UNIVERSITY OF CALIFORNIA SAN DIEGO

Localization of a quiet moving source with modal-MUSIC and range-coherent matched field processing

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

in

Oceanography

by

Franklin Hunter Akins

Committee in charge:

Professor William Kuperman, Co-Chair
Professor Bruce Cornuelle, Co-Chair
Professor William Coles
Professor William Hodgkiss
Professor Florian Meyer

2023

The Dissertation of Franklin Hunter Akins is approved, and it is acceptable in quality and form for publication on microfilm and electronically.

University of California San Diego

2023

DEDICATION

This thesis is dedicated to my parents Paul and Victoria Akins.

EPIGRAPH

Well, once there was only dark. If you ask me, the light's winning.

Rust Cohle

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

ACKNOWLEDGEMENTS

I would like to acknowledge Professor Bill Kuperman for his mentorship. His sense of humor, boundless knowledge, and endless ideas made the work in this thesis a great pleasure. I would like to acknowledge Professor Bill Hodgkiss for his generosity with his time, sharing his impeccable attention to detail, and for bringing me to sea on three separate cruises. I would like to acknowledge Professor Bruce Cornuelle for sharing with me his infectious enthusiasm, fervent curiosity, and rigorous approach. Their comments and critiques will echo in my mind and guide me in my future endeavors.

I would like to acknowledge Heechun Song, whose enthusiasm for underwater acoustics first piqued my interest in the subject. I would like to acknowledge Florian Meyer for being a good shipmate and for many interesting conversations. I would like to acknowledge Gihoon Byun for our collaboration and exchange of information and ideas during my time at Scripps. I would like to acknowledge my peer mentor Ludovic Tenorio, who, after many surf sessions and trips to Baja, became a true friend. I would like to acknowledge my former house mate Momme Hell, who is a great friend and an impressive scientist. I would like to Ryan Saenger with whom I have had the pleasure to share an office, as well as many stimulating and enlightening discussions.

Finally, I would like to acknowledge my parents, Paul and Victoria Akins, for always encouraging and supporting me in my search for knowledge.

Chapter 2, in full, is a reprint of the material as it appears in Akins, F. H. and Kuperman, W. A. (2021). Range-coherent matched field processing for low signal-to-noise ratio localization. *The Journal of the Acoustical Society of America*, 150(1), 270-280. The dissertation author was the primary investigator and author of this paper. The coauthor listed in this publication directed and supervised the research.

Chapter 3, in full, is a reprint of the material as it appears in Akins, F. H. and Kuperman,

W. A. (2022). Modal-MUSIC: A passive mode estimation algorithm for partially spanning arrays. *JASA Express Letters*, 2(7), 074802. The dissertation author was the primary investigator and author of this paper. The coauthor listed in this publication directed and supervised the research.

Chapter 4, in full, has been submitted for publication of the material as it may appear in Akins and Kuperman (2023), *The Journal of the Acoustical Society of America*. The dissertation author was the primary investigator and author of this paper. The coauthor listed in this publication directed and supervised the research.

VITA

2017	Bachelor of Arts, University of California, Berkeley
2017-2018	Scientific engineering associate, Lawrence Berkeley National Laboratory
2018-2023	Graduate student researcher, Scripps Institution of Oceanography University of California, San Diego
2023	Doctor of Philosophy, University of California San Diego

PUBLICATIONS

1. F. Hunter Akins and W. A. Kuperman. Modal-MUSIC: A passive mode estimation algorithm for partially spanning arrays. JASA Express Letters, 2(7):074802, 2022.

2. F. Hunter Akins and W. A. Kuperman. Experimental low signal-to-noise-ratio matched field processing without prior geoacoustic information. (submitted for publication in The Journal of the Acoustical Society of America). 2023.

3. Gihoon Byun, F. Hunter Akins, Kay L. Gemba, H. C. Song, and W. A. Kuperman. Multiple constraint matched field processing tolerant to array tilt mismatch. The Journal of the Acoustical Society of America, 147(2):12311238, 2020.

4. B. Dafflon, S. Wielandt, J. Lamb, P. McClure, I. Shirley, S. Uhlemann, C. Wang, S. Fiolleau, C. Brunetti, F. H. Akins, J. Fitzpatrick, S. Pullman, R. Busey, C. Ulrich, J. Peterson, and S. S. Hubbard. A distributed temperature profiling system for vertically and laterally dense acquisition of soil and snow temperature. The Cryosphere, 16(2):719736, 2022.

5. F. Hunter Akins and W. A. Kuperman. Range-coherent matched field processing for low signal-to-noise ratio localization. The Journal of the Acoustical Society of America, 150(1): 270-280, 2021.

ABSTRACT OF THE DISSERTATION

Localization of a quiet moving source with modal-MUSIC and range-coherent matched field processing

by

Franklin Hunter Akins

Doctor of Philosophy in Oceanography

University of California San Diego, 2023

Professor William Kuperman, Co-Chair
Professor Bruce Cornuelle, Co-Chair

Passive acoustic signal processing has the remarkable capacity to achieve traditional goals of active sonar, such as detection and ranging, simply by listening. Matched field processing (MFP) exploits the physics of wave propagation to design filters tailored to the underwater acoustic environment and solve passive sonar problems of detection and localization. The second chapter of this dissertation introduces range-coherent MFP to increase the coherent integration time of the signal from a moving source. Range-coherent MFP searches over candidate source tracks to compute time-dependent replicas that are then used to coherently combine data over a vertical array and over multiple resolution cells. The method improves on existing methods

by extending the application of MFP to low signal-to-noise ratio (SNR) regimes. The third chapter introduces modal-MUSIC to estimate waveguide properties by using the structure of noise recorded on a vertical line array. Modal-MUSIC estimates the horizontal wavenumbers of the modal field from noise on a partially-spanning array. It relaxes the requirement of state-of-the art passive methods which require a full water column spanning array. In this chapter, the modal field parameters estimated from noise with modal-MUSIC are successfully used to localize a source at high SNR without an environmental model. In the fourth chapter, these two developments are brought together to demonstrate the feasibility of MFP at low SNR in the absence of extensive *a priori* information on the seabed properties. The fourth chapter improves upon the model-free localization results with modal-MUSIC in chapter three by fitting the modal wavenumbers to a geoacoustic model. This dissertation extends the application of MFP to the passive localization of a low SNR moving source in the absence of knowledge of the seabed properties.

Chapter 1

Introduction

Passive sonar uses measurements of the acoustic field to detect, classify, localize, and track sources of acoustic energy and to infer characteristics of the acoustic medium simply by listening. Contrast this with an active sonar system, such as that used by a bat or a dolphin, that uses the properties of the backscattered field of a controlled source to solve such problems. The use of a controlled source "illuminates" any reflectors in the acoustic environment, but with the undesirable consequence of broadcasting one's presence. The ability to infer information by simply listening has the attractive features of simpler hardware requirements (no transmitter), less energy required, and of not impacting the system under observation. This dissertation addresses limitations of matched field processing (MFP), a passive sonar paradigm, in situations with low signal-to-noise ratio and limited environmental knowledge.

1.1 Matched field processing

MFP addresses the passive localization of a point source in an underwater acoustic medium. It combines the physics of wave propagation in layered media with the signal processing tools of filter design and parameter estimation. When applied to an array, MFP computes a bank of spatial filters, or "replicas", whose weights depend on both the source coordinates as well as the characteristics of the environment. It therefore requires a numerical code that can model sound in layered media for various candidate source positions, as well as accurate information

describing the environment.

In free space, far from any sources, the acoustic field is a sum of plane waves: one plane wave from each source. An array of hydrophone receivers can use plane wave beamforming to measure the angles of these sources simply by listening. The broad utility of beamforming, with applications in radar, sonar, communications, geophysics and more, has resulted in the rich development of signal processing methods to overcome challenges associated with its implementation, such as array element position uncertainty and the estimation of a source signal in the presence of loud interfering sources [1].

The underwater acoustic medium provides two additional complications with respect to the free space problem: boundaries and heterogeneity within the medium itself. In this context, a single source no longer produces a single plane wave, but rather a complicated interference pattern dependent on the source-receiver locations and the medium properties. This presents a significant increase in difficulty when compared with plane wave beamforming, since the computation of weights for MFP filters adds the modeling problem as well as the challenge and cost of environmental characterization.

The modeling problem is to numerically solve the wave equation for a given parameterization of the acoustic environment and has been well-studied for the most useful underwater acoustic applications [2, 3, 4, 5]. Once equipped with an accurate forward model, MFP computes beamformer weights for an array at each candidate source position within the waveguide. Compare this to plane wave beamforming, where there is a set of weights for each candidate source angle θ, and the weights are trivial to compute.

Environmental characterization requires that acoustically important medium properties such as index of refraction and boundary qualities be estimated using oceanographic measurements such as temperature depth profiles and geological surveys, or via acoustic remote sensing as in tomography [6, 7]. Environmental mismatch (inaccurate specification of the environment) causes an error in the field prediction due to an inaccurate characterization of environmental parameters, and has proven to be a serious limitation on the practical utility of MFP [8]. When

environmental mismatch is accounted for, MFP promises improved array resolution over plane wave beamforming [9] as well as the ability to discriminate source depth, range, and azimuth (rather than just direction-of-arrival) [10].

Chapters three and four of this dissertation address the challenge of environmental characterization by estimating an acoustic model directly from ambient acoustic noise measured on a vertical line array spanning roughly half of the water column.

1.2 Effect of source motion on narrowband underwater acoustic signals

Chapter two of this dissertation presents a method to exploit signal coherence between resolution cells by modeling source motion. Source motion must be considered when designing passive acoustic systems due to its implications for the stationarity and bandwidth of the received signal. Stationarity (or lack thereof) is important for the stability of covariance matrix estimation required by adaptive methods that use estimates of the signal and noise statistics to optimize system performance [11, 12, 13, 14].

1.2.1 Processing within a resolution cell

Due to the dependence of the acoustic field on source position, moving sources violate the assumption of stationarity required to design constant filters. The general approach to dealing with this problem is to restrict analysis to data collected while a source stays within a resolution cell [14].

The resolution cell for a given array geometry and waveguide is a spatial region over which the structure of samples on the array is approximately constant for a source located anywhere within that region. In the context of MFP, a resolution cell can also be viewed as a region over which a single fixed replica will provide a minimum amount of gain determined by the parameter α.

1.2.2 MFP for a moving narrowband source

In order to apply MFP to a moving narrowband source, pre-processing parameters and covariance integration periods must be selected based on prior information of the source motion. First, the FFT length is fixed to give optimal input SNR for an assumed source bandwidth. The output of the FFT at a single frequency at all array elements is referred to as a "snapshot". Once the FFT length is fixed, the number of snapshots corresponding to a single resolution cell is fixed based on a maximum prior source speed. Within each resolution cell, a single MFP replica computed for a representative stationary source position will be a good match for each snapshot. The power of each filter output (corresponding to each candidate source position) defines an ambiguity surface.

Chapter two introduces range-coherent MFP, which then coherently combines the (complex) filter outputs of each resolution cell according to a hypothesized source track. When the hypothesized source track corresponds to the true track, the filter outputs are reinforced over noise providing gain.

