Blind deconvolution in multipath environments and extensions to remote source localization

By

Shima Hossein Abadi

A dissertation submitted in partial fulfillment
of the requirement for the degree of
Doctor of Philosophy
(Mechanical Engineering)
In the University of Michigan
2013

Doctoral Committee:

Professor David R. Dowling, Chair
Professor Karl Grosh
Assistant Professor Eric Jonhsen
Research Scientist Aaron M. Thode, Scripps Institute of Oceanography
Associate Professor Gregory H. Wakefield

ACKNOWLEDGEMENTS

I would like to sincerely thank my research advisor, Professor David Dowling for his continuous support, patience, and guidance throughout this work. He gave me the freedom to explore on my own. At the same time, he helped me when I had any questions.

I would also thank my dissertation committee members, Professor Karl Grosh, Professor Eric Johnsen, Professor Gregory Wakefield, and Doctor Aaron Thode for agreeing to be on my dissertation committee and providing many valuable comments that improved this dissertation. Especially, I thank Doctor Thode for his great support during my summer 2012 visit to Scripps Institute of Oceanography.

I would like to thank my family and friends for their love and support. I am grateful to my mother who always encouraged me and my husband for his endless love and support during these years.

I would like to thank the U.S. Office of Naval Research for their financial support.

Table of Contents

LIST OF TABLES

LIST OF FIGURES

ABSTRACT

In the ocean, the acoustic signal from a remote source recorded by an underwater hydrophone array is commonly distorted by multipath propagation. Blind deconvolution is the task of determining the source signal and the impulse response from array-recorded sounds when the source signal and the environment's impulse response are both unknown. Synthetic time reversal (STR) is a passive blind deconvolution technique that relies on generic features (rays or modes) of multipath sound propagation to accomplish two remote sensing tasks. 1) It can be used to estimate the original source signal and the source-to-array impulse responses, and 2) it can be used to localize the remote source when some information is available about the acoustic environment. The performance of STR for both tasks is considered in this thesis.

For the first task, simulations and underwater experiments (CAPEx09)[1] have shown STR to be successful for 50 millisecond chirp signals with a bandwidth of 1.5 to 4.0 kHz broadcast to source-array ranges of 100 m to 500 m in 60-m-deep water. Here STR is successful when the signal-to-noise ratio is high enough, and the receiving array has sufficient aperture and element density so that conventional delay-and-sum beamforming can be used to distinguish ray-path-arrival directions. Also, an unconventional beamforming technique (frequency-difference beamforming) that manufactures frequency differences from the recorded signals has been developed. It allows STR to be successful with sparse array measurements where conventional beamforming fails. Broadband simulations and experimental data from the focused acoustic field

---

[1] Experimental data provided by Dr. Daniel Rouseff of the Applied Physics Laboratory of the University of Washington.

experiment (FAF06)[2] have been used to determine the performance of STR when combined with frequency-difference beamforming when the array elements are nearly 40 signal-center-frequency wavelengths apart. The results are good; the cross-correlation coefficient between the source-broadcast and STR-reconstructed-signal waveforms for the simulations and experiments are 98% and 91-92%, respectively.

In addition, the performance of frequency-difference beamforming and conventional beamforming has been simulated for random sparse arrays. These simulation results indicate that frequency-difference beamforming can determine the array-to-source direction when conventional beamforming cannot. However, extension of the frequency-difference concept to frequency-sum beamforming does not yield a robust beamforming technique.

For the source localization task, the STR-estimated impulse responses may be combined with ray-based back-propagation simulations and the environmental characteristics at the array into a computationally efficient scheme that localizes the remote sound source. These localization results from STR are less ambiguous than those obtained from conventional broadband matched field processing in the same bandwidth. However, when the frequency of the recorded signals is sufficiently low and close to modal cutoff frequencies, STR-based source localization may fail because of dispersion in the environment. For such cases, an extension of mode-based STR has been developed for sound source ranging with a vertical array in a dispersive underwater sound channel using bowhead whale calls recorded with a 12-element vertical array (Arctic 2010)[3]. Here the root-mean-square ranging error was found to be 0.31 km from 18 calls with acoustic path lengths of 6.5 to 24.5 km.

---

[2] Experimental data provided by Dr. Heechun Song of the Scripps Institution of Oceanography.

[3] Experimental data provided by Dr. Aaron Thode of the Scripps Institution of Oceanography.

CHAPTER 1

Introduction

1.1 Blind deconvolution and Synthetic Time Reversal

1.1.1 Blind deconvolution

The acoustic signal from a remote source recorded by an underwater hydrophone array is commonly distorted by multipath propagation. Such recordings are the convolution of the source signal and the impulse response of environment at the time of signal transmission. Blind deconvolution is the name given to the task of determining the source signal and the impulse response from array-recorded sounds when the source signal and the environment's impulse response are both unknown. In general, blind deconvolution is ill posed since many possible signal and impulse-response pairs are mathematically possible for a single set of array recordings. Thus, additional information or assumptions are needed to reduce the solution space, and thereby produce unique – and hopefully correct – results.

Blind deconvolution has applications in many research areas such as image processing, radar, and underwater acoustics which are presented below.

- **Blind deconvolution in image processing**

Blind deconvolution is also the name given to a variety of processes for improving digital images. When an image is imperfect (for example not fully focused), experience shows that some information is needed to successfully restore the image. Regular linear and non-linear deconvolution techniques utilize a known intensity point spread function (PSF) which is estimated from the image.

Such deconvolution is performed for image restoration in many applications. For example, blind deconvolution techniques have been used in image processing to reconstruct the original scene from a degraded observation (Kundur & Hatzinakos, 1996). Blind color image deconvolution has been developed to recover edges in color images and reduce color artifacts (Chen, He, & Yap, 2011). Another application of blind deconvolution involves estimating the frequency response of a two-dimensional spatially invariant linear system through which an image has been passed and blurred (Cannon, 1976). Blind image deconvolution has also been used to locate quantum-dot (q-dot) encoded micro-particles in three-dimensional images of an ultra-high density 3-D microarray (Sarder & Nehorai, 2008). It also has been applied to medical ultrasound imaging to recover diagnostically important image details obscured due to the resolution limitations (Michailovich & Adam, 2005). In ultrasonic image processing applications (Taxt and Strand 2001, Yu *et al*. 2012), the goal of blind deconvolution is to enhance image (signal) quality by correcting for an imperfect point spread function. Here the number of receiving elements (i.e. the number of pixels) may greatly exceed the number of temporal samples – perhaps just a single image. Blind image deconvolution techniques also have applications in astronomy in order to recover object and point spread function information from noisy data (Stuart & Christou, 1993).

For this dissertation, the goal of blind deconvolution is similar – improving signal quality – but the form of the input data is different; the number of receiving transducers N (a countable number) is typically much less than the number of temporal signal samples (thousands or even millions). The work reported in this dissertation differs from image-based applications of blind deconvolution in three ways: (i) the primary independent variable is time (not space or angle), (ii) the form of the temporal transfer function (the equivalent of the PSF) may be entirely unknown, and (iii) the duration of this transfer function may exceeds that of the signal. The equivalent situation in image processing would necessitate reconstruction of the intended image using information recorded at vertical or horizontal locations more than an image-height or image-width away.

- **Blind deconvolution in radar**

In recent radar work, blind deconvolution has been pursued for improving the range estimation possible by object restoration from the data observations (Jason, Richard, & Stephen, 2010). Interestingly, the emphasis of this radar effort is closely aligned with that of the thesis investigation proposed here. Three-dimensional (3D) FLASH laser radar (LADAR) is a pulsed radar system for both imaging and ranging. It produces a time sequence of two-dimensional (2D) images due to a fast range gate resulting in a 3D data cube of spatial and range scene data with excellent range resolution. The basic idea is to process the data in the spatial dimensions (x, y) while improving ranging performance in the time dimension (z). The algorithm presented in this article is powerful in that it can perform blind deconvolution via recursive image processing in situations with no extra information about the PSF. This methodology relies on the knowledge that the target produces a waveform peak in the detected returns. However, this algorithm assumes the optimized PSF is the same throughout a data cube, and it involves optimization and is computationally expensive. For comparison, the blind deconvolution method described in this dissertation does not require iteration or optimization, and it can recover different impulse responses for different spatial locations.

- **Blind deconvolution in underwater acoustics**

Several blind deconvolution techniques for underwater acoustics have been developed. In particular, in underwater applications, blind deconvolution involves using N receiving-array recordings to estimate $N + 1$ waveforms: N source-to-receiver transfer-function waveforms, and one source-signal waveform. Thus, a successful technique for blind deconvolution must incorporate additional information to reach unique and correct results. In past blind deconvolution efforts, this extra information has been developed from: Monte-Carlo optimization and a well-chosen cost function (Smith and Finette 1993), additional measurements from a known source (Siderius *et al*. 1997), an adaptive super-exponential algorithm (Weber & Bohme, 2002), higher order statistics (Broadhead *et al*. 2000), information criteria (Xinhua *et al*. 2001), adaptive algorithms (Sibul *et al.* 2002), time-frequency analysis (Martins *et al*. 2002), multiple convolutions (Smith 2003), an assumption about the probability density function of the signal (Roan *et al*. 2003), knowledge of statistical properties of acoustic Green's functions for enhancing the detection and classification performance of active and passive sonar systems (Chapin, Ioup, Ioup, & Smith, 2001), and a least-squares criterion (Zeng *et al*. 2009). The blind deconvolution technique reported in this dissertation does not need any extra information, additional measurements, and iterations.

Although the ill-posed nature of blind deconvolution problems is central, there are other limitations for blind deconvolution as well. One of them is noise. Blind deconvolution methods that work well at high signal-to-noise ratio (SNR) may struggle in the presence of noise (Broadhead & Pflug, 2000). The other limitation is Green's-function (or transfer-function) mismatch. In situations where the Green's function structure is simple (e.g., direct arrival and surface reflection), single-channel deconvolution may provide satisfactory results. When multipath effects (due to interaction with layered bottom sediments for example) are present, it may be difficult to get a good source signal estimate (Broadhead, Field, & Leclere, 1993).

1.1.2 Synthetic Time Reversal

Synthetic time reversal (STR, also known as artificial time reversal, ATR) is a relatively simple technique that may be attractive for performing blind deconvolution in underwater sound channels in the bandwidth of the source signal to 1) determine the original source signal and the source-to-array-element impulse responses, and 2) localize the remote source (Abadi *et al*. 2012). For the first of these two tasks, the additional information used in STR to uniquely estimate the source signal and the environment's impulse response is drawn from the generic characteristics of the acoustic modes (Sabra & Dowling, 2004) or the acoustic rays (Sabra, Song, & Dowling, 2010) that convey sound from the source to the array. Once mode- or ray-based propagation is assumed, no additional assumptions are needed about the form or statistics of the source signal or the environment's impulse response. Furthermore, STR does not require parametric searches or optimization; its computational burden is only marginally greater than forward and inverse fast-Fourier transformation of the recorded signals. When the first-task effort is successful, the second task becomes possible when basic environmental characteristics are known at the receiving array, and the range-dependence of the underwater environment is mild.

1.2 Beamforming

Beamforming techniques are commonly used in array signal processing to find the ray-path-arrival directions (Steinberg 1976, Ziomek 1995). In general, beamforming is a spatial filtering process intended to highlight the propagation direction(s) of array-recorded signals.

When a remote source is near enough to the array or when the acoustic environment causes predictable reflections and scattering – for example in a known sound channel – simple beamforming may be extended to matched-field processing (MFP) and the location of the remote source may be determined (see Jensen *et al*. 1994). Minimum Variance Distortionless Response (MVDR) is an adaptive beamforming technique to suppress side lobes and enhance the spatial resolution of beamforming (Jensen *et al*. 1994).

In this dissertation, beamforming is used to determine the propagation direction(s) of array-recorded sound(s). However, in chapter 5, beamforming is used to localize a single sound source in the near field of a linear array, and the resulting output can be considered representative of the acoustic imaging point spread function of the array at the location of the source.

- **Beamforming in ultrasound imaging**

Specialized beamforming techniques have been developed for applications in medical ultrasound imaging to improve image quality. Conventional delay-and-sum beamforming is a traditional beamforming technique for ultrasound imaging (Karaman *et al*. 1995). Here the spatial filtering is linear because the received field is filtered using weights that depend only on environmental factors and the receiving array's geometry. More recent research has shown that the MVDR beamforming can improve image quality compared to delay-and-sum beamforming (Synnevag *et al*. 2007, Holfort *et al*. 2009). In this case the spatial filtering is nonlinear because the received field is filtered using weights that depend on environmental factors, the receiving array's geometry, and the received field.

1.3 Acoustic source localization

Acoustic source localization is a task of locating a sound source given measurement of the sound field. Remote source localization is one of continuing interest in a variety of sonar applications.

There are many techniques for acoustic source localization. Some techniques such as match-field processing (MFP) match the measured field at the array with simulated replicas of the field for all possible source locations. Some traditional techniques use the time difference of arrivals at the receiving array. Statistical analysis can also be used for acoustic source localization. For instance, a maximum a posteriori estimation method is able to estimates source location and spectral characteristics of multiple sources in underwater environments via Gibbs sampling (Michalopoulou, 2006). A Bayesian formulation is another method to find simultaneous localization of multiple acoustic sources when properties of the ocean environment are poorly known (Dosso & Wilmut, 2011). The relative delay between two (or more) microphone signals for the direct sound can be used to find the position of an acoustic source in a room (Benesty, 2000).

In the last three decades, a variety of match-field processing (MFP) techniques have been shown to localize the sound source successfully when sufficient environmental information is available. MFP calculations were first conducted using normal modes (Bucker, 1976). The review article by Baggeroer, Kuperman, & Mikhalevsky (1993) and the tenth chapter in Jensen, Kuperman, Porter, & Schmidt (1994) provide relevant background. The capability of the different MFP schemes to localize an unknown remote source under conditions of environmental mismatch is presented in Porter & Tolstoy (1994). More recently, the maximum a posteriori (MAP) estimator for MFP has been reported in Harrison, Vaccaro, & Tufts (1998). Matched-field source localization using data-derived modes can be used to estimate both the wave numbers and bottom properties (Hursky, Hodgkiss, & Kuperman, 2001). At higher frequencies, the broadband match-field processing method presented in Hursky *et al*. (2004) is able to localize a remote sound source by cross-correlating measured and modeled impulse response functions and selecting the maximum cross-correlation peak. The coherent match-field processing method proposed in this dissertation is a variation of Hursky's algorithm with one difference: the actual impulse response was not measured; it was estimated by STR.

MFP has also been extended to estimating environmental parameters and the remote source location simultaneously, a technique called focalization (Collins & Kuperman, 1991). Other relevant geoacoustic inversion schemes for using waterborne acoustic propagation data to determine the geoacoustic properties of the sea bottom are provided in Herman and Gerstoft (1996), and Siderius and Hermand (1999). Notably, the MBMF technique can also be used for geoacoustic inversion (Hermand, 1999). Similarly, simultaneous estimation of the local sound speed profile and localization of a target on the ocean bottom in front of the host vehicle is possible using the Adaptive Bathymetric Estimator (ABE) (Cousins, 2005). In addition, Source localization based on eigenvalue decomposition is described in (Benesty, 2000), and source localization with horizontal arrays in shallow water is reported in (Bogart & Yang, 1994). Another acoustic source localization method is a maximum likelihood (ML) acoustic source location estimation which uses acoustic signal energy measurements taken at individual sensors of a wireless sensor network to estimate the locations of multiple acoustic sources (Xiaohong & Yu-Hen, 2005).

The source localization effort presented here is not fully blind; it does rely on knowledge of the environmental characteristics at the receiving array to back propagate impulses along a handful of acoustic rays emanating from the array. However, it does not involve extensive field calculations, and is more robust that matched-field techniques since it does not require precise phase-matching to localize the source.

1.4 Dissertation motivation and organization

This thesis presents the results of a technique (synthetic time reversal, STR) for reconstructing the source signal in an almost unknown multipath environment. Recovering the original sound waveform broadcast from an unknown remote source is of interest for source classification. In particular, one possible application of this technique is identifying, tracking, and monitoring marine mammals that vocalize underwater in unknown, noisy, and dynamic ocean environments. The ocean environment has always included an abundance of natural noises, such as the sounds generated by rain, waves, earthquakes, and sea creatures. However, a growing number of ships and oil rigs, as well as increased use of sonar by navies and researchers, is adding to the natural noise that already surrounds marine life. The potential impacts of increased background noise and specific sound sources, cause marine animals to change their behavior, prevent marine animals from hearing important sounds, cause hearing loss, or even damage tissue. One of the solutions to this problem is to know how the animals are spread throughout the area and whether or not a particular species is found in the area at the time of year when a potentially dangerous man-made source is operating. STR may be a useful technique to localize marine animals and identify their species – may be even identify individuals – from their recorded sounds.

Synthetic Time Reversal may also be helpful for underwater acoustic communication. The ability of time reversal to reduce multipath dispersion and its simplicity of implementation makes it ideal for underwater communication. Passive time reversal processing was introduced some time ago (Dowling, 1994). It uses the first arrival, from a stream of pulses that have traversed a complex refractive medium, as a filter for later pulse arrivals. This method has been used for an experiment conducted in Puget Sound near Seattle (Rouseff *et al*. 2001).

This thesis is organized into seven chapters. The current chapter contains the introductory material and literature review. Chapter 2 provides the foundation for the investigation into blind deconvolution in underwater sound channels. It provides the formulation of the synthetic time reversal for reconstructing the source signal and impulse responses. Then, it describes source localization techniques such as broadband Bartlett matched-field processing (MFP) and how the STR-reconstructed impulse response can be used for simple ray-based back-propagation source localization. Chapter 3 analyzes the performance of ray-based STR in a typical near-shore underwater environment. The purpose of this chapter is to document how ray-based STR signal estimation depends on receiving array size and signal-to-noise ratio, how it can be improved through a coherent combination of results from individual rays, and how the STR-estimated impulse response can be used to for source localization via matched-field processing or a simple ray-path back-propagation scheme. Chapter 4 presents STR blind deconvolution results for source signal estimation when the receiving array is sparse and conventional beamforming is not appropriate for the frequency band of interest (11-19 kHz), and introduce an unconventional beamforming technique based on manufacturing frequency differences from the array recordings that allows STR to be successful with sparse array measurements in the presence of modeling mismatch. Chapter 5 presents the other application of unconventional beamforming (presented in chapter 4) for sparse random array beamforming and an extension of this method for near-field beamforming. Chapter 6 shows how mode-based STR can be used for marine mammals’ ranging with a vertical array in a dispersive shallow-ocean waveguide. The final chapter, summarizes this work and presents the conclusions drawn from it.

CHAPTER 2

Mathematical formulation and foundations

This chapter provides the mathematical foundations for this investigation into blind deconvolution in an underwater sound channel. It provides the formulations for synthetic time reversal (STR), conventional beamforming techniques, and matched field processing (MFP).

2.1 Synthetic Time Reversal

Synthetic time reversal (STR) is a technique for simultaneously estimating the original source signal and the source-to-array transfer functions in an unknown underwater sound channel. The mathematical formulation of propagating-mode-based STR is presented in Sabra and Dowling (2004) and its extension to acoustic rays is outlined in Sabra *et al*. (2010). To illustrate the use of STR in underwater acoustics, a simple simulation with two ray paths is considered in this chapter. The figure below shows the geometry used in this simulation. There is a direct path from sound source located at 30 m depth and a surface reflection that is simulated with an image source. Data is recorded by a 15.5 m vertical receiving array with 32 elements located in the middle of the water column, 500 m away from the source.

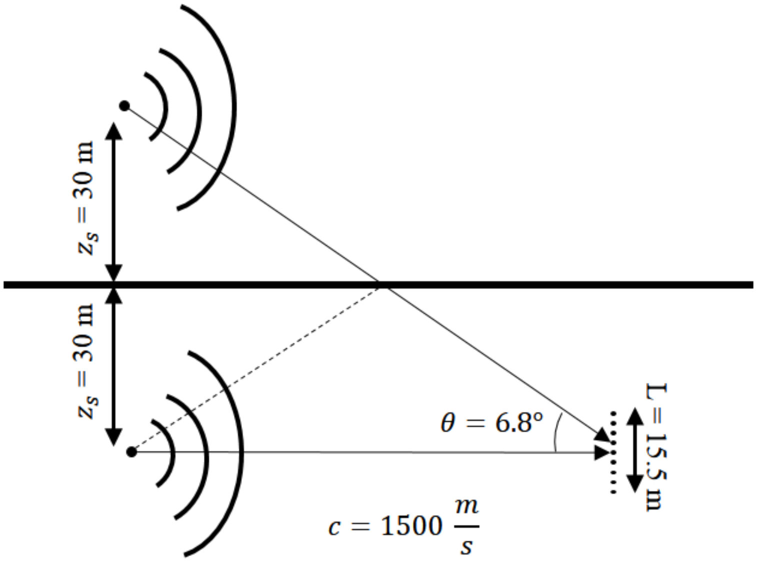


Figure 2-1: Sound channel for the simulations. Here the source depth is z_s = 30 m and the primary source-array ranges for this study are r_s = 500 m.

Consider a point source located at $\vec{r}_s$ that emits a signal $s(t)$ having Fourier transform $S(\omega)$:

$$S(\omega) = \frac{1}{2\pi}\int_{-\infty}^{\infty} s(t)e^{i\omega t}dt = |S(\omega)|e^{i\varphi_s(\omega)} \qquad \text{2-1}$$

where t is time, ω is temporal radian frequency, and $\varphi_s(\omega)$ is the signal's phase as a function of frequency. The broadcast signal used in the simple simulations, 62.5 ms linear frequency chirp from 1.8 kHz to 2.2 kHz, is shown in Figure 2-2(a).

The emitted sound travels through the ocean sound channel where it is recorded by a vertical array of N receiving transducers at locations $\vec{r}_j$ ($1 \le j \le N$), The recordings, $p_j(t)$, are solutions of the wave equation for a stationary point source. Figure 2-2(b) shows simulated signals recorded by the receiver at the shallowest depth for the acoustic environment shown in

Figure 2-1. Here the recording interval is presumed to be longer than the multipath time spread of the sound channel. The Fourier transform of $p_j(t)$ is $P_j(\omega)$, which is a convolution of the sound channel's Green's function and the source signal.

$$P_j(\omega) = G(\vec{r}_j, \vec{r}_s, \omega) S(\omega) \qquad \text{2-2}$$

where $G(\vec{r}_j, \vec{r}_s, \omega)$ is the sound channel's Green's function between the source location and the receiving transducer locations at frequency ω at the time of the source broadcast. The goal of the blind deconvolution signal processing technique is to recover $s(t)$ from the recordings $p_j(t)$ without explicit knowledge of $G(\vec{r}_j, \vec{r}_s, \omega)$.

The formulation of STR begins by developing an estimate of $G(\vec{r}_j, \vec{r}_s, \omega)$ from $P_j(\omega)$ alone. The first step is a simple normalization of $P_j(\omega)$ in equation 2-2 that eliminates the signal amplitude,

$$\tilde{P}_j(\omega) = \frac{P_j(\omega)}{\sqrt{\sum_{j=1}^{N} \left| P_j(\omega) \right|^2}} = \frac{\left| G(\vec{r}_j, \vec{r}_s, \omega) \right|}{\sqrt{\sum_{j=1}^{N} \left| G(\vec{r}_j, \vec{r}_s, \omega) \right|^2}} \exp\{ i(\varphi_s(\omega) + \arg[G(\vec{r}_j, \vec{r}_s, \omega)]) \} \qquad \text{2-3}$$

.

To produce a normalized estimate of $G(\vec{r}_j, \vec{r}_s, \omega)$ from equation 2-3, the signal's phase $\varphi_s(\omega)$ must be estimated and removed from the right side of equation 2-3. This is the pivotal step in STR.

One possible class of phase correction factors $e^{-i\alpha(\omega)}$ can be constructed using a weighted sum of the recordings:

$$\alpha = \arg\{\sum_{j=1}^{N} W_j \tilde{P}_j(\omega)\} = \arg\{\sum_{j=1}^{N} W_j G(\vec{r}_j, \vec{r}_s, \omega)\} + \varphi_s(\omega), \tag{2-4}$$

where W_j are the transducer weights which will be chosen to isolate the propagation phase of a single mode or ray. Currently, there are two approaches to estimate the source signal: 1) Mode-based STR, 2) Ray-based STR which are described at the end of this section.

The product of the phase correction factor and the normalized data vector produces an estimate of Green's function with an unknown time shift:

$$\hat{\tilde{G}}(\vec{r}_j, \vec{r}_s, \omega) = \tilde{P}(\omega) e^{-i\alpha(\omega)} = \frac{G(\vec{r}_j, \vec{r}_s, \omega)}{\sqrt{\sum_{j=1}^{N} \left|G(\vec{r}_j, \vec{r}_s, \omega)\right|^2}} \exp\{-i \arg(\sum_{j=1}^{N} W_j G(\vec{r}_j, \vec{r}_s, \omega))\}\,. \tag{2-5}$$

W_j in equations 2-4 and 2-5 should be chosen so that the extra phase in equation 2-5, $\Gamma(\omega)$, is linearly dependent on frequency, i.e. a desirable weighting produces:

$$\Gamma(\omega) = \arg\{W_j G(\vec{r}_j, \vec{r}_s, \omega)\} = a + b\omega \tag{2-6}$$

where *a* and *b* are real constants. Mode- and ray-based weightings are described in the next two subsections.

The Fourier transform of the STR-estimated signal $\tilde{S}(\omega)$ and the STR-estimated impulse response $\tilde{G}(\vec{r}_j, \vec{r}_s, \omega)$ between the source location $\vec{r}_s$ and array-element locations $\vec{r}_j$ are

$$\hat{S}(\omega) = \sum_{j=1}^{N} \hat{\tilde{G}}^*(\vec{r}_j, \vec{r}_s, \omega) P_j(\omega) = \sum_{j=1}^{N} \left\{\tilde{P}_j(\omega) e^{-i\alpha}\right\}^* P_j(\omega) \tag{2-7}$$

Here α is an extra correction phase which can be computed from equation 2-4, an asterisk denotes a complex conjugate, a tilde denotes a normalized function, and a caret denotes an estimated quantity.

An inverse Fourier transform of equation 2-7 produces the final STR-estimated source signal in the time domain,

$$\hat{s}(t) = \int_{-\infty}^{\infty} \hat{S}(\omega) e^{-i\omega t} dt \,. \qquad \text{2-8}$$

When STR is successful, $\hat{s}(t)$ is a good estimate of the initial source signal $s(t)$, up to a multiplicative constant and an arbitrary time shift. The reconstructed source signal for the simulations is shown in Figure 2-2(c). The maximum of the temporal correlation $C_{\max}$ of the initial source signal with the STR-reconstructed signal, can be used to measure the performance of the blind deconvolution operation by:

$$C_{\max} = \max{}_t \left(\frac{\int_{-\infty}^{\infty} S(\omega)\hat{S}(\omega) e^{-i\omega t} d\omega}{\sqrt{\int_{-\infty}^{\infty} \left|S(\omega)\right|^2 e^{-i\omega t} d\omega} . \sqrt{\int_{-\infty}^{\infty} \left|\hat{S}(\omega)\right|^2 e^{-i\omega t} d\omega}} \right) \qquad \text{2-9}$$

The reconstructed source signal shown in Figure 2-2(c) has a 99% correlation with the broadcast signal. However, it has a 0.33 second delay in comparison with the broadcast signal which is the travel time along the direct path, $\frac{r_s}{c}$.

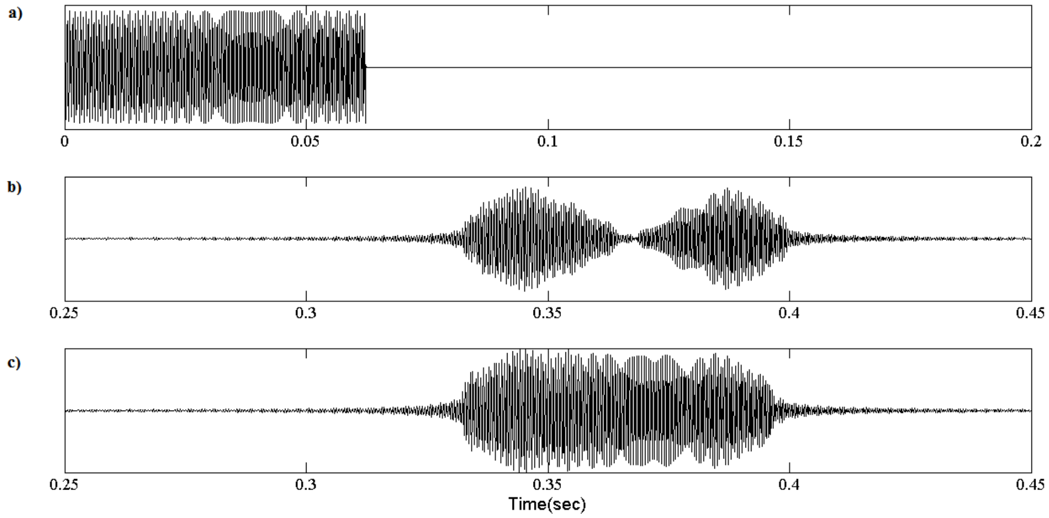


Figure 2-2: Sample input and output signals for ray-based STR. (a) Simulated broadcast signal, a linear frequency chirp from 1800 Hz to 2200 Hz with a duration of 62.5 ms (b) Received signal at the deepest array element at a range of 500 m. The cross correlation coefficient of this signal with the broadcast signal is 70%. (c) Ray-based STR estimated source signal using the direct path (0° ray arrival) as the reference ray shown in Figure 2-3. The cross correlation coefficient of this signal with the broadcast signal is 99%.

For completing the foundations needed for STR, the weights used in equation 2-4 need to be determined. Two approaches for estimating source signal are: 1) Mode-based STR, 2) Ray-based STR.

1. Mode-based STR

Overall, when mode shape and/or mode wave number information is available, the W_j can be selected for either vertical or horizontal arrays. Low-order mode shape estimates can be drawn from sound channel characteristics (Shang, 1985).

The weights W_j used to form the correction phase α in equation 2-4 are chosen empirically based on the receiving array's geometry and the character of the acoustic propagation. For situations involving a vertical line array with elements at depths z_j and modal propagation, the W_j can be selected to match the l^{th} propagating mode:

$$W_j \propto \Psi_l(z_j) \tag{2-10}$$

where $\Psi_l(z_j)$ is the vertical profile of the l^{th} propagating mode. In a range independent ocean sound channel, the Green's function can be represented as a sum over propagating modes (Jensen *et al*., 1994). If mode functions are orthogonal across the array aperture, W_j based on mode shapes can be used to extract the phase of individual modes or group of modes when no extra information is needed about the environment (Sabra & Dowling, 2004). Then, the time delay between the broadcast source signal and the reconstructed source signal (*b* in equation 2-6) will be the phase-speed travel time of the selected mode.

2. Ray-based STR

The mathematical formulation of the ray-based STR follows the prior mode-based formulation. For the same array geometry with propagation along acoustic rays, the W_j can be determined from plane-wave (or more sophisticated) beamforming:

$$W_j \propto \exp(-i\omega\tau(\theta_m, z_j)) \tag{2-11}$$

where τ is the time delay for the m^{th} ray path that arrives at the j^{th} receiver at nominal elevation angle θ_m from the horizontal, and can be computed from plane-wave or more sophisticated beamforming. For this dissertation, plane-wave beamforming has been used to determine the arrival angles respect to the middle of the array. For simple plane-wave beamforming with a linear vertical array, τ can be computed from

$$\tau(\theta, z_j) \cong (j-1)(\frac{d \sin\theta}{\bar{c}}) \qquad \text{2-12}$$

where $\bar{c}$ is the depth-averaged speed of sound across the array, and *d* is the distance between receivers. In this case, the time delay between the broadcast source signal and the reconstructed source signal (*b* in equation 2-6) will be the travel time along the selected ray. For this dissertation, it has been assumed that the array is stationary and has no deviation from the vertical position and the elements locations along the array are known.

The possible values for θ_m are determined from the maxima of the received beamformed energy (equation 2-13) which will be discussed in the following section.

