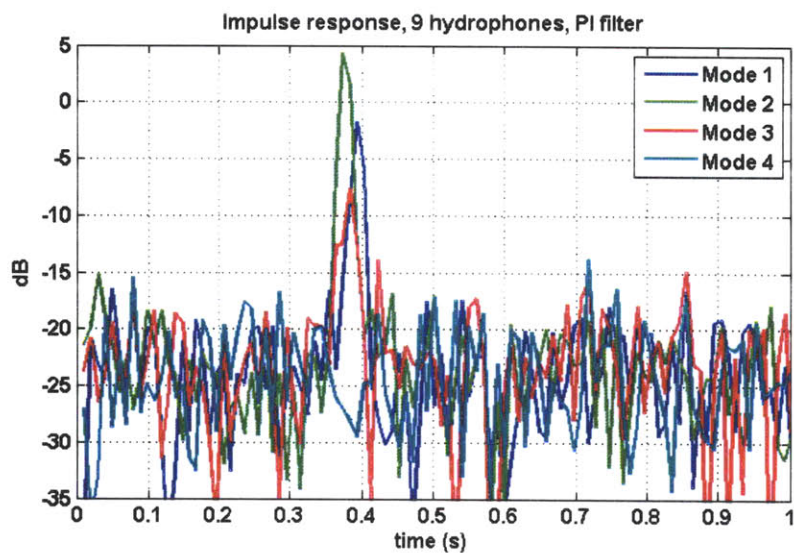

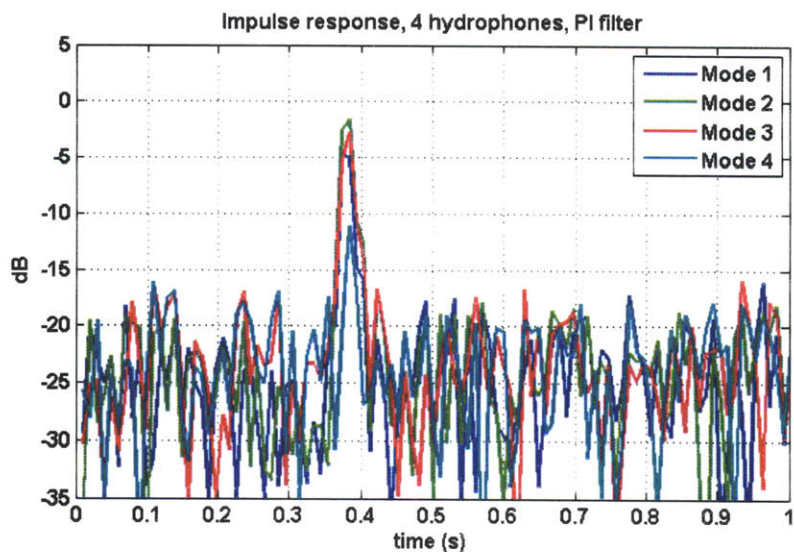Figure 3-27: Estimated channel impulse response using 9 hydrophones.

Figure 3-28: Estimated channel impulse response using 4 hydrophones.

Figure 3-29 shows the mode covariance matrix of the signal at the receiver. The majority of the signal energy is contained in modes 1 and 2 whereas the higher order

54

modes contain almost no signal energy. This is why, as will be seen later, mode filters for modes 1 or 3 often do worse when additional hydrophones are added since more signal energy from mode 2 is filtered out. Furthermore, each mode filter produces a similar estimate of the channel impulse response. The shape of the response therefore depends more on the channel and the number of hydrophones used than it does on the choice of mode filter.
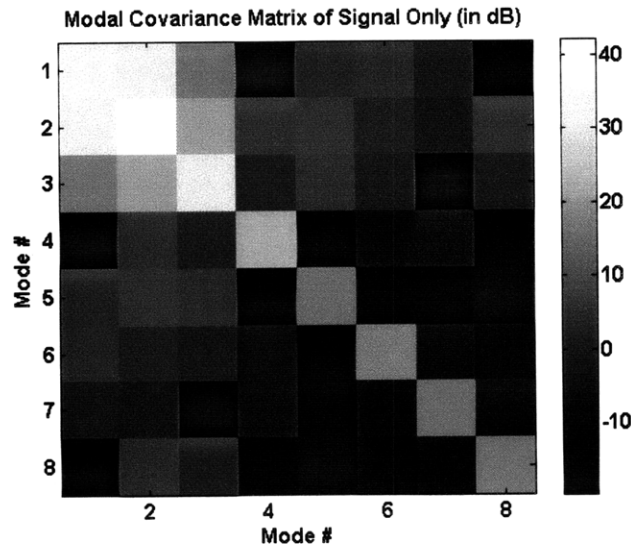

Figure 3-29: Mode covariance matrix of the signal at the receiver location.

Let us now turn our attention to the mean-square-error (MSE) and bit-error-rate (BER) filter performance metrics. Figures 3-30 through 3-32 show the mean-square-error in estimating a given mode using 1 to 9 hydrophones instead of 80. The hydrophones are added from the bottom to the top of the vertical array. The 0 dB level represents the power of the 80-element array's output, so always guessing the output is zero will lead to an MSE of 0 dB. The plots below show that the mode filters using mode 2 are able to accurately approximate the response of the 80-hydrophone array, but not using modes 1 or 3.

Figures 3-33 through 3-35 show the bit error rates for the overall communication system. Mode 2 was the strongest mode and the lowest error rates are obtained using this mode for processing. When filtering with mode 1 or 3, the filters often perform best when only 3 to 6 hydrophones are used, and then the performance decreases as

55

the number of hydrophones increases. One reason for this is because the mode filter is unable to accurately filter for a given mode with only 3 to 6 hydrophones, and energy from mode 2 leaks into the estimates of the other modes. As the number of hydrophones increases, energy from the strong second mode is filtered out better and the overall communication performance degrades. Furthermore, the reduced rank PI filter often performs worse as the number of hydrophones is increased. As was observed in the CGRV simulation, this is due to increased white noise sensitivity as additional singular values are included in the pseudoinversion. When using the strongest mode to filter with (mode 2), the PCML-MAP filter is able to do almost as well as the MAP filter which knows the signal and noise statistics a priori.
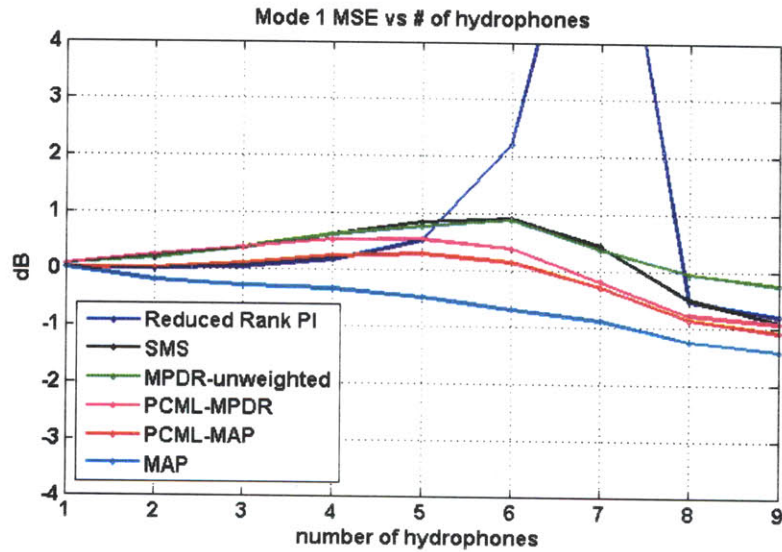

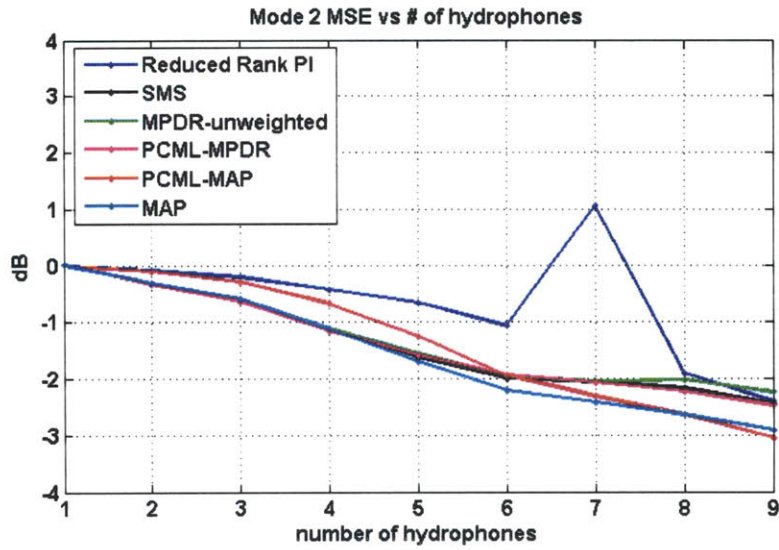Figure 3-30: Mean-square-error in estimating mode 1 with 9 hydrophones instead of 80.