1.3 Environmental model description for MFP

The sensitivity of the acoustic field to the environmental parameters is dependent on both source frequency as well as source-receiver geometry [15, 16, 17]. A "correct" environmental specification for modeling purposes is therefore also frequency dependent. In particular, at low frequency (say, less than 200 Hz) in shallow water, the field is sensitive to the properties of the underlying seafloor in the top tens of meters [18, 19, 20, 21]. The success of MFP in this frequency regime is thus dependent on accurate specification of the seabed near the interface with the ocean.

Methods of acquiring the relevant seabed information include direct in situ sampling with sediment cores and drilling [22] as well as inverse methods based on active or passive acoustic measurements . Examples of acoustic remote sensing methods use modal group speed

[23], noise correlation from surface noise [24], the power reflection coefficient measured from ambient noise [25], complex pressure field measured at known ranges [26], and echosounder measurements.

Chapter three of this dissertation presents a passive method for estimating the modal wavenumbers from noise with a partially-spanning array using knowledge of the SSP from the surface over the array, but no information regarding the properties of the seabed. This is an improvement over existing passive modal estimation methods which require a fully-spanning VLA [27]. The estimated wavenumbers provided sufficient accuracy to perform MFP at high-SNR. Chapter four shows that using the wavenumber estimates to infer seabed properties improves replica accuracy and permits source localization at low SNR.

1.4 Dissertation overview

This dissertation introduces two methods for extending the applicability of MFP to the localization of a low-SNR source with limited environmental information. Chapter two presents range-coherent MFP for low signal-to-noise ratio localization of a coherent source. Range-coherent addresses the limitation on coherent gain imposed by the changing field with source motion. Chapter three presents modal-MUSIC, a passive mode estimation method using a partially-spanning VLA. Modal-MUSIC addresses the challenges associated with obtaining seabed information required for accurate modeling of low-frequency sound in shallow water. Chapter four improves the results of chapters two and three, performing low signal-to-noise ratio localization with range-coherent MFP without a prior geoacoustic model.

Bibliography

[1] B. Van Veen and K. Buckley, "Beamforming: a versatile approach to spatial filtering," *IEEE ASSP Magazine*, vol. 5, no. 2, pp. 4–24, 1988.

[2] M. Porter and E. L. Reiss, "A numerical method for oceanacoustic normal modes," *The Journal of the Acoustical Society of America*, vol. 76, no. 1, pp. 244–252, 1984.

[3] M. B. Porter, "The KRAKEN Normal Mode Program," tech. rep., SACLANT Undersea Research Centre, La Spezia, Italy Rep. SM-245, 1991.

[4] M. D. Collins and W. L. Siegmann, *Parabolic Wave Equations with Applications*. Springer, 2019.

[5] F. B. Jensen, W. A. Kuperman, M. B. Porter, and H. Schmidt, *Computational ocean acoustics*. Springer Science & Business Media, 2011.

[6] B. Cornuelle, W. Munk, and P. Worcester, "Ocean acoustic tomography from ships," *Journal of Geophysical Research: Oceans*, vol. 94, no. C5, pp. 6232–6250, 1989.

[7] W. Munk and C. Wunsch, "Ocean acoustic tomography: A scheme for large scale monitoring," *Deep Sea Research Part A. Oceanographic Research Papers*, vol. 26, no. 2, pp. 123–161, 1979.

[8] M. B. Porter and A. Tolstoy, "The matched field processing benchmark problems," *Journal of Computational Acoustics*, vol. 02, no. 03, pp. 161–185, 1994.

[9] S. Kim, G. F. Edelmann, W. A. Kuperman, W. S. Hodgkiss, H. C. Song, and T. Akal, "Spatial resolution of time-reversal arrays in shallow water," *The Journal of the Acoustical Society of America*, vol. 110, no. 2, pp. 820–829, 2001.

[10] A. B. Baggeroer, W. A. Kuperman, and H. Schmidt, "Matched field processing: Source localization in correlated noise as an optimum parameter estimation problem," *The Journal of the Acoustical Society of America*, vol. 83, no. 2, pp. 571–587, 1988.

[11] H. Song, W. A. Kuperman, W. S. Hodgkiss, P. Gerstoft, and Jea Soo Kim, "Null broadening with snapshot-deficient covariance matrices in passive sonar," *IEEE Journal of Oceanic Engineering*, vol. 28, pp. 250–261, April 2003.

[12] C. A. Zala and J. M. Ozard, "Matchedfield processing for a moving source," *The Journal of the Acoustical Society of America*, vol. 92, no. 1, pp. 403–417, 1992.

[13] L. M. Zurk, N. Lee, and J. Ward, "Source motion mitigation for adaptive matched field processing," *The Journal of the Acoustical Society of America*, vol. 113, no. 5, pp. 2719–2731, 2003.

[14] A. B. Baggeroer and H. Cox, "Passive sonar limits upon nulling multiple moving ships with large aperture arrays," in *Conference Record of the Thirty-Third Asilomar Conference on Signals, Systems, and Computers (Cat. No.CH37020)*, vol. 1, pp. 103–108 vol.1, Oct 1999.

[15] M. V. De Hoop and R. D. van Der Hilst, "On sensitivity kernels for wave-equationtransmission tomography," *Geophysical Journal International*, vol. 160, no. 2, pp. 621–633, 2005.

[16] J. Sarkar, B. D. Cornuelle, and W. A. Kuperman, "Information and linearity of time-domain complex demodulated amplitude and phase data in shallow water," *The Journal of the Acoustical Society of America*, vol. 130, no. 3, pp. 1242–1252, 2011.

[17] B. Sarkar, *Linear acoustic sensitivity kernels and their applications in shallow water environments*. University of California, San Diego, 2011.

[18] C. Yardim, P. Gerstoft, and W. S. Hodgkiss, "Tracking of geoacoustic parameters using kalman and particle filters," *The Journal of the Acoustical Society of America*, vol. 125, no. 2, pp. 746–760, 2009.

[19] C. Yardim, P. Gerstoft, and W. S. Hodgkiss, "Geoacoustic and source tracking using particle filtering: Experimental results," *The Journal of the Acoustical Society of America*, vol. 128, no. 1, pp. 75–87, 2010.

[20] P. A. Baxley, N. O. Booth, and W. S. Hodgkiss, "Matched-field replica model optimization and bottom property inversion in shallow water," *The Journal of the Acoustical Society of America*, vol. 107, no. 3, pp. 1301–1323, 2000.

[21] M. Musil, N. R. Chapman, and M. J. Wilmut, "Range-dependent matched-field inversion of swellex-96 data using the downhill simplex algorithm," *The Journal of the Acoustical Society of America*, vol. 106, no. 6, pp. 3270–3281, 1999.

[22] M. Richardson and K. Briggs, "In situ and laboratory geoacoustic measurements in soft mud and hard-packed sand sediments: Implications for high-frequency acoustic propagation and scattering," *Geo-Marine Letters*, vol. 16, pp. 196–203, 1996.

[23] J. Bonnel, Y.-T. Lin, D. Eleftherakis, J. A. Goff, S. Dosso, R. Chapman, J. H. Miller, and G. R. Potty, "Geoacoustic inversion on the new england mud patch using warping and dispersion curves of high-order modes," *The Journal of the Acoustical Society of America*, vol. 143, no. 5, pp. EL405–EL411, 2018.

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

[25] C. H. Harrison and D. G. Simons, "Geoacoustic inversion of ambient noise: A simple method," *The Journal of the Acoustical Society of America*, vol. 112, no. 4, pp. 1377–1389, 2002.

[26] G. V. Frisk, K. M. Becker, S. D. Rajan, C. J. Sellers, K. von der Heydt, C. M. Smith, and M. S. Ballard, "Modal mapping experiment and geoacoustic inversion using sonobuoys," *IEEE Journal of Oceanic Engineering*, vol. 40, no. 3, pp. 607–620, 2015.

[27] T. B. Neilsen and E. K. Westwood, "Extraction of acoustic normal mode depth functions using vertical line array data," *The Journal of the Acoustical Society of America*, vol. 111, no. 2, pp. 748–756, 2002.

Chapter 2

Range-coherent matched field processing for low signal-to-noise ratio localization

Range-coherent matched field processing (MFP) coherently combines snapshots to localize a moving, narrowband source. This approach differs from existing MFP approaches that treat each snapshot as having a random phase due to both unknown motion through the medium and imprecise knowledge of the source frequency. Range-coherent MFP requires determination of the source phase acquired between snapshots. With that information, MFP can be applied to the cross-spectrum of snapshots acquired at different times, since relative phase between snapshots is determined by the medium properties, source location, and source velocity. Viewed another way, range-coherent MFP is simply MFP applied to a passive synthetic aperture formed from a moving source. The synthetic aperture geometry depends on source velocity, which is included in the MFP search space. Range-coherent MFP produces robust velocity estimates at low signal-to-noise ratio (SNR), which permits the use of a longer FFT in pre-processing. The synthetic aperture array gain plus the increased input SNR afforded by the enhanced pre-processing significantly lower the required signal level for successful localization. In data from the SWellEx-96 experiment, range-coherent MFP successfully localizes a source that is too quiet for conventional methods to localize.

2.1 Introduction

Matched field processing (MFP), a generalized beamforming method, correlates data acquired on acoustic arrays with model outputs to localize sources [1]. For each candidate source position, a physical model provides a prediction which is then compared with the data in either a linear (e.g. the Bartlett beamformer) or nonlinear/adaptive (e.g. maximum-likelihood estimator) fashion to estimate the likelihood of the candidate being the true position. MFP remains a topic of largely theoretical interest because it requires a) high input signal-to-noise ratio (SNR) and b) accurate environmental knowledge, both of which are rarely available in situations of practical interest. In this paper, we confront the limitation due to SNR by applying MFP to a passive synthetic aperture formed from a moving source.

Passive synthetic apertures are a method of increasing system gain and resolution [2, 3]. Several works demonstrate the formation of towed-array synthetic apertures for bearing estimation [4, 3, 5], as well as for improving geoacoustic inversion results in shallow water [6, 7]. The goal of this paper is to incorporate synthetic aperture theory directly into an MFP algorithm in a simple and robust way, using the increased gain to lower the SNR regime in which localization efforts can succeed.

A few conclusions may be drawn from a review of passive synthetic aperture work in underwater acoustics. First, coherence times for stable, low-frequency continuous wave (CW) transmissions in the ocean may be quite long [8]. Second, the main limitation in synthetic aperture implementations come from tow-path irregularities [8, 9, 10, 11]. Third, synthetic aperture source localization efforts have been largely limited to processing data from towed horizontal line arrays (HLAs) to obtain bearing estimates [3, 4, 12]. Range-coherent MFP takes advantage of signal coherence at low frequencies to extend synthetic aperture methods to source localization in multipath shallow water environments.

Range-coherent MFP also has much in common with coherent multi-frequency processors [13, 14, 15, 16]. Those processors form a synthetic array (or "supervector") by stacking

measurements made on an array at various *frequencies*. The proposed processor forms a synthetic array by stacking measurements made on an array at various *ranges*. Those processors require knowledge of the source spectrum at each frequency, either a priori or via an estimator. The proposed narrowband processor uses the source phase accumulated between snapshots. In both cases, additional gain comes at the price of additional knowledge or algorithm complexity.