2.2 Conventional Beamforming Techniques

The arrival directions of ray paths between a sound source and a receiving array can be determined by beamforming the array-recorded signals. Commonly, the field received by the array can be modeled with plane-wave (far-field) or spherical-wave (near-field) approximations. For an array composed of *N* elements with constant spacing between elements *d*, the array's far-field is reached (in free space) when $L_A^2 \big/ 4/ r$ is less than unity where $L_A = (N - 1)d$ is the overall array length, λ is the source-signal center-frequency wavelength, and *r* is the distance between source and array (Kinsler *et al*. 2000, Ziomek 1993). This dissertation will introduce unconventional beamforming techniques to recover out-of-band lower-frequency signal information from finite bandwidth signals in chapter 4.

Bartlett beamforming is one of the standard acoustic beamforming techniques (Jensen et. al., 1994). When the signal bandwidth is $2\pi f_1 < \omega < 2\pi f_2$ and $|\theta| \leq \pi/2$, it can be calculated by:

$$B(\theta) = \int_{2\pi f_1}^{2\pi f_2} \left| b(\theta,\omega) \right|^2 d\omega \qquad 2\text{-}13$$

where,

$$b(\theta,\omega) = \sum_{j=1}^{N} W_j^* P_j(\omega) \qquad 2\text{-}14$$

When the array is far from the source and signal wave-fronts are well modeled as being planar, W should be a plane-wave phase factor:

$$W_{far-field} = \exp\{i\omega\tau\} \qquad 2\text{-}15$$

where τ can be computed from equation 2-12. When the array is near the source, W should be a spherical-wave phase factor:

$$W_{near-field} = \exp\{i\omega \left| \vec{r} - \vec{r}_j \right| / \bar{c}\} \qquad 2\text{-}16$$

The resolution (or transverse spot size) of such conventional beamforming is proportional to $\bar{c}/\omega L$, where L is the dimension of the array perpendicular to the average source-array direction. Thus, higher frequencies hold the promise of higher resolution acoustic imaging.

The Bartlett beamforming output can be written in matrix notation:

$$B_{Bart} = W^{\dagger}(dd^{\dagger})W \qquad 2\text{-}17$$

where † denotes the complex transpose operation, d is the data vector, and W is the plane-wave phase factor vector for far-field calculations and the spherical-wave phase factor vector for near-field calculations. Figure 2-3 shows the Bartlett beamforming output for the geometry presented in section 2.1 as a function of steering angle and frequency. It detects two main arrivals at the receiving array location, 0° and 6.8°.

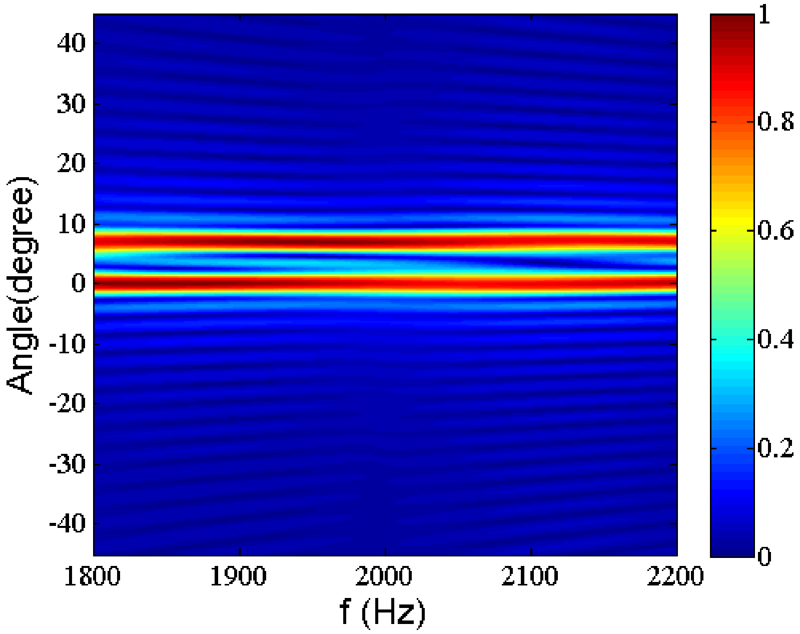


Figure 2-3: : Magnitude of the Bartlett beamformed output from receiving array at a source-array ranges of 500 m

Sometimes, Bartlett beamforming techniques are not successful in suppressing signal energy received from directions other than look direction for each θ (side lobes). The minimum variance distortionless processor (MV or MVDR) is one of the adaptive beamforming techniques which can suppress side lobes more than the Bartlett beamforming technique. The output of the MV processor is:

$$B_{MV} = [W^{\dagger}(dd^{\dagger})^{-1}W]^{-1} \tag{2-18}$$

where W is plane-wave phase factor vector (far-field) or spherical-wave phase factor vector (near-field). When the sound field occurs in a complicated multipath environment where it's not well described as sum of plane or spherical waves, the W vector may be determined from a acoustic propagation simulations that (hopefully) match the actual acoustic propagation in the real environment. In this case, W is referred to as a replica vector and the above beamforming processes are renamed Bartlett matched-field processing (MFP) and MV (or MVD) MFP. Both conventional MFP schemes determine the source location when successful. MFP is further discussed in the next section.

2.3 Matched field processing

In chapter 3, ray-based STR will be used to find the location (range and depth coordinates) of a remote underwater sound source when some environmental information is available at the array, and ray-path arrivals can be separated by beamforming at the array. The accuracy of this technique will be compared with conventional underwater source localization techniques such as incoherent and coherent Bartlett matched field processing (MFP). This section provides the formulas for the incoherent and coherent Bartlett MFP techniques which are used in chapter 3.

Conventional broadband matched field processing (MFP) provides a sophisticated means of source localization. MFP matches the measured field at the array with computed replicas of the expected field for all possible source locations in the region of interest and is successful when sufficient environmental information is available (Bucker 1976, Hinich 1979, Fizell 1987, Baggeroer et. al., 1988 and 1993). The process starts by putting a test point source at each point of a search grid and computing the acoustic field at the array. These computed field values are used as replica vectors in equation 2-17 or 2-18. The MFP output is a cross correlation between the modeled field and recorded data at each test-source location and is known as an ambiguity surface (a term borrowed from radar signal processing). When the test source location is at or close to the actual source location, and the propagation calculations match the actual acoustic propagation, the cross correlation will reach a maximum and the ambiguity surface will display a peak. Two standard MFP routines are Bartlett and MV (minimum variance). Both are extensions of plane-wave and spherical-wave beamforming which were discussed in section 2.2 (see also Jensen et. al., 1994).

To illustrate the performance of Bartlett and Minimum Variance MFP, a simple free-space simulation has been undertaken. Source-array geometry of this simulation is shown in Figure 2-4. The receiver array has 32 elements that are distributed equally in the sound channel with 3 m spacing.

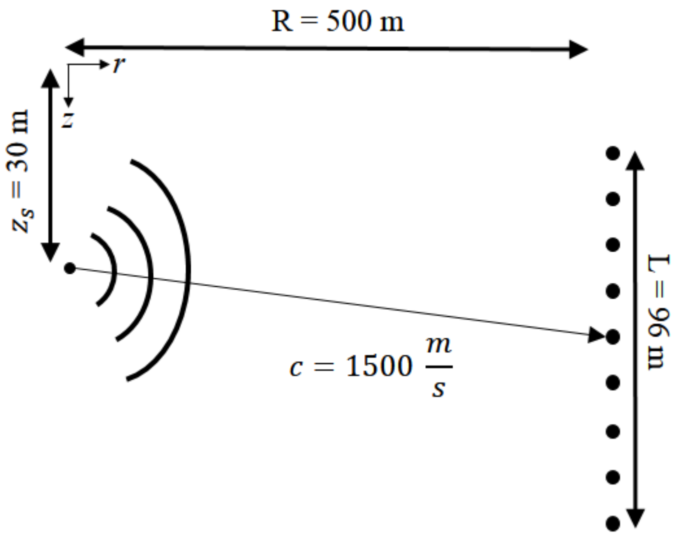


Figure 2-4: Sound channel for the free-space simulations. Here the source depth is $z_s = 30$ m and the primary source-array ranges for this study are $r_s = 500$ m.

Figure 2-5 shows the Bartlett and Minimum Variance MFP ambiguity surface for free-space propagation when the sound source is located at 30m depth and 500m range and broadcasting the signal shown in Figure 2-2(a). In this figure, minimum variance has higher resolution than Bartlett MFP.

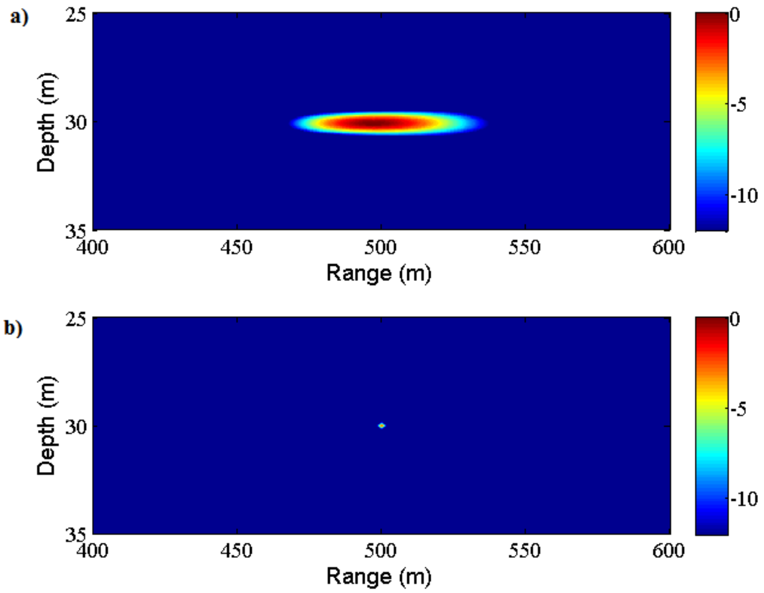


Figure 2-5: (a) Bartlett MFP for a free-space propagation. (b) Minimum Variance MFP for a free-space propagation. Source is located at 30 m depth and 500 m range respect to a linear vertical array. This shows that Minimum Variance MFP has a better resolution (smaller spot size) than Bartlett MFP.

In this dissertation, incoherent and coherent Bartlett MFP schemes will be considered in chapter 3. The incoherent calculations utilized M frequencies, within the signal bandwidth and the overall ambiguity surface $A_i(r,z)$ is determined from an incoherent sum of the single-frequency results:

$$A_i(r,z) = \frac{1}{M}\sum_{k=1}^{M} \frac{\left|\sum_{j=1}^{N} P_j(\omega_k) G_c^*(\vec{r}_j,\vec{r}_s;\omega_k)\right|^2}{(\sum_{j=1}^{N}\left|P_j(\omega_k)\right|^2)(\sum_{j=1}^{N}\left|G_c(\vec{r}_j,\vec{r}_s;\omega_k)\right|^2)} \qquad \text{2-19}$$

where $G_c(\vec{r}_j,\vec{r}_s;\omega_k)$ is the calculated complex acoustic pressure at location $\vec{r}_j$ and frequency ω_k from a harmonic point source located at $\vec{r}_s$. The numerator of equation 2-19 is a frequency-domain correlation between the measurements and the calculated impulse response across the array, while the denominator of equation 2-19 provides the appropriate normalization so that $0 \leq A_i \leq 1$.

The extension of equation 2-19 to broadband coherent Bartlett MFP involves field calculations throughout the signal bandwidth is formulated as a cross correlation between the STR-estimated impulse response and the computed impulse response.

$$A_c(r,z) = \max_t \left(\frac{\left| \int_{-\infty}^{+\infty} \tilde{G}(\vec{r}_j, \vec{r}_s, \omega) G_c^*(\vec{r}_j, \vec{r}_s; \omega_k) e^{-i\omega t} d\omega \right|^2}{\left(\int_{-\infty}^{+\infty} \left| \tilde{G}(\vec{r}_j, \vec{r}_s, \omega) \right|^2 d\omega \right) \left(\int_{-\infty}^{+\infty} \left| G_c(\vec{r}_j, \vec{r}_s; \omega_k) \right|^2 d\omega \right)} \right) \qquad \text{2-20}$$

This formulation of coherent broadband MFP involving $\tilde{G}(\vec{r}_j, \vec{r}_s, \omega)$ is only possible when estimates of the source-to-array impulse responses are available. In section 3.4, the estimates of the source-to-array impulse responses used for calculating the coherent MFP are determined from ray-based STR (equation 2-5).

CHAPTER 3

Broadcast signal reconstruction and source localization at 1.5-4 kHz

This section describes the results from the application of ray-based synthetic time reversal (STR) to simulations and underwater experiments involving source-array ranges of 100 m to 500 m in 60-m-deep water and 50 millisecond chirp signals with a bandwidth of 1.5 to 4.0 kHz (Abadi *et al*. 2012). The correlation coefficient between the original signal and the STR-reconstructed signals are presented as a function of signal-to-noise ratio. Also, the effect of reducing the number of elements of the receiving array and the use of a coherent combination of reconstructed results for various ray arrival directions on cross correlation coefficient are presented. The STR-based localization results are found to be superior to comparable results from coherent and incoherent Bartlett matched field processing (MFP), even though the STR results required only a tiny fraction of the computational effort necessary for MFP.

3.1 Simulations

The simulation results provided here are based calculations using the *MATLAB* version of *BELLHOP*, a Gaussian-beam tracing model for predicting acoustic pressure fields in underwater environments. *BELLHOP* can produce a variety of useful outputs including transmission loss, eigenrays, arrivals, and received time-series. A theoretical description may be found in (Porter & Bucker, 1987). The simulated signal, source-array geometries, and acoustic environment for these simulations match that of the CAPEx09 underwater propagation experiment conducted in

Lake Washington (Rouseff *et al*. 2010), where a 50 ms chirp signal from 1.5 to 4 kHz is broadcast from a single stationary source to a linear vertical receiving array at source-array ranges of 100 m to 500 m in water that is 60 m deep. The sound speed varies with depth but not with range. The bottom properties were not measured during CAPEx09, but Lake Washington is known anecdotally to have a soft lakebed. Thus, the lakebed properties used in the simulations are typical of sandy mud (see APL 1994): sound speed of 1420 m/s, density 1149 kg/m^3, and attenuation 0.2 dB/wavelength.

3.2 CAPEx09 Experiment

The September 2009 Cooperative Array Performance Experiment (CAPEx09) was conducted in Lake Washington near Seattle. Two adjacent vertical receiving arrays of similar length were deployed from the stern of the two-point moored R/V Robertson: a 32-element pressure sensor array and an 8-element vector sensor array. Each element of the vector-sensor array measured the acoustic pressure plus the three components of acoustic particle acceleration. Consequently, the two arrays made exactly the same number of acoustic measurements over similar vertical apertures. The received signals were sampled at 25 kHz per channel. The source-to-arrays range varied between 10 m and 4 km in water nominally 60 m deep. A variety of signals were transmitted, but the present analysis is restricted to 50 ms duration frequency-modulated chirps sweeping linearly from 1.5 to 4 kHz with nominal source depth 30 m. The analysis is further limited to data collected on the pressure-sensor array.

Figure 3-1 shows the measured sound speed profile in the water column together with other parameters from the experiment. The contrast in sound speed between the warm surface water and the cool water below is more than 40 m/s resulting in sharp refraction of acoustic rays; ray traces (not shown) revealed that a direct path from the source to the entire array was lost for ranges beyond approximately 400 m. Figure 3-1 includes the 32-element pressure-sensor array shown to vertical scale. The top element was at nominal depth 30 m with uniform 22.4 cm spacing between the elements. Since the elements were attached to each other by a rope, there is a slight vertical deviation between them. However, this misalignment has been ignored for this dissertation. The dense spacing of the array elements permits successful conventional beamforming of the received chirp signals. The beam-steering angle, θ, is measured from the horizontal and is positive upward as shown.

The recordings at range 10 m permitted time gating of the direct signal to eliminate surface and bottom reflections. The resulting measured signal, $s(t)$, serves as the true signal against which the blind deconvolution results at much greater ranges are compared. Recordings at source-array ranges 100, 250, and 500 m are emphasized in the present study of STR's performance.

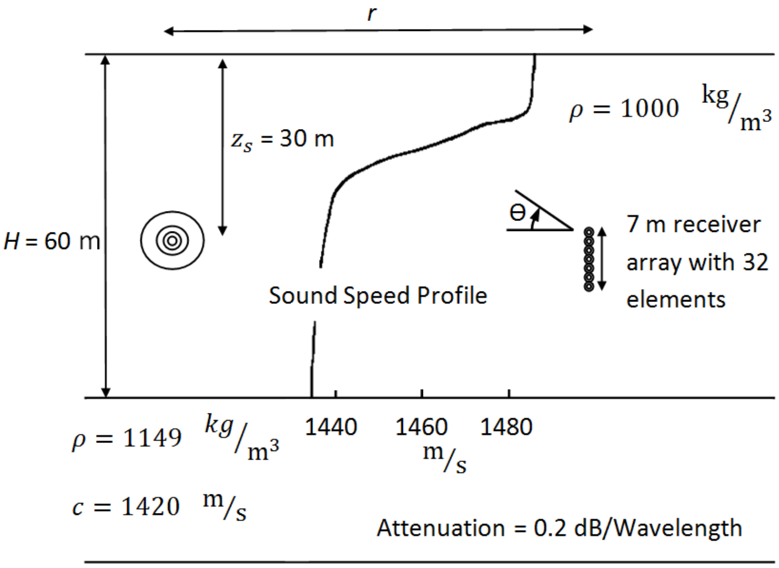


Figure 3-1: Sound channel for the simulations and the experiments. Here the nominal source depth is z_s = 30 m and the primary source-array ranges for this study are r_s = 100 m, 250 m, and 500 m.

Figure 3-2 provides measured propagation results from the CAPEX09 experiment at a source-array ranges of 100 m (a), 250 m (b), 500 m (c), and 1.0 km (d) via beamformed output, $b(\theta,\omega)$, from the receiving array using equation 2-14. The dynamic range shown in the figure covers 50 dB. At the 100 m source-array range, the direct path at 5° and surface-reflected path at 30° show up clearly throughout the signal bandwidth, while a weaker bottom reflection at –34° is also apparent. At 250 m, the direct path with an arrival angle near 7° is the strongest, and several weaker paths exist within ±30° or so of this direct path angle. At 500 m, there are two strong ray-path arrivals with angles that waver around –7° and –12° or so. Here again several weaker paths at larger angles exist intermittently at the range. At 1.0 km, there are no ray-arrival angles that persist throughout the bandwidth of the signal. The recorded data at these four source-array ranges span the possible range of STR performance outcomes.

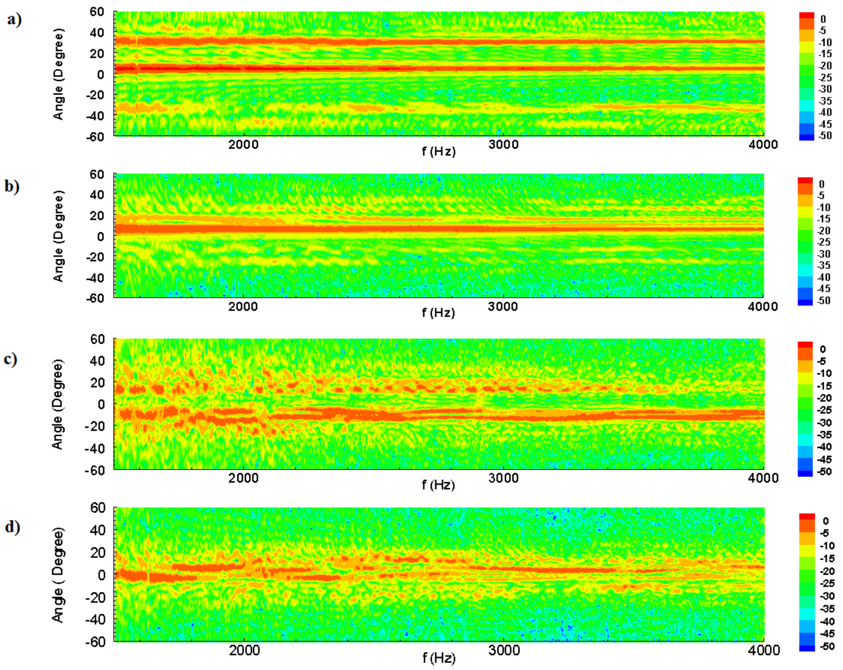


Figure 3-2: Magnitude of the beamformed output $b(\theta,\omega)$ from the CAPEX09 receiving array at a source-array ranges of 100 m (a), 250 m (b), 500 m (c), and 1.0 km (d) as a function of frequency (Hz) and elevation angle (degrees). Ray-based STR is successful when there is at least one distinct propagation path that persists at the same angle throughout the bandwidth of the signal. For these CAPEX09 measurements, STR works well at the shorter two ranges, has some success at 500 m, but fails at 1.0 km.

Figure 3-3d shows sample ray-based STR waveform results (equation 2-8) from the experimental measurements at the 250 m range when the direct path is selected as the reference ray. In this figure, the first waveform (Figure 3-3a) is the measured broadcast signal $s(t)$, the second waveform (Figure 3-3b) is the signal recorded by the first (shallowest) element of the receiving array $p_1(t)$, the third waveform (Figure 3-3c) is the output from delay-and-sum beamforming with a receiving direction of 6.7°, the fourth waveform (Figure 3-3d) is the STR-estimated signal $\hat{s}(t)$ when all 32 $p_j(t)$ are utilized in the processing. Here, STR provides a noticeable improvement in the signal envelope shape over the single-receiver and delay-and-sum beamforming results because it coherently adds signal information from all propagation paths. Thus, ray-based STR can be considered an extension of delay-and-sum beamforming for blind deconvolution since it also provides an estimate of the source-to-array-element impulse responses that delay-and-sum beamforming does not provide.

The fifth waveform shown in Figure 3-3e, which has 25 ms offset compared to the other waveforms, is the amplitude of the STR-estimated impulse response $\hat{\tilde{g}}(\vec{r}_1,\vec{r}_s,t)$ between the source and the first array element at the 250 m range. The first important peak in this sample of $\hat{\tilde{g}}(\vec{r}_1,\vec{r}_s,t)$ occurs at t = 0 and corresponds to the reference ray path (θ_1 = 6.7°). The second, third, and fourth peaks correspond ray paths with arrival angles of 16.3°, –23.3°, and 26.8° and signal-propagation times that are approximately 0.6, 8, and 22 ms longer than that for the reference ray. Also, there is another weak bottom reflection at -11.2° (Figure 3-2b) which generates an arrival path at 40 ms (Figure 3-3e and Figure 3-4). Since it is a weak arrival with a small amplitude in the impulse response compared to other arrivals and may add more error to the calculation, this weak ray path has not been considered for calculation presented in this section. If one of these other rays were chosen as the reference, the associated STR-estimated impulse response would place the peak for that ray at *t* = 0. This time shifting is shown in Figure 3-4. Figure 3-4a is the amplitude of the STR-estimated impulse response $\hat{\tilde{g}}(\vec{r}_1,\vec{r}_s,t)$ between the source and the first array element at the 250 m range when the reference ray path is 6.7° (direct path) and Figure 3-4b shows the same signal when the reference ray path is -23.3° (bottom reflection). It shows that the peak corresponds to the selected reference ray path has been placed at *t* = 0 for both cases and all other peaks have been shifted equally. Although the absolute source-to-array travel time on any of these rays remains unknown, travel-time differences between ray paths are apparent.

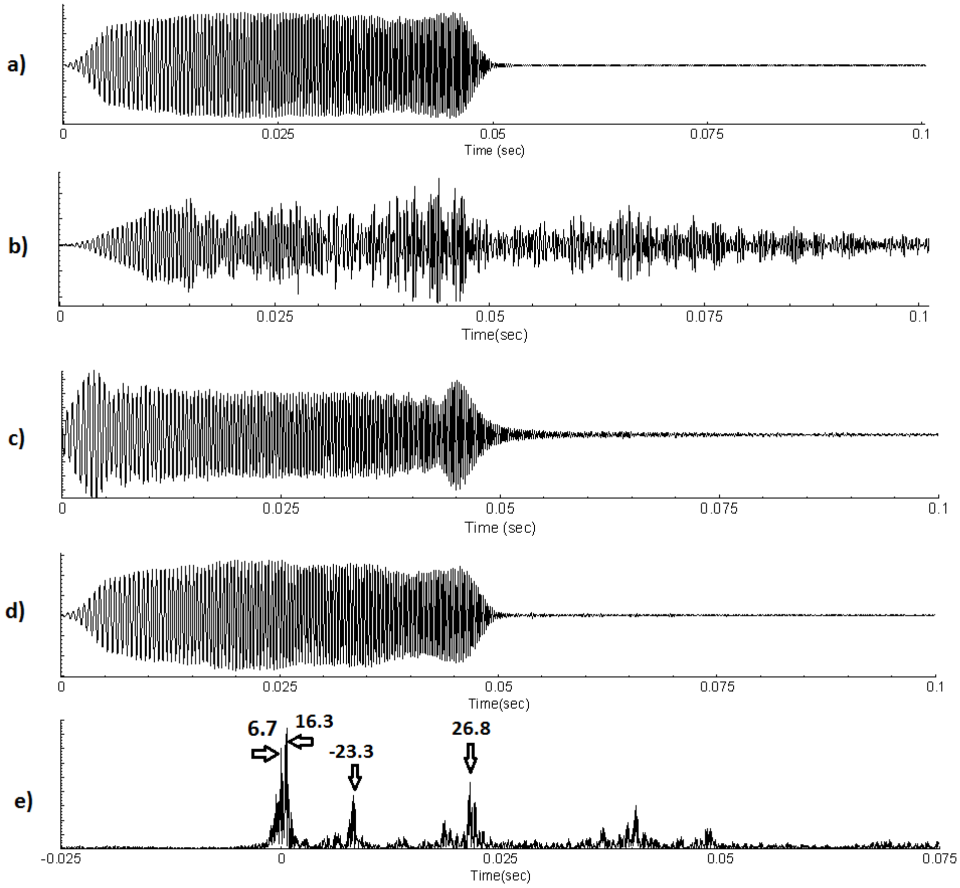


Figure 3-3: Sample input and output signals for ray-based STR. (a) Measured broadcast signal, a nearly uniform amplitude sweep from 1.5 kHz to 4.0 kHz with a duration of 50 ms (b) Received signal at the shallowest array element at a range of 250 m. The cross correlation coefficient of this signal with the broadcast signal is 57%. (c) Delay-and-sum beamformed output using the 6.7° ray-path shown in Figure 3-2b). The cross correlation coefficient of this signal with the broadcast signal is 95%. (d) Ray-based STR estimated source signal using the 6.7° ray arrival shown in Figure 3-2b) as the reference ray. The cross correlation coefficient of this signal with the broadcast signal is 99%. (e) Absolute value of the ray-based STR-estimated impulse response between the source and the shallowest array element using the 6.7° ray-path shown in Figure 3-2b) as the reference ray. The impulse response peak for this ray appears at *t* = 0. The other impulse response peaks correspond to ray-arrival angles of 16.3°, 26.8° and -23.3°.

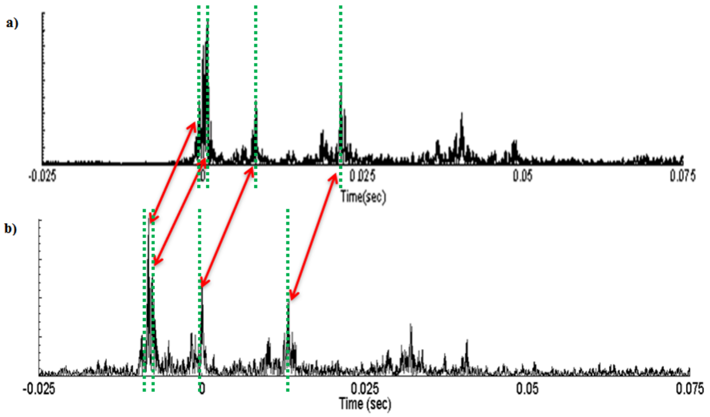


Figure 3-4: Absolute value of the ray-based STR-estimated impulse response between the source and the shallowest array element using the 6.7° ray-path (a), the -23.3° ray-path (b) as the reference ray. The impulse response peak for the selected ray=path appears at *t* = 0. Each ray arrival has been marked by green dashed lines.

3.3 Parametric Dependencies of STR Signal Reconstruction

This section reports a variety of ray-based-STR signal-reconstruction performance results from the CAPEx09 experiment along with range-independent companion simulations. The primary performance metric for signal reconstruction is the maximum cross-correlation coefficient ($C_{\max}$) from equation 2-9, between the broadcast signal $s(t)$ and the STR-estimated signal $\hat{s}(t)$. For the data shown on Figure 3-3, the $C_{\max}$ of the original signal with the sample received signal, the delay-and-sum beamformed signal, and the STR-estimated signal are 57%, 95%, and 99%, respectively. In general, a $C_{\max}$ above 90% is needed for a blind deconvolution technique to be considered useful.

The parametric dependencies of STR's signal estimation performance are provided on Figure 3-5, Figure 3-7, and Figure 3-8. The first of these shows both simulation results (filled symbols) and experimental results (open symbols) for $C_{\max}$ from (equation 2-9) as function of the number of receivers ($2 \le N \le 32$) for source array ranges of 100 m, 250 m, and 500 m. Here, for $N < 32$, contiguous array elements were used starting with the shallowest array element (j = 1); thus, the receiving array's aperture in Figure 3-5 is directly proportional to $N - 1$ and increases downward from the shallowest element with increasing N. In all cases, STR's signal-estimation performance increases with increasing N, an array resolution effect. A longer array can better resolve ray-arrival directions, and thereby produce a better measurement of the requisite correction phase, α in equation 2-4. For the results shown in Figure 3-5, reference ray-path arrival angles have been determined based on $B(\theta)$ from all 32 elements and have not been altered for smaller N. Yet, it is potentially remarkable that greater than 90% signal $C_{\max}$ can be achieved at source-array ranges of 100 and 250 m with as few as 7 or 8 array elements. Furthermore, at these ranges, the simulation and experimental $C_{\max}$ results are within one or two percent of each other and the residual small differences are most likely the mild detrimental effects of finite signal-to-noise ratio, weak random scattering in the experiments, or the limitations of plane-wave beamforming.

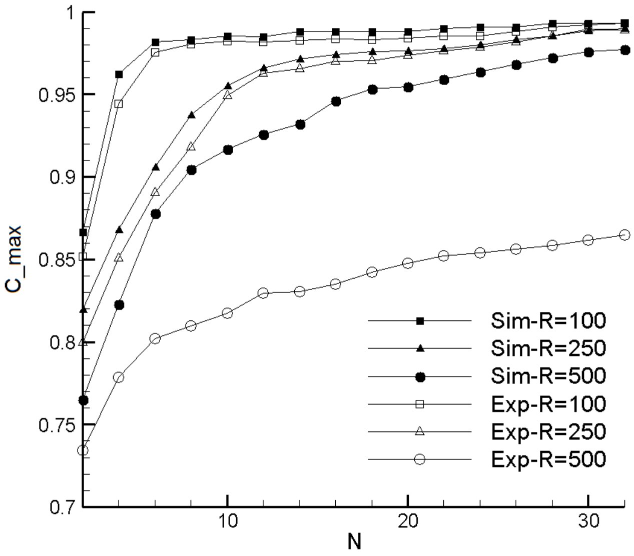


Figure 3-5: Cross correlation coefficient ($C_{\max}$) from equation 2-9 for the simulations (filled symbols) and the CAPEX09 experiments (open symbols) for source-array ranges of 100 m, 250 m, and 500 m vs. the number N of receiving array elements. Here $C_{\max}$ values increase with increasing N. STR's simulated and experimental performance matches at the two shorter ranges, but differs by as much as 10% at the longer range.

However, the simulated and experimental $C_{\max}$ results in Figure 3-5 for the 500 m range differ by as much as 10% when $N > 15$, and this points to a limitation of ray-based STR. Its success depends on there being at least one ray-path arrival that persists with (nearly) constant θ_m across the frequency range of the signal. An examination of the beamformed CAPEx09 signal shown on Figure 3-2 supports this contention. At a source-array range of 100 m, there are two persistent ray-paths. At 250 m (Figure 3-2b), there is certainly one persistent ray-path arrival at $\theta_m = 6.7°$. At 500 m (Figure 3-2c), there are one or possibly two tenuously persistent arrivals that waver and intermittently disappear. At 1 km (Figure 3-2d), there are no ray path arrivals with sufficient persistence across the entire frequency band of the signal for successful ray-based STR.