Figure 3-31: Mean-square-error in estimating mode 2 with 9 hydrophones instead of 80.
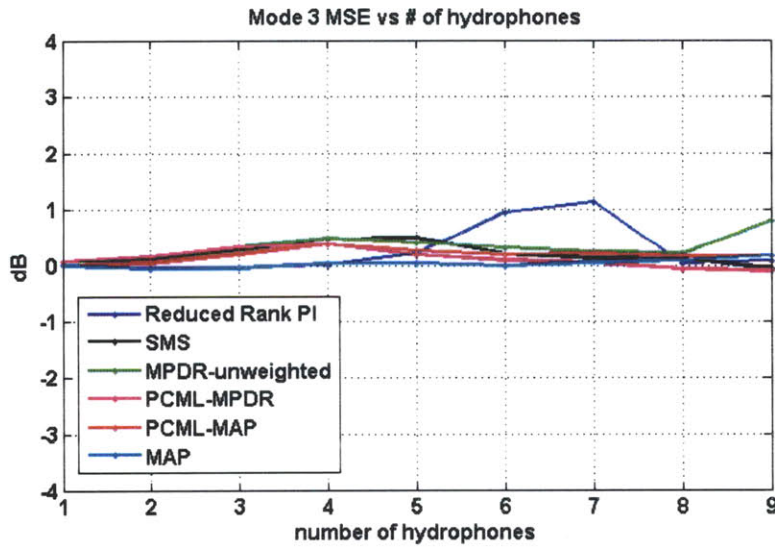

Figure 3-32: Mean-square-error in estimating mode 3 with 9 hydrophones instead of 80.
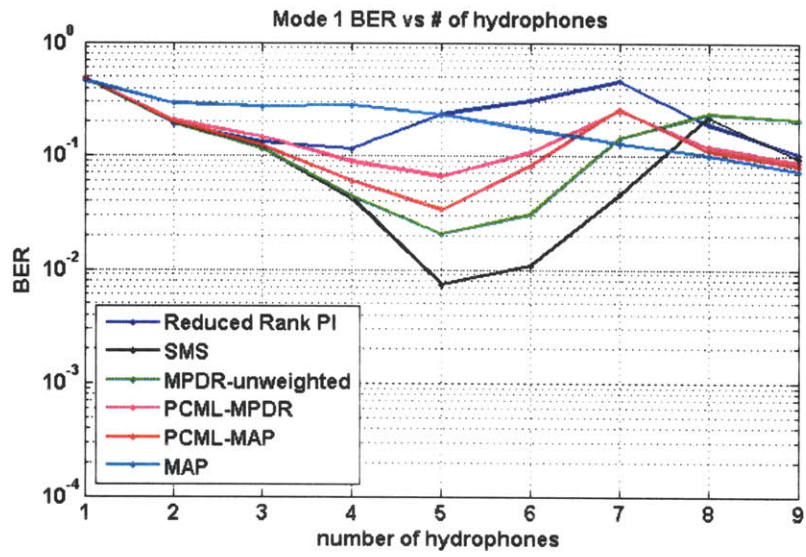
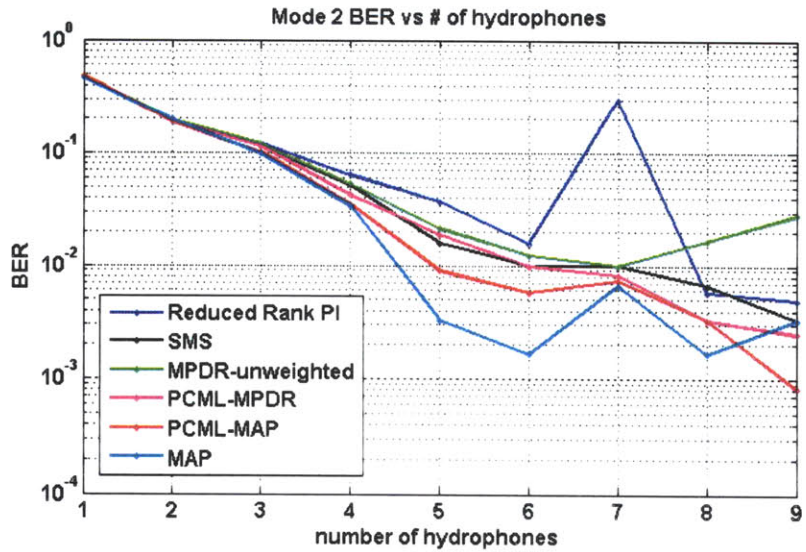Figure 3-33: Bit-error-rates when filtering with mode 1.

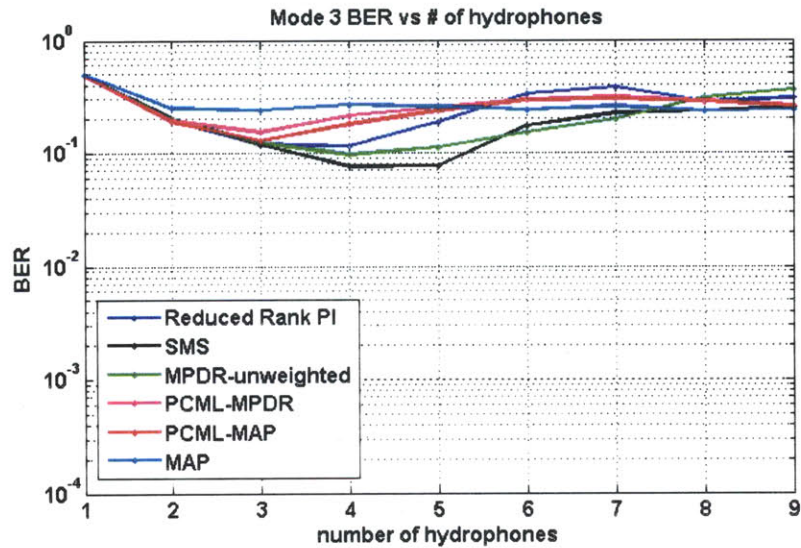Figure 3-34: Bit-error-rates when filtering with mode 2.

Figure 3-35: Bit-error-rates when filtering with mode 3.

3.6 Shallow Water 2006 Experiment Setup

The Shallow Water 2006 experiment was conducted along the continental shelf off the coast of New Jersey. The duration of the experiment analyzed in this thesis was that received from the sound source known as the Miami Sound Machine (MSM). It consisted of a single transmitter at a depth of 56 m in 75 m of water. The signal propagated 19.2 km along the continental shelf to a vertical array of receivers. The 203 Hz BPSK signal is analyzed here, which has a one-sided bandwidth of 50 Hz and a baud rate of 50.9 bits per second. A 127-bit maximum length sequence was repeatedly transmitted from this sound source [3]. Temperature and salinity sensors were located on and near the array which allowed the acoustic modes to be calculated as a function of time during the entire experiment. In this processing, however, the modes are assumed to be constant for each 90 second transmission interval. A map showing the location and bathymetry of the SW06 experiment from the technical report by Newhall et al [3] is found in figure 3-36.
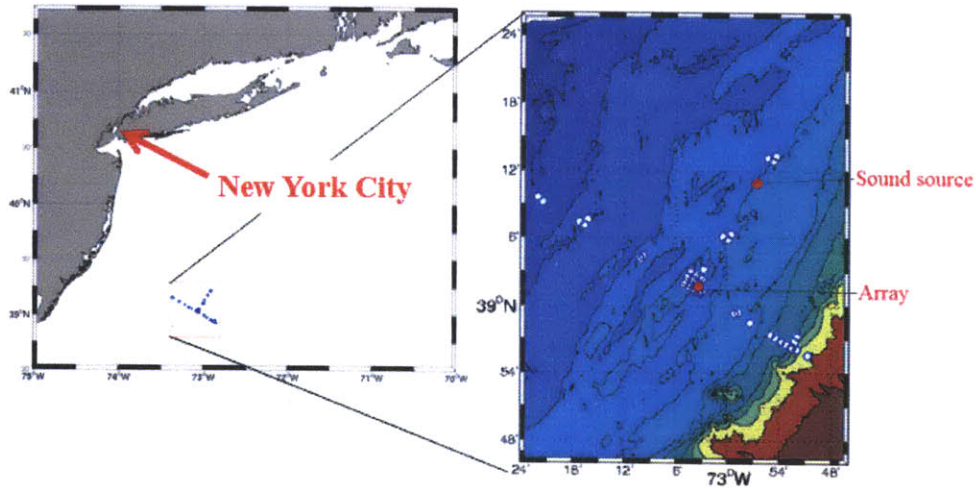
Figure 3-36: Shallow Water 2006 experiment area and bathymetry [3].

3.7 Shallow Water 2006 Data Processing Results

This section describes the processing results from applying the narrowband mode filters in broadband processing of the data collected during the Shallow Water 2006 experiment. Figures 3-37, 3-38, and 3-39 show the estimated impulse response on day 219, which was a typical time period where mode 2 was strongest. All of the mode filters generate a similar estimate of the channel impulse response, as was observed in the RAM simulations.