It is worth distinguishing the proposed method from a third category of research: techniques that use source motion to incoherently combine snapshots from different ranges. Explicit target motion compensation (ETMC) [17, 18] addresses the "blurring" effect that source motion has on sample covariance matrix (SCM) formation by using a velocity hypothesis to correct for source motion. A recently developed method reformulates ETMC using waveguide invariant theory [19], although source velocity still appears implicitly. Matched field tracking (MFT) [20, 21], another incoherent approach, sums ambiguity surfaces from separate covariance integration periods according to a velocity hypothesis.

Whereas ETMC improves estimation of the usual SCM, range-coherent MFP estimates a larger SCM whose elements include the cross-spectrum of the synthetic aperture elements. In this sense, it has more in common with coherent passive synthetic aperture methods and frequency-coherent processors than with methods such as ETMC and incoherent tracking algorithms.

During data pre-processing, range-coherent MFP demonstrates an additional advantage over traditional MFP. At low SNR, traditional MFP is constrained to use an FFT bin width that encompasses the unknown Doppler shift. Range-coherent MFP's capability to produce robust velocity estimates at low SNR removes this constraint, permitting a longer FFT in pre-processing and accordingly improved input SNR.

This paper formulates range-coherent MFP on a vertical line array (VLA) and compares its performance to existing MFP approaches. Section 2.2 presents background theory for a moving, narrowband source in shallow water and shows how to form a passive synthetic aperture. Section 2.3 reviews relevant aspects of linear (Bartlett) and white noise constraint (WNC) matched field processors and SCM estimation. Section 2.4 presents simulations and Cramèr-

Rao lower bound (CRLB) calculations that compare range-coherent MFP and traditional MFP performance. Section 2.5 discusses the SWellEx-96 experiment, along with pre-processing and modeling considerations. Section 2.6 presents the results of applying range-coherent MFP to the SWellEx-96 dataset. The range-coherent WNC processor localizes the source using a received multi-tone signal 36 dB quieter than the pilot tones used for localization in previous studies. To the best of the authors' knowledge, this is the first study demonstrating successful localization on any of the low-level tonal sets transmitted during SWellEx-96.

2.2 Constructing a passive synthetic aperture

This section shows how to form a synthetic aperture from a narrowband source moving at constant range-rate and depth through shallow water. The material presented here will be applied to data in Sec. 2.6.

2.2.1 Uniform range-rate source in shallow water

Modal doppler theory provides a description of the distant field radiated by a narrowband source moving at constant depth z_s and uniform range-rate (radial velocity) v (no acceleration) through a range-independent shallow water waveguide [23, 24, 25]. Even though the source is narrowband, the received signal has a finite bandwidth that is determined by v, along with the waveguide properties. Wavenumbers computed for a stationary source will differ from those for a moving source by a factor of the order of v/c where c is the sound speed. Here we assume the source moves slowly enough that use of the stationary source wavenumbers introduces negligible errors.

The position track of the source is given in cylindrical coordinates as

$$(r(t), z(t)) = (r_0 + vt, z_s)\,, \tag{2.1}$$

and it radiates at angular frequency ω_0.

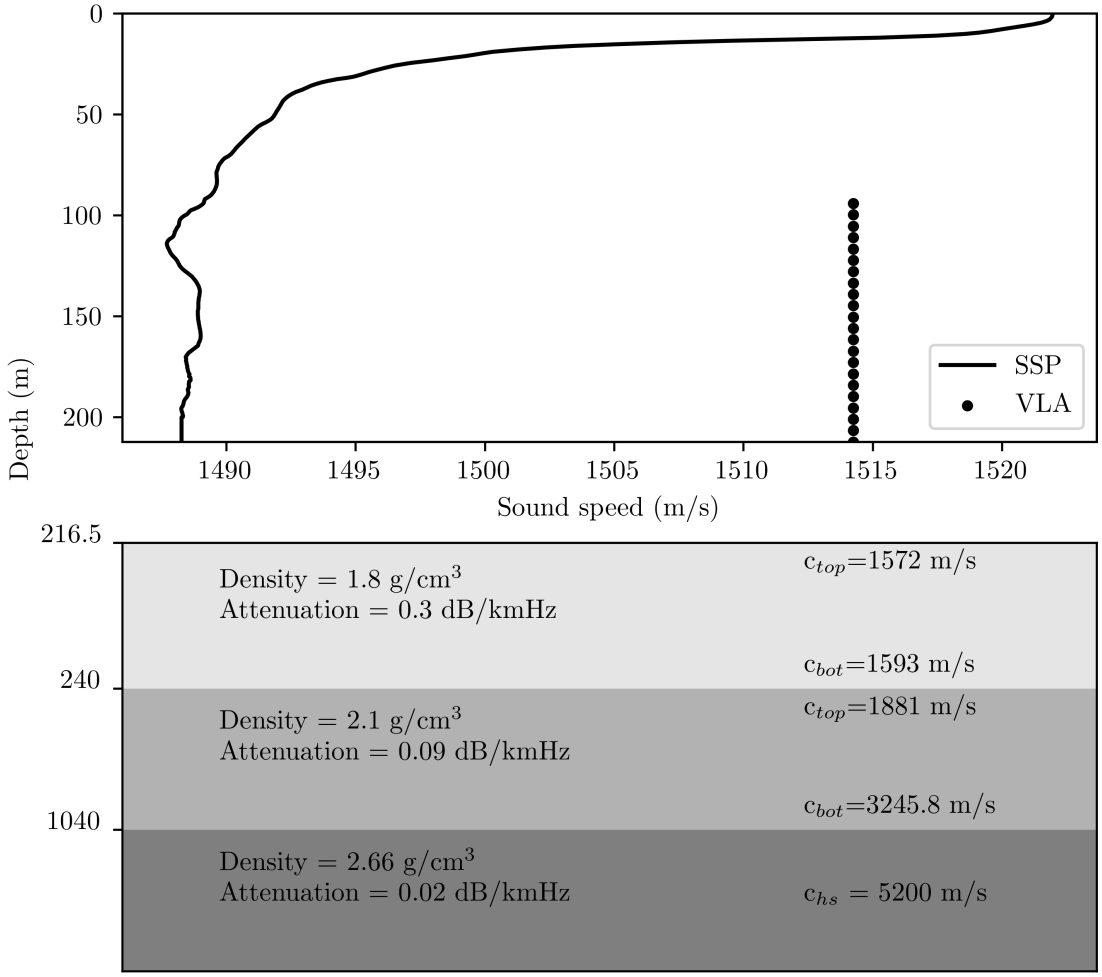


Figure 2.1. Environment used for modeling the SWellEx-96 S5 event. Consists of water column with sound speed profile from a CTD cast during the experiment overlying the geoacoustic model derived in Baxley et al. [22]. Recording system consists of a VLA with 21 elements spaced at 5.625 meters.

For a stationary sensor at a depth z on the z-axis of a cylindrical coordinate system, the received field from the moving source is

$$p(r,z,t) = e^{j\omega_0 t} A(t) \sum_i a_i(z,z_s) e^{-jk_i r_0} e^{-jk_i vt} \ . \tag{2.2}$$

The a_i are the modal amplitudes:

$$a_i = \frac{\Phi_i(z)\Phi_i(z_s)}{\sqrt{k_i}} \ , \tag{2.3}$$

where Φ is the ith modal eigenfunction and k_i is its associated eigenvalue.

$$A(t) = \frac{e^{-j\pi/4 + j\phi_0}}{\sqrt{8\pi(r_0 + vt)}} \tag{2.4}$$

is a slowly varying envelope that contains the source phase term, $e^{j\phi_0}$, and the range spreading term.

It is convenient to introduce the energy-averaged wavenumber $\bar{k} \equiv \sum_i a_i^2 k_i / \sum_i a_i^2$, and the received "carrier frequency" $\omega_d \equiv \omega_0 - \bar{k}v$. Then we can rewrite Eq. 2.2 as

$$p(r,z,t) = e^{j\omega_d t} A(t) \sum_i a_i(z,z_s) e^{-jk_i r_0} e^{-j(k_i - \bar{k})vt} \ . \tag{2.5}$$

The bandwidth of the signal is largely determined by the terms of the sum and is well-approximated by the Doppler spread between the highest and lowest order modes

$$B \approx \frac{1}{2}(\max\{k_i\} - \min\{k_i\})v/2\pi \ . \tag{2.6}$$

The coherence time, τ, is defined as the inverse of B:

$$\tau = \frac{4\pi}{v \times (\max\{k_i\} - \min\{k_i\})} \ . \tag{2.7}$$

As a reference example, for a 100 Hz signal moving through the far-field at $v = 2.5$ m/s in a 100 meter deep Pekeris waveguide with bottom speed $c_b = 1600$ m/s and density $\rho_b = 1800$ kg/m^3, τ is roughly 200 seconds.

If we denote the time at the beginning of a short snapshot window (duration $T \ll \tau$) by t_0, the complex baseband pressure at the Doppler-shifted frequency ω_d is well approximated by

$$\tilde{p}(z,t_0) = We^{j\omega_d t_0} A(t_0) \sum_i a_i(z,z_s) e^{-jk_i r(t_0)} , \tag{2.8}$$

which is the field due to a stationary source at $r(t_0)$ (W, a real number, accounts for the scaling effect of the window on the signal amplitude).

Note the source phase term in front of the modal sum (2.8): $e^{j\omega_d t_0} = e^{j\omega_0 t_0 + j\bar{k}vt_0}$. As pointed out by Bucker in one of the earliest MFP papers [26], this phase term a) naturally rotates with time if the reference frequency isn't precisely matched to ω_0 and b) rotates with (small) changes in source location. This motivates the use of the cross-spectral density matrix in array processing, which only emphasizes the relative phase between array elements. However, with a) precise knowledge of the source frequency ω_0 and b) an estimate of the range-rate v, the relative phase between snapshots can be modeled and a synthetic aperture can be formed.

2.2.2 Forming a passive synthetic aperture

To form a synthetic array, take a set of snapshots at a single array element at depth z for a sequence of initial times $\{t_n\} = \{t_0 + nT_{syn}\}$ to obtain a corresponding sequence of complex values $\{\tilde{d_n}(z)\} = \{\tilde{p}(z,t_n) + \tilde{n}(z,t_n)\}$ that describes the evolution of the pressure field at that sensor. $\tilde{p}$ represents the signal portion of the data, and $\tilde{n}$ is additive noise. From 2.8, note that the signal portion of each snapshot has a source phase component $e^{j\omega_0 t_n}$ related to the source frequency and the start of the snapshot time. Remove the source phase component from each