The loss of persistent ray paths in the CAPEx09 data with increasing range may have both deterministic and random origins. First, the nominal resolution of the receiving array at the signal's band-center frequency is ~3 degrees. Thus, the receiving array may not fully distinguish the two wavering ray-paths with arrival angles near –10° at the 500 m range. Second, based on eigenray calculations, the steep sound speed gradient in the CAPEX09 environment causes different ray paths to reach the top and bottom of the array. Such differences in propagation characteristics were verified by separately beamforming the signal using the top and bottom halves of the array (Figure 3-6). Figure 3-6 shows that the bottom arrivals at bottom half of the receiving array is stronger than the top half of the array. Since the current implementation of ray-based STR is built from plane-wave beamforming, its success is likely to be reduced in an environment where wave-front arrivals do not extend over the full spatial aperture of the receiving array. And finally, some random refraction and scattering is expected in the real underwater waveguide, but was not simulated. Such refraction and scattering increases in importance with increasing source-array range, and is likely to distort the signal wave fronts so they are no longer planar, the net result being a detrimental impact on ray-based STR performance that is only apparent in the experimental results.

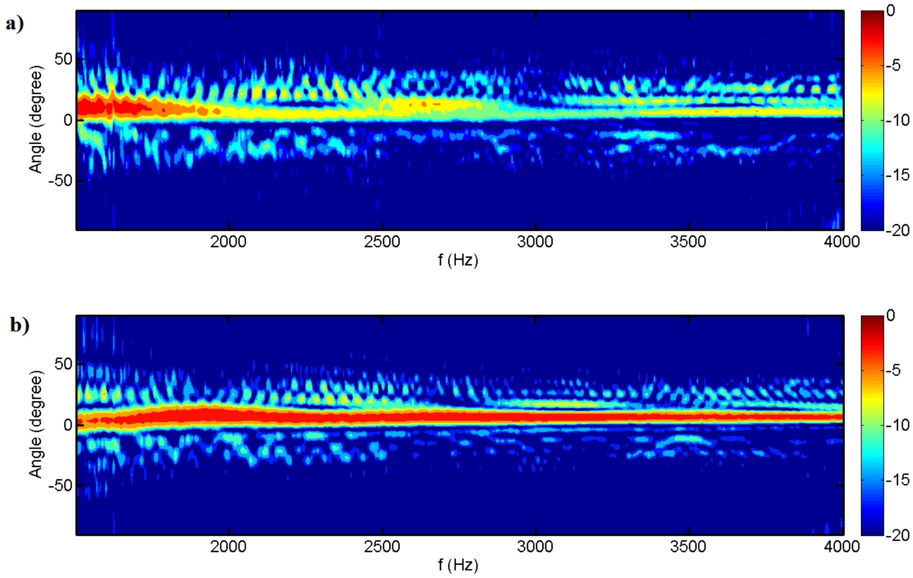


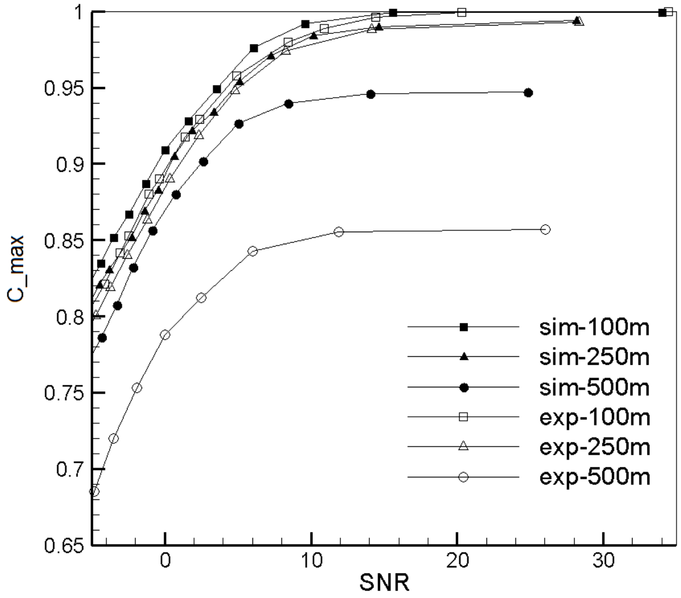
Figure 3-6: Magnitude of the normalized beamformed output $b(\theta,\omega)$ in dB from the top half of CAPEx09 receiving array (a), bottom half of CAPEx09 receiving array (b), at a source-array ranges of 250 m as a function of frequency (Hz) and elevation angle (degrees). This figure shows that different ray paths reach the top and bottom of the array.

A second limitation of STR arises from finite received signal-to-noise ratio (SNR). To quantify the impact of variable SNR on ray-based STR performance, noise samples $n_j(t)$ measured at each receiver approximately one half second after reception of the CAPEX09 signal, and having the same duration as the received signals, were multiplied by a dimensionless coefficient $\gamma > 0$ and added to the measured signal from each receiver. Thus, SNR, as defined by equation 3-1 was varied by increasing γ.

$$\mathrm{SNR} = 10\log\left\{\frac{\int_{2\pi f_1}^{2\pi f_2}\sum_{j=1}^{N}\left[\left|P_j(\omega)\right|^2 - \left|N_j(\omega)\right|^2\right]d\omega}{(\gamma+1)\int_{2\pi f_1}^{2\pi f_2}\sum_{j=1}^{N}\left|N_j(\omega)\right|^2 d\omega}\right\} \qquad \text{3-1}$$

Here $N_j(\omega)$ is the Fourier transform of $n_j(t)$. The measured noise spectra for some of the 32 elements include a single peak just above 2 kHz but otherwise all were nearly flat through the signal bandwidth.

The results of these variable SNR studies are shown in Figure 3-7 where $C_{\max}$ is plotted vs. SNR from equation 3-1 for the simulations and the experiments at source array ranges of 100, 250, and 500 m when all 32 receiving array elements are used. In all cases, $C_{\max}$ increases monotonically with increasing SNR. The simulation and experimental results at the shorter two ranges all fall within ±1% of each other, and STR achieves a $C_{\max}$ of greater than 90% at an SNR of +2 dB at these two ranges. The longer-range simulation and the experimental results fall below the others because the resolution requirements for achieving any fixed $C_{\max}$ value increase with increasing range, and because the received field in the experiment has only tenuously persistent ray-path arrivals.



.

Figure 3-7: Cross correlation coefficient ($C_{\max}$) from equation 2-9 for the simulations (filled symbols) and the CAPEX09 experiments (open symbols) for source-array ranges of 100 m, 250 m, and 500 m vs. the SNR from equation 3-1. Here $C_{\max}$ values increase with increasing SNR. STR's simulated and experimental performance again matches at two shorter ranges, but differs by nearly 10% at the longer range.

In a variety of other underwater sound propagation scenarios, more than one ray-path arrival can typically be identified at the receiving array. Thus, the possibility exists that the final STR output may be improved in finite signal-to-noise ratio situations by separately using each persistent ray-path arrival as the reference and then coherently combining the various STR results. This possibility was considered for the experimental data at the 100 m range where the direct and surface-reflected ray paths are well resolved, persistent, and of comparable strength. The STR $C_{\max}$ results using each path as the reference path are shown along with $C_{\max}$ results for a coherent combination of the path-specific results as a function of SNR on Figure 3-8. Although the percentage differences are small, the coherent combination provides the highest $C_{\max}$ for all

SNRs. Thus, improving STR results via a coherent combination of results from different reference rays is a promising possibility.

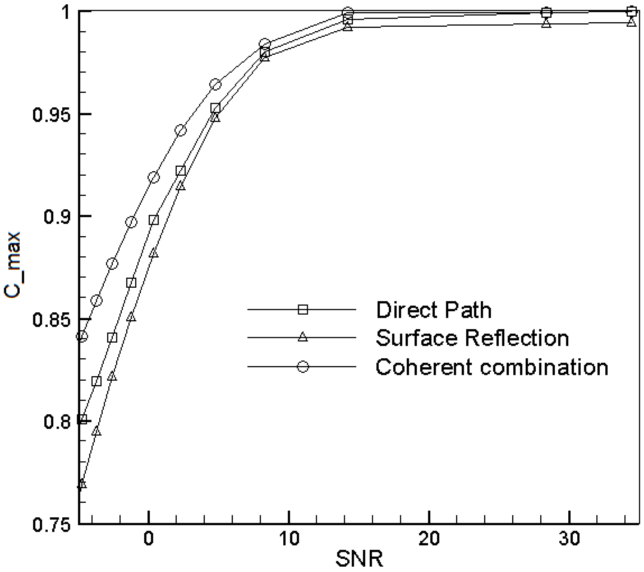


Figure 3-8: Cross correlation coefficient ($C_{\max}$) from equation 2-9 vs. SNR from equation 3-1 for the CAPEX09 experimental data at a source array range of 100 m. Here there are two persistent ray-path arrivals corresponding to direct and surface-reflected paths. A coherent combination of results from separate STR computations using each path as the reference is superior to that from either path alone.

3.4 STR and Source Localization

STR can be used to estimate simultaneously the source signal and impulse response waveforms from a remote unknown source. Unfortunately, the unknown time shift in the reconstructed waveforms prevents elementary distance-equals-speed-times-time estimation of the source-array range. However, the relative timing between peaks in the STR impulse response can be used to estimate the source range and depth when some environmental information is available. This possibility was explored using ray paths determined from *BELLHOP* and the CAPEx09 data set.

As a preliminary step, the correspondence between ray-path arrival angles and impulse response peaks must be determined. In the present study, this was done by inspecting $B(\theta)$ to select possible ray-path arrival angles, θ_m with $1 \leq m \leq M$, where $B(\theta)$ showed a local maximum, and then completing M STR calculations to determine the impulse response peak corresponding to each θ_m. When a valid θ_m is used as the reference-path arrival angle, the impulse response peak associated with that path appears at the time origin when $\hat{\tilde{g}}(\vec{r}_j, \vec{r}_s, t)$ is plotted vs. t. For example, Figure 3-3e) displays $\left|\hat{\tilde{g}}(\vec{r}_j, \vec{r}_s, t)\right|$ for the shallowest receiver at a source-array range of 250 m when an angle of 6.7° is used for the reference ray path. Here, the first peak of $\left|\hat{\tilde{g}}(\vec{r}_j, \vec{r}_s, t)\right|$ occurs at $t_1 = 0$ (the time origin) and this allows the identification $\theta_1 =$ 6.7°. The other impulse-response peaks occurring at t_m in Figure 3-3e) represent later arriving ray paths. If $\theta_2 = 16.8°$ had been chosen as the reference ray-path, then the second impulse response peak in Figure 3-3e) would have appeared at the time origin. Once all possible ray path arrival angles have been considered, the arrival angles θ_m and STR-estimated relative time shifts $t_m - t_1$ for the various path connecting the source and the array are known. In a multipath environment, this angle and timing information is a signature of the source location, and this location may be estimated when there is enough environmental information for ray path calculations.

Three possible schemes for source localization are considered here: simple ray-based back propagation along identified rays, and incoherent and coherent Bartlett matched field processing. For all three techniques, the environment is assumed to be range independent and the environmental information is limited to receiver depths, water column sound speed profile at the array, water depth at the array, and generic bottom type at the array. Thus, all three techniques are equally challenged by mismatch between the computational and actual environments. The formulation of the three techniques used in the present study is described in the next two paragraphs.

The ray-based back-propagation technique is based on acoustic time reversal (or phase conjugation in the frequency domain). First, the environmental information and the ray arrival angles are used to compute M rays launched at angles θ_m starting from the center of the array and extending out to the largest array-source range of interest, about 600 m in the current investigation. Next, the STR-determined impulse response is idealized as a series of perfect impulses that occur with the STR-determined arrival-time differences. This series of impulses is then time reversed and each impulse is launched along its associated ray path from the array. As the various impulses, located at range-depth coordinates (r_m, z_m) propagate away from the array along their corresponding rays, the root-mean-square (rms) impulse position, $\zeta = \left[(1/M)\sum_{m=1}^{M}\left((r_m - R)^2 + (z_m - Z)^2\right)\right]^{1/2}$, based on Euclidian distances from the impulse centroid $(R,Z) = (1/M)\sum_{m=1}^{M}(r_m, z_m)$, is monitored. The centroid location with the minimum ζ within the domain of interest provides an estimate of the source location. An example of such a ray-based back-propagation calculation is shown on Figure 3-9 where the impulse positions are shown for three different times. In this figure, the array is on the left at $r = 0$ and the three rays emerge from the array-center depth of 33.5 m. In this case, a global minimum of ζ is occurs when the impulse centroid is located at (27m, 100m) when the source was nominally located at (30m, 100m). Although such a simple scheme can be refined and enhanced, its current formulation is computationally efficient since it merely requires back-propagation calculations along a few ray paths once θ_m and $t_m - t_1$ are determined.

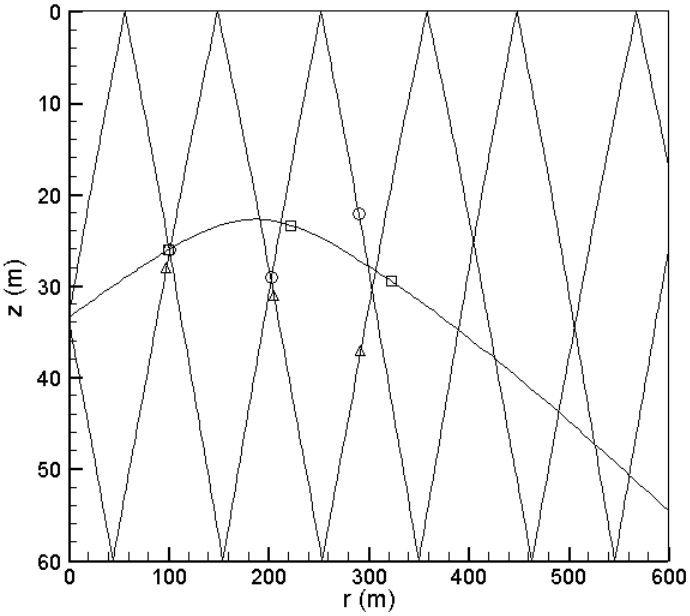


Figure 3-9: Sample ray trace back propagation calculation. The rays emerge from the center of receiving array at $r = 0$ and $z = 33.5$ m. Here symbols are shown at impulse locations at several different times when the rms impulse location, ζ, achieves a local minimum. The actual source range and depth is 100 m and 30 m, respectively.

Broadband matched field processing (MFP) provides a more sophisticated means of source localization but involves a significant increase in computational effort. Here both incoherent and coherent MFP schemes were considered. The incoherent calculations utilized six frequencies (1.5, 2.0, 2.5, 3.0, 3.5, and 4 kHz) within the signal bandwidth and the ambiguity surface $A_i(r,z)$ was determined from an incoherent sum of single-frequency results:

$$A_i(r,z) = \frac{1}{6}\sum_{k=1}^{6} \frac{\left|\sum_{j=1}^{N} P_j(\omega_k) G_c^*(\vec{r}_j, \vec{r}_s; \omega_k)\right|^2}{\left(\sum_{j=1}^{N} \left|P_j(\omega_k)\right|^2\right)\left(\sum_{j=1}^{N} \left|G_c(\vec{r}_j, \vec{r}_s; \omega_k)\right|^2\right)} \quad \text{3-2}$$

where $G_c(\vec{r}_j, \vec{r}_s; \omega_k)$ is the calculated complex acoustic pressure at location $\vec{r}_j$ and frequency ω_k from a unity strength harmonic point source located at $\vec{r}_s$. The numerator of equation 3-2 amounts to a correlation between the measurements and the calculated impulse response across the array, while the denominator of equation 3-2 provides the appropriate normalization so that $0 \leq A_i \leq 1$.

The extension of equation 3-2 to broadband coherent Bartlett MFP involved field calculations throughout the signal bandwidth, and was formulated as a cross correlation between the STR-estimated impulse response and the computed impulse response.

$$A_c(r,z) = \max_t \left(\frac{\left| \int_{-\infty}^{+\infty} \tilde{G}(\vec{r}_j, \vec{r}_s, \omega) G_c^*(\vec{r}_j, \vec{r}_s; \omega_k) e^{-i\omega t} d\omega \right|^2}{\left(\int_{-\infty}^{+\infty} \left| \tilde{G}(\vec{r}_j, \vec{r}_s, \omega) \right|^2 d\omega \right) \left(\int_{-\infty}^{+\infty} \left| G_c(\vec{r}_j, \vec{r}_s; \omega_k) \right|^2 d\omega \right)} \right) \tag{3-3}$$

This formulation of coherent broadband MFP involving $\tilde{G}(\vec{r}_j, \vec{r}_s, \omega)$ is only possible when estimates of the source-to-array impulse responses are available.

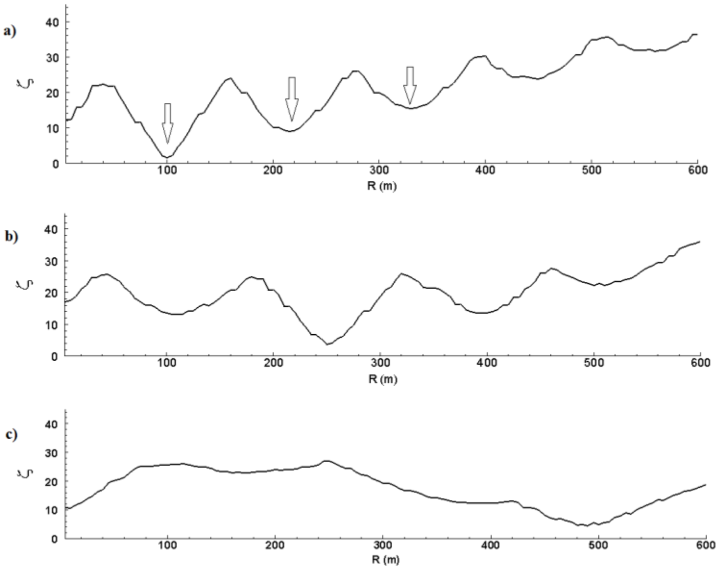


Figure 3-10: Root-mean-square impulse location ζ vs. range for source-array ranges of 100 m (a), 250 m (b), and 500 m (c). Here the minimum ζ unambiguously occurs near the actual source-array range. The arrows on the top panel lie at the ranges corresponding to the marker locations shown in Figure 3-9.

The source localization results from the three techniques using the CAPEX09 data are provided on Figure 3-10 for ray-based back-propagation, Figure 3-11 for incoherent MFP from equation 3-2, and Figure 3-12 for coherent MFP from equation 3-3. In each figure, the a), b), and c) panels are for source-array ranges of 100 m, 250 m, and 500 m, respectively. The MFP results on Figure 3-11 and Figure 3-12 are presented in decibels, 10log10 (A_i) and 10log10 (A_c), so that a perfect MFP localization result would produce a peak of 0 dB. In addition, the MFP results were computed with a range and depth resolutions of 5 m and 1 m, respectively. The search domain was the same in each case: $R \leq 600$ m, $0 \leq z \leq 60$ m.

Figure 3-10 shows the rms centroid distance ζ from the ray-based back-propagation calculations as a function of source-array range. Although the results at each range display several local minima, the global minimum rms distance, ζ_{min}, in each case occurs unambiguously near the actual source range. The alternative minima correspond to chance coalescence of the back-propagating impulses. For example, the three arrows shown on the top panel of Figure 3-10 correspond to the marker locations shown on Figure 3-9.

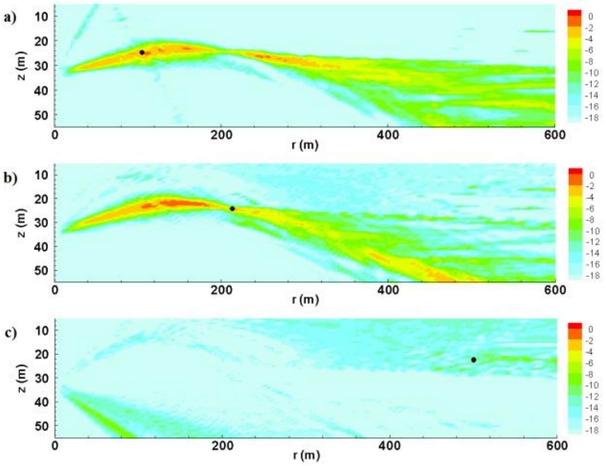


Figure 3-11: Ambiguity surface, A_i, for incoherent Bartlett matched field processing from equation 3-2 vs. range and depth for source-array ranges of 100 m (a), 250 m (b), and 500 m (c). In this case, one or more peaks occur in A_i near the actual source range and depth. The small black circle marks the ambiguity surface peak nearest to the source location.

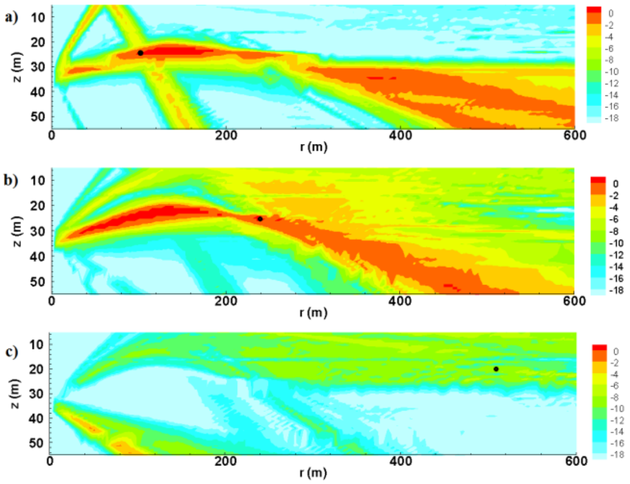


Figure 3-12: Same as Figure 3-11 except this figure shows the ambiguity surface, A_c, for coherent Bartlett matched field processing from equation 3-3. Here again, peaks occur in A_c near the actual source range and depth. These results are based on the STR-estimated impulse response.

The MFP results on Figure 3-11 and Figure 3-12 are less decisive. Ambiguity surface values near or above that found close to the source location occur at multiple places within the spatial region considered. Mismatch between the actual and simulated environments is the likely reason for this indeterminacy. However, in all cases, an ambiguity-surface peak, $(A_i)_p$ and $(A_c)_p$, can be found near the actual source location and these are marked by small black circle on Figure 3-11 and Figure 3-12. These marked peaks have dB levels of –3 to –10. The downward curving streak of high ambiguity surface values in parts a) and b) of Figure 3-11 and Figure 3-12 follows the direct ray path linking the source and the array. This streak is absent in part c) of these two figures because surface- and bottom-reflected ray paths primarily link the source and the array at the 500 m range.

Table 3-1: Performance comparision of ray-based localization, incoherent MFP, and coherent MFP.

	R=100	R=250	R=500
Ray-based localization			
ζ_{min}/ζ_{sl}	1.63/9.06	3.80/13.26	4.58/10.86
Range Error (m)	0 m	0 m	-10 m
Depth Error (m)	3.3 m	1.5 m	6 m
Incoherent MFP			
$(A_i)_p/(A_i)_{sl}$	0.63/0.58	0.37/0.51	0.15/0.21
Range Error (m)	5 m	10 m	0 m
Depth Error (m)	6 m	5 m	7 m
Coherent MFP			
$(A_c)_p/(A_c)_{sl}$	0.60/0.62	0.53/0.68	0.25/0.41
Range Error (m)	0 m	0 m	35 m
Depth Error (m)	5 m	5 m	9 m

Performance comparisons using ζ_{min} and the marked MFP peaks are provided in Table 3-1 which lists extrema-to-side lobe ratios, ζ_{min}/ζ_{sl}, $(A_i)_p/(A_i)_{sl}$, and $(A_c)_p/ (A_c)_{sl}$; range error, and depth error. The subscript "*sl*" in the prior listing and in Table 3-1 stands for side lobe. Overall, the three techniques provide comparable localization accuracy for the chosen ambiguity surface peaks. The tabulated source localization errors are acceptably small for the likely amount of mismatch between the actual and simulation environments. However, there are at least two reasons to prefer the simple ray-based back-propagation approach. It is the least burdensome computationally, and, more importantly, it provides correct and unambiguous localization results when the MFP results do not. In particular, the peak-to-side-lobe ratio is near or below unity for all six MFP calculations, while the dynamic range between ζ_{min} and ζ_{sl} is more than a factor of two for all three back-propagation calculations.

3.5 Summary and conclusions

Synthetic time reversal (STR) is simple means for performing blind deconvolution in multipath environments that relies on generic characteristics of the acoustic rays (or modes) that connect the source and a receiving array to produce unique estimates of source-signal and impulse response waveforms. The current study focused on how array size, signal-to-noise ratio, and propagation characteristics influence the STR output; and on how STR impulse responses might be used to estimate the location of a remote unknown source. The results provided here are based on propagation simulations and underwater sound measurements involving a single source and a linear vertical receiving array at signal frequencies of several kilohertz and source-array ranges of 100 m to 500 m in water that is 60 m deep.

CHAPTER 4

Broadband sparse-array blind deconvolution using frequency-difference beamforming

Underwater acoustic communication relies on sending and receiving sound below water to transmit messages. Underwater communication is complicated by factors like multi-path propagation, time variations of the environment, small available bandwidth and strong signal attenuation, especially over long ranges (more than 10 km). These factors cause distortion in the received signal and limit channel capacity. Synthetic time reversal (STR) can be effective in reducing the influence of multipath propagation and thereby produce useful reconstructions of the original signal. For this chapter, STR has been applied to broadband signal pulses (11-19 kHz) recorded with a vertical 16-element receiving array having a 3.75-m-spacing between elements and 50 kHz sampling rate. Both simple propagation simulations and measured results from the FAF06[4] experiment involving 2.2 km of down slope propagation from 46 m to 92 m water depth are considered. The sparse recording array and the high frequency source signal caused conventional beamforming to fail. Being able to determine consistent ray paths is the key to successful utilization of STR. Hence, we developed a novel beamforming technique (Frequency-Difference Beamforming) which works well with sparse arrays at high frequencies. Here, the source-signal's phase is estimated by beamforming a nonlinear product of complex signal amplitudes at the difference frequency $\omega_2 - \omega_1$. This chapter describes how STR is

---

[4] Focused Acoustic Field 2006

implemented even when the receiving-array elements are many wavelengths apart and conventional beamforming is inadequate. The cross-correlation coefficient between the source-broadcast and STR-reconstructed-signal waveforms for the simulations and experiments are 98% and 91-92%, respectively. In addition, frequency-difference beamforming can be used to determine signal-path-arrival angles that conventional beamforming cannot. Both sound source and receiver array are stationary and any vertical deviation between receivers has been ignored. Also, the environment has been assumed time invariant during the recording of the signals.

4.1 Introduction

When the receiver is a sparse array of hydrophones, it may not be suitable for conventional delay-and-sum plane-wave beamforming techniques. In addition, the received field might not be modeled precisely as a superposition of plane waves propagating across the array aperture (i.e., modeling mismatch). In previous chapter, it was shown that ray-based STR performs well when the acoustic propagation is well described by a ray-path sum in a mid-frequency region (1.5 - 4 kHz) and the receiving array is vertical with sufficient element density so that conventional delay-and-sum beamforming can be used to distinguish ray-path-arrival directions (Sabra *et al*. 2010, Abadi *et al*. 2012). However, ray-based STR fails when no persistent ray-arrival appears in the beamforming output.

The purpose of this chapter is two-fold: first, present STR blind deconvolution results for source signal estimation when the receiving array is sparse and conventional beamforming is not appropriate for the frequency band of interest (11-19 kHz), and second, introduce an unconventional beamforming technique based on manufacturing frequency differences from the array recordings that allows STR to be successful with sparse array measurements in the presence of modeling mismatch. The technique presented in this chapter is similar to directional spectral analysis (Bennett, 1985) with a difference: directional spectral analysis estimated the principal direction and spectral amplitude; however, the unconventional beamforming technique presented in this chapter manufactures lower-frequency signal phase information from the higher-frequency broadband signal recordings.

STR signal reconstruction results are reported for simple simulations involving three acoustic paths, and then for comparable propagation measurements made during the FAF06 experiment conducted in the Mediterranean Sea (Song *et al*. 2009, 2010). In both cases the signal is a tapered linear-frequency-modulation (LFM) chirp (11-19 kHz), and the receiving array is vertical with 16 elements spaced almost 40 signal-center-frequency wavelengths apart. Interestingly, the unconventional frequency-difference beamforming technique is successful finding ray-path directions when conventional beamforming is not, and likely has applications beyond STR blind deconvolution.

4.2 Mathematical Formula

Material in this section presents the formal development of frequency-difference beamforming technique.

4.2.1 Frequency-Difference Beamforming

The Fourier transforms, $P_j(\omega)$, of the received signals can be written in terms of the Fourier transform of the signal, $S(\omega)=|S(\omega)|e^{i\varphi_s(\omega)}$, and the environment's Green's function, $G(\vec{r}_j,\vec{r}_s,\omega)$ (the Fourier transform of the environment's impulse response),

$$\frac{1}{2\pi}\int_{-\infty}^{+\infty} p_j(t)e^{i\omega t}dt = P_j(\omega) = G(\vec{r}_j,\vec{r}_s,\omega)|S(\omega)|e^{i\varphi_s(\omega)} \qquad 4.1$$

when the acoustic propagation is independent of time for the duration of the signal. Here, $\vec{r}_s$ is the unknown source location, and $\varphi_s(\omega)$ is the source signal's phase as a function of frequency ω.

A simple normalization of $P_j(\omega)$ in equation 4.1 eliminates the signal amplitude,

$$\tilde{P}_j(\omega) = \frac{P_j(\omega)}{\sqrt{\sum_{j=1}^{N} \left| P_j(\omega) \right|^2}} = \frac{G(\vec{r}_j, \vec{r}_s, \omega)}{\sqrt{\sum_{j=1}^{N} \left| G(\vec{r}_j, \vec{r}_s, \omega) \right|^2}} e^{i\varphi} \qquad 4.2$$

Frequency-difference beamforming stems from the following ray-path approximation for the sound channel's impulse response:

$$G(\vec{r}_j, \vec{r}_s, \omega) = \sum_{l=1}^{L} A_{lj} \exp\left\{ i\omega r_{lj} / \bar{c} \right\} \qquad 4.3$$

Here L is the number of ray paths between the sound source and receiving array, $1 \leq l \leq L$, A_{lj} is amplitude for each ray to each receiver, r_{lj} is the effective length of each ray path to each receiver, and $\bar{c}$ is an appropriate average sound speed. In general, A_{lj} is a complex number and may depend on frequency but such dependence is neglected here. An equivalent formulation based on a modal sum, instead of equation 4.3, is likely possible but is not discussed here. Combining equation 4.2 and 4.3 produces:

$$\tilde{P}_j(\omega) = \frac{e^{i\varphi_s(\omega)}}{\sqrt{\sum_{j=1}^{N} \left| G(\vec{r}_j, \vec{r}_s, \omega) \right|^2}} \sum_{l=1}^{L} A_{lj} \exp\left\{ i \frac{\omega r_{lj}}{\bar{c}} \right\} \qquad 4.4$$

This equation explicitly shows how the frequency ω influences phase, even though the path amplitudes A_{lj}, path lengths r_{lj}, and average sound speed $\bar{c}$ are unknown. Equation 4.4 can be developed into an expression that includes a frequency difference that is small enough for plane-wave beamforming to be effective. These steps are: evaluate equation 4.4 at two different frequencies $\omega_2 > \omega_1$, complex conjugate the ω_1-evaluation, and form the normalized field product,

$$\tilde{P}_j^*(\omega_1)\tilde{P}_j(\omega_2) = \frac{\exp\{i(\varphi_s(\omega_2)-\varphi_s(\omega_1))\}}{\sqrt{\sum_{j=1}^{N}\left|G(\vec{r}_j,\vec{r}_s,\omega_1)\right|^2}\sqrt{\sum_{j=1}^{N}\left|G(\vec{r}_j,\vec{r}_s,\omega_2)\right|^2}}\sum_{l=1}^{L}\sum_{m=1}^{L}A_{lj}^*A_{mj}\exp\left\{i\frac{(\omega_2 r_{mj}-\omega_1 r_{lj})}{\bar{c}}\right\} \tag{4.5}$$

The phase relationship embodied in equation 4.5 is of interest because the source phase difference, $\varphi_s(\omega_2) - \varphi_s(\omega_1)$, appears on the right side, and because the exponential phase inside the double sum is proportional to $\omega_2 - \omega_1$ when $m = l$. In equation 4.5 the square-root factors are real functions, so equating phases in equation 4.5 leads to:

$$\begin{aligned}&\arg\left(\tilde{P}_j^*(\omega_1)\tilde{P}_j(\omega_2)\right)=\\&\varphi_s(\omega_2)-\varphi_s(\omega_1)+\arg\left[\sum_{l=1}^{L}A_{lj}^*A_{lj}\exp\left\{i\frac{(\omega_2-\omega_1)r_{lj}}{\bar{c}}\right\}+\sum_{l\neq m}^{L,L}A_{lj}^*A_{mj}\exp\left\{i\frac{(\omega_2 r_{mj}-\omega_1 r_{lj})}{\bar{c}}\right\}\right]\end{aligned} \tag{4.6}$$

Here, the double sum over ray paths has been separated into diagonal ($l = m$) and off-diagonal ($l \neq m$) terms. The diagonal terms in equation 4.6 explicitly include the frequency difference $\omega_2 - \omega_1$ and take the following form:

$$\textit{diagonal terms} = \sum_{l=1}^{L}B_{lj}\exp\left\{i\frac{(\omega_2-\omega_1)r_{lj}}{\bar{c}}\right\} \tag{4.7}$$

where $B_{lj} = A_{lj}^*A_{lj}$ (no sum implied). Interestingly, equation 4.7 is functionally the same as equation 4.3 with ω replaced by $\omega_2 - \omega_1$. In both equation 4.3 and 4.7 the r_{lj} correspond to L signal-paths having arrival angles θ_l at the receiving array. Thus, conventional delay-and-sum beamforming of the field product $\tilde{P}_j^*(\omega_1)\tilde{P}_j(\omega_2)$ at the difference frequency $\omega_2 - \omega_1$ may yield a useful estimate of the signal phase difference $\varphi_s(\omega_2) - \varphi_s(\omega_1)$ when the beam steering angle is equal to θ_l, and the off-diagonal terms in equation 4.6 are unimportant as will be described below.