For comparison, figure 3-40 shows the impulse response estimates directly to three of the hydrophones. These impulse responses are more dispersed and less intense than the mode filtered response for mode 2. This highlights how mode filtering reduces the effective delay spread of the channel.
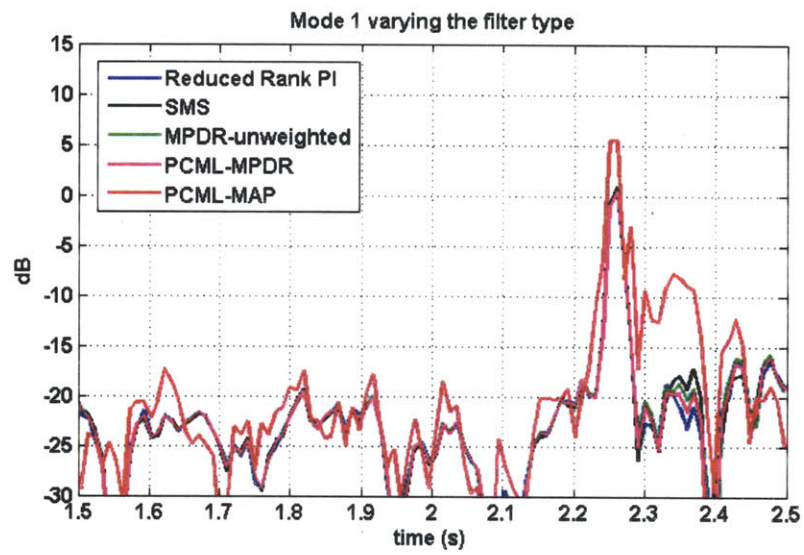
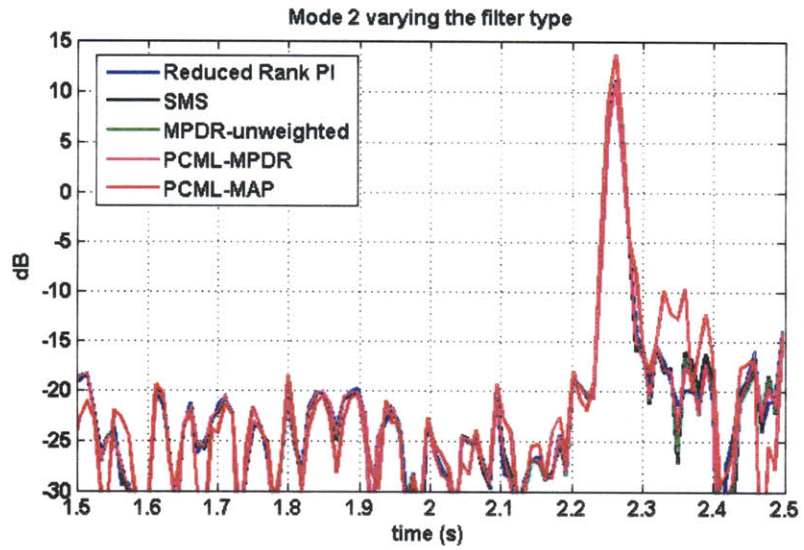Figure 3-37: SW06 impulse response estimate day 219, mode 1.

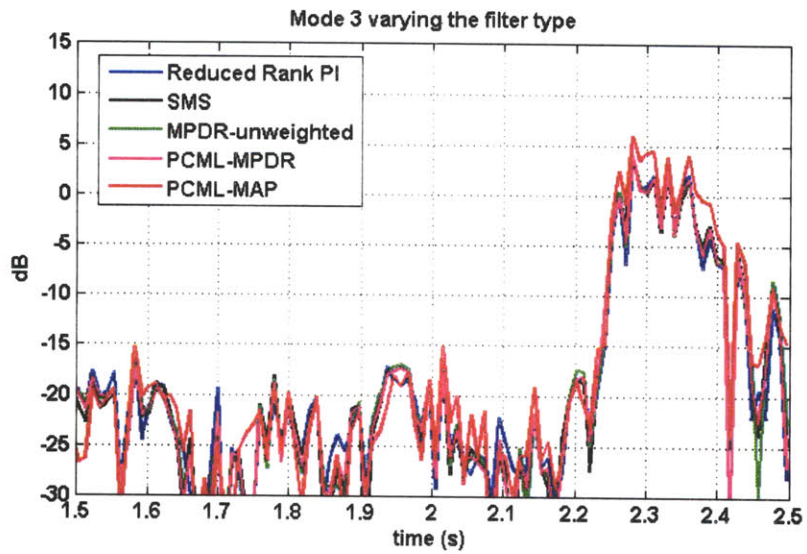Figure 3-38: SW06 impulse response estimate day 219, mode 2.

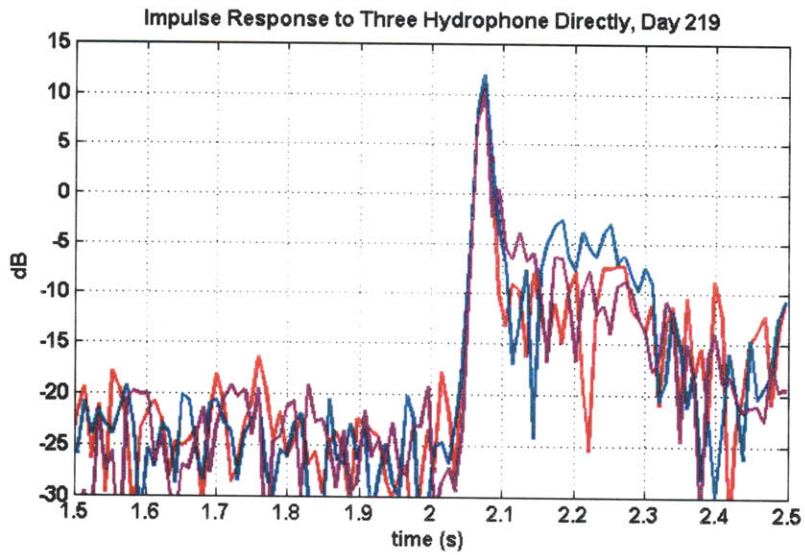Figure 3-39: SW06 impulse response estimate day 219, mode 3.

Figure 3-40: SW06 impulse response estimate day 219.

The following figures show the communication bit-error-rates for day 219. The hydrophones used are located at equally spaced depths of 17.25, 24.75, 32.25, 39.75, 47.25, 54.75, 62.25, 69.75, and 77.25 meters. The hydrophones were added from the bottom to the top of the array. The unweighted MPDR and SMS filters have the lowest BER out of all the filter types when filtering for mode 1. Since these filters likely do worse in estimating mode 1, energy from mode 2 leaks into the filter's estimate of mode 1 and this results in a lower BER relative to the other filters. Also, we again see the error spike for the reduced rank PI filter as the number of hydrophones is increased. This is because of increased sensitivity to white noise as more singular values are included in the pseudoinversion. Figure 3-41 shows a plot of the L2 norm of the PI weights for mode 2 as a function of the number of hydrophones. The larger the L2 norm of the filter weights, the smaller the array gain in the presence of spatially white noise, which results in worse filter performance. The spike in the L2 norm coincides with the spike in the BER curve for mode 2 in figure 3-43.
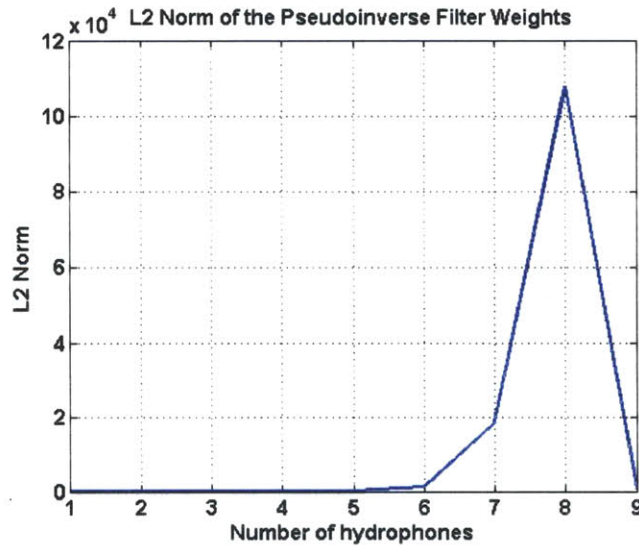

Figure 3-41: L2 norm of PI weights for mode 2 on day 219.
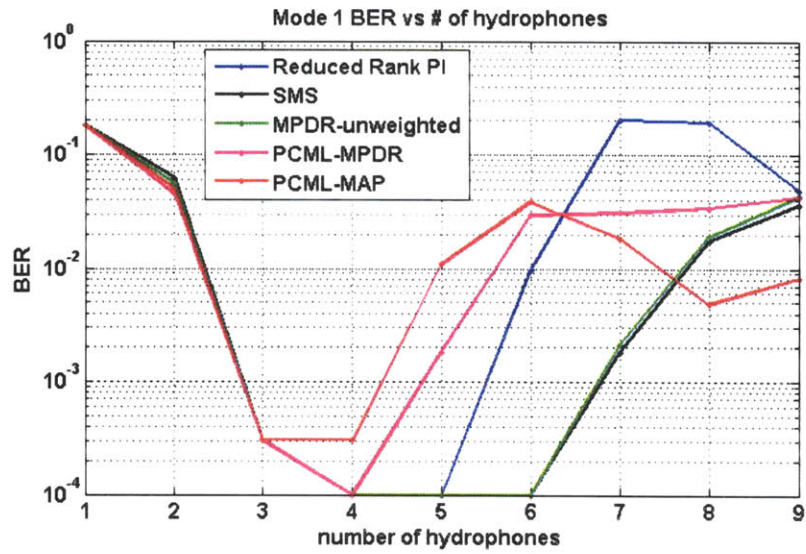
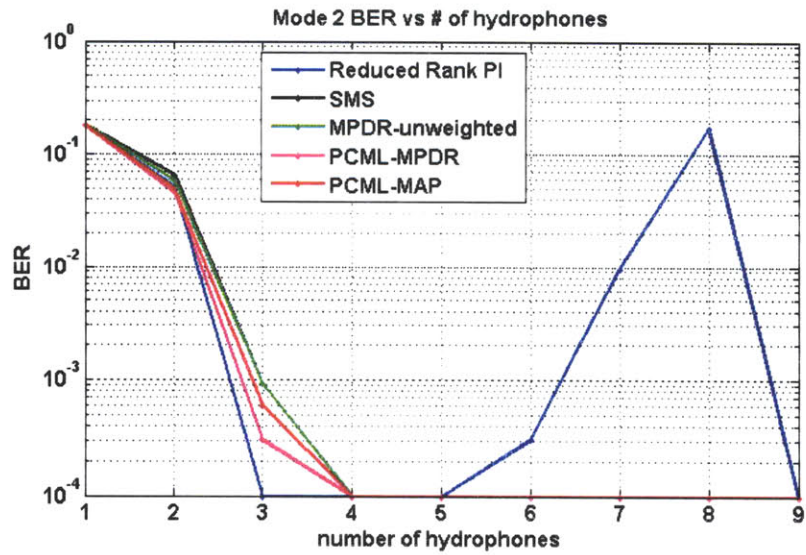Figure 3-42: SW06 bit-error-rates, day 219, mode 1.