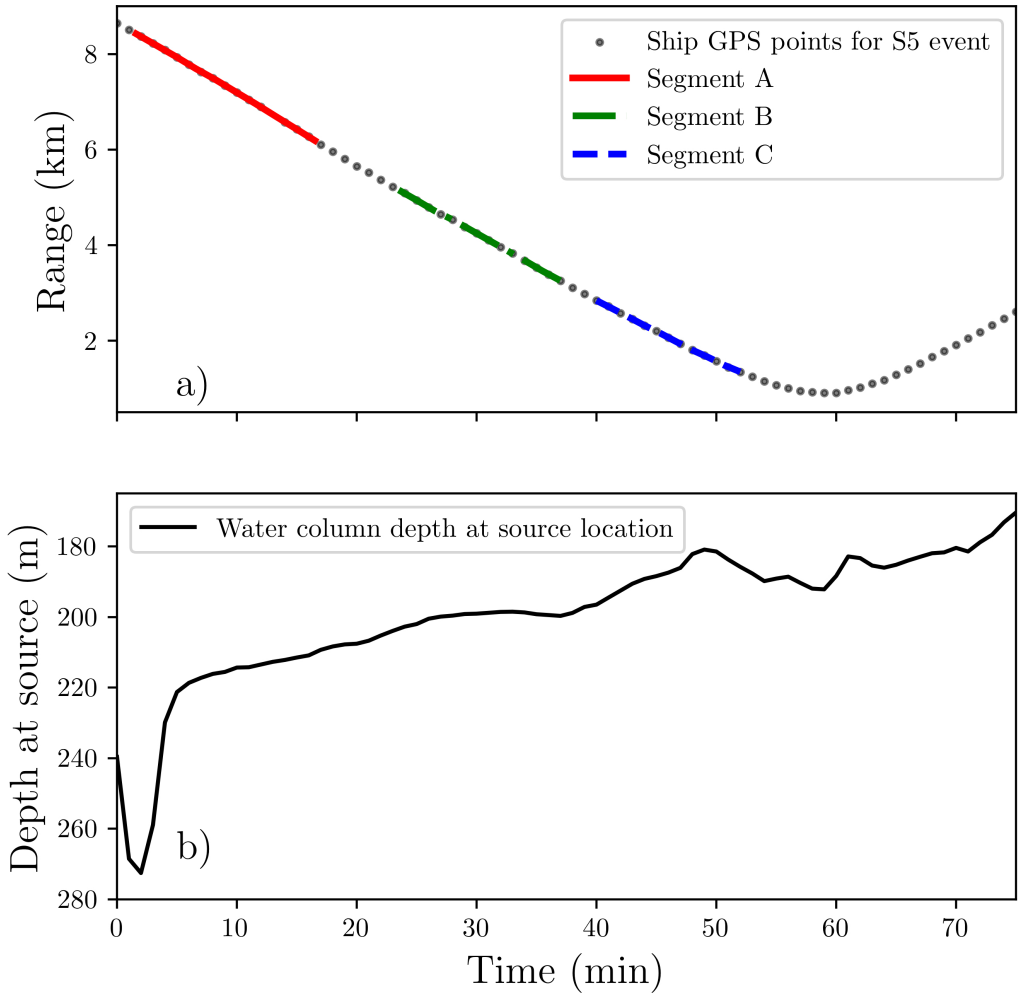


Figure 2.2. (Color online) S5 event. The ship's range from the VLA is shown in a). The three segments of the track that are analyzed, labeled A, B, and C, are superimposed over the GPS track. b) shows the water column depth at the source location over the course of the event.

snapshot by multiplying by its conjugate:

$$\tilde{d}_n^{'}(z) \equiv e^{-j\omega_0 t_n}\tilde{d}_n(z)\ . \tag{2.9}$$

From Eq. (2.5), the modified snapshots will then be described by

$$\tilde{d}_n^{'}(z) = WA(t_n)\sum_i a_i(z,z_s)e^{-jk_i r(t_n)} + \tilde{n}_n'\ . \tag{2.10}$$

The noise has been modified to $\tilde{n}_n' = \tilde{n}_n e^{-j\omega_0 t_n}$.

Now the difference between $\tilde{d}_m^{'}$ and a later measurement $\tilde{d}_n^{'}$ is only dependent on the range traversed by the source from time t_m to time t_n. In other words, the modified $\{\tilde{d}_n^{'}\}$ are identical to the signal received from a stationary source on a stationary HLA with element spacing vT_{syn}.

If the original measurement system is an M-element VLA with sensors at depths $z_0 + m\Delta z$, $m = 0, 1, ..., M-1$, then the moving source generates a planar synthetic array oriented endfire to a source. The array "elements" are indexed by l, with a column-major indexing. With this indexing, the lth array element corresponds to a sensor at the $(l \bmod M)$th sensor depth and the time sample corresponding to the integer part of l/M. Put more simply, the N_{syn} column vectors of modified VLA measurements are stacked into a "supervector" $\tilde{\mathbf{d}}$ of length $M \cdot N_{syn}$, which is then fed into an MFP algorithm.

2.3 Matched Field Processing

As shown in the previous section, data from sequential times may be formed into a single array of measurements, $\tilde{\mathbf{d}}$. A sample covariance matrix (SCM), $\hat{\mathbf{K}} = \frac{1}{L}\sum_i \tilde{\mathbf{d}}_i \tilde{\mathbf{d}}_i^{\dagger}$, can be formed from L samples $\{\tilde{\mathbf{d}}_i\}$ of $\tilde{\mathbf{d}}$. Note that for range-coherent MFP, the SCM is formed from the "supervectors" described in the previous section.

MFP algorithms compute a weighting vector $\mathbf{w}(\boldsymbol{\theta})$ for each potential vector of source

parameters $\boldsymbol{\theta}$ and combine them with $\hat{\mathbf{K}}$ to produce an ambiguity surface

$$P(\hat{\mathbf{K}}, \boldsymbol{\theta}) = \mathbf{w}^{\dagger}(\boldsymbol{\theta})\hat{\mathbf{K}}\mathbf{w}(\boldsymbol{\theta}) \, . \tag{2.11}$$

In general, the weighting vector $\mathbf{w}(\boldsymbol{\theta})$ is some function of the steering vector $\mathbf{s}(\boldsymbol{\theta})$, computed from an acoustic model for each search $\boldsymbol{\theta}$. In the absence of any mismatch, the steering vector $\mathbf{s}(\boldsymbol{\theta})$ will match the source signal $\tilde{\mathbf{p}}(\boldsymbol{\theta}_T)$ received on the array when $\boldsymbol{\theta}$ matches $\boldsymbol{\theta}_T$, the true source parameter value. Steering vectors $\mathbf{s}$ are normalized to the number of array elements M: $\sqrt{\mathbf{s}^{\dagger}\mathbf{s}} = M$.

In a range-independent environment, MFP uses the source position in cylindrical coordinates as the source parameter vector: $\boldsymbol{\theta} = (r \;\; z)^T$. Range-coherent MFP for a constant depth, uniform range-rate source includes source range-rate v in the parameter vector: $\boldsymbol{\theta} = (r \;\; z \;\; v)^T$.

2.3.1 Bartlett processor

The Bartlett (also called "linear" or "conventional") processor [26] uses a scaled steering vector as the weighting vector, $\mathbf{w}(\boldsymbol{\theta}) = \mathbf{s}(\boldsymbol{\theta})/M$, so the Bartlett ambiguity surface is simply

$$P_{BART}(\hat{\mathbf{K}}, \boldsymbol{\theta}) = \frac{1}{M^2}\mathbf{s}^{\dagger}(\boldsymbol{\theta})\hat{\mathbf{K}}\mathbf{s}(\boldsymbol{\theta}) \, . \tag{2.12}$$

The Bartlett processor is cheap to compute and robust to model mismatch, but performs poorly in the presence of interferers and significant sidelobe ambiguity structure.

2.3.2 White noise constraint processor

The white noise constraint (WNC) processor [27, 28] strives to balance the interference rejection and sidelobe suppression of a minimum-variance (MV) beamformer (also called Capon's method and the maximum-likelihood method) with the robustness of the Bartlett processor. To accomplish this, it seeks an adaptive weighting vector $\mathbf{w}_a(\theta)$ that minimizes total output power, maintains a borehole constraint, and satisfies an inequality constraint on the white

noise gain (defined below):

$$\begin{aligned} &\min \mathbf{w}_a^\dagger(\boldsymbol{\theta})\hat{\mathbf{K}}\mathbf{w}_a(\boldsymbol{\theta}) \\ &\text{subject to } \mathbf{w}_a(\boldsymbol{\theta})^\dagger\mathbf{s}(\boldsymbol{\theta}) = 1 \\ &\left(\mathbf{w}_a(\boldsymbol{\theta})^\dagger\mathbf{w}_a(\boldsymbol{\theta})\right)^{-1} \geq \delta^2 \,. \end{aligned} \tag{2.13}$$

The white noise gain is defined as

$$G_W = \frac{|\mathbf{w}_a^\dagger\mathbf{s}|^2}{\mathbf{w}_a^\dagger\mathbf{w}_a} \,. \tag{2.14}$$

Its inverse $S_W = 1/G_W$ is a measure of sensitivity to errors between the actual signal $\tilde{\mathbf{p}}$ and the steering vector $\mathbf{s}$ [27]. From the Cauchy-Schwarz inequality and the normalization of $\mathbf{s}$,

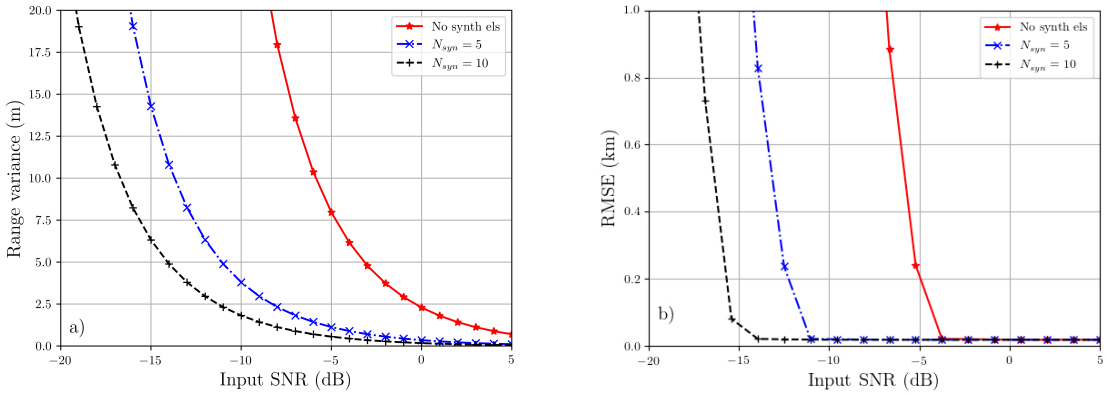
$$G_W \leq M \tag{2.15}$$

with equality when $\mathbf{w} = \alpha\mathbf{s}$ for any $\alpha \in \mathbb{R}$. Hence, the Bartlett weight vector $\mathbf{w} = \mathbf{s}/M$ maximizes white noise gain and minimizes sensitivity. Removing the white noise constraint entirely (or, equivalently, setting $\delta^2 = 0$) results in the MV processor, which maximizes sensitivity. The user's choice of δ^2 controls the tradeoff between robustness and sidelobe suppression/interference rejection.

Solving the constrained optimization problem posed in Eq. 2.13 yields the corresponding WNC weighting vector

$$\mathbf{w}_a(\hat{\mathbf{K}}, \boldsymbol{\theta}, \delta^2) = \frac{(\hat{\mathbf{K}} + \varepsilon\mathbf{I})^{-1}\mathbf{s}(\boldsymbol{\theta})}{\mathbf{s}(\boldsymbol{\theta})(\hat{\mathbf{K}} + \varepsilon\mathbf{I})^{-1}\mathbf{s}(\boldsymbol{\theta})} \,, \tag{2.16}$$

where $\mathbf{I}$ is the identity matrix and $\varepsilon(\hat{\mathbf{K}}, \boldsymbol{\theta}, \delta^2)$ is a regularization parameter. An eigenvector decomposition of $\hat{\mathbf{K}}$ permits efficient calculation of 2.16 (see 10.3.4 of Jensen et al. [24]) to find, for each search location, an ε that satisfies the white noise gain constraint.



a) b)

Figure 2.3. (Color online) Performance comparison of a simple VLA, a five element planar array, and a ten element planar array. a) shows the CRLB for range estimators as a function of input SNR for the three arrays. In b) we show the RMSE for Bartlett range estimates from 1000 simulation runs for a range of input SNR. Note the different units between a) and b). For each simulation, four snapshots are used to form the SCM (four realizations of noise + signal).