In the current investigation, such a signal-phase-difference estimate is developed from

$$b(\theta,\omega_1,\omega_2) = \sum_{j=1}^{N} \tilde{P}_j^*(\omega_1)\tilde{P}_j(\omega_2)\exp\{-i(\omega_2-\omega_1)\tau(\theta,\vec{r}_j)\} \qquad 4.8$$

where b is the frequency-difference beamforming output, τ is the time delay, and θ is the beam steering angle defined with respect to broadside ($\theta = 0$). If $\tilde{P}_j^*(\omega_1 = 0)$ is a non-zero constant that is independent of j, then equation 4.8 reduces to conventional delay-and-sum beamforming in the limit $\omega_1 \to 0$,

$$\lim_{\omega_1\to 0} b(\theta,\omega_1,\omega_2) = b(\theta,0,\omega_2) \propto \sum_{j=1}^{N} \tilde{P}_j(\omega_2)\exp\{-i\omega_2\tau(\theta,\vec{r}_j)\} \qquad 4.9$$

For evenly spaced elements along a linear vertical array (the array geometry of interest here), the time delays in an iso-speed sound channel are simply related to the beam steering angle,

$$\tau(\theta,\vec{r}_j) = \tau(\theta,z_j) \cong (j-1)\left(d/\bar{c}\right)\sin\theta \qquad 4.10$$

where d is the distance between array elements. When there is significant vertical variation in the channel's sound speed $c(z)$, the time delays $\tau(\theta, z_j)$ can be selected in accordance with ray group velocities to account for the curvature of the incoming wave-fronts (Dzieciuch *et al*. 2001, Roux *et al*. 2008). However, In blind deconvolution, $c(z)$ is considered unknown. So, an appropriate constant value of $\bar{c}$ will be used to generate the results.

4.2.2 Structure of the Field Product

The mathematical structure of the field product in equation 4.8 can be illustrated by using equation 4.3 in the simple case of two ray paths (L = 2) when the A_{lj} are real coefficients. First combine equation 4.1 and 4.3 with L = 2, to find

$$P_j(\omega) = S(\omega)\left(A_{1j}\exp\{i\,\omega r_{1j}/\bar{c}\} + A_{2j}\exp\{i\,\omega r_{2j}/\bar{c}\}\right) \qquad 4.11$$

Thus, the field product becomes:

$$\begin{aligned} &P_j^*(\omega_1)P_j(\omega_2) = \\ &S^*(\omega_1)S(\omega_2)\times\left(A_{1j}\exp\left\{\frac{i\omega_1 r_{1j}}{\bar{c}}\right\} + A_{2j}\exp\left\{\frac{i\omega_1 r_{2j}}{\bar{c}}\right\}\right)^*\left(A_{1j}\exp\left\{\frac{i\omega_2 r_{1j}}{\bar{c}}\right\} + A_{2j}\exp\left\{\frac{i\omega_2 r_{2j}}{\bar{c}}\right\}\right) \end{aligned} \qquad 4.12$$

which, after some algebra, reduces to

$$\begin{aligned} &P_j^*(\omega_1)P_j(\omega_2) = \\ &S^*(\omega_1)S(\omega_2)\left(\begin{aligned} &A_{1j}^2\exp\left\{i\frac{\Delta\omega}{\bar{c}}r_{1j}\right\} + A_{2j}^2\exp\left\{i\frac{\Delta\omega}{\bar{c}}r_{2j}\right\} + \\ &2A_{1j}A_{2j}\cos\left[\left(\frac{\omega_1 + \frac{1}{2}\Delta\omega}{\bar{c}}\right)\left(r_{2j} - r_{ij}\right)\right]\exp\left\{i\frac{\Delta\omega}{\bar{c}}\left(\frac{r_{1j} + r_{2j}}{2}\right)\right\} \end{aligned}\right) \end{aligned} \qquad 4.13$$

where $\Delta\omega = \omega_2 - \omega_1$. The first two terms inside the big parentheses on the right side of (4.13) are the diagonal terms of the field product. They follow the form of (4.7) and their phases only depend on $\Delta\omega$, $\bar{c}$ and the two ray path lengths. When $P_j^*(\omega_1)P_j(\omega_2)$ from equation 4.13 is beamformed at the difference frequency, $\Delta\omega$, these diagonal terms will make a contribution to $b(\theta,\omega_1,\omega_2)$ that does not depend on ω_1.

On the other hand, the third term inside the big parentheses on the right side of equation 4.13 results from combining the two off-diagonal terms of the field product. It depends on $\Delta\omega$,

ω_1, $\bar{c}$ and the sum and difference of the two ray path lengths. Thus, when $P_j^*(\omega_1)P_j(\omega_2)$ from equation 4.13 is beamformed at the difference frequency, $\Delta\omega$, this term will change as ω_1 is varied. Such ω_1-dependent contributions to $b(\theta,\omega_1,\omega_2)$ can be considered structured interference or noise as will be evident in the simulation results in Sec. 4.3. In general, when L ray paths connect the source and the receiving array, the number of desired ω_1-independent diagonal terms (*signal*) increases like L while the number of undesired ω_1-dependent off-diagonal-term contributions (*noise*) increases like $L(L-1)/2$. Thus, for an arbitrary L, there may be an inherent limit to frequency-difference beamforming's utility since its signal-to-noise ratio may decrease like $(L-1)^{-1}$ with increasing L. However, this limit – if it exists – has not been reached by either the simulation or experimental results presented in the following sections.

4.2.3 Implementation of STR with Frequency-Difference Beamforming

With this understanding of frequency-difference beamforming, the phase relationship embodied in (4.8) is:

$$\arg\big(b(\theta,\omega_1,\omega_2)\big) = \\ \varphi_s(\omega_2) - \varphi_s(\omega_1) + \arg\left(\sum_{j=1}^{N}\sum_{l=1}^{L}\sum_{m=1}^{L} A_{lj}^* A_{mj}\ \exp\left\{ i\left(\frac{(\omega_2 r_{mj} - \omega_1 r_{lj})}{\bar{c}} - (\omega_2-\omega_1)\tau(\theta,\vec{r}_j)\right)\right\}\right) \qquad 4.14$$

where the final term involving the triple sum can be assumed approximately a constant angle, K, independent of ω_1 and ω_2 when θ coincides with θ_l, a signal-path arrival angle, and the off-diagonal terms of the field product are unimportant. Under these conditions, an estimate of the source-signal phase can then be developed throughout the signal bandwidth by recursively applying equation 4.14 after assuming an initial signal phase of zero. For example, if $P_j(\omega)$ is obtained from an FFT of $p_j(t)$ at evenly spaced frequencies ω_q, $1 \le q \le Q$, then the source-signal phase estimate $\hat{\varphi}(\omega_q)$ produced by equation 4.14 at these frequencies when $\theta = \theta_l$ is:

$$\begin{aligned}
&\hat{\varphi}_s(\omega_1) = 0 \\
&\hat{\varphi}_s(\omega_2) = \arg\big(b(\theta_l, \omega_1, \omega_2)\big) - K \\
&\hat{\varphi}_s(\omega_3) = \arg\big(b(\theta_l, \omega_2, \omega_3)\big) + \varphi_s(\omega_2) - K = \arg\big(b(\theta_l, \omega_2, \omega_3)\big) + \arg\big(b(\theta_l, \omega_1, \omega_2)\big) - 2K \\
&\quad\vdots \\
&\hat{\varphi}_s(\omega_q) = \arg\big(b(\theta_l, \omega_{q-1}, \omega_q)\big) + \varphi_s(\omega_{q-1}) + K = \sum_{p=2}^{q} \arg\big(b(\theta_l, \omega_{p-1}, \omega_p)\big) - (q-1)K \\
&\quad\vdots \\
&\hat{\varphi}_s(\omega_Q) = \arg\big(b(\theta_l, \omega_{Q-1}, \omega_Q)\big) + \varphi_s(\omega_{Q-1}) = \sum_{p=2}^{Q} \arg\big(b(\theta_l, \omega_{p-1}, \omega_p)\big) - (Q-1)K.
\end{aligned} \tag{4.15}$$

Although the terms on the right involving K remain unknown, they lead to a phase contribution to $\hat{\varphi}_s(\omega)$ that is proportional to frequency, and such a phase contribution merely offsets the time origin of the STR-reconstructed waveforms. Therefore, STR cannot be used to determine absolute timing information, but the STR-determined waveform shapes are independent of K when it is constant.

The source-signal phase estimate $\hat{\varphi}_s(\omega)$ can be combined with equation 4.2 to produce a normalized estimate of the environment's Green's function (similar to equation 2-5):

$$\hat{\tilde{G}}(\vec{r}_j, \vec{r}_s, \omega) = \tilde{P}_j(\omega) \exp\{-i\hat{\varphi}_s(\omega)\} = \frac{G(\vec{r}_j, \vec{r}_s, \omega)}{\sqrt{\sum_{j=1}^{N} \left|G(\vec{r}_j, \vec{r}_s, \omega)\right|^2}} \exp\{i\varphi_s(\omega) - i\hat{\varphi}_s(\omega)\} \tag{4.16}$$

From here, an estimate of the Fourier transform of the source signal, $\hat{S}(\omega)$, can be obtained by using $\hat{\tilde{G}}(\vec{r}_j, \vec{r}_s, \omega)$ from equation 4.16, for back-propagation (similar to equation 2-7),

$$\hat{S}(\omega) = \sum_{j=1}^{N} \hat{\tilde{G}}^{*}(\vec{r}_j, \vec{r}_s, \omega) P_j(\omega) \qquad 4.17$$

or inverse filtering,

$$\hat{S}(\omega) = \frac{1}{N} \sum_{j=1}^{N} P_j(\omega) \Big/ \hat{\tilde{G}}(\vec{r}_j, \vec{r}_s, \omega) \qquad 4.18$$

The estimated signal spectra from back-propagation and inverse filtering are identical for STR (Sabra & Dowling, 2004). Finally, an inverse Fourier transform of equation 4.16, and 4.17 or 4.18 recovers the source-to-array impulse responses and estimated source-signal waveform in the time domain, for example:

$$\hat{s}(t) = \int_{-\infty}^{+\infty} \hat{S}(\omega) e^{-i\omega t} d\omega \qquad 4.19$$

This formulation of STR benefits in three important ways from the inclusion of frequency-difference beamforming. First, the frequency difference used in equation 4.8 is only limited by the bandwidth of the signal and the spacing between FFT frequency samples; thus, it may often be chosen to suit the situation at hand. For example, in terms of cyclic frequency, *f*, these limits are 12.2 Hz $\leq f_2 - f_1 \leq$ 8 kHz for the STR performance results shown in the next two sections. In fact, the frequency difference can be varied to resolve multiple signal arrival directions having nearly the same arrival angle, even when these paths are not identifiable with conventional beamforming (see Figure 4-4 and Figure 4-12). And second, when several signal propagation paths reach the receiving array at nearly the same angle, θ', a small value of the frequency difference may be chosen so that the receiving array does not distinguish these paths. In this case, when the beam steering angle is θ', the frequency-difference beamforming estimate, $\hat{\varphi}_s(\omega)$, of the source-signal phase from equation 4.12 is based on the average of these propagation paths, and STR blind deconvolution results may be excellent even though no single propagation path is resolved (see Figure 4-7 and Figure 4-11). Finally, the lower frequency ($\omega_2 - \omega_1$) produced from frequency-difference beamforming makes the STR process and beamforming more robust to potential mismatch between the actual measured field and the common plane-wave assumption for a sparse large-aperture array (see Figure 4-1 and Figure 4-2).

4.3 STR Results from Simple Broadband Propagation Simulations

To determine the possible performance of STR in conjunction with frequency-difference beamforming, broadband simulations are undertaken that approximately mimic the signals and geometry of the FAF06 experiment. The acoustic environment is a 92-m-deep range-independent ideal waveguide with a uniform sound speed of 1500 m/s having a flat surface and bottom (see Figure 4-1). The source at 39 m depth broadcasts a 60 ms cosine-tapered LFM chirp from 11 kHz to 19 kHz. The linear vertical receiving array is located 2.2 km away from the source and centered at a depth of 52.8 m. It is composed of 16 elements spaced d = 3.75 m apart for an overall array length of $L_A = (N - 1)d$ = 56.25 m. At the source-signal center frequency (15 kHz), the element spacing corresponds to 37.5 wavelengths, making the array extremely sparse.

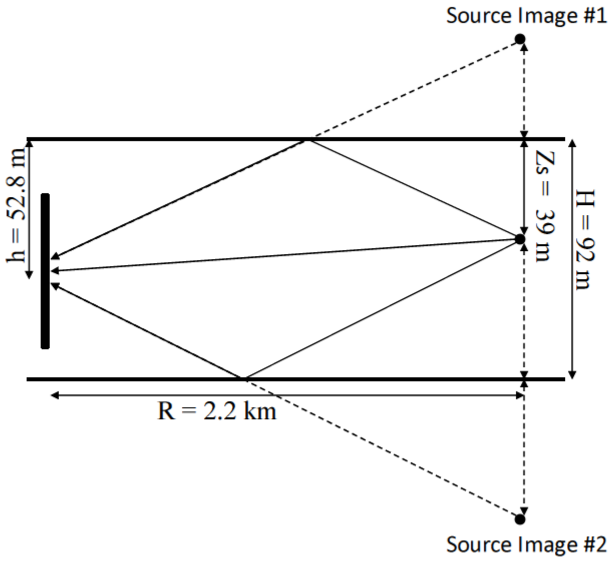


Figure 4-1: Ideal sound channel that supports three propagation paths using the method of images. The source signal is a cosine tapered linear frequency modulation (LFM) sweep from 11 kHz to 19 kHz and is broadcast from a depth of zs = 39 m.

Only the three ray paths shown in Figure 4-1 are considered in these simulations using the method of images to ensure that the channel's delay spread is much smaller than the signal duration, as in the case of the FAF06 experiment. The results of the simulations are provided in Figure 4-2 through Figure 4-7. Note that the field received by the sparse array cannot be modeled precisely as three propagating plane waves for the current geometry and high frequency signal because the source is in the near field of the array at the signal's center frequency wavelength of λ = 10 cm. Following Kinsler *et al*. (2000) or equivalently Ziomek (1993), the array's far-field is reached (in free space) when $L_A^2/4/r$ is less than unity, but for the geometry considered here with r = 2.2 km this parameter is greater than one, $L_A^2/4\lambda r = (56.25m)^2/[4(0.1m)2200m] \approx 3.6$

.

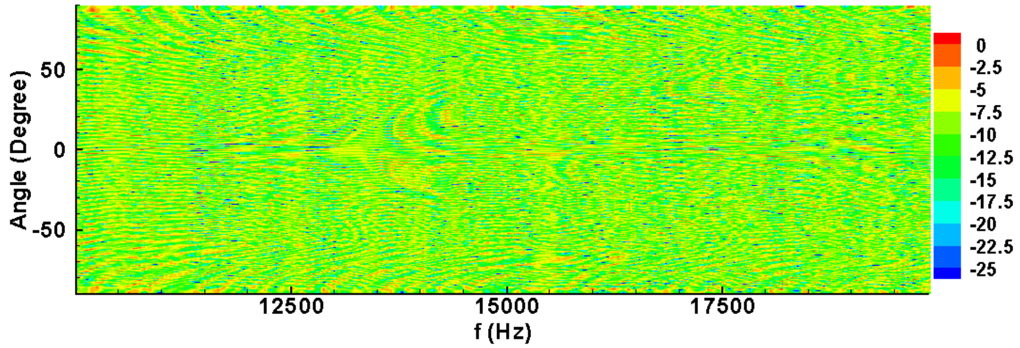


Figure 4-2: Conventional plane-wave beamforming output for the simulated signals as a function of look angle and frequency in the signal bandwidth (11-19 kHz) in dB scale.

Figure 4-2 through Figure 4-5 illustrate the performance of conventional and frequency-difference beamforming with $\bar{c}$ = 1500 m/s. First, Figure 4-2 shows the results of conventional plane-wave delay-and-sum beamforming of the array-recorded signals as a function of frequency, $f = \omega/2\pi$ (in Hz), over the bandwidth of the signal (11-19 kHz) and the beam steering angle (from -90° to 90°). As expected for a sparse receiving array along with mismatch in modeling (i.e., the existence of wave front curvature), the results are featureless. By comparison, the frequency-difference beamforming results, developed from equation 4.8 and shown in Figure 4-3, display a clearer structure. The five panels in Figure 4-3 show the same range of frequency (horizontal axis), $f_1 = \omega_1/2\pi$, each representing five increasing frequency differences, $\Delta f = (\omega_2 - \omega_1)/2\pi$: a) 12.21 Hz, b) 48.83 Hz, c) 195.3 Hz, d) 781.25 Hz, and e) 1562.5 Hz. Since the sampling rate is 50 kHz and the FFT size is chosen as 4096 points, the result for the lowest possible frequency difference, Δf = 12.21 Hz, is shown in Figure 4-3a). The angular width of the broad central ridge in this panel is consistent with a linear array having $kL_A = 2.87$, where k is the wave number based on the frequency difference, $k = 2\pi\Delta f/\bar{c}$ = 0.0511 m^{-1}, and L_A is 56.25 m. Figure 4-3b) shows a similar result with Δf = 48.83 Hz, a factor of four increase in the frequency difference over that in Figure 4-3a. Here, the central ridge has narrowed in accordance with an increase in frequency. And, as anticipated from the discussion of equation 4.13, some intermittent side lobes from the off-diagonal terms of the field product emerge. Figure 4-3c shows a result with Δf = 195.3 Hz, a factor of four increase in the frequency difference above that in Figure 4-3b. Here again, the central ridge in the frequency-difference beamforming result shows a corresponding increase in resolution. However, this resolution is not yet sufficient to identify the angles of the three simulation ray paths. Perhaps more interesting in Figure 4-3c is the increased prominence of the side lobes from the off-diagonal terms of the field product, which now appear as curving structures that enhance or distort the central ridge in Figure 4-3c. At this frequency difference (195.3 Hz), the array-element spacing is approximately $\lambda/2$, so the remaining panels of Figure 4-3 display an increasingly narrower angular range to prevent repetition of the side lobe pattern (i.e., spatial aliasing). Figure 4-3d) shows results with Δf = 781.25 Hz, a factor of four increase in the frequency difference above that shown in Figure 4-3c. At this frequency difference, the side lobes from the off-diagonal terms of the field product are stronger, and the horizontal central ridge is more uneven. Figure 4-3e shows a result with Δf = 1562.50 Hz, a factor of two increase in the frequency difference above that of Figure 4-3d). At this frequency difference, kL_A (as defined above) is almost 370 and the receiving array's angular resolution is about a degree near broadside. Although it is difficult to identify in the mix of slanted side lobes in Figure 4-3e), there are three horizontal bands of varying magnitude due to the diagonal terms of the field product that lie within ±3° and correspond to the three simulation ray paths (see Figure 4-4).

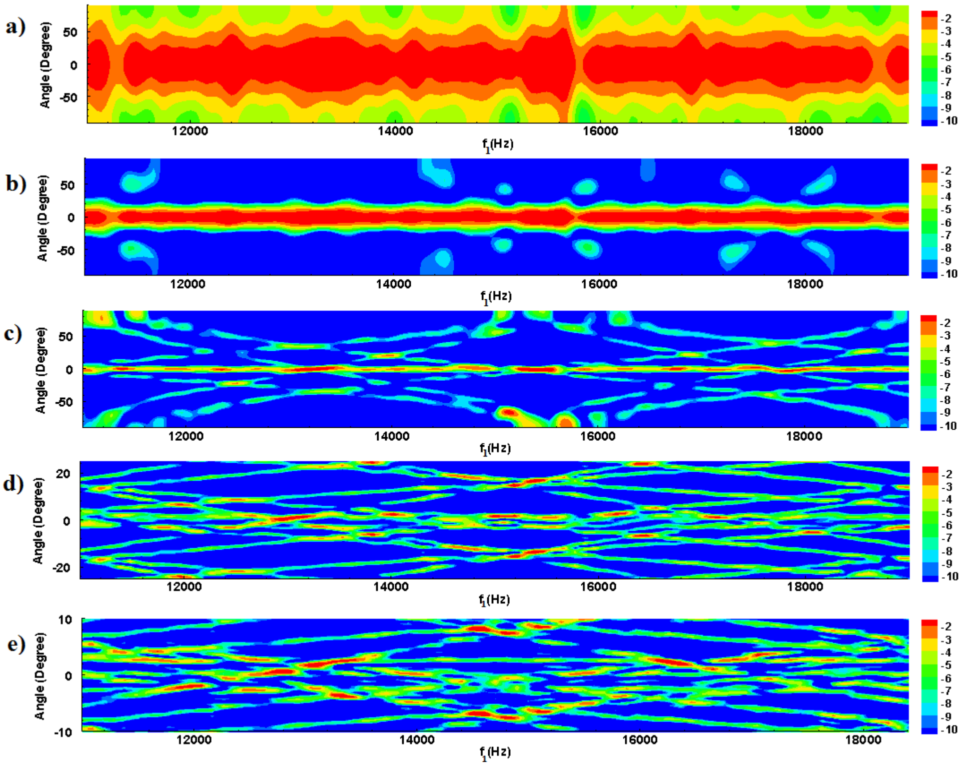


Figure 4-3: Unconventional frequency-difference beamforming output from equation 4.8 for the simulated signals as a function of look angle θ and frequency $f_1 = \frac{\omega_1}{2\pi}$ from 11 kHz to 19 kHz with various frequency-differences: (a) Δf = 12.21 Hz, (b) Δf = 48.83 Hz, (c) Δf = 195.31 Hz, (d) Δf = 781.25 Hz, and (e) Δf = 1562.5 Hz. Note that the angular range of panels (d) and (e) is reduced to capture the output structure of frequency-difference beamforming near $\theta = 0°$ for the high frequency-difference cases.

The angles of the three simulation ray paths may be recovered from the frequency-difference beamforming result shown in Figure 4-3e when the individual frequency output is integrated over the bandwidth of the signal. Within such an integration, the persistent contribution of the field product's diagonal terms reinforces their prominence while the slanting side lobes seen in Figure 4-3 from the field product's off-diagonal terms disperses their impact. The results of such an integration are shown in Figure 4-4 where the dashed curve is for conventional beamforming while the solid curve is for frequency-difference beamforming at Δf = 1562.50 Hz. The conventional beamforming result obtained by integrating horizontally through Figure 4-2 fails to show any prominent ray path direction. On the other hand, the integrated frequency-difference beamforming result clearly displays three peaks at each of the arrival angles of the propagation simulations, –2.4°, 0.3°, and 2.6°, which correspond to bottom-reflected, direct, and surface-reflected ray paths, respectively.

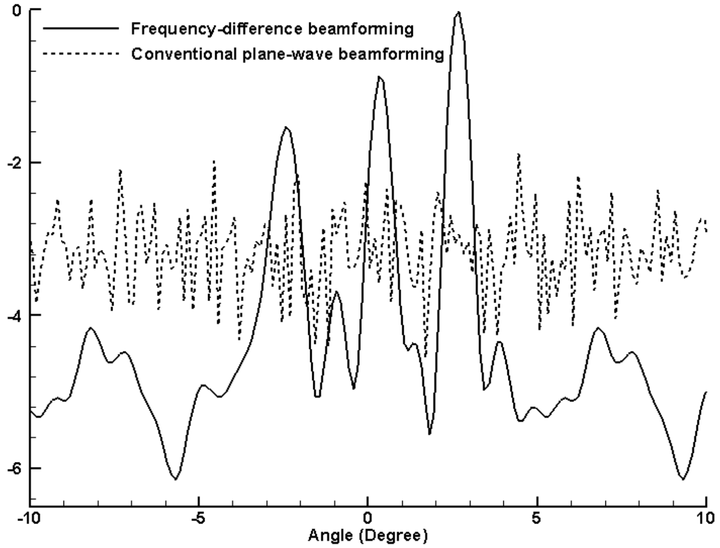


Figure 4-4: Beamforming results incoherently summed over the frequency band using the simulated signals: Frequency-difference beamforming with Δf = 1562.5 Hz (solid line) from Figure 4-3e clearly indicates signal-arrival directions whereas conventional plane-wave beamforming (dash line) from Figure 4-2 does not. The three peaks correspond to the three ray-path arrival angles (–2.4°, 0.3°, and 2.6°) of the propagation simulations, respectively.

For a more complete picture, a comparison of integrated frequency-difference beamforming results is shown in Figure 4-5 for a factor of ten change in Δf for a fixed range of beam steering angle $-20° \leq \theta \leq +20°$. The ten curves proceed from Δf = 195.3 Hz (bottom) to Δf = 1953 Hz (top) in uniform steps of 195.3 Hz. The third curve from the top is equivalent to the one shown in Figure 4-4 (solid). Here, the vertical axis of Figure 4-5 is linear (not dB), and each curve is normalized by its maximum value and offset upward by one unit from the curve below. At the lowest frequency-difference (bottom), there is a single broad peak near the broadside $\theta = 0°$ without side lobes. As Δf increases, the angular resolution of the beamforming improves (two or more ray arrivals close to each other can be distinguished by this beamforming technique) and the central peak separates into three peaks at the simulation ray-path angles in the middle of this figure. The three ray paths are first distinct at a frequency difference of 976.5 Hz and remain so at all higher frequency differences. However, as Δf further increases, side lobes appear on both sides and march toward $\theta = 0°$. Interestingly, this behavior is analogous to that expected for conventional narrowband beamforming when the signal and processing frequency equals Δf. However, the frequency range of the curves shown in Figure 4-5, 195.3 Hz to 1953 Hz, lies well below the frequency band of the broadcast signal (11 kHz to 19 kHz), making it more robust to potential mismatch in modeling. Thus, frequency-difference beamforming expands the possibilities for ray-path angle determination beyond those of conventional beamforming.

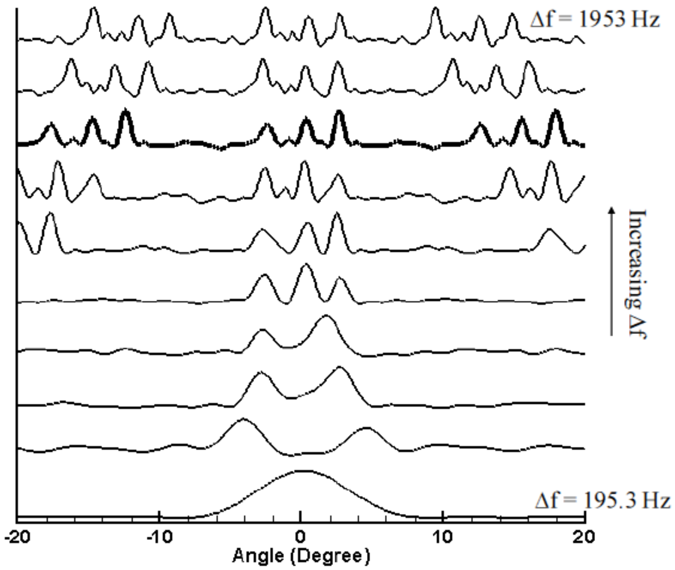


Figure 4-5: Frequency-difference beamforming results from the simulated signals integrated over the source signal's bandwidth, 11 kHz ≤ f_1 ≤ 19 kHz, vs. beam steering angle θ for ten different values of Δ*f*. The curves proceed from Δ*f* = 195.3 Hz (bottom) to Δ*f* = 1953 Hz (top) in uniform steps of 195.3 Hz. The vertical axis is linear. Each curve is normalized by its maximum value and offset vertically by one unit from the curve below. The trade-off between angular resolution and spatial aliasing is observed as Δ*f* increases.

Since the frequency-difference beamforming technique needs to integrate the beamforming output over the source signal's bandwidth to suppress the off-diagonal terms and while keeping the persistent contributions of the field product, its performance degrades when less bandwidth is available. Figure 4-6 shows integrated frequency-difference beamforming over four different source signal bandwidths. This figure indicates that frequency-difference beamforming requires at least 4 kHz of bandwidth (approximately three times the difference frequency) to resolve the arrival angles in the propagation scenario (see Figure 4.1) considered here.

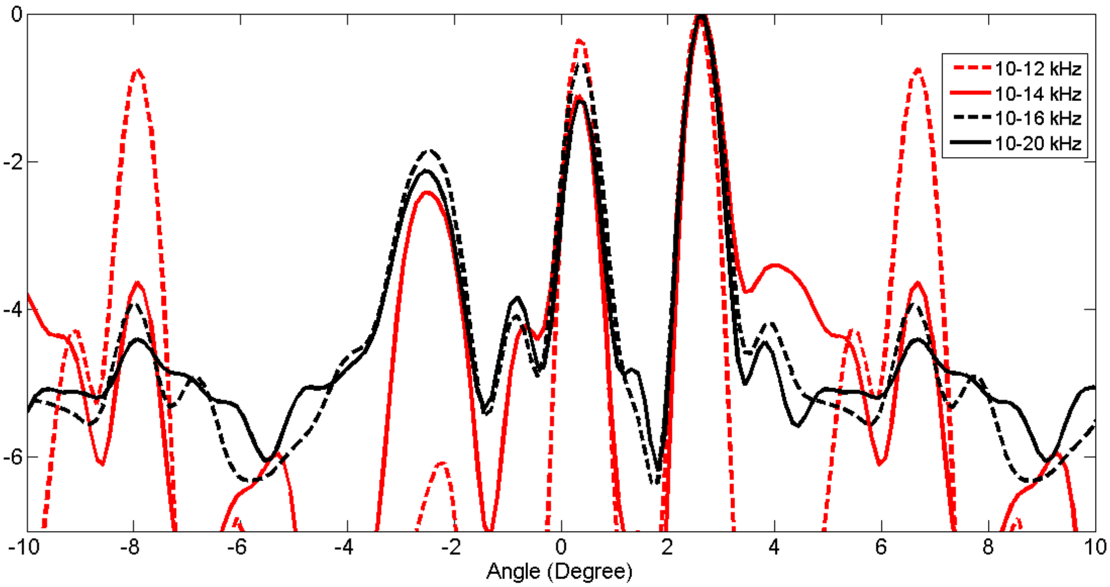


Figure 4-6: Frequency-difference beamforming results from the simulated signals integrated over the different source signal's bandwidth vs. beam steering angle θ with Δf = 1562.5 Hz.