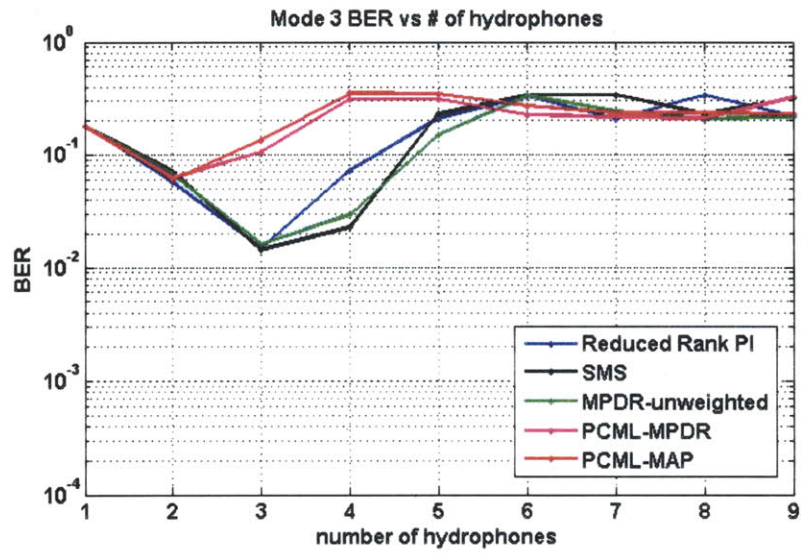Figure 3-43: SW06 bit-error-rates, day 219, mode 2.

Figure 3-44: SW06 bit-error-rates, day 219, mode 3.

Lastly, the following figures show the results from day 231 during an atypical transmission period when there was a large amount of internal wave activity [25]. Mode 1 is strongest mode during this time period. Figure 3-48 shows a plot of the impulse response directly to three of the hydrophones without mode filtering for comparison. As expected, the direct impulse response has a longer dispersion and is not as intense as the mode filtered version.

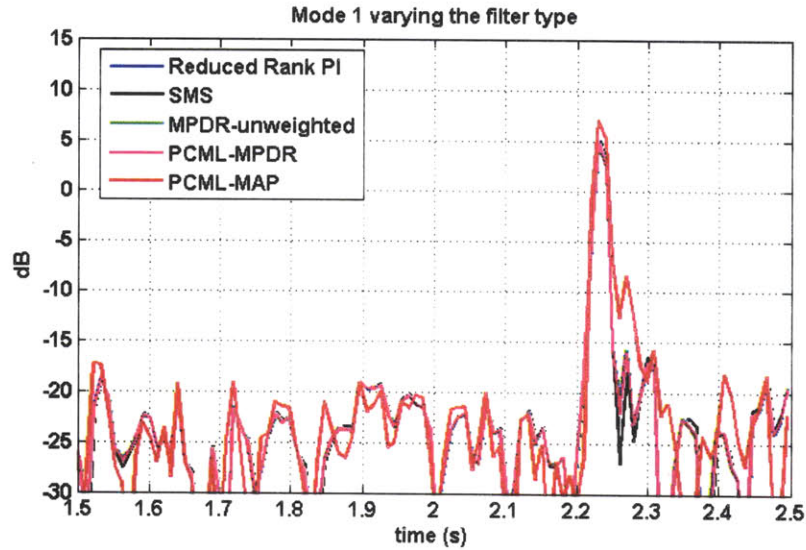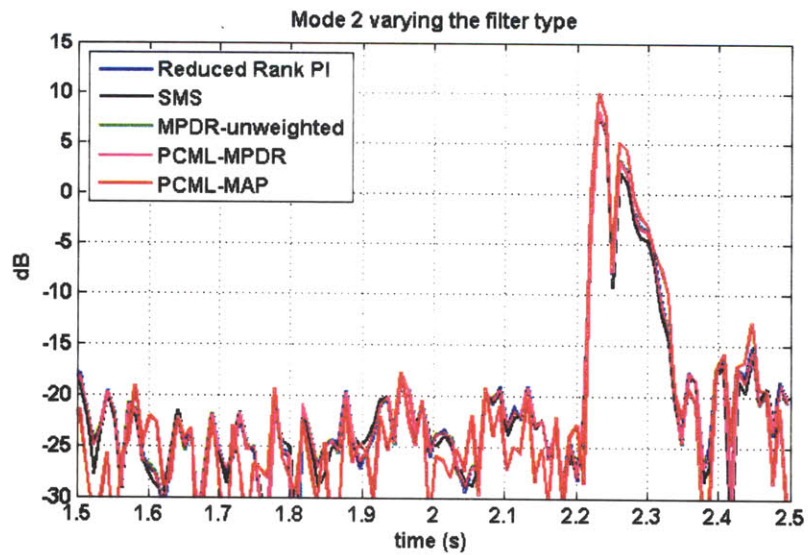Figure 3-45: SW06 impulse response estimate day 231, mode 1.

Figure 3-46: SW06 impulse response estimate day 231, mode 2.
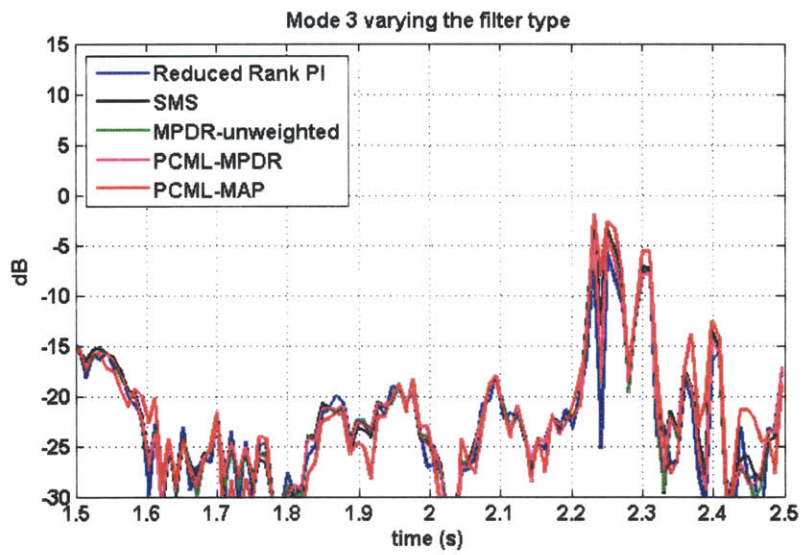
66

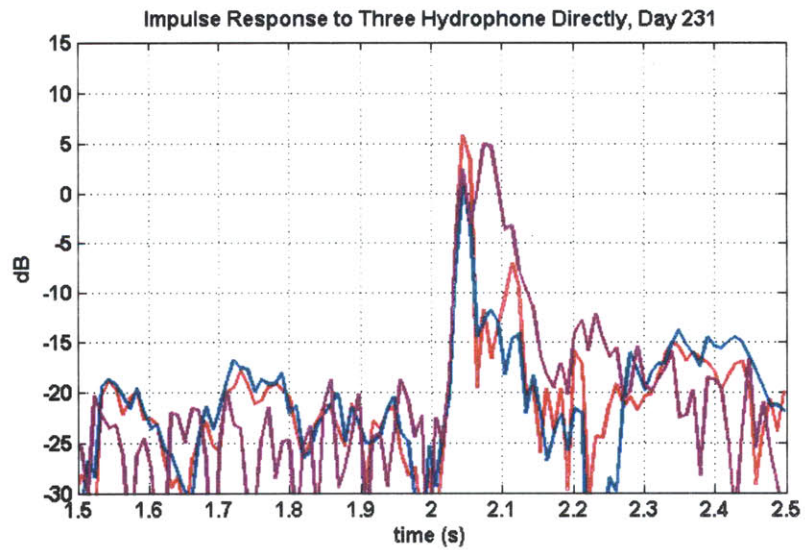Figure 3-47: SW06 impulse response estimate day 231, mode 3.