Often δ^2 is measured as some number of dB down from the maximum white noise gain, so the expression "the white noise gain is constrained to X dB", translates to $10\log_{10}(\delta^2/M) = X$. Thus, 0 dB WNC processor is a Bartlett processor, and $-\infty$ dB WNC processor is the MV processor.

The output of the WNC processor is, from Eq. 2.11,

$$P_{WNC}(\hat{\mathbf{K}}, \boldsymbol{\theta}, \delta^2) = \mathbf{w}_a(\hat{\mathbf{K}}, \boldsymbol{\theta}, \delta^2)\hat{\mathbf{K}}\mathbf{w}_a(\hat{\mathbf{K}}, \boldsymbol{\theta}, \delta^2)\,. \tag{2.17}$$

2.3.3 Snapshot deficiency

Adaptive methods presuppose an accurate estimate of the SCM $\hat{\mathbf{K}}$. Snapshot deficiency occurs when a limited number of snapshots available to form $\hat{\mathbf{K}}$ results in a poor, possibly singular estimate of the SCM. For a moving source, the number of available snapshots is limited by the amount of time spent by the source in a range cell (the "beamwidth" of a Bartlett range estimator [18]). The synthetic array SCM dimension is a factor of N_{synth} larger than the dimension of the

simple VLA SCM, range-coherent MFP requires N_{synth} times more snapshots than "traditional" MFP. As a result, snapshot deficiency is likely to occur for all but the slowest moving sources.

There are a number of methods for mitigating snapshot deficiency. The simplest method is diagonal loading [29]. This strategy will fail to null a loud interferer creating colored noise along the array, but solves the singularity issue of a severely deficient $\hat{\mathbf{K}}$. To deal with a loud interferer or colored environmental noise, a full-rank sample covariance matrix can be formed by averaging over nearby frequency bins [30] or by averaging sequences of shorter snapshots [31]. Alternatively, subspace methods can be applied to the snapshot-deficient SCM [32]. The SWellEx-96 event analyzed in this paper has no loud interferers, so the simplest approach is taken: the SCMs used are diagonally loaded

$$\hat{\mathbf{K}}' \equiv \hat{\mathbf{K}} + \varepsilon \mathbf{I} \,, \tag{2.18}$$

with $\varepsilon = 10^{-20}$ (around eight orders of magnitude lower than the noise power on a single phone). However, in a more complicated noise environment, one of the above strategies would be necessary.

2.3.4 Multi-tone MFP

When a source transmits at multiple frequencies, a single ambiguity surface may be obtained by performing a geometric mean of the ambiguity surfaces from processing each transmitted frequency [1]. The result is a single ambiguity surface:

$$P = \frac{1}{N_f} \sum_{k=1}^{N_f} 10 \log_{10} P(f_k) \,. \tag{2.19}$$

More optimal approaches that compensate for the frequency-dependence of background noise exist. However, the simplicity and effectiveness of the naive summation are deemed sufficient for the present work. Whenever MFP results are presented in this paper, they will be formed

from the incoherent frequency sum in Eq. 2.19.

2.4 Performance of range-coherent MFP

A synthetic array demonstrates two advantages over the simple array from which it is formed: array gain and improved resolution. Both of these advantages follow from the near equivalence between the passive synthetic array formed from a moving source and a VLA, and a planar endfire array recording a stationary source. The only difference between these two measurement scenarios is that a planar endfire array recording a stationary source doesn't suffer from Doppler broadening and limited covariance integration time.

The Cramèr-Rao lower bound (CRLB) provides a tool to quantify both the resolution and threshold effect for both simple VLAs and synthetic planar arrays. The CRLB provides a lower bound on the variance of any statistical estimator of a given parameter [33, 1]. The CRLB shows that although there is an increase in resolution from adding synthetic array elements, the main benefit of the synthetic aperture method is the lowered required input SNR.

We computed the CRLB for the case of a source transmitting at 100 Hz in an ideal waveguide (pressure release surface and rigid bottom) with sound speed $c = 1500$ m/s and depth $D = 216$ meters. The parameter vector to be estimated is $\mathbf{a}_T = (r_s, z_s)$, the source range and depth.

The data model consists of the signal from the source with uniformly distributed random source phase $e^{j\phi_0}\mathbf{p}(r_s, z_s)$ plus additive, white complex Gaussian noise received on the VLA shown in Fig. 2.1. The simulated data follow a zero-mean, complex Gaussian distribution with covariance matrix $\mathbf{K}(r_s, z_s)$ determined by the waveguide properties, source location and noise power σ_W^2 :

$$\mathbf{K}(r_s, z_s) = \sigma_W^2 \mathbf{I} + \mathbf{p}(r_s, z_s)\mathbf{p}^\dagger(r_s, z_s) \; . \tag{2.20}$$

Input power is defined to be $\sigma_S^2 = \frac{1}{N}\text{trace}(\mathbf{p}\mathbf{p}^\dagger)$, and input SNR in dB is $10\log_{10}(\sigma_S^2/\sigma_W^2)$.

From Kay [33] Eq. 15.52, the Fisher information matrix (FIM) is

$$\mathbf{J}_{ij} = \text{trace}\left(\mathbf{K}^{-1}\frac{\partial \mathbf{K}}{\partial \mathbf{a}_{\mathbf{T}i}}\mathbf{K}^{-1}\frac{\partial \mathbf{K}}{\partial \mathbf{a}_{\mathbf{T}j}}\right) . \tag{2.21}$$

The resulting bounds on the variance of any estimator, $\hat{\mathbf{a}}_i$, of $\mathbf{a}_{Ti}$ follow from the diagonal elements of the inverse of the FIM:

$$E((\hat{\mathbf{a}}_i - \mathbf{a}_{Ti})^2) \geq [\mathbf{J}^{-1}]_{ii} . \tag{2.22}$$

The results from applying the CRLB to three array geometries are shown in Fig. 2.3 a). As expected, the array gain of additional VLA staves lowers the input SNR threshold required for accurate localization.

CRLB calculations for SNR thresholds are also consistent with Monte Carlo simulations of the (inefficient) Bartlett processor applied to a snapshot-deficient SCM. For each of the three array geometries and input SNR values, we generated 1000 realizations of the SCM. The SCM for each realization was formed from four independent snapshots (signal plus complex white Gaussian noise). Applying the Bartlett range estimator to each realization generates an ensemble of 1000 range estimates. From these we form an estimate of the root mean-squared error (RMSE). The RMSE estimates are shown in the Fig. 2.3, and the SNR thresholds at which they degrade match the CRLB-derived thresholds.

The CRLB and simulations show that range-coherent MFP can localize a source in SNR regimes where traditional MFP cannot.

2.5 SWellEx-96 Experiment

Here we provide a brief orientation to the SWellEx-96 data set [34] (and more specifically the S5 Event), as well as the pre-processing and modeling required to perform MFP and range-coherent MFP.

2.5.1 S5 Event

SWellEx-96 refers to a well-known shallow water MFP experiment that was conducted off the coast of San Diego in 1996. The S5 event consists of a 75 minute source tow at roughly 2.5 m/s under a smooth sea state and over variable bathymetry. The S5 event has a number of features which make it a good candidate for demonstrating our proposed method. It contains a slowly moving source transmitting CW signals, so the model in Sec. 2.2 matches the scenario. A geoacoustic model that optimizes the match between replicas and data was developed by Bachman [35] et al. and later refined by Baxley et al. [22], so accurate replica calculation is not an issue. It also contains low SNR data that have never been used successfully for localization.

The environment is shown in Fig. 2.1, and the range of the ship, which towed the two acoustic sources, for the S5 event is shown in Fig. 2.2. The water column depth at the ship's location as a function of experiment time is also shown. The two sources consist of a shallow source at 9 m depth and a deep source at around 55m depth.

The deep source simultaneously transmitted 5 sets of tonals, which are labeled here as TSL ('L' for 'loud'), TSQ1, TSQ2, TSQ3, and TSQ4 ('Q' for 'quiet'). As shown in Table 2.1, TSL is 26 dB above TSQ1, and each subsequent quiet tonal set is 4 dB below the previous set. Each tonal set contains thirteen tones spaced between 49 and 400 Hz, and the tones of each set differ from those of the adjacent sets by 3 Hz.

The data are received on a VLA with 64 elements spanning the water column from 100 to 200 meters with a sampling rate of 1500 Hz. Only every third element from the array was included in the processing, resulting in a 21 element VLA with 5.625 meter spacing.

2.5.2 FFT Length, SNR and SCM formation

FFT length T simultaneously sets the input SNR to the matched field processor and limits the number of snapshots available for forming the SCM. When processing data from a narrowband source moving at unknown velocity, a wide bin width is required to account for

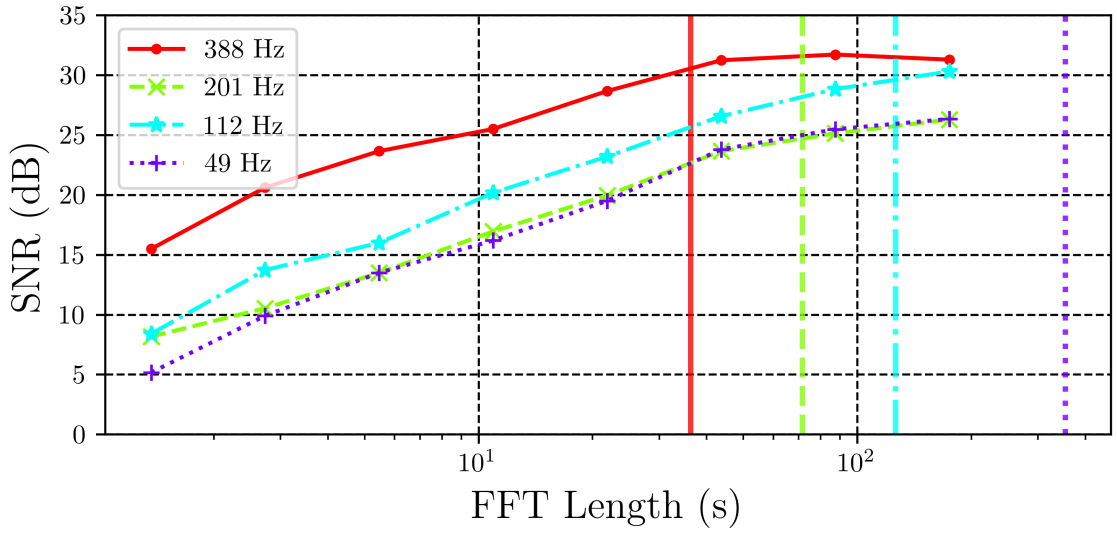


Figure 2.4. (Color online) SNR estimates as a function of FFT length shown for four tones from the loud tonal set. The dash-dot lines show the theoretical coherence times τ calculated from Eq. 2.7 for the SWellEx environment with $v = -2.5$ m/s. For a coherent source, doubling the FFT length should increase SNR should $10\log_{10}2$ dB. Both 49 and 112 Hz SNR estimates follow this trend quite accurately. The SNR for the higher frequencies follow a linear trend until they exceed their coherence time, at which point they taper off.

the unknown Doppler shift. If the possible velocity is constrained to 2.5 m/s and the highest frequency considered is 400 Hz, then from the maximum possible Doppler shift will be 0.66 Hz (as calculated from $\bar{k}v/2\pi$). Calculations (not shown for brevity) show that for a Hamming window longer than one second, the leakage loss for the 400 Hz band will exceed 3 dB. The lack of source velocity knowledge limits FFT length, and, hence, input SNR. Since range-coherent MFP includes velocity as a search parameter, the leakage constraint derived above no longer applies. For each search range-rate the Doppler shift can be computed and the appropriate bin can be selected for input to the matched field processor. Of course, one could add search range-rate and implement this bin-selection algorithm in traditional MFP. However, we found that simply using search range-rate to estimate the proper bins for traditional MFP completely fails at low SNR. Therefore, this strategy is only available to the range-coherent method which can accurately estimate range-rate at low SNR.