Sample signal waveforms that illustrate the STR blind deconvolution in the simulation are shown in Figure 4-7 .The top panel, Figure 4-7a, is the original signal waveform. The middle panel, Figure 4-7b, shows the simulated received signal at the shallowest array element. The bottom panel, Figure 4-7c, shows the signal waveform that is reconstructed from the simulated array measurements using STR and frequency-difference beamforming with Δf = 12.21 Hz (see Figure 4-3a. Although the envelope of the reconstructed signal is not perfect, the cross correlation coefficient between Figure 4-7a and Figure 4-7c is surprisingly good, 98%, and this result is independent of the value of $\bar{c}$. For these simulations, comparable signal reconstruction results are obtained for frequency differences as high as 48.83 Hz when $\bar{c}$ = 1500 m/s.

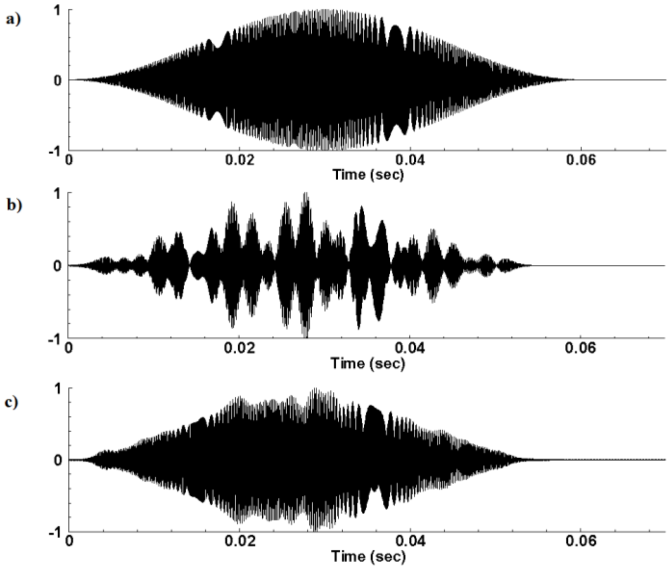


Figure 4-7: Sample input and output STR signals for the simulation results. (a) broadcast signal, a cosine-tapered 60-ms, LFM sweep (11-19 kHz), (b) received signal captured at the shallowest array element, and (c) STR estimated source signal using frequency-difference beamforming with $\Delta f = 12.21$ Hz. The cross correlation coefficient of received signal with the broadcast signal is 57%. The cross correlation coefficient of the STR reconstructed signal with the broadcast signal is 98%.

These simulation results indicate that STR, in conjunction with frequency-difference beamforming, can provide a successful means of blind deconvolution for sparse-array recordings. Furthermore, frequency-difference beamforming allows signal-arrival angles to be determined from sparse-array recordings of broadband signals in the presence of modeling mismatch. The next section explores the extent to which these findings persist with sparse-array experimental data.

4.4 STR Results from FAF06 Broadband Propagation Measurements

The focused acoustic field experiment (FAF06) was conducted off the west coast of Italy in July of 2006 (Song *et al*. 2009, 2010). For the receiving-array measurements used in this study, the source depth, source signal, source-array range, and vertical receiving array (VRA) geometry nominally match the simulations discussed in section 4.3. However, the actual geometry involved down slope propagation from a water depth of 46 m to 92 m as shown in Figure 4-8a. The sound speed profile measured near the VRA is displayed in Figure 4-8b, indicating that most of the array elements are located in a nearly constant sound speed region below the thermocline. The FAF06 broadcast signals were composed of channel-probing pulses followed by communication sequences. The current blind deconvolution and beamforming results are developed from probe-pulse broadcasts with $\bar{c}$ = 1510 m/s, and are provided in Figure 4-9 through Figure 4-12.

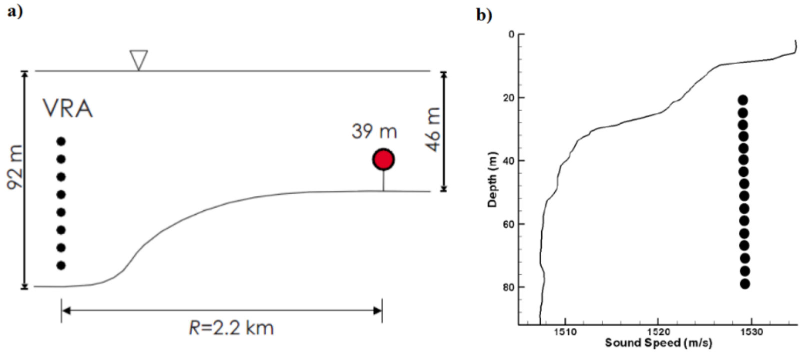


Figure 4-8: FAF06 experimental geometry (a) and a sound speed profile measured near the receiving array (b). The source was deployed near the bottom in 46 m of water. The receiving array was centered at a depth of 52.8 m in 92 m of water at a source-array range of 2.2 km. A sound speed profile measured near the receiving array shows a typical downward refracting profile during the summer.

Figure 4-9 shows the measured signals along the array, $p_j(t)$, for a source at 42.6 m depth. Here the signal from the shallowest receiver appears as the upper-most pressure time trace and the signal from the deepest receiver appears as the lowest pressure time trace. The uneven envelopes of the signal recordings show that there was sufficient multipath propagation to distort the broadcast signal at every receiver depth. The maximum cross correlation coefficient between the FAF06 broadcast signal and any of these signals varies from 37% to 75%. In addition, there is a short-duration noise pulse that follows the main transmission in the data shown in Figure 4-9 in the time interval 0.05s < t < 0.07s. It has an apparent arrival angle of 25°, and was not a part of the FAF06 signal broadcast. Such steep angle noises were observed occasionally during the experiment that were apparently due to the interaction of the long array cable with a moored surface buoy.

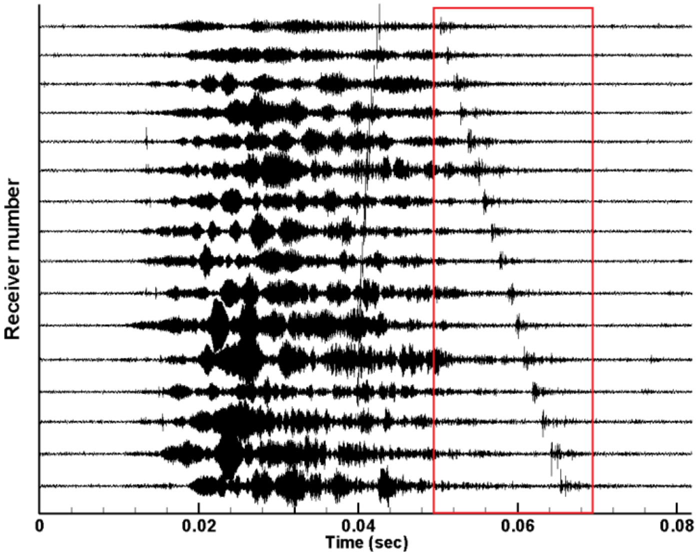


Figure 4-9: Measured FAF06 waveforms along the vertical array for a probe-signal broadcast at a source depth of 42.6 m and a source-array range of 2.2 km. The maximum cross correlation coefficient between the FAF06 broadcast signal and any of these signals varies from 37% to 75%. Note the presence of a short auxiliary noise pulse contained in the box for $t >$ 0.05s. This pulse was not a part of the FAF06 probe-signal broadcast.

Figure 4-10 provides a comparison of beamforming results for the FAF06 signals shown in Figure 4-9. Similar to Figure 4-2, the conventional beamforming result in Figure 4-10a is featureless. In contrast, the frequency-difference beamforming result shown in Figure 4-10b with Δf = 48.83 Hz displays a broad central ridge with an angular width of approximately 20° and resembles the corresponding simulation results in Figure 4-3b. Thus Figure 4-10 confirms that the primary features of the simulation results shown in Figure 4-2 and Figure 4-3 persist with measured propagation data.

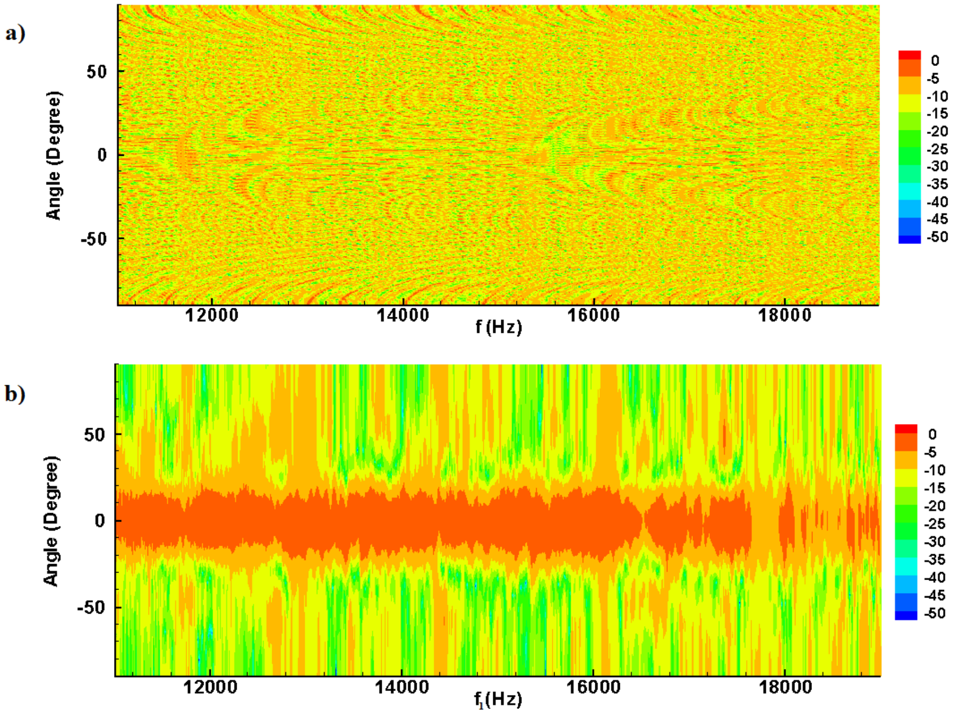


Figure 4-10: Beamforming output for the measured and signals as a function of look angle θ and frequency in the signal bandwidth (11-19 kHz) in dB scale: (a) conventional beamforming (b) frequency-difference beamforming with Δf = 48.83 Hz. Similar to Figure 4-2, conventional beamforming (a) is not useful while frequency-difference beamforming (b) shows signal structure centered on $\theta = 0$.

Figure 4-11 shows a comparison of the source signal waveforms from the FAF06 experiment. The top waveform, Figure 4-11a, is the ideal signal of Figure 4-7a corrected for the transmit amplitude response of the FAF06 sound projector. This is the best available estimate of the FAF06 broadcast pulse without a monitoring hydrophone that could actually measure the amplitude and phase response of the FAF06 sound projector. The lower three panels of Figure 4-11 show blind deconvolution results for the broadcast pulse using STR and frequency-difference beamforming with Δf = 12.21 Hz. Figure 4-11b and Figure 4-11c were recorded an hour apart with a source depth of 39 m and a source-array range of 2.2 km. The cross correlation coefficients between these two pulses and the intended broadcast pulse are 92% and 91%, respectively. This is a significant improvement over the array recordings which show an average cross correlation coefficient with the intended broadcast pulse of only 50%. Even higher cross correlations might be possible if the broadcast pulse could be more accurately compensated for the phase response of the FAF06 sound projector. The bottom waveform shown in Figure 4-11 is obtained from the array recordings shown in Figure 4-9. Although it came from near identical conditions that led to the pulses shown in Figure 4-11b and Figure 4-11c, the Figure 4-11d pulse has only an 80% correlation with the intended signal pulse. This drop in the cross correlation coefficient is caused by the presence of the short-duration noise pulse contained in the box in Figure 4-9. When the Figure 4-9 recordings are temporally trimmed to remove the noise pulse, the reconstructed signal's cross correlation increases to 85%. Although a cross correlation closer to 90% was expected after trimming, it was not feasible with this data because another short-duration noise pulse, at $t \approx 0.04$s, can be seen within main FAF06 recordings shown in Figure 4-9. Thus, as currently implemented, STR loses its effectiveness when the array recordings involve multiple sources.

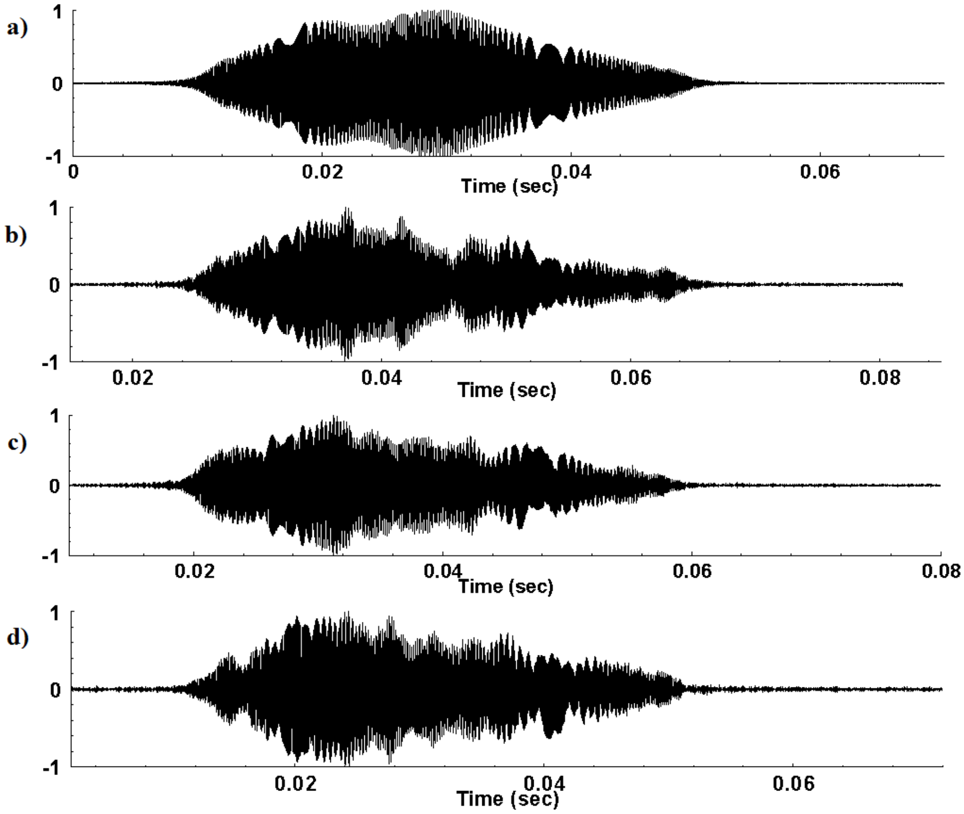


Figure 4-11: Reconstructed FAF06 waveforms. (a) intended broadcast signal, a cosine-tapered 60-ms LFM sweep corrected for the projector's amplitude response, (b) STR estimated source signal using frequency-difference beamforming with Δf = 12.21 Hz at the source depth of 39 m, (c) same as (b) but recorded an hour earlier, and (d) same as (b) but for the array recordings shown in Figure 4-9, which was collected 5 hours later than (b) at a source depth of 42.6 m. The cross correlation coefficients between the intended broadcast signal (a) and reconstructed signals (b), (c), and (d) are 92%, 91%, and 80%, respectively.

Figure 4-12 shows beamforming results integrated over the bandwidth of the signal for the portion of the FAF06 recordings shown in Figure 4-9 *within* the box, $t > 0.05$s. Similar to Figure 4-4, Figure 4-12 shows a dashed curve for conventional beamforming and a solid curve for frequency-difference beamforming with Δf = 195.3 Hz. As expected, the conventional beamforming result does not indicate any ray-path arrival angles. However, the frequency-difference beamforming result correctly indicates the arrival angle of the signal-coda, just below 0°, and the arrival angle of the short auxiliary noise pulse, 25°. Although signal arrival angles for the FAF06 probe-pulse broadcasts can be estimated with frequency-difference beamforming, they are not reported here because the downward refracting sound speed profile shown in Figure 4-8 causes all ray paths to interact with sea floor and there is insufficient knowledge of the FAF06 environment (e.g., bathymetry and geoacoustic properties) to perform accurate ray-trace calculations for comparison. Nevertheless, Figure 4-12 indicates that frequency-difference beamforming potentially has general applicability for identifying signal-arrival directions for acoustic data gathered with sparse receiving arrays since the 25° arrival angle for the short noise pulse can be determined graphically from Figure 4-9 with the known array geometry and $\bar{c} \approx$ 1510 m/s.

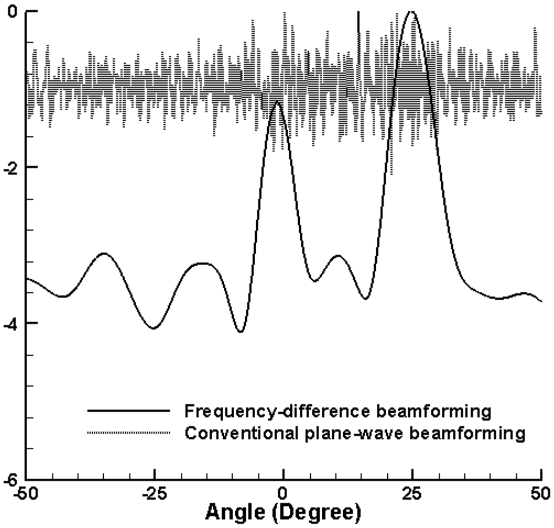


Figure 4-12: Beamforming output integrated over the frequency band using the measured signals shown in Figure 4-9 for 0.05s < *t* < 0.08s. It shows conventional plane-wave beamforming (dashed line) and frequency-difference beamforming with Δf = 195.3 Hz (solid line). The frequency-difference beamforming result displays peaks at 25° and just below 0° that correspond to the short auxiliary noise pulse and the signal's coda, respectively.

4.5 Summary and Conclusions

Synthetic time reversal (STR) is a computationally simple means for blind deconvolution of array-recorded sounds that have propagated from an unknown source to a receiving array through an unknown multipath environment. This study has explored the performance of STR with simulated and measured underwater propagation data when the receiving array was extremely sparse and conventional beamforming of the array recordings was not useful.

To achieve these sparse-array blind deconvolution results with STR, unconventional frequency-difference beamforming was developed to estimate the phase signature of the unknown source signal. STR blind deconvolution was successful, with simulation and experimental data recorded by sparse receiving array, the cross correlation coefficient between the broadcast signal and the STR reconstructed signal are routinely above 90%.

CHAPTER 5

Extensions of Unconventional Beamforming

This chapter investigates some extensions of the unconventional beamforming presented in the previous chapter for different underwater circumstances. Simulations of frequency-difference beamforming with a sparse random receiving array are presented in the first section. Then, another unconventional beamforming strategy (frequency-sum beamforming), an extension of frequency-difference beamforming that manufactures higher, not lower, frequency signal information, is presented and assessed in the second and third sections.

5.1 Frequency-difference beamforming with sparse random arrays

Randomly distributed receivers have been used for some applications in underwater acoustics where an orderly array is not feasible. For example, sonobuoys may be distributed from a passing aircraft to form random linear or area arrays. Or, the array may have non-trivial aperture in all three dimensions (a volumetric array) depending on the type of experiment (Hodges, 2010). However, if the receivers are acoustically far from each other (large element spacing and/or high frequency signal), the random receiver array is sparse and conventional plane-wave beamforming may fail. For such situations, frequency-difference beamforming may be a good technique for determining signal-path arrival angles.

To assess this possibility, a simple three-dimensional range-independent simulation was been undertaken to investigate the utility of frequency-difference beamforming with a sparse random array. The simulated ocean environment was 92 m depth and included ten receivers distributed randomly (300 m average spacing) near the origin of coordinates in the horizontal x-y plane at 30 m depth. The source was located at (x_s, y_s) = (10 km, 10 km) with respect to the center of the array at the same depth as the receivers. The source broadcast a linear frequency chirp centered at 2 kHz with 500 Hz bandwidth. Only three paths were been considered between the source and each receiver for this simulation (direct path, surface reflection, and bottom reflection). Figure 5-1 shows the geometry used for this simulation.

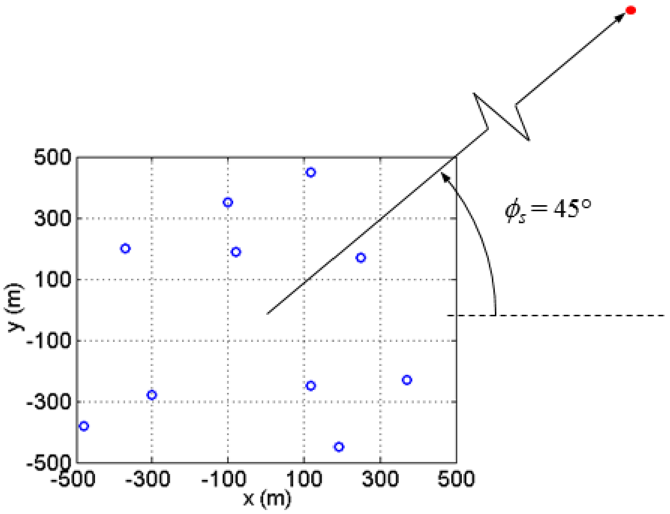


Figure 5-1: Sparse random receiver array with 300 m average spacing. Source is at (x_s, y_s) = (10 km, 10 km) with 45º respect to the origin of x-y plane which is shown by a red circle.

Figure 5-2 illustrates the performance of conventional and frequency-difference beamforming with $\bar{c}$ = 1500 m/s. Figure 5-2a shows the results of conventional plane-wave delay-and-sum beamforming of the array-recorded signals as a function of frequency, $f = \omega/2\pi$ (in Hz), over the bandwidth of the signal (1750-2250 Hz) and the beam steering angle (from -90° to 90°). As expected for a sparse random receiving array, the results are featureless. By comparison, the frequency-difference beamforming results Δf = 12.2 Hz, shown in Figure 5-2b, display a clearer structure.

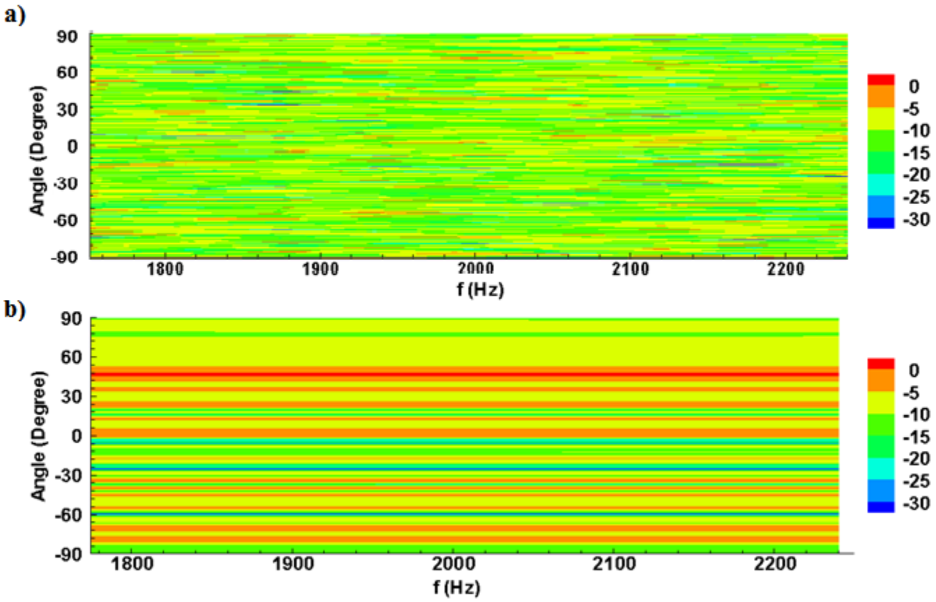


Figure 5-2: Beamforming output for the simulated signals for a sparse random receiver array. It is a function of look angle θ and frequency f in the signal bandwidth (1750-2250 kHz) in dB scale: (a) conventional beamforming (b) frequency-difference beamforming with Δf = 12.2 Hz.

Figure 5-3 shows beamforming results integrated over the bandwidth of the signal. Figure 5-3 shows a red curve for conventional beamforming and a blue curve for frequency-difference beamforming with Δf = 12.2 Hz. As expected, the conventional beamforming result does not indicate any ray-path arrival angles. However, the frequency-difference beamforming result correctly indicates the arrival angle of the direct path at 45°.

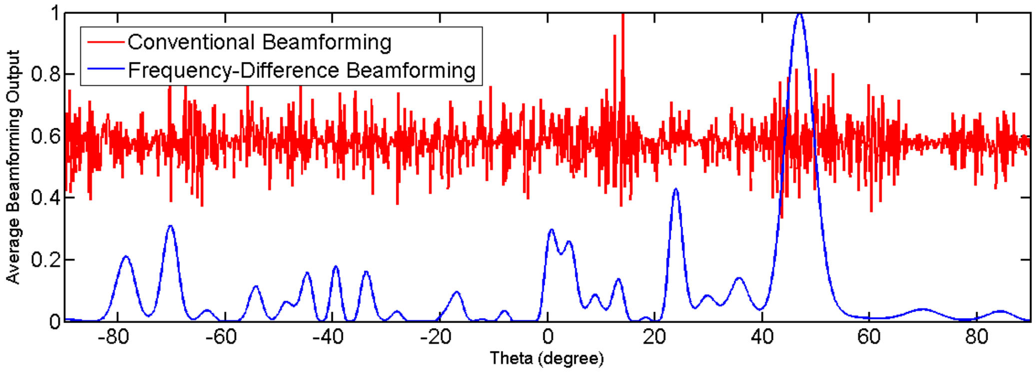


Figure 5-3: Normalized beamforming output integrated over the frequency band: conventional plane-wave beamforming (red line) and frequency-difference beamforming with Δ*f* = 12.2 Hz (blue line). The frequency-difference beamforming result displays peaks at 45° that correspond to the direct path between source and averaged receivers' locations.

This simple simulation shows that frequency-difference beamforming may be successful for beamforming with a sparse random array in a multipath environment when the signals are coherent at both the broadcast frequency and the difference frequency. To show the performance of frequency-difference beamforming when signals are only coherent at the difference frequency, a Gaussian-distributed random time shift with a standard deviation of 1 ms is added to each of the propagation paths to make the signals incoherent at the recorded frequencies. Figure 5-4 shows the beamforming results integrated over the bandwidth of the signal in this circumstance. It again shows that the conventional beamforming result does not indicate any ray-path arrival angles. However, the frequency-difference beamforming result correctly indicates the arrival angle of the direct path at 45°.

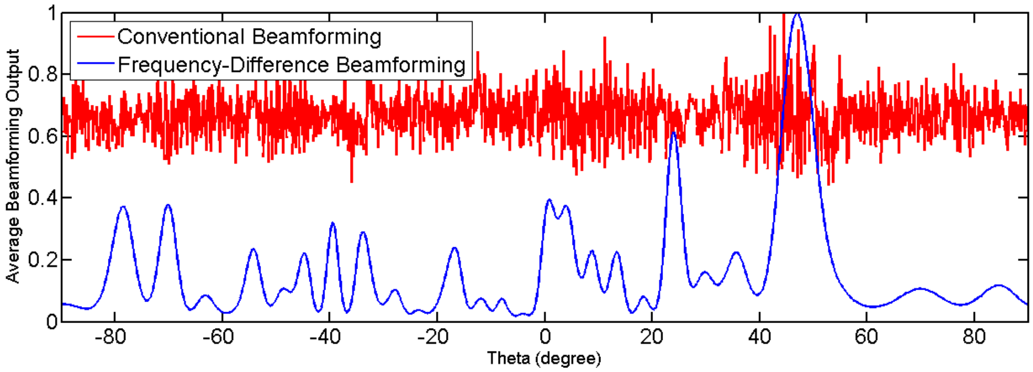


Figure 5-4: Normalized beamforming output integrated over the frequency band where a multiple-millisecond random time shift is added to each propagation paths: conventional plane-wave beamforming (red line) and frequency-difference beamforming with Δf = 12.2 Hz (blue line). The frequency-difference beamforming result displays peaks at 45° that correspond to the direct path between source and averaged receivers' locations.

Thus, a positive assessment of the performance of frequency difference beamforming can be offered for a three-dimensional environment containing fluctuations that degrade temporal coherence in the bandwidth of the signal. Based on this, the frequency-difference concept might be a worthwhile means to similarly increase the robustness of matched-field source localization routines that commonly suffer from model-environment propagation mismatch.

5.2 Frequency-sum beamforming for free-space propagation

Frequency-sum beamforming is an extension of frequency-difference beamforming technique that manufactures higher frequency signal information by from nonlinear product of field amplitudes that sums the frequencies from lower-frequency signal components. It is intended for acoustic environments with one source where a free-space propagation model is expected to be useful but perhaps slightly imperfect. Similar to frequency-difference beamforming, frequency-sum beamforming involves linear and nonlinear spatial filtering of a nonlinear (quadratic or higher) product of complex received-field amplitudes. This section describes frequency-sum beamforming with linear and nonlinear weight functions, and then illustrates it with simulation results for a point-source in free space. These simplest simulations provide alluring but fragile results.

5.2.1 Frequency-sum beamforming formulation for free-space propagation

For the near-field acoustic imaging geometry shown in Figure 5-5, the environment's Green's function is approximately:

$$G(\vec{r}_j, \vec{r}_s, \omega) = \frac{\exp\left\{i\omega\left|\vec{r}_j - \vec{r}_s\right| / \bar{c}\right\}}{4\pi\left|\vec{r}_j - \vec{r}_s\right|} \qquad \text{5-1}$$

where $\vec{r}_s$ is the source location, $\vec{r}_j$ is a receiver location, and $\bar{c}$ is an appropriate average sound speed since inhomogeneities may cause mild variations in sound speed.

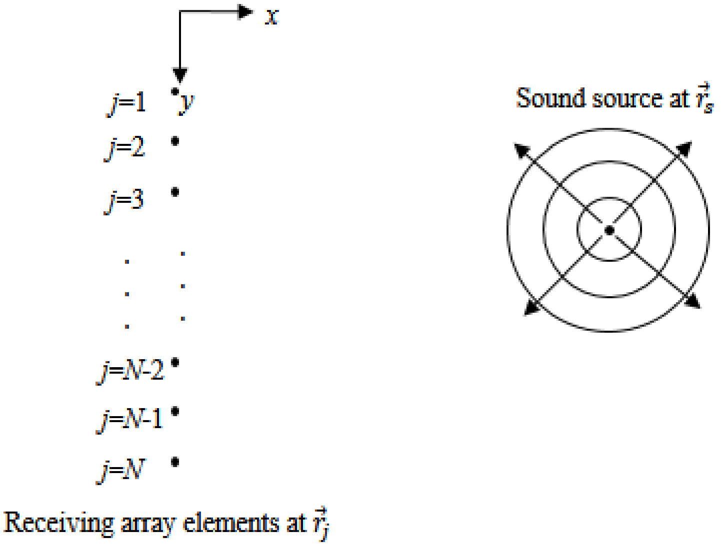


Figure 5-5: The simulated generic geometry. A linear recording array receives signals broadcast by a near-field sound source. The origin of coordinates is at 5 cm above the highest receiver and the array elements lie along the *y*-axis.

It has been shown in chapter 2 that the temporal Fourier transform, $P_j(W)$, of the signal recorded at the j^{th} receiver ($1 \le j \le N$), can be modeled as:

$$P_j(\omega) = S(\omega)G(\vec{r}_j, \vec{r}_s, \omega) = S(\omega)\frac{\exp\{i\omega|\vec{r}_j - \vec{r}_s|/\bar{c}\}}{4\pi|\vec{r}_j - \vec{r}_s|} \tag{5-2}$$

where $S(\omega)$ is the Fourier transform of the broadcast signal and the conventional narrowband near-field delay-and-sum beamforming output $B_1(\vec{r}, \omega)$ at temporal frequency ω is

$$B_1(\vec{r}, \omega) = \frac{1}{\upsilon_1}\left|\sum_{j=1}^{N} P_j(\omega)\exp\left\{-i\omega|\vec{r} - \vec{r}_j|/\bar{c}\right\}\right|^2 \tag{5-3}$$

where $\vec{r}$ is the search location, and υ_1 is a normalization factor that can be chosen in variety of ways.