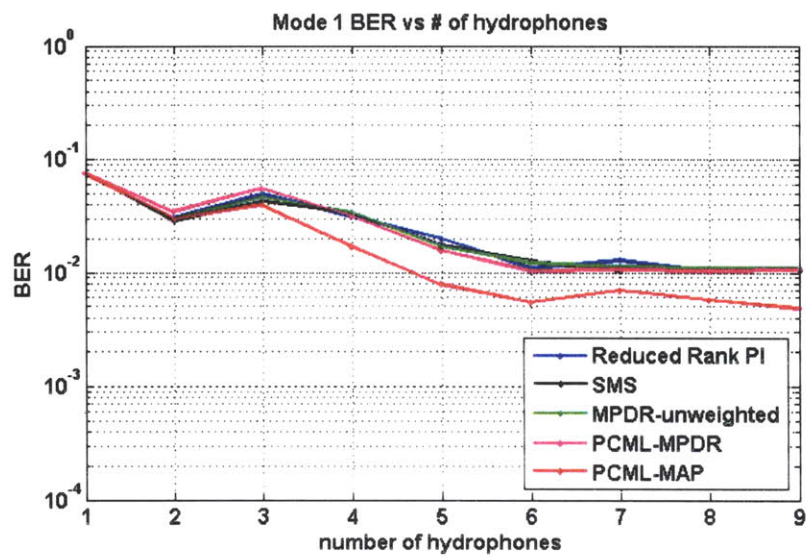Figure 3-48: SW06 impulse response estimate day 231.

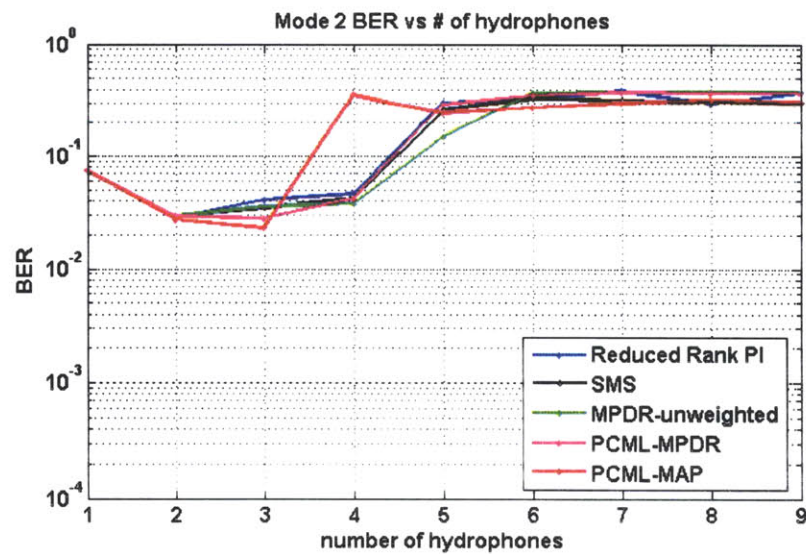Figure 3-49: SW06 bit-error-rates, day 231, mode 1.

Figure 3-50: SW06 bit-error-rates, day 231, mode 2.
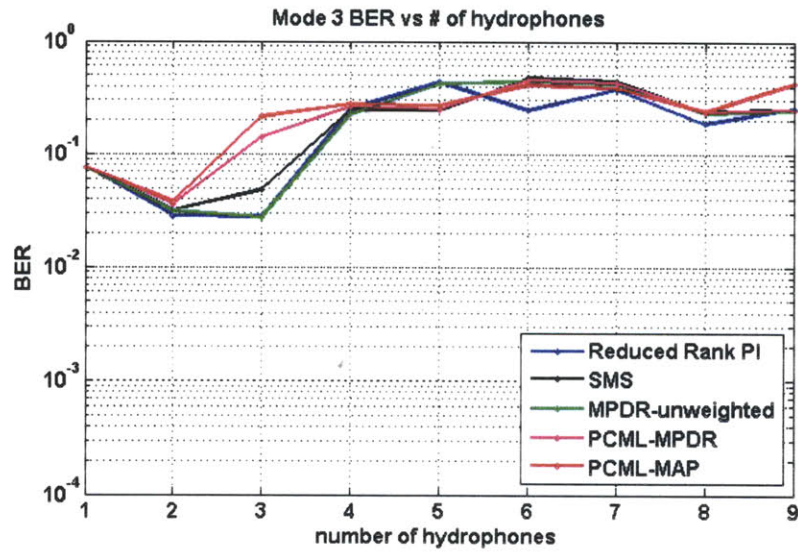
Figure 3-51: SW06 bit-error-rates, day 231, mode 3.

Chapter 4

Mode Equalization

The mode filtering results from the previous chapter only make use of one the energy contained in one particular mode of propagation at a time. This chapter proposes a linear equalizer for combining the outputs of several mode filters together so signal energy contained in several of the first few modes is used. The mode equalizer is compared with an equalizer directly applied to the received data at each hydrophone. An exponentially weighted recursive least squares (RLS) algorithm was used to update the equalizer weights and is discussed in [26, 27].

4.1 Equalizer Setup

The equalizer setup is shown in figure 4-1. $\mathbf{X}$ is the transmitted binary data which is passed through a noisy channel, $\mathbf{Y}$ is the received data, and $\hat{\mathbf{X}}$ is the hard decision estimate. Two versions of the equalizer were used. The first was a direct equalizer on the received data at each of the hydrophones and the second was an equalizer on the outputs of the mode filters for each mode. In the first case, the input data to the equalizer consists of the past N data points from each hydrophone stacked in a

71

vector. In other words, the input vector is

$$\mathbf{Y}[n] = \begin{bmatrix} \mathbf{y}_1 \\ \mathbf{y}_2 \\ \vdots \\ \mathbf{y}_{nh} \end{bmatrix} \tag{4.1}$$

where each $\mathbf{y}_i$ is a vector of the last N temporal samples from the ith hydrophone:

$$\mathbf{y}_i[n] = \begin{bmatrix} y_i[n] \\ y_i[n-1] \\ \vdots \\ y_i[n-N+1] \end{bmatrix} \tag{4.2}$$
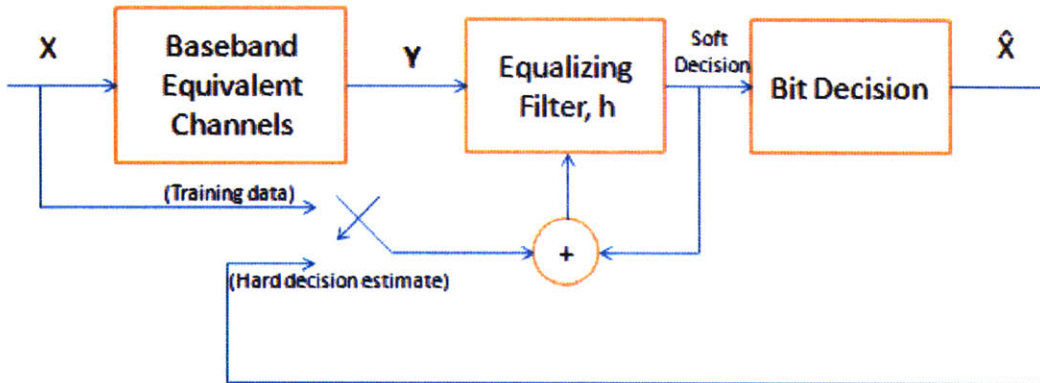
Figure 4-1: Equalizer setup.

In the second case of the mode equalizer, the input is instead a vector of the past N taps from the re-aligned output of each mode filter (for mode 1, 2, 3, etc). The mode filter outputs are re-aligned by introducing a delay into each output based on its offset relative to the offset of mode 1. This compensates for the different group velocities of each mode and allows for fewer temporal taps to be used. The offset is determined by the maximum of the circular cross-correlation of the filter output with

the transmitted m-sequence. **Y** is formed from the outputs of each mode filter and not from the hydrophone signals directly. An exponentially weighted RLS algorithm [27, 28] is used to update the filter coefficients. It is important to use enough modes in the equalizer to capture most of the transmitted energy, but beyond a certain point there is no additional benefit from including more modes. Using four modes in the equalizer was generally sufficient to capture enough of the signal energy. The number of hydrophones was varied prior to mode filtering and equalization, and these results are discussed in section 4.2.

The input data, **Y**, advances by two points for every one transmitted symbol, since the received data is downsampled to twice the bit rate. Three temporal taps were used on each channel (1.5 symbols) for both the mode equalizer and the direct equalizer. 200 bits of training data are sent before the equalizer switches to decision directed mode. The BER was very low for the cases studied, and so little error was introduced by using the decision directed mode. The transmitted data is cyclic every 127 bits since the same m-sequence was repeatedly sent. The coarse alignment of the transmitted data with the hard decision estimates was manually set to be that which produced the smallest soft decision error.

4.2   Equalizer Results

Figures 4-2 and 4-3 show the bit error rates and the steady-state soft decision error for the mode equalizers and the direct equalizer in the RAM simulation. The direct equalizer on the hydrophones performs best when fewer hydrophones are used, and the mode equalizer performs best when all 9 hydrophones are used. This is likely because the mode filters have a difficult time accurately estimating the response of a particular mode when few hydrophones are used, but do better as more are added. Also, there is a dimensionality advantage with the mode equalizer as the number of hydrophones is increased. The number of taps to be equalized remains fixed at the number of modes times the number of temporal taps per mode, whereas the direct equalizer must adjust the number of hydrophones times the number of temporal taps

per hydrophone. Four modes were used in the mode equalizer, whereas up to nine hydrophones were used in the direct equalizer.