SNR estimations performed for TSL, shown in Fig. 2.4, demonstrate the received signal

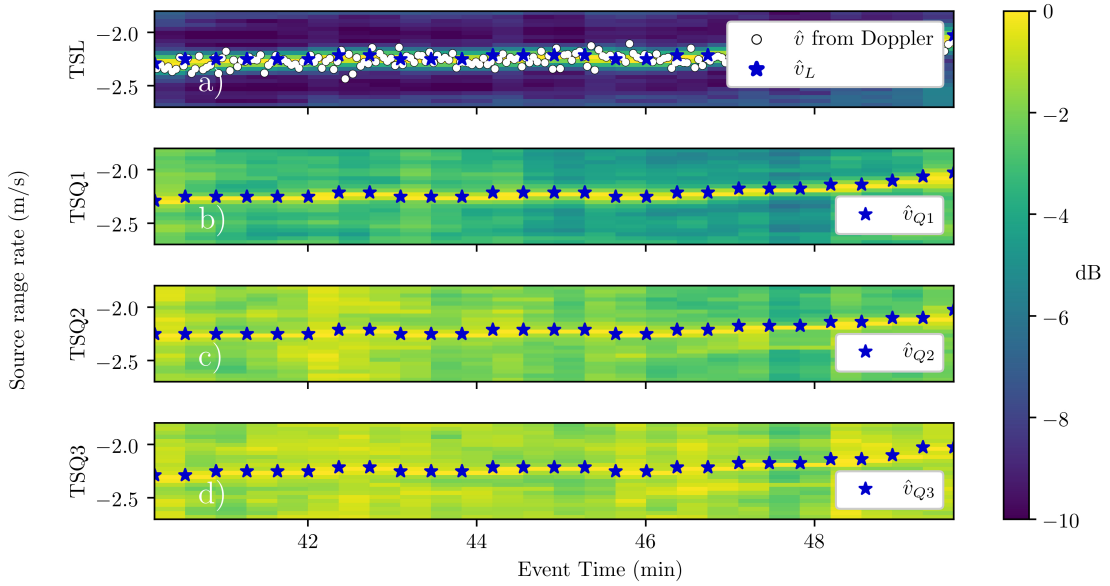


Figure 2.5. (Color online) Multi-tone Bartlett range-rate/time ambiguity surfaces for Segment C for a five synthetic element array ($N_{syn} = 5$). The results using TSL are shown in a), TSQ1 in b), TSQ2 in c), and TSQ3 in d). In each figure, a blue star marks the location of the maximum of the ambiguity surface for each time step. In a), range-rate estimates from Doppler shifts of TSL are superimposed. In b) through d), the SNR is too low to estimate range-rate from Doppler, so only v_{Qi}, estimates from range-coherent MFP applied to TSQi, are shown.

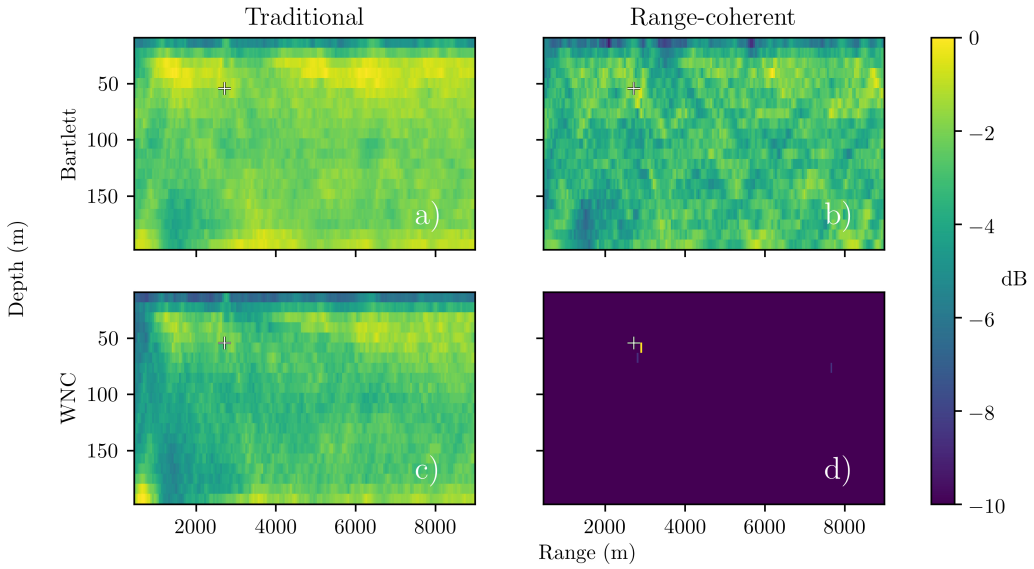


Figure 2.6. (Color online) Comparison of benchmark and range-coherent ambiguity surfaces for a sample interval for TSQ3. a) and c) show the Bartlett and WNC results on a simple VLA. b) and d) show the Bartlett and WNC range-coherent MFP for a five synthetic element array ($N_{syn} = 5$). The cross shows the ship range from GPS and nominal source depth of 55 meters.

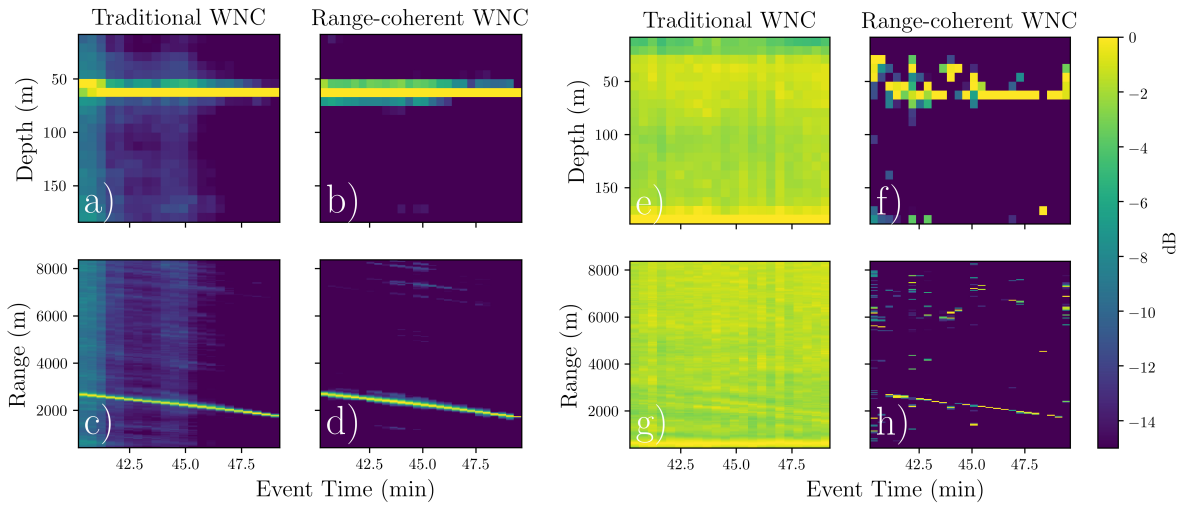


Figure 2.7. (Color online) Comparing "traditional" WNC to range-coherent WNC. Tracks are shown for Segment C for TSL (left panel) and TSQ3 (right panel) tonal sets. Top row a), b), e), and f) show depth/time ambiguity surfaces. Bottom row c), d), g) and h) show range/time ambiguity surfaces. a), c) and e), g) show "traditional" WNC. b), d) and f) and h) show range-coherent WNC with five synthetic elements. Both methods localize the source over the whole segment on TSL. Only range-coherent WNC localizes the source using TSQ3.

coherence properties and inform FFT length selection when the Doppler shift is known. For a coherent signal in white noise, doubling FFT length T should increase SNR by around 3 dB until T approaches the coherence time τ (the time at which the source leaves its range cell). Once T reaches τ, modal leakage out of the FFT bin cancels out the coherent gain, and the SNR reaches a plateau.

With these considerations in mind, the data were processed in two ways: one with a 2048 point FFT ($T = 1.4$ s) with 50% overlap, and another with a 16384 point FFT ($T = 10.9$ s) with 50% overlap. Table 2.1 shows the resulting input SNR for these two settings for each tonal set. Bins corresponding to the nominal source frequency are selected from the 2048-point FFT to feed into traditional, "velocity-ignorant" matched-field processors. For range-coherent processing, the bin is selected by estimating the Doppler shift from the known source frequency f_0 and the given search range-rate v. For example, the signal from source frequency f_0 will be associated with the bin closest to $f_0(1 - v/c)$ for each search range-rate v (c is average modal group velocity). Range-coherent MFP's capacity to estimate source range-rate therefore adds

Table 2.1. Signal level and maximum input SNR, averaged over tones in each set, for processed segments of S5 event for two different FFT lengths. The maximum SNR occurs at the closest range analyzed, around 1.3 km. As the source moves away, SNR drops approximately by the cylindrical spreading law $10\log_{10}(r/r_{min})$. For example, 8 km range, the SNR will be 8 dB lower than the max value shown here.

Tonal set	Source Level (dB ref. 1μPa)	Max. SNR (dB) $N_{FFT} = 2048$ (dB)	Max. SNR (dB) $N_{FFT} = 16384$ (dB)
TSL	158	11	21.5
TSQ1	132	-15	-4.5
TSQ2	128	-19	-8.5
TSQ3	124	-23	-12.5
TSQ4	120	-27	-16.5

gain at the outset by simply increasing the input SNR.

The covariance integration time is set to the time spent in a range cell of a 2.5 m/s source transmitting at the maximum frequency of 400 Hz: around 25 seconds [36]. For the 2048 point FFT sequence, 36 snapshots are used to form the SCM. For the 16384 point FFT sequence, four snapshots are used to form the SCM. The synthetic element spacing T_{syn} is chosen to equal the covariance integration time. The result is two sequences of covariance matrices for each frequency, $\{\hat{\mathbf{K}}_i(f)\}_{2048}$ formed from simple snapshots, and $\{\hat{\mathbf{K}}_i(f,N_{syn})\}_{16384}$ formed from "supervectors" of N_{syn} stacked snapshots.