Frequency-sum beamforming increases the resolution of B_1 from equation 5-3 by manufacturing a higher frequency from a quadratic or higher field product that is used in place of $P_j(\omega)$ in equation 5-3. For example, consider two nearby frequencies, $\omega_1 = \omega_0 + \Delta\omega$ and $\omega_2 = \omega_0 - \Delta\omega$, that lie in the signal's bandwidth, and form the quadratic product:

$$P_j(\omega_1)P_j(\omega_2) = S(\omega_1)S(\omega_2)G(\vec{r}_j, \vec{r}_s, \omega_1)G(\vec{r}_j, \vec{r}_s, \omega_2) = S(\omega_1)S(\omega_2)\frac{\exp\left\{i2\omega_0|\vec{r}_j - \vec{r}_s|/\bar{c}\right\}}{16\pi^2|\vec{r}_j - \vec{r}_s|^2} \tag{5-4}$$

The phase of the final form in equation 5-4 depends on the sum frequency $2\omega_0$. Thus, the quadratic field product $P_j(\omega_1)P_j(\omega_2) = P_j(\omega_0 + \Delta\omega)P_j(\omega_0 - \Delta\omega)$ might be profitably used for delay-and-sum beamforming at this higher frequency,

$$B_2(\vec{r}, \omega_0) = \frac{1}{\upsilon_2}\left|\sum_{j=1}^{N} P_j(\omega_0 + \Delta\omega)P_j(\omega_0 - \Delta\omega)\exp\left\{-i2\omega_0|\vec{r} - \vec{r}_j|/\bar{c}\right\}\right|^2 \tag{5-5}$$

in the hope of obtaining an acoustic image of the source with twice the resolution of B_1 from equation 5-3 evaluated at ω_0. Here, $\Delta\omega$ might be zero or it might be the frequency increment between neighboring complex amplitudes calculated from a fast-Fourier transform (FFT) of the array-recorded signals. In general, $\Delta\omega$ should be chosen to optimize the beamformed output, but such an optimization effort is not considered here.

The quadratic nonlinearity that leads to B_2 is readily extended to higher powers of the recorded field. For example, a fourth-order nonlinear field product can be constructed as follows:

$$B_4(\vec{r},\omega_0) = \frac{1}{\upsilon_4}\left|\sum_{j=1}^{N} P_j(\omega_0+2\Delta\omega)P_j(\omega_0+\Delta\omega)P_j(\omega_0-\Delta\omega)P_j(\omega_0-2\Delta\omega)\exp\left\{-i4\omega_0\left|\vec{r}-\vec{r}_j\right|/\bar{c}\right\}\right|^2 \qquad \text{5-6}$$

This construction may have four times the resolution of B_1 from equation 5-3 evaluated at ω_0 since the sum frequency manufactured by the nonlinear product is $4\omega_0$. Again, $\Delta\omega$ can be chosen to be zero or set to another appropriate value. The subscripts 1, 2, and 4 in equation 5-3, 5-5, and 5-6, respectively, denote the number of the complex field amplitudes used in the beamforming.

Similarly, this frequency-sum concept can also be applied to nonlinear beamforming schemes such as the Minimum Variance Distortionless Response (MVDR) (see Jensen et. al., 1994). Conventional narrowband MVDR beamforming has been described in section 2-2. Here the possibility of using the quadratic and quartic field products for MVDR beamforming at the higher frequencies $2\omega_0$ and $4\omega_0$ is investigated too.

In the next section, results from linear and nonlinear spatial filtering of a nonlinear product of complex received-field amplitudes are compared with usual delay-and-sum and MVDR beamforming outputs using simulated signals. For all the results in the following section, ω_0 is the signal's center frequency, $\Delta\omega$ is the frequency difference between FFT samples, and υ_i (i = 1, 2, or 4) is the maximum value of the beamformed output. This choice for the normalization factors allows beamforming outputs to be plotted in decibels with 0 dB being the maximum beamformed value.

5.2.2 Results from Simple Propagation Simulations

To investigate the best possible performance of frequency-sum beamforming, simple simulations of a point sound source in free space were undertaken. The coordinate system used is shown in Figure 5-1, and the array and source lie in the plane defined by $z = 0$. The sound source was located at $\vec{r}_s = (x_S, y_S)$ = (30 cm, 10 cm). The linear array was composed of 16 elements spaced 3.81 cm apart along the x-axis. For this simple geometry and free-space environment, equation 5-3, 5-5, and 5-6 are equivalent to spherical-wave beamforming. The broadcast signals were 100 micro-second-duration, Gaussian-shaded, sinusoidal pulses having center frequencies of 30 kHz, 60 kHz, and 120 kHz. In water with a nominal sound speed of 1500 m/s, the corresponding center-frequency wavelengths are 4.9 cm, 2.5 cm, and 1.2 cm, respectively. Thus, the fixed-geometry array becomes sparse as the center frequency increases.

Simulation results for B_1, B_2, and B_4 are shown in Figure 5-6 and Figure 5-7. Conventional beamforming results for B_1 from equation 5-3 at the 30 kHz, 60 kHz, and 120 kHz are shown in the three panels of Figure 5-6. A white cross marks the source location in each panel. As expected, the beamformed output reaches a maximum at the source location, and the resolution improves (the image spot size shrinks) with increasing frequency.

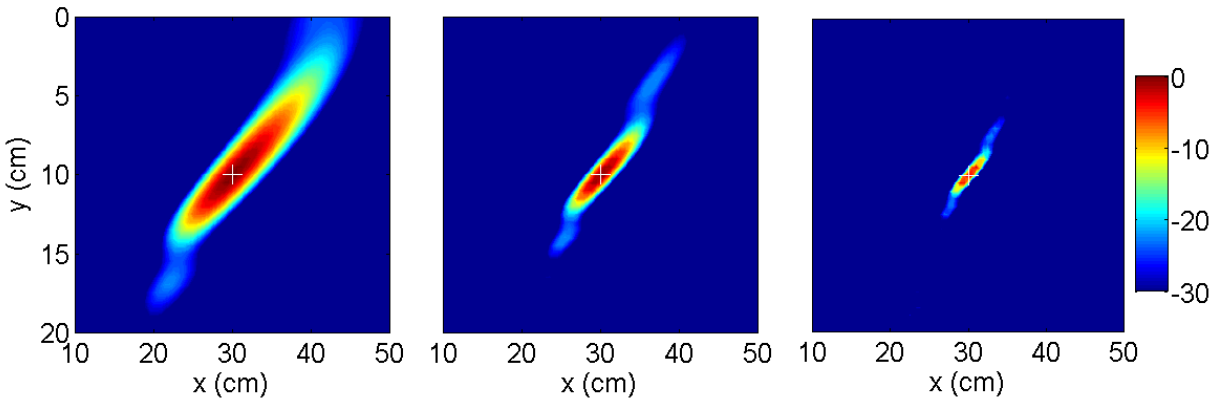


Figure 5-6: Simulated beamforming output for B_1 from equation 5-3 at 30 kHz (a), 60 kHz (b), and 120 kHz (c). The colors scale is in decibels. The actual source location is marked with a white cross. As expected, the spot size at the source location decreases with increasing frequency.

Figure 5-7 shows a comparison of frequency-sum and conventional beamforming results when the sum frequency and the signal center frequency are 120 kHz. Figure 5-7(a) is B_4 from (5-6) using the 30 kHz signal. Figure 5-7(b) is B_2 from (5-5) using the 60 kHz signal. And, Figure 5-7(c) is B_1 from equation 5-3 using the 120 kHz signal [it is the same as Figure 5-6(c)]. The three panels of Figure 5-7 are nearly identical and this clearly shows that frequency-sum beamforming can be used to improve the resolution of acoustic source images under (very!) ideal conditions. Here, frequency-sum beamforming using ideal 30 kHz signals has been used to generate localization (imaging) results having the same resolution as conventional beamforming of ideal 120 kHz signals.

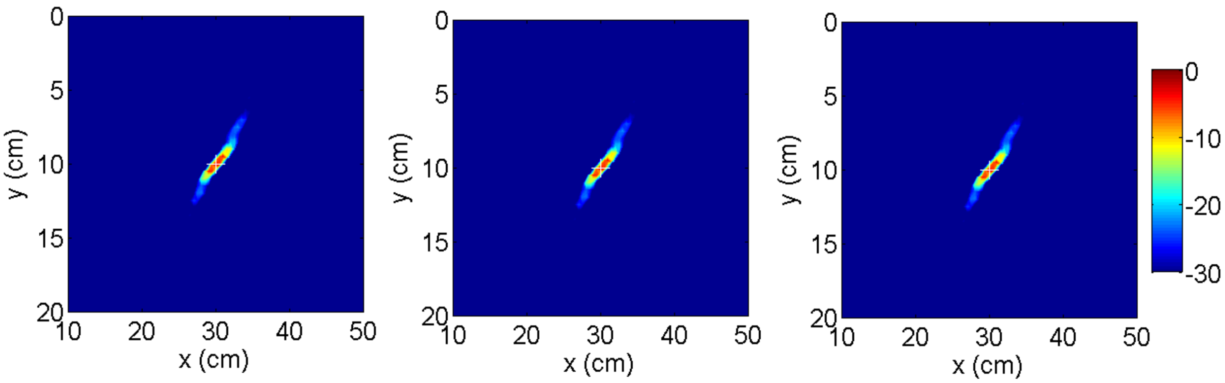


Figure 5-7: Same as Figure 5-6 except (a) shows B_4 from (5-6) using the 30 kHz signal, and (b) shows B_2 from (5-5) using the 60 kHz signal. The three panels are essentially identical and this shows that frequency-sum beamforming can be used to improve the resolution of acoustic source images under ideal conditions.

Figure 5-8 shows MVDR beamforming results from equation 2-18 at the 30 kHz, 60 kHz, and 120 kHz. A white cross marks the source location in each panel. As expected, the beamformed output reaches a maximum at the source location, and the resolution improves (the image spot size shrinks) with increasing frequency. Here the spot sizes are significantly smaller than those of the delay-and-sum beamforming results in Figure 5-6. Such spot-size shrinkage is an advantage of nonlinear spatial filtering of the recorded signals at high SNR. Since the image spots are very small in Figure 5-8 and Figure 5-9, they have been zoomed in.

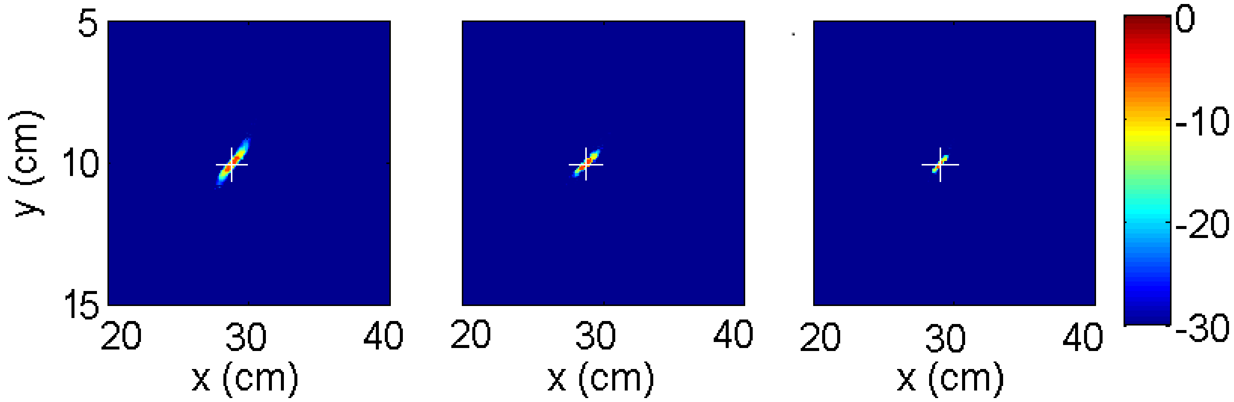


Figure 5-8: Simulated beamforming output for minimum variance distortionless beamforming at 30 kHz (a), 60 kHz (b), and 120 kHz (c). The colors scale is in decibels. The actual source location is marked with a white cross. As expected, the spot size at the source location decreases.

Figure 5-9 shows a comparison of frequency-sum and conventional beamforming results from the MVDR technique when the sum frequency and the signal center frequency are 120 kHz. Figure 5-9(a) is for the fourth-order nonlinear field product in equation 2-18 using the 30 kHz signal. Figure 5-9(b) is for the second-order nonlinear field product in equation 2-18 using the 60 kHz signal. And, Figure 5-9(c) is the usual MVDR output from equation 2-18 using the 120 kHz signal [it is the same as Figure 5-8(c)]. The three panels of Figure 5-9 are nearly identical and this shows that MVDR beamforming can be used to improve the resolution of acoustic source images under ideal conditions. Here, frequency-sum beamforming using ideal 30 kHz signals has been used to generate localization (imaging) results having the same resolution as conventional MVDR beamforming of ideal 120 kHz signals.

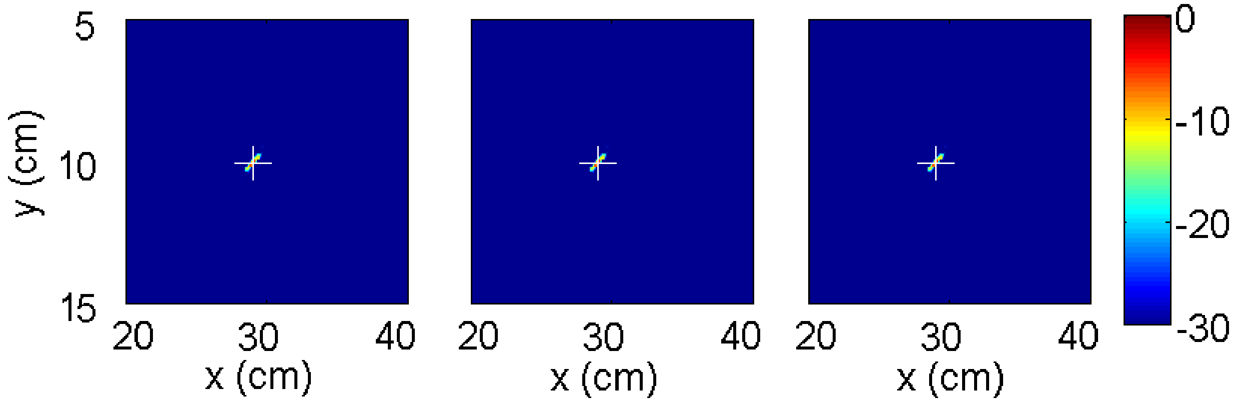


Figure 5-9: Same as Figure 5-8 except (a) shows the fourth-order nonlinear field product in equation 2-18 using the 30 kHz signal, and (b) shows the second-order nonlinear field product in equation 2-18 using the 60 kHz signal. The three panels are essentially identical and this shows that frequency-sum beamforming can be used to improve the resolution of acoustic source images under ideal conditions.

Results in this section show that frequency-sum beamforming with a linear and nonlinear weight function can be used to improve the resolution of beamforming output for a single point-source in free space. Here it must be stated clearly that these results provide an overly optimistic forecast of the performance frequency-sum beamforming in more complicated environments or with more than a single source. The next section presents the performance of frequency-sum beamforming in a more complicated environment.

5.3 Frequency-sum beamforming for multipath environments

Frequency-sum beamforming works well in an ideal circumstances: a single point-source, no noise, and free-space propagation. In this section, the performance of frequency-sum beamforming in a simple multipath environment is presented. For this purpose, a surface reflection has been added to the environment used in previous section (shown in Figure 5-10). To see the beamforming peak of the image source, the *y*-axis for figures in this section has been changed to -20 to 20 cm.

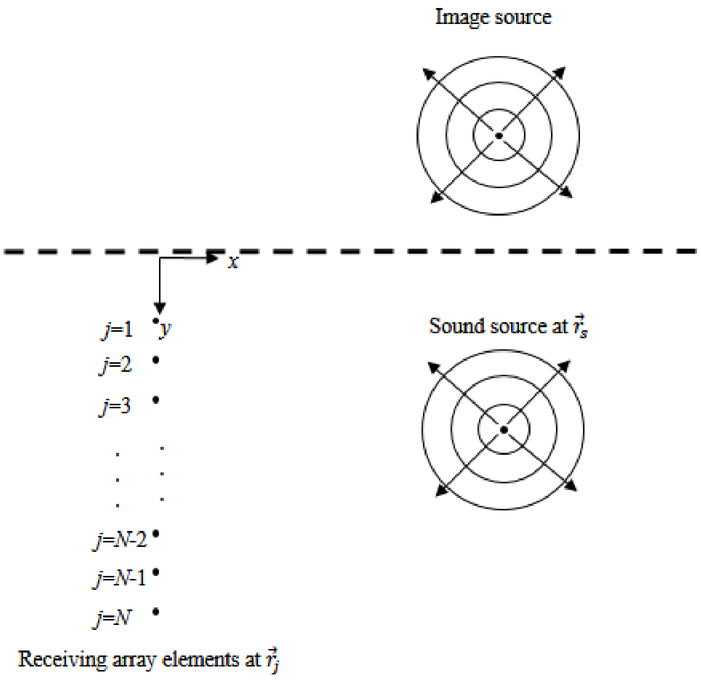


Figure 5-10: This is the generic geometry. An images source has been added to Figure 5-5. A linear recording array receives signals broadcast by a near-field sound source and an image source. The origin of coordinates coincides with the surface and the array elements lie along the *y*-axis.

Simulation results for B_1, B_2, and B_4 are shown in Figure 5-11 and Figure 5-12. Conventional beamforming results for B_1 from equation 5-3 at the 30 kHz, 60 kHz, and 120 kHz are shown in the three panels of Figure 5-11. A white cross marks the source location in each panel. As expected, the beamformed output reaches a maximum at the source location, and the resolution improves (the image spot size shrinks) with increasing frequency. For Figure 5-11 through Figure 5-14, a white plus and white cross marks the actual source and image source location, respectively.

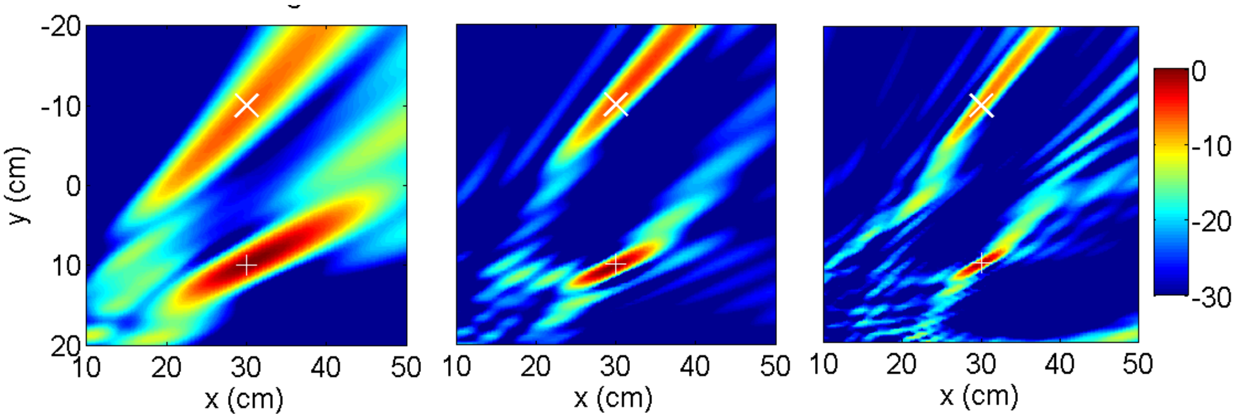


Figure 5-11: Simulated beamforming output for B_1 from equation 5-3 at 30 kHz (a), 60 kHz (b), and 120 kHz (c). The colors scale is in decibels. The actual source location is marked with a white plus and the image source is marked with a white cross. As expected, the spot size at the source location decreases with increasing frequency.

Figure 5-12 shows a comparison of frequency-sum and conventional beamforming results when the sum frequency and the signal center frequency are 120 kHz. Figure 5-12a is B_4 from (5-6) using the 30 kHz signal. Figure 5-12b is B_2 from equation 5-5 using the 60 kHz signal. And, Figure 5-12c is B_1 from equation 5-3 using the 120 kHz signal [it is the same as Figure 5-11c]. The three panels of Figure 5-12 were expected to be nearly identical (similar to Figure 5-7). However, the frequency-sum beamforming output using ideal 30 kHz and 60 kHz signals has become confusing and maxima found at the wrong locations. The peaks that appear at the wrong locations are caused by the cross terms between the actual and image sources that arise in the nonlinear product that is beamformed to create B_2 and B_4. Unlike the frequency difference case, these cross terms occur at the same frequency as the intended diagonal terms and cannot be averaged-away by integrating through the signal bandwidth.

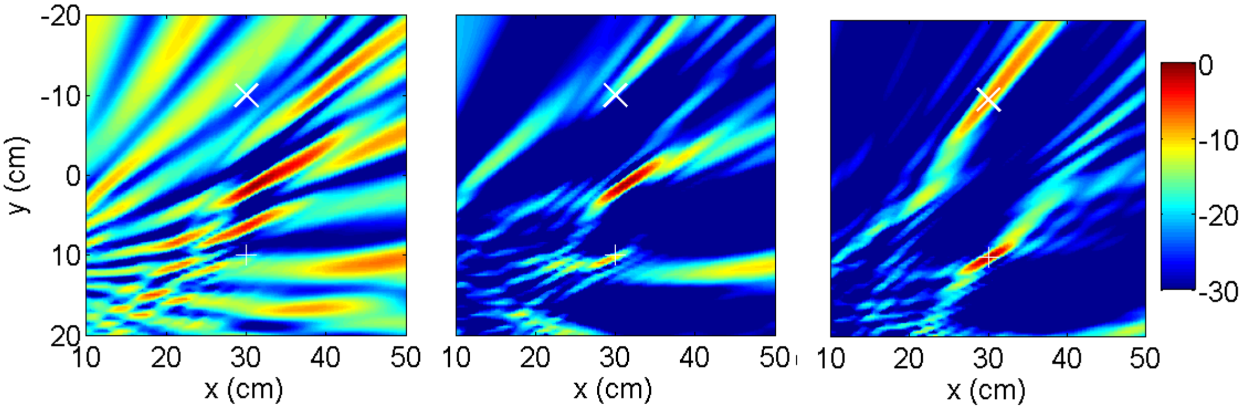


Figure 5-12: Same as Figure 5-11 except (a) shows B_4 from equation 5-6 using the 30 kHz signal, and (b) shows B_2 from equation 5-5 using the 60 kHz signal. The three panels should essentially be identical. But, this shows that frequency-sum beamforming does not work for acoustic source imaging in multipath environments.

Figure 5-13 shows minimum variance distortionless beamforming results from equation 2-18 at the 30 kHz, 60 kHz, and 120 kHz. As expected, the beamformed output reaches a maximum at the source location, and the resolution improves with increasing frequency. The resolutions are slightly higher than the delay-and-sum beamforming results in Figure 5-11 (the image spot sizes are slightly smaller) which (again) illustrates the advantage of nonlinear spatial filtering of recorded signals at high SNR.

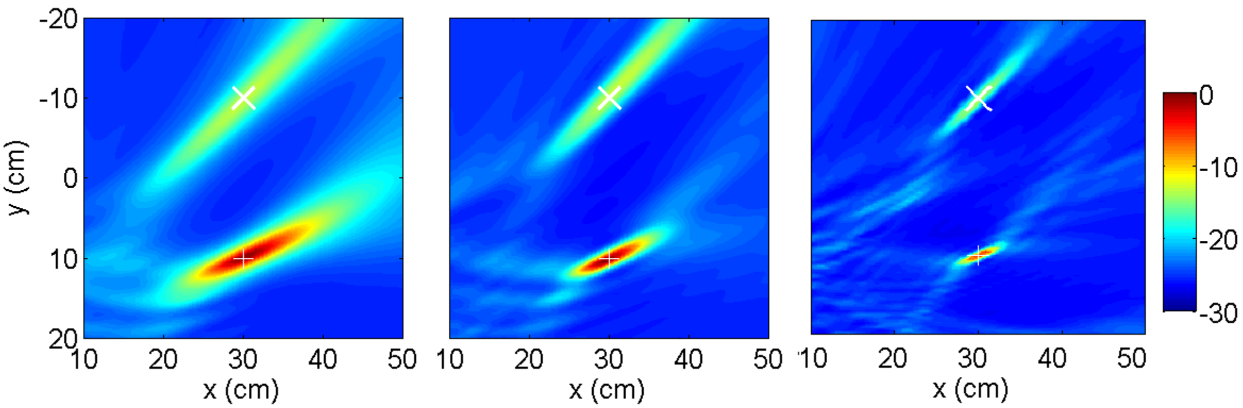


Figure 5-13: Simulated beamforming output for minimum variance distortionless beamforming at 30 kHz (a), 60 kHz (b), and 120 kHz (c). The colors scale is in decibels. The actual source location is marked with a white plus and the image source is marked with a white cross. As expected, the spot size at the source location decreases with increasing frequency.

Figure 5-14 shows a comparison of frequency-sum and conventional beamforming results from minimum variance distortionless beamforming technique when the sum frequency and the signal center frequency are 120 kHz. Figure 5-14a is the fourth-order nonlinear field product in equation 2-18 using the 30 kHz signal. Figure 5-14b is the second-order nonlinear field product in equation 2-18 using the 60 kHz signal. And, Figure 5-14c is the usual MVDR output from equation 2-18 using the 120 kHz signal [it is the same as Figure 5-13c]. The three panels of Figure 5-14 were expected to be identical. However, frequency-sum beamforming using ideal 30 kHz and 60 kHz signals are again confusing and maxima occur the wrong locations.

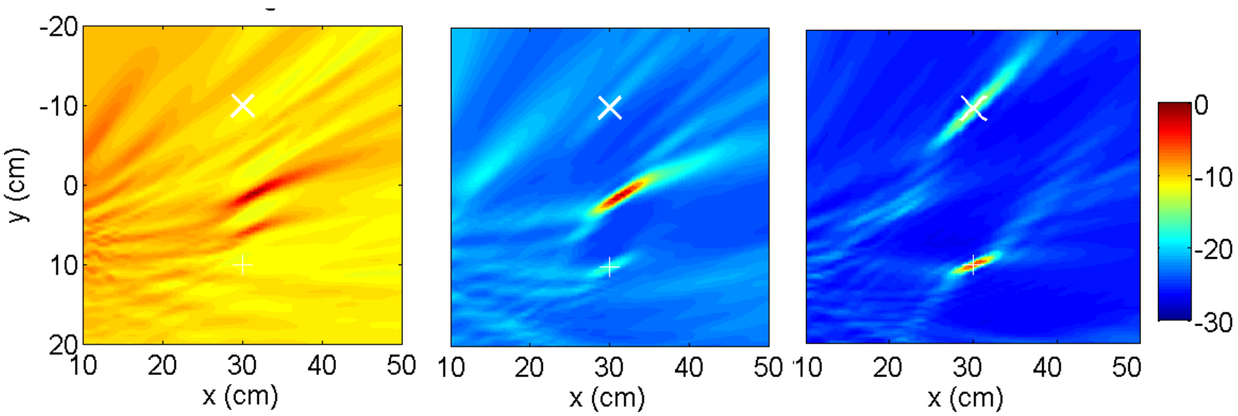


Figure 5-14: Same as Figure 5-13 except (a) shows the fourth-order nonlinear field product in equation 2-18 using the 30 kHz signal, and (b) shows the second-order nonlinear field product in equation 2-18 using the 60 kHz signal. The three panels should be identical. However, this shows that frequency-sum beamforming does not work for acoustic source imaging in multipath environments.

Thus, a positive assessment of the performance of frequency sum beamforming can only be offered for the most ideal environmental circumstance (one source, no noise, free space), a circumstance where little improvement of existing beamforming techniques is needed. The addition of a single reflecting surface, or a second (image) source causes frequency sum beamforming to produce nonsense. The results in this section illustrate that frequency-sum beamforming is not suitable for beamforming underwater acoustic signals that propagate through a multipath sound channel. Therefore, frequency-sum beamforming should be abandoned unless it can be modified to mitigate or eliminate the deleterious effects of the unwanted cross terms arising in the nonlinear field product. It will not be considered further in this thesis.

CHAPTER 6

Remote ranging of bowhead whale calls in a dispersive underwater sound channel

This chapter describes how mode filtering (MF) or the blind-deconvolution technique synthetic time reversal (STR) can be used to determine the range of bowhead whale calls from a single linear vertical array in a dispersive underwater sound channel. Here the whale-call bandwidth is nominally 50 Hz to 500 Hz, and some environmental knowledge at the array location is needed for the ranging calculations. This study's primary findings are based on simulations and passive listening experiments conducted in the coastal waters near Kaktovik, Alaska. A 12-element vertical array nominally spanning the middle 60% of the water column was deployed in 55-m-deep water alongside a distributed array of Directional Autonomous Seafloor Acoustics Recorders (DASAR) arranged in triangular grids. A total of 18 whale calls were considered, and all of them could be used for ranging and comparisons. The estimated call-location-to-array ranges determined from mode filtering and STR are compared with triangulation results from the DASAR. The vertical-array ranging results are generally within ±10% of the DASAR results with the STR results being slightly more accurate than those from mode filtering. The whale call data set and the DASAR measurements were provided to the author of this thesis by Dr. Aaron Thode of the Scripps Institution of Oceanography.

6.1 Introduction

Remote monitoring of marine mammals is important for biological studies and for assessment of anthropogenic activities in the marine environment. For example, the presence or absence of marine mammals may be a critical factor for undersea activities that involve sound generation. Here, reliable means for remote ranging of marine mammals from a relatively-easily-deployed vertical array located near the sound generation site would be advantageous for the safe conduct of active acoustic activities.

Even at high signal-to-noise ratios, passive undersea ranging of remote unknown sound sources may be challenging because the source signal is unknown, the acoustic environment supports multipath propagation, and the requisite environmental information for matched field methods is unavailable (Bucker 1976, Zurk *et al*. 2003, Collison & Dosso 2000, Yoo & Yang 1999, Shang *et al*. 1985, Lin *et al*. 2012). The two ranging techniques considered are based on mode filtering (Krolik 1992, Porter & Reiss 1985, Lo *et al*. 1983, Yang 1989, Buck *et al*. 1998) and synthetic time reversal (Sabra & Dowling 2004, Sabra *et al*. 2010, Abadi *et al*. 2012).

These techniques share many similarities. Both techniques involve maximizing the cross correlation of mode-specific signals (whale calls) to determine source-array range. Both are suitable for use in incompletely-known underwater sound channels with modal dispersion in the bandwidth of interest. Both require a modest amount of environmental information at the receiving array. And, both perform best in range independent environments. However, the two techniques also have one important difference. Mode filtering relies on mode-shape weighting of the array measurements to determine signal amplitude and phase in the signal bandwidth, while the STR-based technique relies on mode-shape weighting to determine only signal phase in the signal bandwidth. This difference is of no consequence in simulations of whale-call ranging, but it confers a slight accuracy advantage to the STR-based technique for ocean-recorded whale calls.

This chapter describes and compares these two techniques for remote ranging of bowhead whale calls recorded with a vertical transducer array. The performance of the two techniques is evaluated using natural calls recorded in coastal waters north of Alaska. The ranging performance of the two techniques is benchmarked against results obtained from separate widely-distributed bottom mounted direction sensors (Directional Autonomous Seafloor Acoustics Recorders; DASAR) which determine mammal call locations in two horizontal dimensions via triangulation.

The DASAR record three channels continuously at a sampling rate of 1 kHz, using a suspended sensor protected by a latex "sock" covering an aluminum frame. One channel samples data from an omni-directional (pressure) hydrophone with a sensitivity of -149 dB re 1 V/μPa @ 100 Hz, with a high-pass frequency response of around 20 dB per decade, in order to pre-whiten the expected ambient noise spectrum. The other two channels measure particle velocity in orthogonal directions, with a sensitivity of 97 dB re 1 V/(m/s) @ 100 Hz (with translates into -146 dB re 1 V/μPa @ 100 Hz sensitivity for an acoustic plane wave). After a net 0 dB amplification the data are converted to binary 16-bit samples, stored to flash memory, and then periodically dumped to a laptop hard drive. The electronic noise floor of the omnidirectional sensor is equivalent to a background noise level of 43 dB re 1 μPa^2/Hz @ 100 Hz, with a 6 dB/octave spectral slope. These values generally lie 15 dB below the Knudsen noise spectrum estimated for sea state 0. Specific details on the internal electronics are given in (Greene *et al.*, 2004), which describes an earlier version of the DASAR that uses a different sensor.