Even though the mode filter that uses one hydrophone is unable to resolve any modes, it still produces an output that is different from its input because of the broadband processing performed on the signal, described in section 2.4. This is the reason for the performance difference between the direct equalizer and the mode equalizer for the case of one hydrophone.
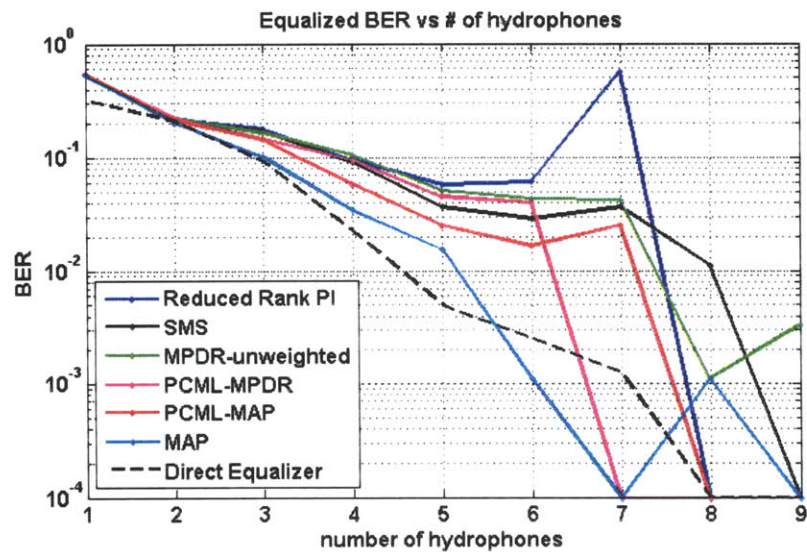

Figure 4-2: BER for RAM simulation with 3-tap equalizer.

Figures 4-4 and 4-5 show the BER and soft decision error when filtering with the mode equalizers and the direct equalizer for a typical day during the SW06 experiment. Both the mode equalizer and the direct equalizer do significantly better than straight decoding of a mode filter output does, and reach zero BER within the first few hydrophones. However, the direct equalizer on the hydrophones has a smaller soft decision error than the mode equalizer does.

Figures 4-6 and 4-7 show the results for day 231 during a period of abnormally high internal wave activity. The plots are similar to those from day 219. The reduced rank PI filter suffers from conditioning issues when 6 or 8 hydrophones are used.
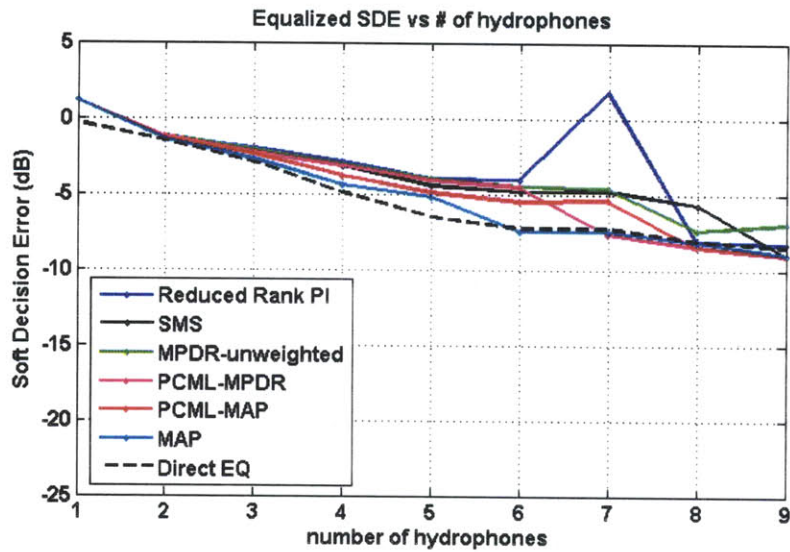
Figure 4-3: Soft Decision Error (in dB) for RAM simulation with 3-tap equalizer.
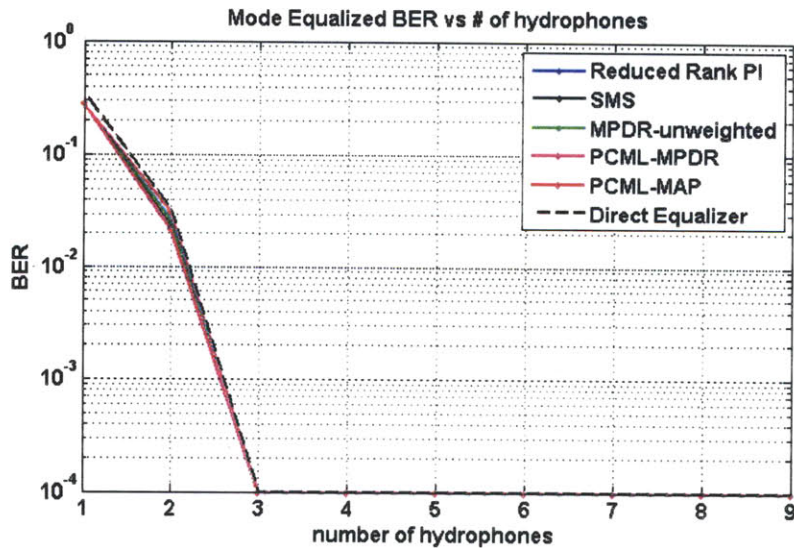

Figure 4-4: BER for SW06 Day 219 with 3-tap equalizer.
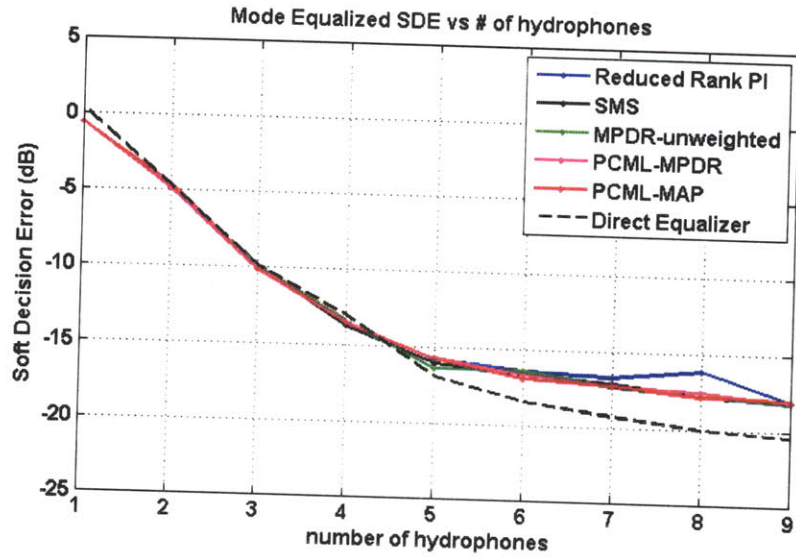
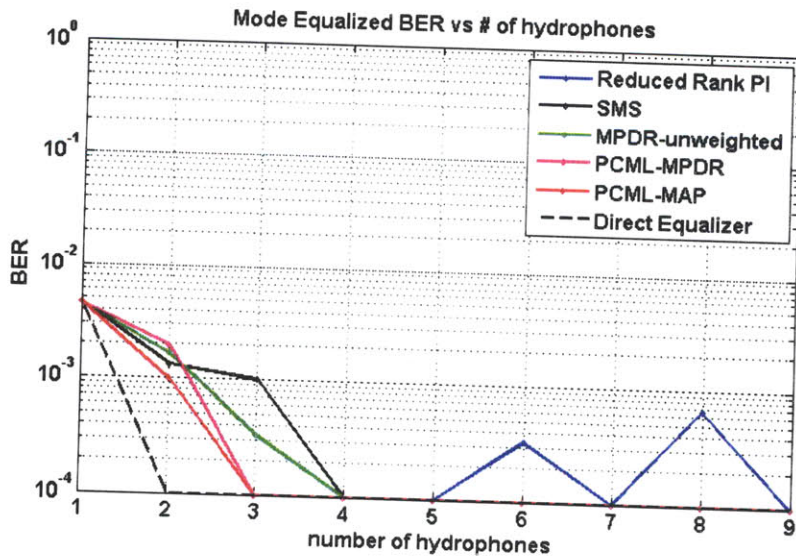Figure 4-5: Soft Decision Error (in dB) for SW06 Day 219 with 3-tap equalizer.

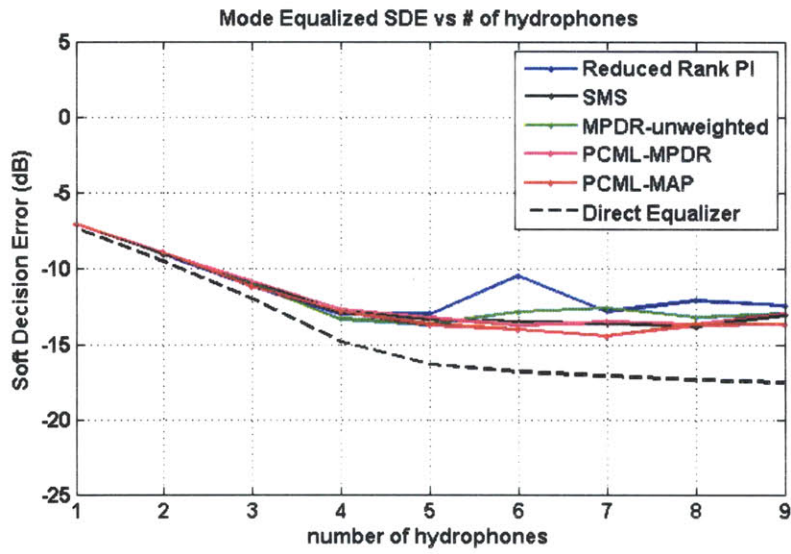Figure 4-6: BER for SW06 Day 231 with 3-tap equalizer.