An explicit example for computing $\hat{\mathbf{K}}_0(f,3)$ follows. Let $\{\tilde{\mathbf{d}}_i(f)\}$, denote the sequence of snapshots that results from applying the 16384 point FFT to the VLA and selecting the bin corresponding to frequency f. First form a sequence of supervectors $\{\tilde{\mathbf{d}}_i^{(3)}(f)\}$:

$$\tilde{\mathbf{d}}_i^{(3)}(f) = [\tilde{\mathbf{d}}_i^T(f)\ \ \tilde{\mathbf{d}}_{i+4}^T(f)\ \ \tilde{\mathbf{d}}_{i+8}^T(f)]^T\ . \tag{2.23}$$

Note that the supervector is formed by stacking every fourth snapshot. Then,

$$\hat{\mathbf{K}}_0(f,3) = \frac{1}{4}\sum_{i=0}^{3}\tilde{\mathbf{d}}_i^{(3)}(f)\tilde{\mathbf{d}}_i^{(3)}(f)^\dagger\ . \tag{2.24}$$

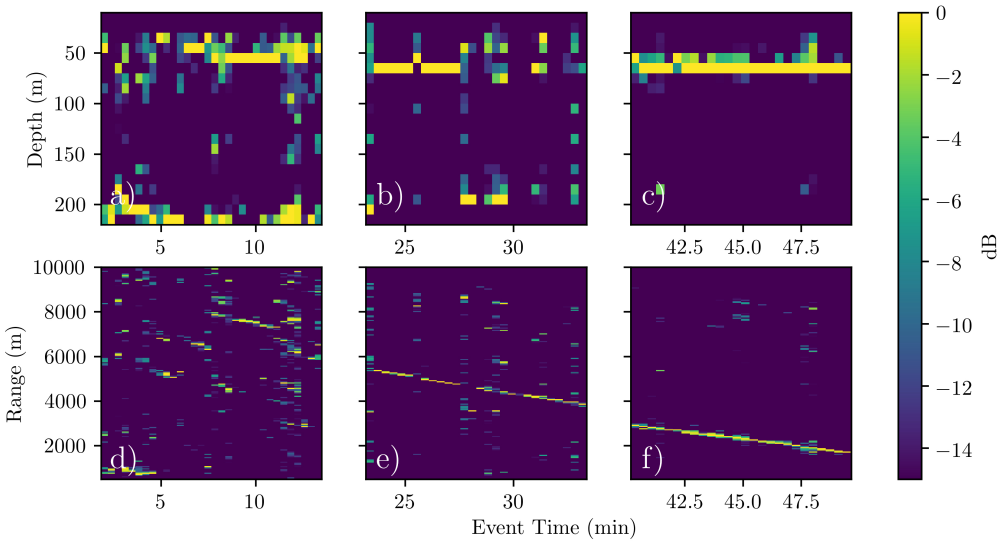


Figure 2.8. (Color online) Range-coherent WNC depth/time and range/time ambiguity surfaces for TSQ1 for all three data segments. Segment C, shown in c) and f), uses a five-element synthetic array. Segments A and B, shown in a) and d) and b) and e) use an eight-element synthetic array. Further from the closest-point of approach (CPA), constant velocity translates to a nearly uniform range-rate. This permits formation of a longer synthetic aperture in Segment A and B than in Segment C.

2.5.3 Modeling

The KRAKENC normal mode program [37] is used to compute the steering vectors $\mathbf{s}(r,z)$ for the simple VLA. The steering vectors account for the known array tilt of 1 degree[34]. For a synthetic array with N_{syn} synthetic elements, the steering "supervector" $\mathbf{s}_{syn}$ is formed by stacking a collection of VLA replicas that depend on the search parameters r, z, v, and the time T_{syn} between snapshots stacked into the data supervector:

$$\mathbf{s}_{syn}(r,z,v) = \begin{bmatrix} \mathbf{s}(r,z) \\ \mathbf{s}(r+vT_{syn},z) \\ \vdots \\ \mathbf{s}(r+(N_{syn}-1)vT_{syn},z) \end{bmatrix}. \tag{2.25}$$

As shown in Fig. 2.1, the bathymetry varies by almost 100 meters over S5. Range-dependent modeling can account for this. However, the results of D'Spain et al. on mirages in shallow water [38] show that a range independent model will still produce a significant correlation maximum, but with the location (range and depth) of maximum correlation shifted by an amount related to the error in assumed source depth. By modeling the bottom as a simple linear segment between the receiver and source, the range and depth of the correlation maximum will be distorted from the true source range and depth by a factor of $\gamma \equiv D/d_s$ where d_s is the true depth at the source location and D is the depth at the receiver. Therefore the receiver depth is used as the depth of a range-independent normal mode model, and the factor γ is computed from the known bathymetry and used to correct the ambiguity surfaces (this same method is used by Booth et al. [39]).

2.6 MFP Results

Traditional MFP and range-coherent MFP were applied to three segments, labeled A, B, and C of the approaching track (see Fig. 2.2). We restrict our analysis to these segments to avoid

portions of the track during which the deep source was turned off. The ship approaches the VLA over the three segments, so segment A is the most distant segment and segment C is the closest segment.

For each tonal set, the ambiguity surfaces from each tone are summed incoherently as in Eq. 2.19. None of the implementations managed to consistently localize TSQ4, so we focus on TSL, TSQ1, TSQ2, and TSQ3 (see Table 2.1 for a review of the source level and SNR of the tonal sets).

To provide an example of "traditional" MFP, Bartlett and WNC processors were applied to the SCM sequences $\{K_i(f)\}_{2048}$. The result for each tonal set is a three-dimensional ambiguity surface over range, depth, and time for each of the two processors. After varying δ^2 for the WNC, we settled on using white noise gain of -2 dB for the "traditional" method.

Then, Bartlett and WNC processors were applied to the range-coherent SCM sequences $\{K_i(f,N_{syn})\}_{16384}$ for various values of N_{syn} to investigate range-coherent MFP performance. For the range-coherent results, a four-dimensional Bartlett ambiguity surface over source parameter vector $\boldsymbol{\theta} = (r \;\; z \;\; v)^T$ and time t_i was formed for each tonal set

$$P_B(\boldsymbol{\theta},t_i) = \frac{1}{13}\sum_{k=1}^{13} 10\log_{10} P_{BART}\left(\hat{\mathbf{K}}_i(f_k',N_{syn}),\boldsymbol{\theta}\right) \;, \tag{2.26}$$

where $f_k' = f_k(1-v/c)$ is the received frequency of the kth source tone f_k in the tonal set. The range-coherent Bartlett estimate is $\hat{\boldsymbol{\theta}}(t) = \arg\max_{\boldsymbol{\theta}} P_B(\boldsymbol{\theta},t)$. The range-rate/time ambiguity surfaces for the different tonal sets are shown in Fig. 2.5.

Then, for various choices of the white noise gain δ^2, a three-dimensional WNC surface is formed over $\boldsymbol{\theta} = (r \;\; z \;\; \hat{v}(t_i))^T$ and time t_i

$$P_W(\boldsymbol{\theta},t_i) = \frac{1}{13}\sum_{k=1}^{13} 10\log_{10} P_{WNC}\left(\hat{\mathbf{K}}_i(f_k',N_{syn}),\boldsymbol{\theta}\right) \;. \tag{2.27}$$

The search range-rate is constrained to the Bartlett estimate $\hat{v}$. All presented results for range-

coherent processing use a white noise gain of -0.5 dB, which seemed to work best.

The number of synthetic elements N_{syn} that produced the best results depended on the experimental segment. In Segment C, the proximity to the CPA created a non-uniform range-rate that limited the synthetic aperture length to around 300 meters ($N_{syn} = 5$). In Segment A and Segment B, the range-rate is more uniform. As such, range-coherent MFP worked on synthetic aperture lengths up to 600 meters ($N_{syn} = 10$).

A representative comparison of ambiguity surfaces for the four implementations is shown in Fig. 2.6 for TSQ3. The signal level is too low for the simple Bartlett or WNC to localize the source. In contrast, range-coherent WNC accurately estimates the source range and depth while also providing an accurate range-rate estimate.

To present comparisons over time graphically, we form two-dimensional ambiguity surfaces by holding time and a single parameter fixed and take the max over the remaining parameters. For example, to form the range-rate/time ambiguity surface shown in Fig. 2.5, take the max over range and depth for each range-rate and time value. Doppler shift range-rate estimates from TSL agree with the range-coherent Bartlett range-rate estimates for each tonal set. The quiet tonal sets are too quiet for Doppler shift estimation, so the only estimates of range-rate come from range-coherent MFP.

We also compare range/time and depth/time ambiguity surfaces to assess the relative performance of the four implementations over the segments of the experiment. A comparison between range/time and depth/time ambiguity surfaces over Segment C for traditional WNC and the range-coherent WNC with five synthetic elements are shown for TSL and TSQ3 in Fig. 2.7. Both processors easily localize the source over the full segment using TSL. For TSQ3, traditional WNC completely breaks down, but range-coherent WNC produces a nice track. In fact, traditional WNC fails on all three segments using any of the quiet tonal sets. We show the range-coherent result for TSQ3 because it was the lowest set that produced a clear source track.

Range-coherent WNC range/time and depth/time ambiguity surfaces for all three segments for TSQ1 are shown in Fig. 2.8. For Segments A and B, the synthetic array consists of

eight elements, whereas for Segment C the synthetic array had only five elements. The ship track in Fig. 2.2 shows the slight curvature of Segment C. The curvature manifests as a non-uniform range-rate. The assumption of uniform range-rate limits the synthetic array length to intervals over which the acceleration creates minimal mismatch. The uniform range-rate at greater range permitted the longer synthetic aperture used in Segments A and B.

Range-coherent MFP demonstrates superior performance to the benchmark MFP implementations on the quiet tonal sets. Using a 500 meter synthetic aperture, range-coherent WNC localizes TSQ1 out to a range of 8 km and TSQ2 out to a range of 5 km.

2.7 Conclusion

Range-coherent MFP can localize quieter sources than traditional methods. It does this by combining MFP with a passive synthetic aperture that provides both spatial array gain and gain from increased integration time. In the presented application to the SWellEx-96 data, these two features result in a total gain of roughly 16 dB over previous MFP results.

The assumption of uniform source motion limits the maximum synthetic aperture length. Future work could incorporate a higher order motion model, for example by applying the array shape matched field inversion of Hodgkiss et al. [40] to the synthetic array, or could minimize the penalty for mismatch by incorporating an array geometry robustness constraint like that used in Byun et al. [41]. Fusing velocity, range, and depth measurements produced by range-coherent MFP in a tracking algorithm could further improve results.

Acknowledgments

This work was funded by the TFO program of the United States Office of Naval Research. Ryan Saenger, Ludovic Tenorio and Paul Akins provided helpful comments on an early draft.

This chapter, in full, is a reprint of the material as it appears in Akins, F. H. and Kuperman, W. A. (2021). Range-coherent matched field processing for low signal-to-noise ratio localization.

The Journal of the Acoustical Society of America, 150(1), 270-280. The dissertation author was the primary investigator and author of this paper. The coauthor listed in this publication directed and supervised the research.

Bibliography

[1] A. B. Baggeroer, W. A. Kuperman, and H. Schmidt, "Matched field processing: Source localization in correlated noise as an optimum parameter estimation problem," *The Journal of the Acoustical Society of America*, vol. 83, no. 2, pp. 571–587, 1988.

[2] S. W. Autrey, "Passive synthetic arrays," *The Journal of the Acoustical Society of America*, vol. 84, no. 2, pp. 592–598, 1988.