A variety of prior efforts to range and locate marine mammal calls based on horizontal and bottom-mounted transducer arrays have been reported. A passive acoustic technique with towed linear arrays has been used for marine mammals' localization (Thode *et al.*, 2000). A portable matched field processing system for tracking marine mammals has also been reported (Thode *et al.* 2006). An automated procedure has been developed for localizing bowhead whale calls in the presence of seismic airgun surveys (Thode *et al.*, 2012). Ocean bottom seismometer networks have been used for tracking fin and blue whales in the northeast Pacific Ocean (Wilcock 2012). Passive acoustic localization based on arrival-time differences of transient bowhead sounds detected on a sparse array has been presented by Clark and Ellison (2000). Migratory and mating behavior of western Arctic bowhead whales have been discussed by Delarue *et al*. (2009). Also, time differences of arrival and feeding them into 2D and 3D hyperbolic localization algorithms is another method for passive acoustic localization (Morrissey *et al.*, 2006).

The present research effort extends these prior efforts to dispersive shallow ocean waveguides, imperfect (i.e. not water column spanning) vertical arrays, and whale call ranging at distances exceeding 20 km from a single array. The purpose of this study is to document how mode filtering (MF) and mode-based STR can be used for whale call ranging with a vertical array in a dispersive shallow-ocean waveguide. The results are based on simulations and experiments involving 50 to 500 Hz bowhead whale calls propagating in a nominally 55-m-deep sound channel.

The remainder of this chapter is divided into four sections. The next section presents the mathematical formulations of MF, mode-based STR, and DASAR ranging techniques in dispersive media. Section 6.3 presents MF and STR results from simulated acoustic propagation based on *KRAKEN* (*KRAKEN* is a normal-mode model for predicting acoustic pressure fields in ocean environments; see Porter 1984) in a simple environment that mimics the Arctic Ocean environment of the companion experiments. Section 6.4 presents MF and STR results from bowhead whale calls recorded in 2010 using a 12-element vertical array deployed in the central 60% of a 55m-deep water column. Section 6.5 summarizes this research effort, and states the conclusions drawn from it.

6.2 Three whale call ranging techniques

The mathematical formulation of three whale call ranging techniques are provided in this section. The first technique is based on mode filtering which has been used in low frequency underwater sounds propagation for many years. The second technique is an extension of mode-based synthetic time reversal (STR). It has been shown in chapter 3 that STR can be used for sound source localization when the medium is not dispersive and basic environmental information at the receiving array is available (Abadi *et al*. 2012). Here, an extension of mode-based STR will be used for ranging whales in a dispersive medium. The last technique, which is used as a reference, is the localization technique based on triangulation of the signals recorded by DASAR (Greene *et al.*, 2004 and Thode *et al*., 2012).

6.2.1 Mode Filtering (MF)

Mode filtering technique uses the measured waveforms $p_j(t)$ from the N elements of a vertical receiving array to reconstruct the source signal $\hat{S}_m(\omega)$ in the frequency domain using array-element weights determined from the vertical shape of the m^{th} propagating acoustic mode:

$$\hat{S}_m(\omega) = \sum_{j=1}^{N} W_{j,m} P_j(\omega), \qquad \text{6-1}$$

where the simplest choice of weight function $W_{j,m}$ is the m^{th} mode shape:

$$W_{j,m} = \Psi_m(z_j)\,. \tag{6-2}$$

In this case, the array-element weights, $W_{j,m}$, are able to isolate the propagation phase of a single mode as in Sabra & Dowling (2004). More sophisticated weighting schemes are possible for the same task (Buck *et. al.*, 1998). However, for the current investigation, results from such schemes were largely the same as those from the above mode-filtering approach, so equation 6-1 and 6-2 were used in this investigation for simplicity.

Equation 6-1 is a reconstruction of the Fourier transform of the source signal and can be written as:

$$\hat{S}_m(\omega) = \left|\hat{S}_m(\omega)\right| e^{\hat{\varphi}_s(\omega) + i\Gamma_m(\omega)}, \tag{6-3}$$

where $\Gamma_m(\omega)$ is a residual phase from source-to-array propagation and it can be found from equation 2-6:

$$\Gamma_m(\omega) = \arg\{\sum_{j=1}^{N} W_{j,m} G_m(\vec{r}_j, \vec{r}_s, \omega)\} \approx \arg\{\frac{\Psi_m(z_s)}{\rho(z_s)\sqrt{8\pi k_m R}} e^{ik_m R}\} = \mathrm{Re}\{k_m\}R \tag{6-4}$$

where,

$$k_m = \frac{\omega}{c_{p,m}}, \tag{6-5}$$

$c_{p,m}$ is the phase speed of m^{th} mode, and *R* is the range from the remote source to the array. Here, if the cutoff frequency of the m^{th} mode is well below the lowest signal frequency (Figure 6-2 shows phase speed vs. frequency for the environment shown in Figure 6-1), the phase speed of the m^{th} mode can be assumed independent of frequency and the extra phase of STR will be a linear function of frequency for a fixed source to receiver geometry. This linear-in-frequency phase is equivalent to a time shift in the time domain. However, in the ocean region of interest, the frequency range of most bowhead whale calls extend down to near the cutoff frequencies of the low-order propagating acoustic modes. Thus, the phase speed of the m^{th} mode will be a function of frequency, and this reduces the performance of mode filtering for reconstructing the source signal.

In this study, an extension of mode filtering and mode-based STR has been developed to localize whale calls when the modal propagation is dispersive. Here, the nonlinear dependence of the correction phase can be removed by mode-filtering using two different modes, and then searching for a frequency-dependent phase correction using wave numbers determined from a modal-sum propagation model evaluated at the receiving array location. To see how this works, start from equation 6-3 and 6-4 for two different modes ($m \neq n$) to construct the following relationship.

$$\arg\{\hat{S}_m(\omega)\hat{S}_n^*(\omega)\} = \{\hat{\varphi}_m(\omega) - \hat{\varphi}_n(\omega)\} + i\{k_m(\omega) - k_n(\omega)\}R \qquad \text{6-6}$$

If the medium is not dispersive and both modes have been excited by the same sound source, $\hat{S}_m(\omega)$ and $\hat{S}_n(\omega)$ should be identical. However, $k_m(\omega) - k_n(\omega)$ is a function of frequency in dispersive environments and this dispersion decreases the cross correlation coefficient between $\hat{S}_m(\omega)$ and $\hat{S}_n(\omega)$. If $\left\{k_m(\omega) - k_n(\omega)\right\}R$, the dispersive portion of the phase, can be estimated and removed from equation 6-6, then the reconstructed source signals should have maximum correlation.

A means for estimating the dispersive phase proceeds from this contention. The wave numbers of each mode, when close to that mode's cut-off frequency, is proportional to the inverse of frequency (Grachev, 1993). So, the dispersive phase can be adequately fit by assuming it follows:

$$\{k_m(\omega) - k_n(\omega)\}R \approx \frac{2\pi C}{\omega}. \qquad \text{6-7}$$

Here, *C* is a constant that maximizes the cross-correlation between $\hat{S}_m(\omega)$ and $\hat{S}_n(\omega)$ including the phase adjustment given by equation 6-7:

$$\hat{S}_m(\omega)\hat{S}_n^*(\omega)e^{-i\frac{2\pi C}{\omega}} = \left|\hat{S}_m(\omega)\right|\left|\hat{S}_n(\omega)\right|e^{\{\hat{\phi}_m(\omega)-\hat{\phi}_n(\omega)\}+i\{k_m(\omega)-k_n(\omega)\}R-\frac{2\pi C}{\omega}}. \qquad \text{6-8}$$

This procedure leads to a frequency-dependent estimate of $\{k_m(\omega) - k_n(\omega)\}R$. The range of the calling whale may then be found by minimizing the difference or error between $\frac{2\pi C}{\omega}$ and the difference between the two wave numbers multiplied by an estimated range, $\{k_m(\omega) - k_n(\omega)\}R_e$, calculated from a propagation model for the array-location environment. If the model is accurate enough, the value of mean-square error in equation 6-9 should be minimum when R_e is at the actual range.

$$Error(R_e) = \sum_{f=f_1}^{f_2}[\{k_m(\omega) - k_n(\omega)\}R_e - \frac{2\pi C}{\omega}]^2 \qquad \text{6-9}$$

The ranging results from this process are shown in sections 6.3 and 6.4 for simulations and ocean recordings, respectively.

6.2.2 Mode-based STR

STR is a simple array-signal-processing technique for simultaneously estimating the waveforms of the unknown source signal and the unknown source-to-array transfer functions in an unknown multipath environment. The mathematical formulation of mode-based STR is presented in section 2. Thus, for the present purposes, only final formulas are provided. Mode-based STR uses the same dispersive-phase fitting and correction technique as mode filtering. However, it generates signal amplitude information differently. The reconstructed source signal $\hat{S}_m(\omega)$ can be found from equation 2-7. After the source signal is reconstructed from two different modes with STR, the remaining steps for ranging are exactly the same as in the previous section. Thus, the only change from MF-based to STR-based whale call ranging is the estimated Fourier transform of the signal given by equation 6-1 is replaced by equation 2-7. The remainder of the MF ranging technique described in the previous section stays the same. The ranging results presented in sections 6.3 and 6.4, show that the two techniques perform identically in simulation but STR works slightly better when results from ocean-whale call recordings are compared to the established triangulation technique.

6.2.3 DASAR

For a given DASAR, the bearing of a transient acoustic signal consisting of *L* samples was estimated by taking the arctangent of the ratio of the mean acoustic intensities obtained from the two directional channels (Greene *et al.*, 2004):

$$\phi_{true} = \phi_{cal} + \tan^{-1}[I_y / I_x] \tag{6-10}$$

where the arctangent is defined over four quadrants, and ϕ_{cal} is a "calibration" bearing that relates the orientation of the DASAR directional axes to true north. Here mean acoustic intensity, I_k, is defined as:

$$I_k = \frac{1}{L}\sum_{l=1}^{L} p(t_l) v_k(t_l) \tag{6-11}$$

with $p(t_l)$ being the time series from the pass-band-filtered omnidirectional channel and $v_k(t_l)$ being the filtered time series from one of the directional channels.

To determine the value of ϕ_{cal} for each DASAR, a series of calibrated transmissions were made from three positions around each DASAR deployment. The projected waveform consisted of a 2-s tone at 400 Hz, a 2-s linear sweep from 400 to 200 Hz, a 2-s linear sweep from 200 to 400 Hz, a 2-s linear sweep from 400 to 200 Hz, and finally a 4-s long section of pseudo-random noise, i.e., an m-sequence with 255 chips, repeated once every second and on a 255 Hz carrier frequency. A calibration is conducted just after a deployment and just before a pickup, to confirm that the instrument did not rotate during the deployment.

The bearings from multiple DASARs are used to estimate a robust maximum-likelihood position of the animal, along with the 90% confidence ellipse (Lenth, 1981a; Greene *et al.*, 2004). A Huber weighting function (Huber, 1964) was used to suppress directional outliers, using a tuning parameter of 1.5 and incorporating 100 bootstrapped estimates of the concentration parameter κ (Lenth, 1981b). The final output is a location, bounded by a 90% confidence ellipse. Further details on how the bearings from individual DASAR are combined to generate a position and position uncertainty ellipse are provided in (Greene *et al.*, 2004).

6.3 Mode filtering and STR ranging results from simulation

To predict the performance of MF and STR in ranging bowhead whale calls, simple simulations were undertaken that approximately mimic the signals and geometry of the Arctic environment where actual whale calls were recorded. Two 6-element vertical arrays were located under each other to make a 12-element vertical array which was deployed in 55 m deep water to record signals with 6250 Hz sampling rate. The deepest element depth was at 50.9 m, and the shallowest element depth was at 16.5 m (Figure 6-1). The element spacing was generally 3 m, except for a portion of the array where the two sub array cables overlapped, yielding 6 phones with 1.5 m spacing.

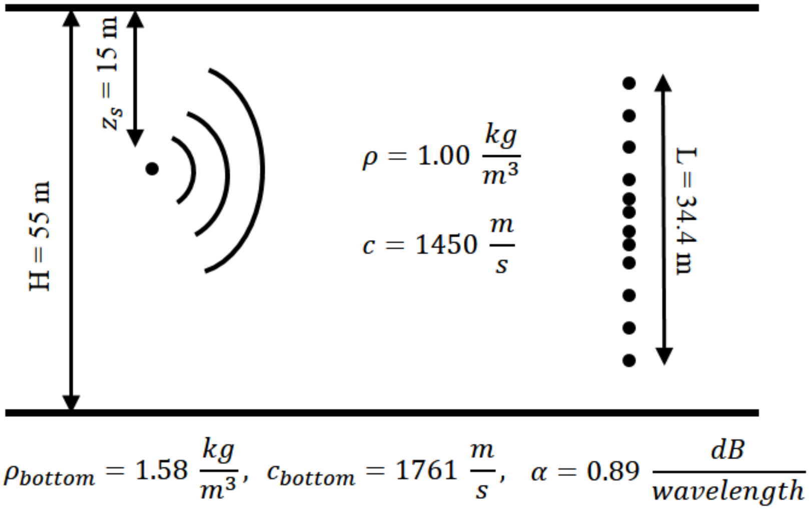


Figure 6-1: Simulation and experiment array geometry.

For the simulations, a chirp signal with 100-300 Hz bandwidth was broadcast 4 km from the recording array from a depth of 15 m. Figure 6-2 shows phase speed for the first four modes calculated from simple two layer Pekeris waveguide in the environment shown in Figure 6-1. It shows that the phase speed of all excited modes in this simulated bandwidth is a function of frequency, i. e., all the modal propagation are dispersive.

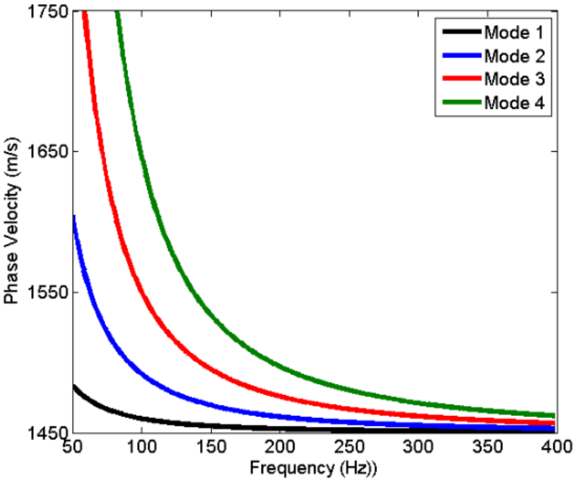


Figure 6-2: Phase velocity of the first four modes vs. frequency

Figure 6-3 shows the simulated array recorded signals, $p_j(t)$. Here the signal from shallowest receiver appears as the upper-most pressure time trace and the signal from the deepest receiver appears as the lowest pressure time trace.

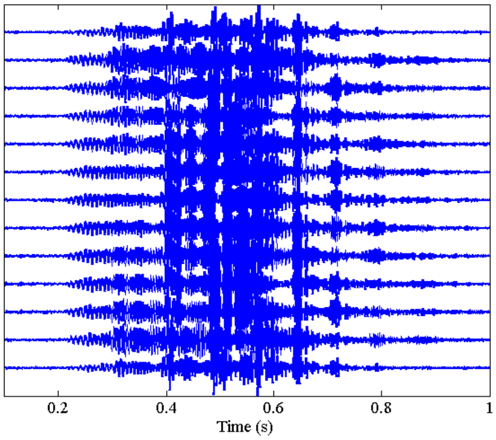


Figure 6-3: Simulated array-recorded signals. The upper signal is the signal recorded from shallowest receiver and the lowest pressure time shows the signal recorded from the deepest receiver.

From these simulated signal recordings, MF and STR can be used to reconstruct the source signal and search for the appropriate dispersion-phase fitting constant, C in equation 6-7, to maximize cross-correlation between $\hat{S}_m(\omega)$ and $\hat{S}_n(\omega)e^{i2\pi C/\omega}$. Figure 6-4 shows the cross-correlation coefficient between $\hat{S}_m(\omega)$ and $\hat{S}_n(\omega)e^{i2\pi C/\omega}$ vs. C for MF and STR, and C has units of radians per second.

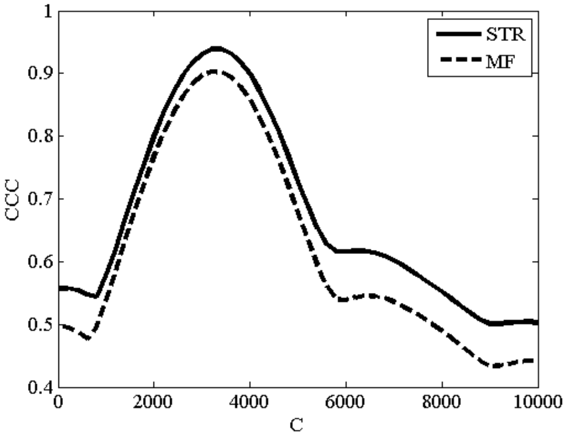


Figure 6-4: Cross-correlation coefficient between $\hat{S}_m(\omega)$ and $\hat{S}_n(\omega)e^{i2\pi C/\omega}$ vs. C. STR and MF results have been shown by solid line and dashed line respectively.

Figure 6-5 is the mean-square error calculated from equation 6-9 for different estimated range. This figure shows that the mean-square error for both MF and STR has minimum at 4 km which is the actual range of simulation.

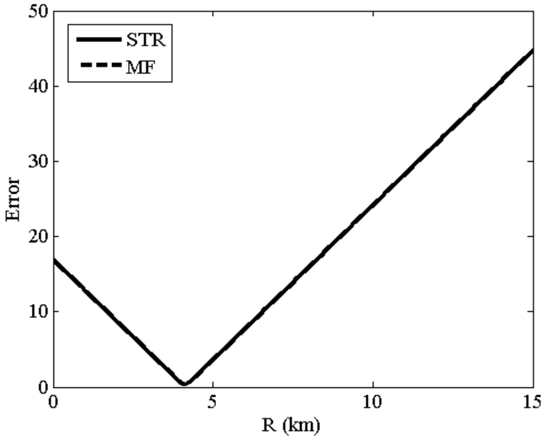


Figure 6-5: Mean-square error calculated form equation 6-9 vs. range. STR and MF results have been shown by solid line and dashed line respectively. However, these lines are coincident in this case.

Different pairs of modes (like mode 2 and 3) can be successfully used instead of mode 1 and 2 if the mode pair are orthogonal across the array aperture and if both modes are excited by the sound source. If either of the modes is not excited by sound source, MF and STR is not able to isolate that mode from other propagating modes. Figure 6-6 shows the shapes of modes 1, 2, and 3 as a function of depth. For example, if the sound source is at a depth of 29 m, it does not excite mode 2, and if it is at a depth of 37 m, it does not excite mode 3. Figure 6-7 and Figure 6-8 show the ranging results when source is located 10 km from the array at 37 m depth. Figure 6-8 shows that the mode 2 and 3 pair is not able to range the sound source. However, the mode 1 and 2 pair ranges the source correctly.

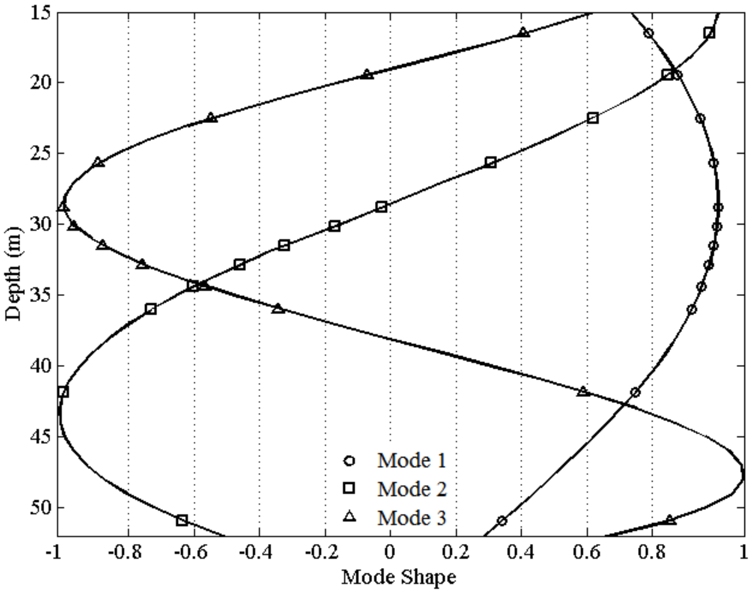


Figure 6-6: Normalized mode shapes at receivers' depths

Figure 6-7 shows the cross-correlation coefficient between $\hat{S}_1(\omega)$ and $\hat{S}_2(\omega)e^{i2\pi C/\omega}$ in blue and the cross-correlation coefficient between $\hat{S}_2(\omega)$ and $\hat{S}_3(\omega)e^{i2\pi C/\omega}$ in red. The mode 1 and 2 pair is able to generate a higher cross correlation coefficient than the mode 2 and 3 pair.

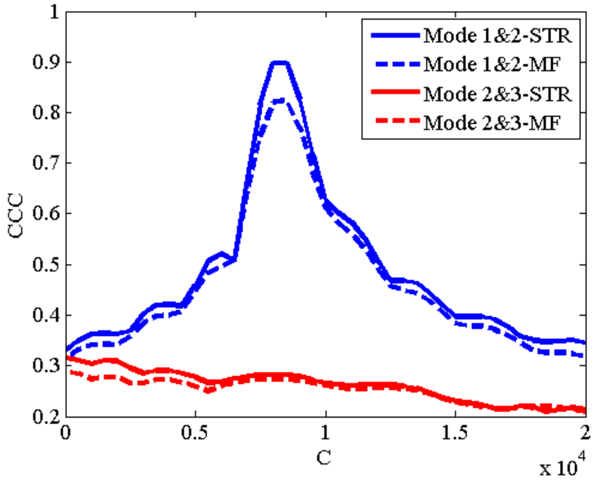


Figure 6-7: Cross-correlation coefficient vs. *C* using mode 1&2 and mode 2&3 with MF and STR calculation.

Figure 6-8 shows the mean-square error calculated from equation 6-9 for both pairs of modes, 1&2 and 2&3. Since mode 3 was not excited, the 2&3 mode pair does not correctly range the source. However, mode 1 and mode 2 have been excited by the source at this depth and the correct range is found.

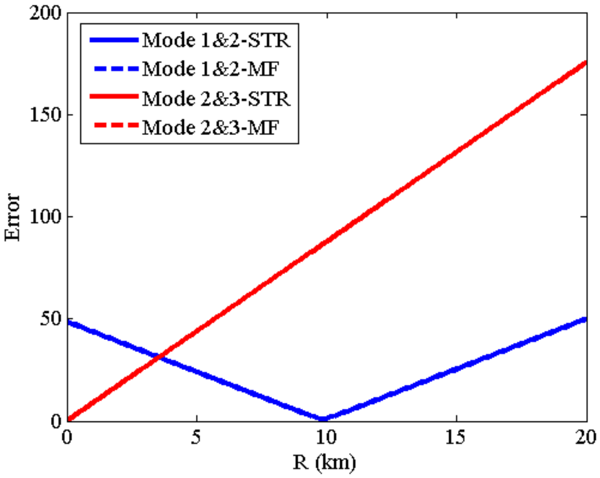


Figure 6-8: Mean-square error calculated form equation 6-9 vs. range using mode 1&2 and mode 2&3 with MF and STR calculation.

6.4 Mode filtering and STR ranging results from 2010 Arctic Experiment

The experiment data used in this section comes from the 2010 Arctic Experiment which was conducted near Kaktovik, Alaska. The specific recordings we made on August 31, 2010 near midnight in 55 m water depth (Figure 6-9). Several marine mammals' calls have been recorded with two instrument packages. This first is a distributed array of Directional Autonomous Seafloor Acoustics Recorders (DASAR) arranged in a triangular grid divided among five sites. The second is a vertical array with 12 elements that was deployed a couple of kilometers away to north-east side of the last DASAR. MF and STR results in this section use the vertical array data. An automated procedure has been developed for localizing bowhead whale sounds in this experiment using DASAR (Thode *et al*., 2012) which was briefly reviewed in section 6.2.3. The results from MF and STR ranging are compared with the results from the established DASAR technique.

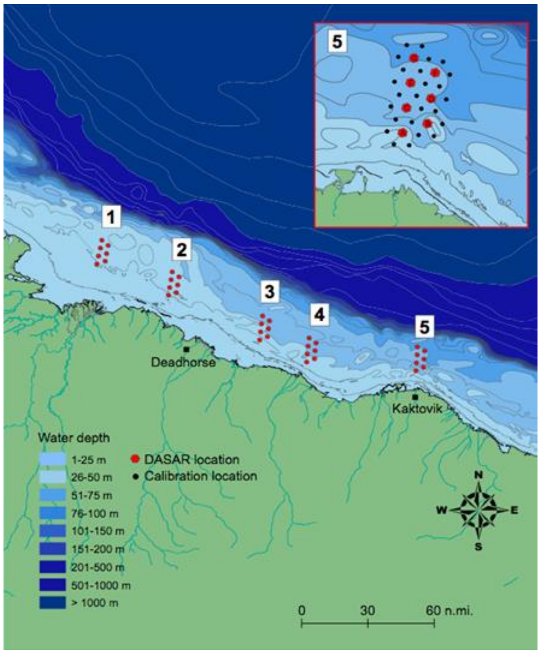


Figure 6-9: DASAR packages deployed in Beaufort Sea. The vertical array with 12 elements has been deployed a couple of kilometer away on north-east side of the last DASAR on each site.

A total of 18 bowhead whale calls have been studied. They were recorded on Aug. 31, 2010 between 00:07:15 p.m. and 01:38:58 a.m. All calls have frequencies between 50-500 Hz, which – at the low frequency end – is very close to the cut-off frequencies of most important modes for this study, modes 1, 2 and 3. Thus, $\Gamma_m(\omega)$ in 6-4 is not a linear function of frequency but this nonlinearity can be used for ranging as described in section 6.2.

Figure 6-10 and Figure 6-11 show the MF and STR ranging results for two sample whale calls. For both figures, part (a) shows the spectrogram of the whale call; part (b) shows the cross-correlation coefficient between $\hat{S}_m(\omega)$ and $\hat{S}_n(\omega)e^{i2\pi C/\omega}$ vs. the dispersive-phase fitting constant C; and part (c) shows the error from equation 6-9 vs. the estimated range, R_e. Finally, the ranges estimated by the MF-based and STR-based ranging techniques are compared in Table 6-1 with the results derived from the DASAR recorded signals for all 18 whale calls.

Figure 6-10 shows a call between two black lines, recorded at 01:01:40 a.m. However, finite SNR and other phenomena present in the ocean (but not in the simulations) have an impact on the magnitude of cross-correlation coefficient in part (b); it is much lower than in the simulations. However, the MF- and STR-ranging techniques are still able to produce adequate range estimates. The range estimated by MF and STR are 8.0 km and 11.4 km, respectively. The DASAR-determined range is 10.18 km for this whale call.

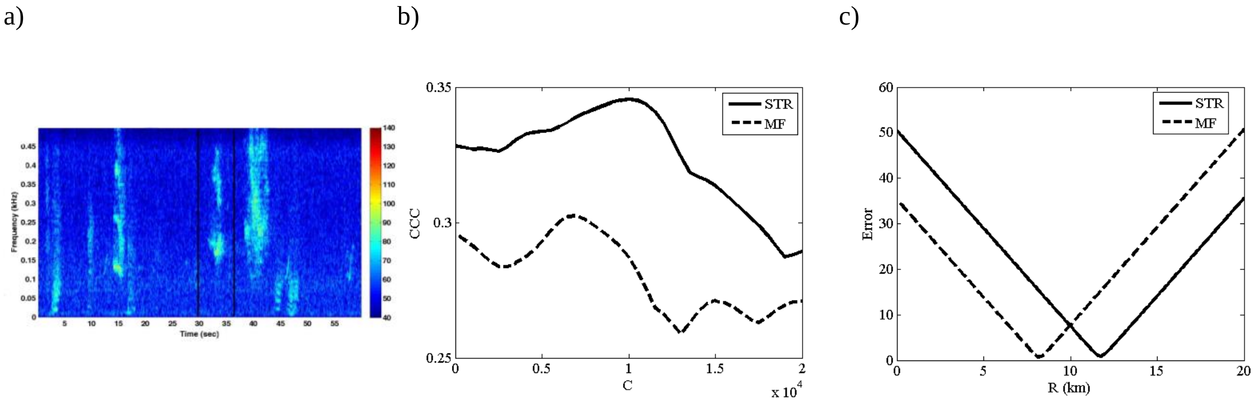


Figure 6-10: a) Spectrogram at 01:01:40 a.m., b) Cross-correlation coefficient vs. *C* using mode 1&2 for MF (dashed line) and STR (solid line), c) Mean-square error calculated form equation 6-9 vs. range using mode 1&2 with MF and STR calculation.

The second example, shown in Figure 6-11, is a call recorded at 00:52:25 a.m. Here, the MF- and STR-determined ranges are both 6.4 km while the DASAR-determined range is 7.37 km.

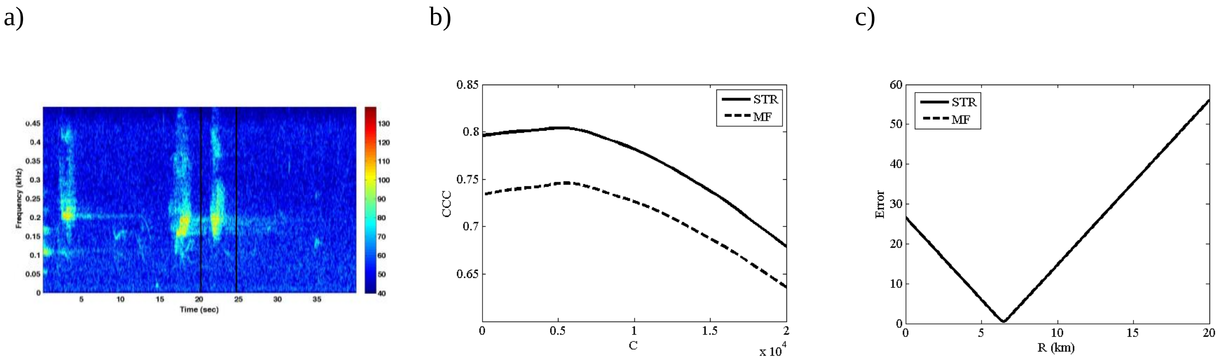


Figure 6-11: a) Spectrogram at 00:52:10 a.m., b) Cross-correlation coefficient vs. *C* using mode 1&2 for MF (dashed line) and STR (solid line), c) Mean-square error calculated form equation 6-9 vs. range using mode 1&2 with MF and STR calculation. MF and STR give the same result for this call.

The same calculation routines were applied to all 18 whale calls recorded on Aug. 31, 2010 between 00:07:15 p.m. and 01:38:58 a.m the next day. The overall comparison between the MF-based and STR-based ranging techniques is presented in Table 6-1. The reference location for these comparisons is the call-array range calculated from the DASAR measurements. Ranging based on the mode filtering technique failed for 1 of these 18 calls (call number 9, shown in Table 6-1). Cross-correlation coefficient vs. dispersion-phase fitting constant, *C*, for this call is shown in Figure 6-12.

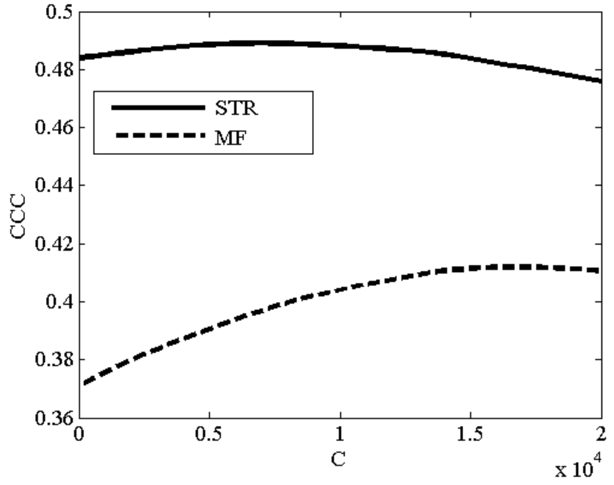


Figure 6-12: Cross-correlation coefficient vs. *C* using mode 1&2 for call number 9 in Table 6-1. MF and STR results are shown by dashed line and solid line, respectively.

Figure 6-12 shows the small difference between the peak value and other values of cross-correlation coefficients which indicates a marginal case for ranging. Since the STR ranging technique is able to generate slightly higher cross-correlation coefficient, it is able to estimate the range better than mode filtering in this case. The physical reason leading to these call-9 results may be that the calling whale was located near the middle of the sound channel where mode 2 cannot be excited. Unfortunately, the Mode 1 & 3 pair cannot be used because of the lack of modal orthogonality across the imperfect vertical aperture of the experimental array.