76

Figure 4-7: Soft Decision Error (in dB) for SW06 Day 231 with 3-tap equalizer.

The forgetting factor (λ) in the RLS algorithm determines the effective window length of the data used for the equalizer [29]. Smaller forgetting factors represent shorter effective window lengths. A general rule of thumb is that the effective window length is approximately $\frac{1}{1-\lambda}$ samples for the exponentially weighted RLS algorithm. The following figures show the effect of the choice of λ on the soft decision error when filtering with 9 hydrophones. For the RAM simulation, the optimal forgetting factor is as large as possible since the channel is not time-varying. This results in a longer averaging window and a better estimate of the RLS inverse covariance matrix since more data snapshots are averaged in. For the SW06 data, the optimal forgetting factor is about 0.99, corresponding to an effective window length of about 100 samples. This suggests that, for the time scales on which we are adapting for this experiment, the channel is very slowly varying. In a rapidly fluctuating environment, the optimal forgetting factor would be much lower.
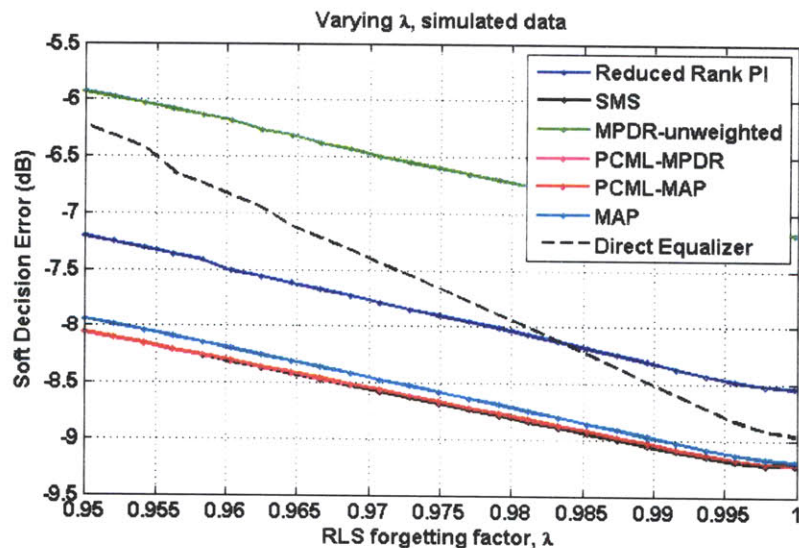

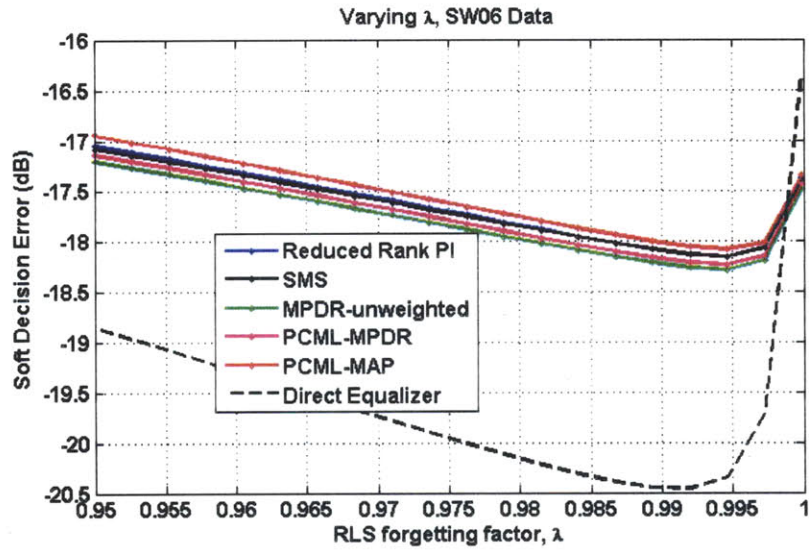Figure 4-8: Soft Decision Error as a function of λ, RAM simulation using all 9 hydrophones.

Figure 4-9: Soft Decision Error as a function of λ, SW06 Day 219 using all 9 hydrophones.
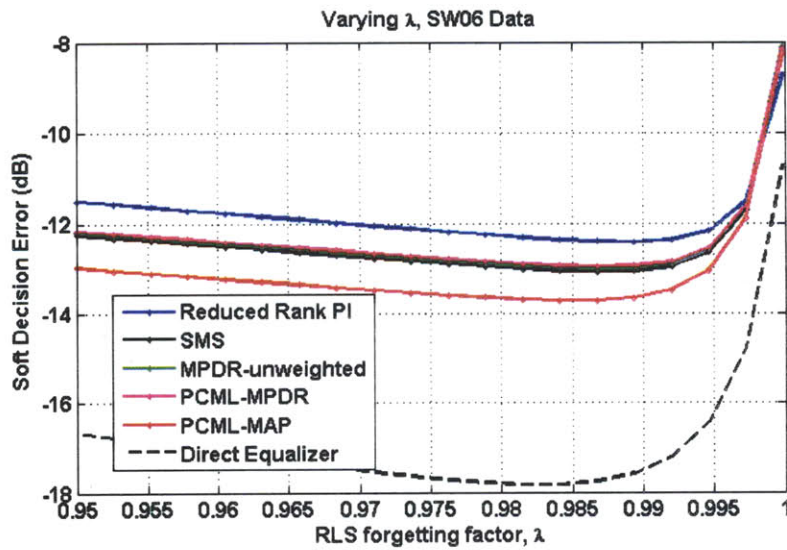
Figure 4-10: Soft Decision Error as a function of λ, SW06 Day 231 using all 9 hydrophones.

Lastly, figures 4-11, 4-12, and 4-13 show the condition number of the inverse covariance matrix used in the RLS algorithm as a function of the number of hydrophones used. For the direct equalizer on the hydrophones, the condition number increases as the number of hydrophones is increased. This is because the number of taps in the equalizer increases with the number of hydrophones, and since the effective number of data samples going into the covariance matrix remains constant, there are fewer snapshots with which to estimate the matrix. For the mode equalizer, the number of equalizer taps remains constant as hydrophones are added. We have a better estimate of the response of each mode as hydrophones are added, making the mode estimates more independent and the inverse covariance matrix better conditioned.

Figure 4-11: Inverse Covariance Matrix Condition Number, RAM Simulation.

Figure 4-12: Inverse Covariance Matrix Condition Number, Day 219.

Figure 4-13: Inverse Covariance Matrix Condition Number, Day 231.

Chapter 5

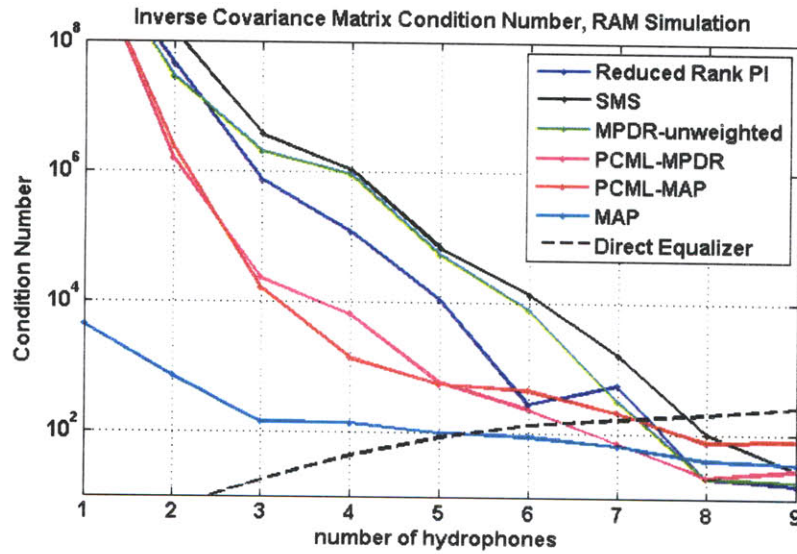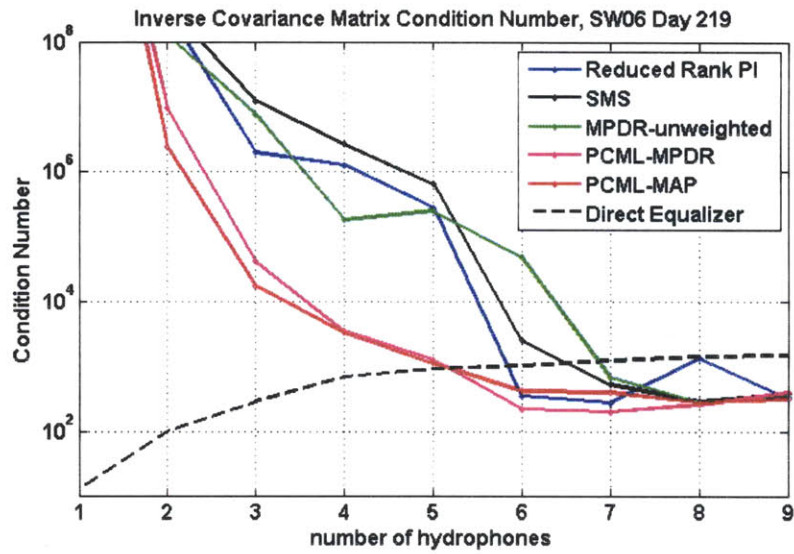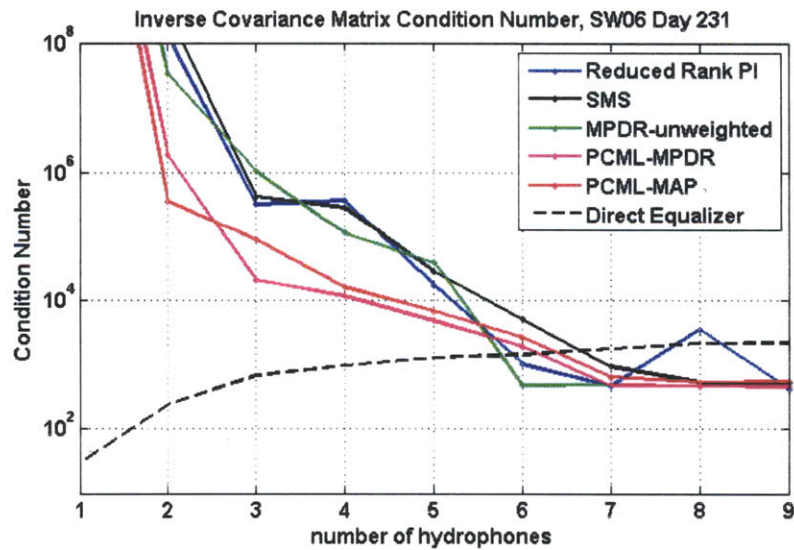# Conclusion