Table 6-1: Comparison between the estimated ranges from Mode filtering, STR and DASAR techniques

		MF (km)	STR (km)	DASAR (km)	DASAR Uncertainty (km)	Peak CCC (MF)	Peak CCC (STR)	Modes
1	31-Aug-2010 00:46:40 a.m.	8.5	7.2	7.3	±0.40	50.58	61.27	2&3
2	31-Aug-2010 00:46:45 a.m.	8.6	7.0	7.0	±0.67	50.47	61.36	2&3
3	31-Aug-2010 00:48:58 a.m.	7.9	7.9	7.2	±0.48	70.05	74.89	1&2
4	31-Aug-2010 00:49:30 a.m.	6.1	8.7	7.2	±0.67	72.04	77.91	1&2
5	31-Aug-2010 00:49:45 a.m.	8.3	6.5	7.0	±0.62	76.65	77.37	1&2
6	31-Aug-2010 00:52:10 a.m.	5.8	5.8	6.5	±1.58	74.72	80.62	1&2
7	31-Aug-2010 00:52:25 a.m.	6.4	6.4	7.4	±0.48	74.65	80.32	1&2
8	31-Aug-2010 00:55:04 a.m.	7.5	7.5	7.5	±0.58	60.99	61.53	1&2
9	31-Aug-2010 00:55:10 a.m.	19.0	7.8	7.4	±0.67	41.28	49.54	1&2
10	31-Aug-2010 00:55:15 a.m.	7.5	7.5	7.5	±0.43	84.46	88.59	1&2
11	31-Aug-2010 00:55:21 a.m.	7.9	7.9	7.5	±0.53	77.79	82.22	1&2
12	31-Aug-2010 00:59:35 a.m.	8.0	8.0	7.6	±0.33	74.38	75.60	1&2
13	31-Aug-2010 01:00:20 a.m.	7.8	5.2	7.1	±0.29	46.88	53.16	2&3
14	31-Aug-2010 01:01:40 a.m.	8.0	11.4	10.2	±1.15	31.27	37.96	1&2
15	31-Aug-2010 01:01:48 a.m.	11.5	10.5	10.2	±0.91	35.65	42.84	1&2
16	31-Aug-2010 01:26:15 a.m.	17.1	17.1	15.7	±0.08	58.86	70.45	1&2
17	31-Aug-2010 01:28:50 a.m.	18.5	17.9	15.7	±0.08	64.53	90.55	2&3
18	31-Aug-2010 01:38:58 a.m.	29.4	28.2	24.5	±0.53	74.26	79.24	1&2

Table 6-1 shows that MF and STR are able to range the calls as far as 24.5 km away from the recorded array with 4.9 km and 3.7 km differences, respectively, from the range calculated from the DASAR measurements. The eight calls which have been underlined, have the same MF-based and STR-based ranging results. However, MF-based and STR-based results are slightly different for the other 10 calls. When all 18 calls are taken together, there is a slight difference in the accuracy of the two vertical-array ranging techniques. The root-mean square distance, R_{rms}, between the estimated ranges, R_e, and DASAR-determined ranges, R_{DASAR}, computed from:

$$R_{rms} = \left[\frac{1}{18} \sum_{n=1}^{18} \left(R_e - R_{DASAR} \right)_n^2 \right]^{1/2} \tag{6-12}$$

are 0.75 km and 0.31 km, for the MF-based and STR-based techniques, respectively, and these are less than 8% and 3% of the average (DASAR) range for the 18 calls, 9.5 km. These numbers show that STR ranging technique is slightly better than MF when compared with DASAR ranging results. When call 9 is removed, the MF-determined rms range error for the remaining 17 calls improves to 0.41 km. Here, any error between DASAR ranging results and (unobtainable) true ranging information has not been considered.

6.5 Summary and Conclusions

The remote source ranging performance of two techniques, based on mode filtering and mode-based STR, has been investigated using simulated and experimental data from the Arctic Ocean. The simulations mimic the signals and geometry of the Arctic experiment. The ranging performance of both techniques is good, and nearly identical, in simulations at high SNR. However, STR performs slightly better with the experimental data from the Arctic Ocean environment with finite SNR. In the experimental portion of this investigation, both techniques were applied to 18 bowhead whale calls. STR successfully produced an estimated range for all calls; however, mode filtering failed for one call. The failure for that call most likely occurred because the call depth did not fully excite one of the modes used for ranging. For successful ranging using the technique developed here, whale calls should have sufficient bandwidth and also be close to the modal cut-off frequencies so that modal propagation is dispersive over the signal bandwidth. If the recorded calls are far from the cut-off frequencies, the reconstructed signals from mode 1 and 2 will be the same and the constant C in equation 6-7 cannot be reliably found. In this case, alternative ranging methods based differences in modal or ray-path propagation time-delays may apply.

These results suggest that whale-call-to-array ranges in a dispersive underwater sound channel can be robustly and reliably determined by: (*i*) using mode filtering or synthetic time reversal to estimate the original whale-call waveforms using two different propagating modes, (*ii*) fitting of a dispersion constant to maximize the cross correlation of the two estimated waveforms, and (iii) estimating the whale-call-to-array range by minimizing the difference between intermodal dispersion curves determined from the fitting and from a model of the ocean environment.

CHAPTER 7

Summary and conclusions

7.1 Summary

The acoustic signal from a remote source recorded by an underwater hydrophone array is commonly distorted by multipath propagation. And, the ocean's ever-changing acoustic environment is seldom known with sufficient fidelity to predict the details of this distortion. Thus, robust means for determining the location of an unknown remote source (source localization) and estimating its original broadcast waveform (blind deconvolution) in a poorly-known or unknown environment are enduring underwater remote sensing priorities. This dissertation describes how ray-based STR can be used to estimate original-signal waveforms and how ray-based STR sound-channel impulse-response estimates may be exploited for approximate source localization in underwater environments.

It has been shown that STR is successful when the receiving array has sufficient aperture and element density so that conventional delay-and-sum beamforming can be used to distinguish ray-path-arrival directions. In addition, this thesis has described how the basic physics of underwater sound propagation can be combined with novel nonlinear array-signal processing to recover out-of-band lower-frequency signal information from finite bandwidth signals. This manufactured signal information can be used for source localization to surpass the usual spatial Nyquist limit of the receiving array at in-band signal frequencies.

In addition, the manufactured below-band signal information can be exploited to overcome the ill-posed character of blind deconvolution, even when the receiving array is sparse in the signal's frequency band and ordinary beamforming is not useful.

When the frequency of the recorded signals is sufficiently low and close to modal cutoff frequencies (such as bowhead whale calls), an extension of mode-based STR has been developed for sound source ranging in dispersive underwater sound channels.

7.2 Conclusions

The major conclusions and unique contributions of this dissertation are described and listed below:

1. **Ray-based STR may be robust and effective with small vertical arrays that do not span the water column.**

The original description of STR (Sabra and Dowling, 2003) assumed a water-column-spanning vertical array. However, this thesis shows that STR performs well ($C_{\max} \geq 90\%$) with a 7-m-long 32-element vertical array in 60-m-deep water with measured additive noise at SNRs approaching 0 dB (Figure 3-7). It has been assumed that the array is stationary and has no deviation from the vertical position and the elements have equally distributed through the recording array. And, it performs similarly at high SNR with less than 10 array elements. In fact, as few as four or five array elements may be adequate in some circumstances (Figure 3-5). The contribution expands the realm of application for STR.

2. **Ray-based STR is successful when at least one ray-path persists at a constant arrival angle throughout the bandwidth of the signal.**

Identification of this limitation is important since ray-based STR fails when there is no reliable ray-path. Fortunately, the likelihood of STR failure because of this limitation can be anticipated by the presence or absence of persistent (angle independent) ray-path arrivals in the beam-formed array output as a function of frequency. Thus, in an application of ray-based STR, the user can independently determine whether or not STR results should be computed or trusted. This internal means for self-assessment may increase the chances of STR adoption for sonar applications.

3. **Ray based STR results can be improved by a coherent addition of results from different ray arrivals**.

When two (or more) persistent ray-arrivals are apparent in the beam-formed signal, improvement to the STR output is possible when each persistent ray is used as the reference ray and these STR results are coherently combined (Figure 3-8). This finding allows for additional signal-reconstruction improvements in multipath environments that support well defined paths.

4. **The STR impulse responses and simple ray-based back-propagation provide a relatively robust means for source localization when some environmental information is available at the receiving array.**

The relative timing between peaks in the STR impulse response can be used to estimate the source location when there is enough environmental information for ray path calculations (Figure 3-10, Table 3-1). Section 3.4 shows that STR based localization is much less sensitive to environmental mismatch than equivalent match-field processing techniques (Figure 3-10 and Figure 3-11). And, since STR-based localization only requires tracing a few rays (the ones identified recording array) into the acoustic environment beyond the array, it is orders of magnitude less computationally burdensome than matched-field techniques. This finding is potentially important for sonar applications involving source localization since it addresses the two primary objections to matched-field techniques (lack of robustness and computational effort). For these calculations, the environment is assumed to be range independent and the environmental information is limited to receiver depths, water column sound speed profile at the array, water depth at the array, and generic bottom type at the array.

5. **Blind deconvolution with STR is possible even when the receiving array is sparse.**

To achieve the sparse-array blind deconvolution results with STR presented in chapter 4, unconventional frequency-difference beamforming was developed to estimate the phase signature of the unknown source signal. Except for possible mathematical similarities with the quadratic non-linearity exploited for the parametric acoustic array (Westervelt, 1963), frequency-difference beamforming appears to be unique. It may be useful in applications of array signal processing beyond blind deconvolution. This finding extends the parametric range where STR may be successfully applied.

6. **When STR blind deconvolution is used with frequency-difference beamforming and the receiving array is too sparse for conventional beamforming, these STR results are robust because a single propagating mode or ray path does not need to be isolated by the receiving array.**

In fact, the highest cross correlations between the source-broadcast and STR-estimated signals in the study in chapter 4 were commonly obtained from frequency-difference beamforming when the angular resolution of the receiving array was the lowest possible, that is, when the frequency difference was the lowest possible. This result increases the robustness of STR since it essentially allows the requisite correction phase to be estimated from multiple acoustic paths.

7. **Frequency-difference beamforming can be used with a sparse receiving array in the presence of modeling mismatch to isolate signal-path arrival angles when conventional beamforming fails to do so.**

This result is genuinely interesting because it might change the way that acoustic arrays are designed. Frequency-difference beamforming manufactures lower-frequency signal information from the higher-frequency broadband signal recordings, thus, it allows beamforming of broadband array recorded signals that surpasses the spatial Nyquist limit (Figure 4-4 and Figure 4-12). In practice this finding is important because it extends the utility any transducer array to broadband signals having frequencies well above the array's design frequency. For example, the results in chapter 4 show that a receiving array that is half-wavelength spaced at 200 Hz can be used to determine high-frequency signal-path arrival angles when the array's transducers are separated by several tens of wavelengths.

8. **Frequency-difference beamforming can be used for beamforming a sparse random array in a multipath environment**

The simulations presented in section 5.1 show that frequency-difference beamforming can be used to find the arrival angles of signals recorded with a random sparse array within a simple three-dimensional range-independent simulation (Figure 5-2, Figure 5-3, and Figure

5-4). Thus, the chapter 4 frequency-difference beamforming results are not limited to linear arrays.

9. **Frequency-sum beamforming, an extension of frequency-difference beamforming, is not useful unless circumstances are ideal (free-space conditions, no noise, and one source).**

The possibility of manufacturing higher-frequency signal information from the lower-frequency broadband signal recordings (frequency-sum beamforming) has been investigated in section 5.2 and 5.3 with simple free-space and multipath simulations. These simulations show that frequency-sum beamforming may reduce the image-spot size of a single ideal acoustic point source in a quiescent free-space environment. However, there are other more robust means (MVDR beamforming for example) to achieve the same, or nearly the same, performance. Plus, the frequency-sum technique is not suitable for beamforming underwater acoustic signals that propagate through a multipath channel or that arise from more than one source (Figure 5-7, Figure 5-9, Figure 5-12, and Figure 5-14).

10. **Mode-based STR can be used for whale-call ranging when modal propagation is dispersive over signal bandwidth.**

Chapter 6 provides an extension of mode-based STR which can be used for ranging whales in a dispersive medium. This technique involves simple fitting of the nonlinear dependence of the correction phase by reconstructing and comparing source-signal estimates from two different modes. Here, 18 whale calls recorded with a partial-water-column vertical array in the Arctic Ocean were considered and mode-based STR was able to range all calls from 6.5 km to 24.5 km with a ±0.31 km root-mean-square error. (Figure 6-10, Figure 6-11, and Table 6-1). To achieve these results, the environment between the whales and the receiving array has been assumed range independent and the array has been assumed not tilted. This finding is important for monitoring marine mammals and their proximity to potentially-disturbing anthropogenic noise sources. Application scenarios for this technique involve Navy active sonar testing, seismic surveying, and pile-driving in lakes, rivers, and oceans.

11. **Mode-based STR whale call ranging provides a slight improvement over mode-filtering-based whale call ranging.**

Ranging results from mode-based STR has been compared to mode-filtering ranging. Simulations results are identical for both methods. However, experimental results show that mode-based STR ranging is slightly better than mode-filtering technique (Table 6-1). The advantage is slight but it is identifiable in the results from the 18 whale calls.

7.3 Suggestions for future research

The research presented in this thesis is about blind deconvolution and source localization in multipath environments. Most of the results are presented in the form of feasibility studies conducted with acoustic propagation simulations, and then with appropriate propagation measurements. Thus, the conclusions in this dissertation are largely based on incomplete investigation(s) of the possible parameter space(s), so a wealth of possible future work involves fuller investigations of the topics, concepts, and techniques presented here.

Several extensions to this research study can be proposed:

- Assess the use of STR for reconstructing long-duration source signals in dynamic multipath environments. The current research effort only considered static and time invariant multipath environments.
- Extend the mode-based STR to recover the waveforms of actual calls from whales and other marine mammals when recorded signals from a linear array are available. Although this task was considered for this thesis, it was not attempted because there is no routine way of independently knowing or measuring the broadcast waveforms of natural marine-mammal calls from which to assess the suitability of STR for this task.
- Compare the performance of STR with other blind deconvolution techniques. Again, this task was attempted as part of the current effort, but no comparable technique was found.
- Extend STR to recording arrays that are not linear and vertical. This thesis took one small simulation step in this direction in Section 5.1. However, the prevalence and practicality of towed horizontal arrays suggests that more could be done with potentially significant impact for remote sensing applications in the ocean.
- Study the effects of noise, signal characteristics, and array configuration parameters on frequency-difference beamforming performance. Here, in addition to simulation with infinite SNR and ~10 kHz bandwidth, experimental data with high SNR has been used. However, the performance of frequency-difference beamforming technique has not been investigated for low SNR and smaller bandwidths.
- Determine if frequency-difference beamforming together with STR can be used for source localization in a manner analogous to that developed for conventional beamforming and STR.

Bibliography

Abadi, S. H., Rouseff, D., & Dowling, D. R. (2012). Blind deconvolution for robust signal estimation and approximate source localization. *J. Acoust. Soc. Am., 131*(4), 2599-2610.

Abadi, S. H., Song, H. C., & Dowling, D. R. (2012). Broadband sparse-array blind deconvolution using frequency-difference beamforming. *J. Acoust. Soc. Am., 132*(5), 3018-3029.

APL. (1994). *APL-UW high-frequency ocean environmental acoustic models handbook.* Seattle, WA.: No. APL-UW TR 9407, Applied Physics Laboratory, University of Washington.

Baggeroer, A. B., Kuperman, W. A., & Mikhalevsky, P. N. (1993). An overview of matched field methods in ocean acoustics. *IEEE J. Ocean. Eng., 18*, 401-424.

Baggeroer, A. B., Kuperman, W. A., & Schmidt, H. (1988). Matched field processing: Source localization in correlated noise as an optimum parameter estimation problem. *J. Acoust. Soc. Am. , 83*, 571–587 .

Benesty, J. (2000). Adaptive eigenvalue decomposition algorithm for passive acoustic source localization. *J. Acoust. Soc. Am., 107*, 384-391.

Bennett, C. M. (1985). Adaptive Frequency-Domain Beam Forming and Spectral Analysis Using Small Sparse Arrays. *IEEE Journal Of Oceanic Engineering*, 241-247.

Bogart, C., & Yang, T. (1994). Source localization with horizontal arrays in shallow water: Spatial sampling and effective aperture. *J. Acoust. Soc. Am., 96*, 1677–1686.

Broadhead, K., Pflug, L., & Field, R. (2000). Use of higher order statistics in source signature estimation. *J. Acoust. Soc. Am., 107*, 2576-2585.

Broadhead, M., & Pflug, L. (2000). Performance of some sparseness criterion blind deconvolution methods in the presence of noise. *J. Acoust. Soc. Am., 107*, 885-893.

Broadhead, M., Field, R., & Leclere, J. (1993). Sensitivity of the deconvolution of acoustic transients to Green's function mismatch. *J. Acoust. Soc. Am., 94*, 994-1002.

Buck, J. R., Preisig, J. C., & Wage, K. E. (1998). A unified framework for mode filtering and the maximum a posteriori mode filter. *J. Acousti. Soc. Am., 103*, 1813-1824.

Bucker, H. P. (1976). Use of calculated sound fields and matched-field detection to locate sound sources in shallow water. *J. Acoust. Soc. Am., 59*, 368-373.

Cannon, M. (1976). Blind deconvolution of spatially invariant image blurs with phase. *IEEE Transactions on Acoustics, Speech and Signal Processing, 24*, 58-63.

Chapin, S., Ioup, J., Ioup, G., & Smith, G. (2001). Statistical analysis of Green's functions in a shallow ocean environment. *J. Acoust. Soc. Am., 109*, 2295.

Chen, C., He, Y., & Yap, K. (2011). New regularization scheme for blind color image. *Journal of Electronic Imaging, 20*.

Clark, C. W., & Ellison, W. T. (2000). Calibration and comparison of the acoustic location methods used during the spring migration of the bowhead whale, Balaena mysticetus, off Pt. Barrow, Alaska 1984–1993. *J. Acoust. Soc. Am., 107,* 3509-3517.

Clark, C. W., & Ellison, W. T. (2000). Calibration and comparison of the acoustic location methods used during the spring migration of the bowhead whale, Balaena mysticetus, off Pt. Barrow, Alaska, 1984–1993. *J. Acoust. Soc. Am*, 3509–3517.

Collins, M., & Kuperman, W. (1991). Focalization: Environmental Focusing and Source Localization. *J. Acoust. Soc. Am., 90*, 1410–1422.

Collison, N. E., & Dosso, S. E. (2000). Regularized matched-mode processing for source localization. *J. Acoust. Soc. Am., 107*, 3089-3100.

Cousins, D. B. (2005). A model-based algorithm for environmentally adaptive bathymetry and sound velocity profile estimation. *PhD. Dissertation, University of Rhode Island*.

Delarue, J., Laurinolli, M., & Martin, B. (2009). Bowhead whale (Balaena mysticetus) songs in the Chukchi Sea between October 2007 and May 2008. *J. Acoust. Soc. Am., 126*(6), 3319-3328.

Dosso, S., & Wilmut, M. (2011). Bayesian multiple-source localization in an uncertain ocean environment. *J. Acoust. Soc. Am., 129*, 3577-3589.

Dowling, D. R. (1994). Acoustic pulse compression using passive phase conjugate processing. *J. Acoust. Soc. Am., 95*, 1450-1458.

Dzieciuch, M., Worcester, P., & Munk, W. (2001). Turning point filters: Analysis of sound propagation on a gyre-scale. *J. Acoust. Soc. Am., 110*, 135–149 .

Fizell, R. G. (1987). Application of high-resolution processing to range and depth estimation using ambiguity function methods. *J. Acoust. Soc. Am. , 82*, 606–613 .

Grachev, G. A. (1993). Theory of acoustic field invariants in layered waveguides. *Acoust. Phys., 39*, 67-71.

Greene, C. R., McLennan, M. W., Norman, R. G., Mcdonald, T. L., Jakubczak, R. S., & Richardson, W. J. (2004). Directional frequency and recording (DIFAR) sensors in seafloor recorders to locate calling bow-head whales during their fall migration. *J. Acoust. Soc. Am., 116*, 799-813.

Harrison, B., Vaccaro, R., & Tufts, D. (1998). Robust matched-field localization in uncertain ocean environments. *J. Acoust. Soc. Am., 103*, 3721-3724.

Hermand, J. P. (1999). Broad-Band Geoacoustic Inversion in Shallow Water from Waveguide Impulse Response Measurements on a Single Hydrophone: Theory and Experimental Results. *IEEE J. Ocean. Eng., 24*, 41-66.

Hermand, J. P., & Gerstoft, P. (1996). Inversion of broad-band multitone acoustic data from the YELLOW SHARK summer experiments. *IEEE J. Ocean. Eng., 21*, 324-346.

Hinich, M. J. (1979). Maximum likelihood estimation of the position of a radiating source in a waveguide-guide. *J. Acoust. Soc. Am. , 66*, 480–483 .

Hodges, R. P. (2010). *Underwater Acoustics: Analysis, Design and Performance of Sonar.* John Wiley & Sons, Ltd.

Holfort, I. K., Grain, F., & Jensen , J. A. (2009). Broadband minimum variance beamforming for ultrasound imaging. *IEEE transactions on ultrasonic, ferroelectrics, and frequency control, 56*, 314-325.

Huber, P. J. (1964). Robust estimation of a location parameter. *Ann. Math. Stat., 35*, 73–78.

Hursky, P., Hodgkiss, W., & Kuperman, W. (2001). Matched field source localization with data derived modes. *J. Acoust. Soc. Am., 109*, 1355–1366.

Hursky, P., Siderius, M., Porter, M. B., & McDonald, V. K. (2004). High-frequency (8–16 kHz) model-based source localization. *J. Acoust. Soc. Am., 115*, 3021-3032.

Jason, R., Richard, K., & Stephen, C. (2010). Three-dimensional FLASH laser radar range estimation via blind deconvolution. *Journal of Applied Remote Sensing, 4*, 043517.

Jensen, F., Kuperman, W., Porter, M., & Schmidt, H. (1994). *Computational Ocean Acoustics.* New York: American Institute of Physics.

Karaman, M., Li, P. C., & O’Donnell , M. (1995). Synthetic aperture imaging for small scale systems. *IEEE transactions on ultrasonic, ferroelectrics, and frequency control, 42*, 429–442.

Kinsler, L. E., Frey, A. R., Coppens, A. B., & Sanders, J. V. (2000). *Fundamentals of Acoustics.* New York: Wiley.

Krolik, J. L. (1992). Matched-field minimum variance beamforming in a random ocean channel. *J. Acoust. Soc. Am., 92*(3), 1408-1419.

Kundur, D., & Hatzinakos, D. (1996). Blind image deconvolution. *IEEE, Signal Processing Magazine, 13*, 43-64.

Lenth, R. V. (1981a). On finding the source of a signal. *Technometrics, 23*, 149–154.

Lenth, R. V. (1981b). Robust measures of location for directional data. *Technometrics, 23,* 77–81.

Lin, Y. T., Newhall, A. E., & Lynch, J. F. (2012). Low-frequency broadband sound source localization using an adaptive normal mode back-propagation approach in a shallow-water ocean. *J. Acoust. Soc. Am., 131*, 1798-1813.

Lo, E., Zhou, J., & Shang, E. (1983). Normal mode filtering in shallow water. *J. Acoust. Soc. Am, 74*(6), 1833-1836.

Martins, M., Jesus, S., Gervaise, C., & Quinqui, A. (2002). A time-frequency approach to blind deconvolution in multipath underwater channels. *ICASSP'02, Orlando, USA*.

Michailovich, O., & Adam, D. (2005). A novel approach to the 2-D blind deconvolution problem in medical ultrasound. *IEEE Transactions on Medical Imaging, 24*, 86 - 104.

Michalopoulou, Z. (2006). Multiple source localization using a maximum a posteriori Gibbs sampling approach. *J. Acoust. Soc. Am., 120*, 2627-2634.

Morrissey, R. P., DiMarzio, N., Jarvis, S., & Moretti, D. J. (2006). Passive acoustic detection and localization of sperm whales (Physeter macrocephalus) in the tongue of the ocean. *Applied Acoustics, 67*(11-12), 1091-1105.

Morrissey, R. P., Ward, J., DiMarzio, N., Jarvis, S., & Moretti, D. J. (2006). Passive acoustic detection and localization of sperm whales (Physeter macrocephalus) in the tongue of the ocean. *Applied Acoustics*, 1091–1105.

Porter, M. B. (1984). *A numerical method for computing ocean acoustic modes.* Northwestern University.

Porter, M. B., & Bucker, H. P. (1987). Gaussian beam tracing for computing ocean acoustic fields. *J. Acoust. Soc. Am., 82*, 1349-1359.

Porter, M. B., & Reiss, E. L. (1985). A numerical method for bottom interacting ocean acoustic normal modes. *J. Acoust. Soc. Am., 77*(5), 1760-1767.

Porter, M. B., & Tolstoy, A. (1994). The matched-field processing benchmark problems. *Journal of Computational Acoustics, 2*, 161–185.

Roan, M., Gramann, M., Erling, J., & Sibul, L. H. (2003). Blind deconvolution applied to acoustical systems identification with supporting experimental results. *J. Acoust. Soc. Am., 114*, 1988–1996.

Rouseff, D., Jackson, D., Fox, W., Jones, C., Ritcey, J., & Dowling, D. (2001). Underwater acoustic communication by passive phase conjugation theory and experimental results. *IEEE Ocean. Engin., 26*, 821-831.

Rouseff, D., Wang, Z., Zhou, S., & Zhou, L. (2010). The 2009 Cooperative Array Performance Experiment (CAPEx09) comparing vector- and pressure-sensor arrays. *in Proc. Eur. Conf. Underwater Acoustics (ECUA).* Istanbul, Turkey.

Roux, P., Cornuelle, B. D., Kuperman, W. A., & Hodgkiss, W. S. (2008). The structure of raylike arrivals in a shallow-water waveguide. *J. Acoust. Soc. Am., 124*(6), 3430-3439.

Sabra, K. G., & Dowling, D. R. (2004). Blind deconvolution in oceanic waveguides using artificial time reversal. *J. Acoust. Soc. Am., 116*, 262–271.

Sabra, K. G., Song, H. C., & Dowling, D. R. (2010). Ray-based blind deconvolution in ocean sound channels. *J. Acoust. Soc. Am., 127*, EL42-EL47.

Sarder, P., & Nehorai, A. (2008). Estimating Locations of Quantum-Dot-Encoded Microparticles From Ultra-High Density 3-D Microarrays. *IEEE Transactions on NanoBioscience, 4*, 284 - 297.

Shang, E. (1985). Source depth estimation in waveguides. *J. Acoust. Soc. Am., 77*, 1413-1418.

Shang, E. C., Clay, C. S., & Wang, Y. Y. (1985). Passive harmonic source ranging in waveguides by using mode filter. *J. Acoust. Soc. Am., 78*(1), 172-175.

Shang, E. C., Clay, C. S., & Wang, Y. Y. (1985). Passive harmonic source ranging in waveguides by using mode filter. *J. Acoust. Soc. Am.*, 172-175.

Sibul, L. H., Roan, M. J., & Erling, J. (2002). Deconvolution and signal extraction in geophysics and acoustics. *J. Acoust. Soc. Am., 112*, 2389.

Siderius, M., & Hermand, J. P. (1999). Yellow Shark Spring 1995: Inversion results from sparse broadband acoustic measurements over a highly range-dependent soft clay layer. *J. Acoust. Soc. Am., 106*, 637-651.

Siderius, M., Jackson, D. R., & Rouseff, D. (1997). Multipath compensation in shallow water environments using a virtual receiver. *J. Acoust. Soc. Am., 102*, 3439–3449.

Smith , J. F., & Finette, S. (1993). Simulated annealing as a method of deconvolution for acoustic transients measured on a vertical array. *J. Acoust. Soc. Am., 94*, 2315-2325.

Smith, G. B. (2003). Multipath compression by multiple convolutions. *J. Acoust. Soc. Am., 113*, 2213.

Song, H. C., Hodgkiss, W. S., Kuperman, W. A., & Stevenson, M. (2010). High-rate multiuser communications in shallow water. *J. Acoust. Soc. Am. , 128*, 2920-2925.

Song, H. C., Hodgkiss, W. S., Kuperman, W. A., Akal, T., & Stevenson, M. (2009). High frequency acoustic communications achieving high bandwidth efficiency (L). *J. Acoust. Soc. Am. , 126*, 561-563.

Steinberg , B. D. (1976). *Principles of Aperture and Array System Design* . New York: Wiley.

Stuart, J., & Christou, J. (1993). Restoration of Astronomical Images by Iterative Blind Deconvolution. *Astrophysical Journal, 415*, 862.

Synnevag, J. F., Austeng, A., & Holm , S. (2007). Adaptive beamforming applied to medical ultrasound imaging. *IEEE transactions on ultrasonic, ferroelectrics, and frequency control, 54*, 1606-1613.

Taxt, T., & Strand, J. (2001). Two-Dimensional Noise-Robust Blind Deconvolution of Ultrasound Images. *IEEE Transactions on Ultrasonic, Ferroelectrics, and Frequency Control, 48*(4), 861-866.

Thode, A. M., Gerstoft, P., Burgess, W. C., Sabra, K. G., Guerra, M., Stokes, D. M., . . . Cato, D. H. (2006). A Portable Matched-Field Processing System Using Passive Acoustic Time Synchronization. *J. Acoust. Soc. Am.*, 696-710.

Thode, A., Kim, K. H., Blackwell, S. B., Greene, C. R., McDonald, T. L., & Macrander, M. A. (2012). Automated detection and localization of bowhead whale sounds in the presence of seismic airgun surveys. *J. Acoust. Soc. Am., 131*, 3726-3747.

Thode, A., Skinner, J., Scott, P., Roswell, J., Straley, J., & Folkert, K. (2000). Tracking sperm whales with a towed acoustic vector sensor. *J. Acoust. Soc. Am.*, 2681-2694.

Weber, R., & Bohme, J. F. (2002). Adaptive super-exponential methods for blind. *Sensor Array and Multichannel Signal Processing Workshop* , 585–589.

Westervelt, P. J. (1963). Parametric acoustic array. *J. Acoust. Soc. Am., 35*, 535–537.

Wilcock, W. S. (2012). Tracking fin whales in the northeast Pacific Ocean with a seafloor seismic network. *J. Acoust. Soc. Am., 132*(4), 2408-2419.

Xiaohong, S., & Yu-Hen, H. (2005). Maximum likelihood multiple-source localization using acoustic energy measurements with wireless sensor networks. *EEE Transactions on Signal Processing* , 44 - 53.

Xinhua, Z., Jiang, X., Li, X. L., & Zhang, X. D. (2001). Deconvolution of sparse underwater acoustic multipath channel with a large time-delay spread. *J. Acoust. Soc. Am., 127*, 909–919.

Yang, T. C. (1989). Effectiveness of mode filtering: A comparison of matched-field and matched-mode processing. *J. Acoust. Soc. Am., 87*(5), 2072-2084.

Yoo, K., & Yang, T. C. (1999). Broadband source localization in shallow water in the presence of internal waves. *J. Acoust. Soc. Am., 106*, 3255-3269.

Yu, C., Zhang, C., & Xie, L. (2012). A Blind Deconvolution Approach to Ultrasound Imaging. *IEEE Transactions on Ultrasonic, Ferroelectrics, and Frequency Control, 59*(2), 271-280.

Zeng, W. J., Jiang, X., Li, X. L., & Zhang, X. D. (2009). Deconvolution of sparse underwater acoustic multipath channel with a large time-delay spread. *J. Acoust. Soc. Am., 127*, 909–919.

Ziomek , L. J. (1995). *Fundamentals of Acoustic Field Theory and Space-Time Signal Processing* . Boca Raton: CRC Press.

Ziomek, L. J. (1993). Three necessary conditions for the validity of the Fresnel phase approximation for the near-field beam pattern of an aperture. *IEEE J. Oceanic Eng., 18*, 73-75.

Zurk, L. M., & Ward, J. (2003). Source motion mitigation for adaptive matched field processing. *J. Acoust. Soc. Am., 113,* 2719-2731.