5.1 Summary

This thesis examined mode filtering as a preprocessing method for the equalization of underwater acoustic communication signals. A physically constrained, maximum likelihood method for adaptive mode filtering was presented. Chapter 2 reviewed acoustic normal modes in the context of an underwater communication system and presented the sampled mode shapes and pseudoinverse mode filters. The MAP mode filter requires statistics of the signal and noise to be known a priori. The MPDR filter was derived for the case where the spatial covariance matrix is known. The final section of chapter 2 presented the PCML method for covariance matrix estimation developed by Kraay [4] for the case of a spatial beamformer.

Chapter 3 applied Kraay's PCML method to the underwater acoustic pressure field model through the MPDR and MAP mode filters. A simulation similar to the one performed by Buck et al [1] was used to compare the performance of the adaptive MPDR and MAP mode filters with the PI and SMS filters. With the PCML method, the estimated covariance matrix converged in orders of magnitude fewer snapshots than the unconstrained sample covariance matrix did. Results from the RAM simulation and Shallow Water 2006 experimental data were used to compare the performance of the mode filtering techniques, and demonstrated the strong performance of the PCML-MPDR and PCML-MAP algorithms.

Chapter 4 applied a linear, decision directed equalizer to the outputs of several mode filters to make use of the signal energy contained in multiple modes. Because the mode filters reduce the delay spread of the channel and since there are typically less energetic modes than there are hydrophones, there are fewer degrees of freedom to equalize in the mode equalizer than in the direct equalizer.

5.2   Suggestions for Future Research

One area of future work is to better explore the noise model used in the PCML algorithm. It may be possible to use statistics from a noise recording before the transmission begins instead of assuming the noise is spatially white. The covariance matrix model would be modified as follows:

$$\mathbf{R} = \sum_{i=1...M} P(w, \psi_i)\psi_i\psi_i^H + \sigma^2\mathbf{R}_n \tag{5.1}$$

where $\mathbf{R}_n$ is the covariance matrix from a nearby noise recording, instead of the identity matrix.

Lastly, the mode processing techniques could be expanded if a vertical array of sound sources was available. With multiple vertically aligned transducers, it would be possible to use mode filtering algorithms to transmit orthogonal signals on several modes in a MIMO application. Instead of sending different signals on each transducer, different signals would be transmitted on each mode. Since the delay spread of a given mode is shorter than the overall delay spread of the channel, each signal transmitted on each mode would have a shorter delay spread than if all the modes were excited at once.

Bibliography

[1] J.C. Preisig J.R. Buck and K.E. Wage. A unified framework for mode filtering and the maximum a posteriori mode filter. *J. Acoust. Soc. Am.*, 103(4):1813–1824, 1998.

[2] A.L. Kraay and A.B. Baggeroer. A physically constrained maximum-likelihood method for snapshot-deficient adaptive array processing. *IEEE Transactions on Signal Processing*, 55(8):4018–4063, 2007.

[3] A. Newhall et al. Acoustic and oceanographic observations and configuration information for the whoi moorings from the sw06 experiment. *Technical Report, Woods Hole Oceanographic Institution*, 2007.

[4] A.L. Kraay. Physically constrained maximum-likelihood method for snapshot-deficient adaptive array processing. Master's thesis, Massachusetts Institute of Technology, 2003.

[5] A.B. Baggeroer K.E. Wage and J.C. Preisig. Modal analysis of broadband acoustic receptions at 3515-km range in the north pacific using short-time fourier techniques. *J. Acoust. Soc. Am.*, 113(2):801–817, 2003.

[6] M.B. Porter F.B. Jensen, W.A. Kuperman and H. Schmidt. *Computational Ocean Acoustics*. Springer-Verlag New York, Inc., 2000.

[7] C.W. Miller C. Chiu and J.F. Lynch. Optimal modal beamforming of bandpass signals using an undersized sparse vertical hydrophone array: Theory and a

shallow-water experiment. *IEEE Journal of Oceanic Engineering*, 22(3):522–532, 1997.

[8] G. Strang. *Linear Algebra and Its Applications*. Thompson Brooks/Cole, fourth edition, 2006.

[9] H.L. Van Trees. *Optimum Array Processing, Part IV of Detection, Estimation, and Modulation Theory*. Wiley-Interscience, 2002.

[10] B.D. Vanveen and K.M. Buckley. Beamforming: A versatile approach to spatial filtering. *IEEE ASSP Magazine*, 1988.

[11] J. Capon. High-resolution frequency-wavenumber spectrum analysis. *Proceedings of the IEEE*, 57(8):1408–1418, 1969.

[12] O.L. Frost. An algorithm for linearly constrained adaptive array processing. *Proceedings of the IEEE*, 60:926–935, 2003.

[13] D.H. Brandwood. A complex gradient operator and its application in adaptive array theory. *IEEE Proceedings*, 130(1):11–16, 1983.

[14] J.R. Guerci. *Space-Time Adaptive Processing for Radar*. Artech House, Inc., 2003.

[15] D. Tse and P. Viswanath. *Fundamentals of Wireless Communication*. Cambridge University Press, 2005.

[16] J.R. Guerci. Theory and application of covariance matrix tapers for robust adaptive beamforming. *IEEE Trans. Signal Process.*, 47(4):977–985, 1999.

[17] F.J. Harris. On the use of windows for harmonic analysis with the discrete fourier transform. *Proceedings of the IEEE*, 66(1):51–83, 1978.

[18] J.R. Buck R.W. Schafer and A.V. Oppenheim. *Discrete Time Signal Processing*. Prentice Hall, second edition, 1999.

[19] G.H. Golub and C.F. Van Loan. An analysis of the total least squares problem. *SIAM Journal on Numerical Analysis*, 17(6):883–893, 1980.

[20] W.A. Kuperman and F. Ingenito. Spatial correlation of surface generated noise in a stratified ocean. *J. Acoust. Soc. Am.*, 62(6):1988–1996, 1980.

[21] P.M. Pardalos and M.G. Resende. *Handbook of Applied Optimization.* Oxford University Press, 2002.

[22] *Ram Matlab Code.* http://909ers.apl.washington.edu/twiki/bin/view/Main/RamMatlabCode. retrieved August 8, 2008.

[23] S.W. Golomb and G. Gong. *Signal Design for Good Correlation: For Wireless Communication, Cryptography, and Radar.* Cambridge University Press, 2005.

[24] K.E. Wage. *Broadband Modal Coherence and Beamforming at Megameter Ranges.* PhD thesis, Massachusetts Institute of Technology and Woods Hole Oceanographic Institution, 2000.

[25] J.C. Preisig A.K. Morozov and J.C. Papp. Modal processing for acoustic communications in shallow water experiment. *J. Acoust. Soc. Am.*, 124(3):177–181, 2008.

[26] J.G. Proakis and M. Salehi. *Digital Communications.* McGraw Hill, fifth edition, 2008.

[27] M.H. Hayes. *Statistical Digital Signal Processing and Modeling.* John Wiley and Sons, 2002.

[28] G.V. Moustakides. Study of the transient phase of the forgetting factor rls. *IEEE Transactions on Signal Processing*, 45(10):2468–2467, 1997.

[29] S. Haykin. *Adaptive Filter Theory.* Prentice Hall, 2001.