[3] N. Yen and W. Carey, "Application of synthetic-aperture processing to towed-array data," *The Journal of the Acoustical Society of America*, vol. 86, no. 2, pp. 754–765, 1989.

[4] S. Stergiopoulos and E. J. Sullivan, "Extended towed array processing by an overlap correlator," *The Journal of the Acoustical Society of America*, vol. 86, no. 1, pp. 158–171, 1989.

[5] E. J. Sullivan, J. D. Holmes, W. M. Carey, and J. F. Lynch, "Broadband passive synthetic aperture: Experimental results," *The Journal of the Acoustical Society of America*, vol. 120, no. 4, pp. EL49–EL54, 2006.

[6] B. A. Tan, P. Gerstoft, C. Yardim, and W. S. Hodgkiss, "Broadband synthetic aperture geoacoustic inversion," *The Journal of the Acoustical Society of America*, vol. 134, no. 1, pp. 312–322, 2013.

[7] J. Unpingco, W. A. Kuperman, W. S. Hodgkiss, and R. Hecht-Nielsen, "Single sensor source tracking and environmental inversion," *The Journal of the Acoustical Society of America*, vol. 106, no. 3, pp. 1316–1329, 1999.

[8] R. E. Williams, "Creating an acoustic synthetic aperture in the ocean," *The Journal of the Acoustical Society of America*, vol. 60, no. 1, pp. 60–73, 1976.

[9] T. C. Yang, "Source depth estimation based on synthetic aperture beamfoming for a moving source," *The Journal of the Acoustical Society of America*, vol. 138, no. 3, pp. 1678–1686, 2015.

[10] T. C. Yang, "Data-based matched-mode source localization for a moving source," *The Journal of the Acoustical Society of America*, vol. 135, no. 3, pp. 1218–1230, 2014.

[11] G. V. Frisk, K. M. Becker, and J. A. Doutt, "Modal mapping in shallow water using synthetic aperture horizontal arrays," in *OCEANS 2000 MTS/IEEE Conference and Exhibition. Conference Proceedings (Cat. No.00CH37158)*, vol. 1, pp. 185–188 vol.1, 2000.

[12] S. Stergiopoulos and H. Urban, "An experimental study in forming a long synthetic aperture at sea," *IEEE Journal of Oceanic Engineering*, vol. 17, no. 1, pp. 62–72, 1992.

[13] Z. . Michalopoulou and M. B. Porter, "Matched-field processing for broad-band source localization," *IEEE Journal of Oceanic Engineering*, vol. 21, pp. 384–392, Oct 1996.

[14] G. J. Orris, M. Nicholas, and J. S. Perkins, "The matched-phase coherent multi-frequency matched-field processor," *The Journal of the Acoustical Society of America*, vol. 107, no. 5, pp. 2563–2575, 2000.

[15] Z. Michalopoulou, "Robust multi-tonal matched-field inversion: A coherent approach," *The Journal of the Acoustical Society of America*, vol. 104, no. 1, pp. 163–170, 1998.

[16] C. Debever and W. A. Kuperman, "Robust matched-field processing using a coherent broadband white noise constraint processor," *The Journal of the Acoustical Society of America*, vol. 122, no. 4, pp. 1979–1986, 2007.

[17] L. M. Zurk, N. Lee, and J. Ward, "3d adaptive matched field processing for a moving source in a shallow water channel," in *Oceans '99. MTS/IEEE. Riding the Crest into the 21st Century. Conference and Exhibition. Conference Proceedings (IEEE Cat. No.99CH37008)*, vol. 2, pp. 728–731 vol.2, Sep. 1999.

[18] L. M. Zurk, N. Lee, and J. Ward, "Source motion mitigation for adaptive matched field processing," *The Journal of the Acoustical Society of America*, vol. 113, no. 5, pp. 2719–2731, 2003.

[19] F. Li, F. Zhu, Y. Zhang, B. Zhang, W. Li, and W. Luo, "Synthetic adaptive matched field processing for moving source with a horizontal line array," *The Journal of the Acoustical Society of America*, vol. 149, no. 2, pp. 1138–1146, 2021.

[20] H. Bucker, "Matched-field tracking in shallow water," *The Journal of the Acoustical Society of America*, vol. 96, no. 6, pp. 3809–3811, 1994.

[21] L. T. Fialkowski, J. S. Perkins, M. D. Collins, M. Nicholas, J. A. Fawcett, and W. A. Kuperman, "Matched-field source tracking by ambiguity surface averaging," *The Journal of the Acoustical Society of America*, vol. 110, no. 2, pp. 739–746, 2001.

[22] P. A. Baxley, N. O. Booth, and W. S. Hodgkiss, "Matched-field replica model optimization and bottom property inversion in shallow water," *The Journal of the Acoustical Society of America*, vol. 107, no. 3, pp. 1301–1323, 2000.

[23] A. N. Guthrie, R. M. Fitzgerald, D. A. Nutile, and J. D. Shaffer, "Long-range low-frequency CW propagation in the deep ocean: Antigua-Newfoundland," *The Journal of the Acoustical Society of America*, vol. 56, no. 1, pp. 58–69, 1974.

[24] F. B. Jensen, W. A. Kuperman, M. B. Porter, and H. Schmidt, *Computational ocean acoustics*. Springer Science & Business Media, 2011.

[25] K. E. Hawker, "A normal mode theory of acoustic doppler effects in the oceanic waveguide," *The Journal of the Acoustical Society of America*, vol. 65, no. 3, pp. 675–681, 1979.

[26] H. P. Bucker, "Use of calculated sound fields and matched-field detection to locate sound sources in shallow water," *The Journal of the Acoustical Society of America*, vol. 59, no. 2, pp. 368–373, 1976.

[27] H. Cox, R. Zeskind, and M. Owen, "Robust adaptive beamforming," *IEEE Transactions on Acoustics, Speech, and Signal Processing*, vol. 35, no. 10, pp. 1365–1376, 1987.

[28] J. N. Maksym, "A robust formulation of an optimum cross-spectral beamformer for line arrays," *The Journal of the Acoustical Society of America*, vol. 65, no. 4, pp. 971–975, 1979.

[29] B. D. Carlson, "Covariance matrix estimation errors and diagonal loading in adaptive arrays," *IEEE Transactions on Aerospace and Electronic Systems*, vol. 24, pp. 397–401, July 1988.

[30] K. Hsu and A. B. Baggeroer, "Application of the maximum-likelihood method (MLM) for sonic velocity logging," *GEOPHYSICS*, vol. 51, no. 3, pp. 780–787, 1986.

[31] H. Cox, "Multi-rate adaptive beamforming (MRABF)," in *Proceedings of the 2000 IEEE Sensor Array and Multichannel Signal Processing Workshop. SAM 2000 (Cat. No.00EX410)*, pp. 306–309, March 2000.

[32] I. Kirsteins and D. Tufts, "On the probability density of signal-to-noise ratio in an improved adaptive detector," in *ICASSP '85. IEEE International Conference on Acoustics, Speech, and Signal Processing*, vol. 10, pp. 572–575, April 1985.

[33] S. Kay, *Fundamentals of Statistical Signal Processing: Estimation theory*. Fundamentals of Statistical Signal Processing, PTR Prentice-Hall, 1993.

[34] J. Murray and D. Ensberg, "The swellex-96 experiment," *URL: http://www. mpl. ucsd. edu/swellex96*, 1996.

[35] R. T. Bachman, P. W. Schey, N. O. Booth, and F. J. Ryan, "Geoacoustic databases for matched-field processing: Preliminary results in shallow water off San Diego, California," *The Journal of the Acoustical Society of America*, vol. 99, no. 4, pp. 2077–2085, 1996.

[36] A. B. Baggeroer and H. Cox, "Passive sonar limits upon nulling multiple moving ships with large aperture arrays," in *Conference Record of the Thirty-Third Asilomar Conference on Signals, Systems, and Computers (Cat. No.CH37020)*, vol. 1, pp. 103–108 vol.1, Oct 1999.

[37] M. B. Porter, "The KRAKEN Normal Mode Program," tech. rep., SACLANT Undersea Research Centre, La Spezia, Italy Rep. SM-245, 1991.

[38] G. L. DSpain, J. J. Murray, W. S. Hodgkiss, N. O. Booth, and P. W. Schey, "Mirages in shallow water matched field processing," *The Journal of the Acoustical Society of America*, vol. 105, no. 6, pp. 3245–3265, 1999.

[39] N. O. Booth, A. T. Abawi, P. W. Schey, and W. S. Hodgkiss, "Detectability of low-level broad-band signals using adaptive matched-field processing with vertical aperture arrays," *IEEE Journal of Oceanic Engineering*, vol. 25, pp. 296–313, July 2000.

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

[41] G. Byun, F. Hunter Akins, K. L. Gemba, H. C. Song, and W. A. Kuperman, "Multiple constraint matched field processing tolerant to array tilt mismatch," *The Journal of the Acoustical Society of America*, vol. 147, no. 2, pp. 1231–1238, 2020.

Chapter 3

Modal-MUSIC: A passive mode estimation algorithm for partially spanning arrays

State-of-the-art mode estimation methods either utilize active source transmissions or rely on a full-spanning array to extract normal modes from noise radiated by a ship-of-opportunity. Modal-MUSIC, an adaptation of the MUSIC algorithm (best-known for direction-of-arrival estimation), extracts normal modes from a moving source of unknown range recorded on a partially-spanning vertical line array (VLA), given knowledge of the water column sound speed profile. The method is demonstrated on simulations, as well as on data from the SWellEx-96 experiment. Extracted normal modes from ship noise during the experiment are used to successfully localize a multitone source without any geoacoustic information.

3.1 Introduction

The normal modes of a waveguide are determined by the sound speed profile (SSP) and boundary conditions. In underwater acoustics, the surface boundary condition is always pressure-release, and the SSP in the water column is easy to measure. The set of all possible modes that satisfy the surface condition and the water-column SSP can be explored via the shooting method [1], which re-frames the boundary value problem (BVP) as an initial value problem (IVP). The true modes for a given waveguide are the subset of IVP solutions that

also satisfy the bottom geoacoustic parameters. This paper presents a mode extraction method, modal-MUSIC, that picks the true modes from the set of IVP solutions by comparison with data from noise, thus replacing the need for a geoacoustic model. As a result, modal-MUSIC allows matched field processing (MFP) to be performed without any information below the sea floor.

Mode estimation methods can be divided into two categories: active methods that require the use of a controlled source; and passive methods, that rely on ambient acoustic noise to perform the extraction. Active source methods fall into two subcategories: wavenumber estimation based methods applied to a horizontal line array/synthetic aperture [2, 3, 4, 5, 6] and modal group/phase speed estimation from broadband impulsive source transmissions [7, 8, 9]. Passive mode extraction on a VLA has thus-far been limited to the special case where the receiving array fully spans and sufficiently samples the water column (in which case the singular value decomposition (SVD) provides mode estimates) [10, 11]. An effective passive mode extraction method for more general experimental configurations would be beneficial; the present work extends passive mode estimation to partially-spanning arrays.

This paper presents modal-MUSIC, a method for extracting normal modes in an underwater waveguide using noise from a moving source of unknown range received on a partially-spanning array, and further demonstrates its application to model-free MFP without needing knowledge of the bottom properties.

3.2 MUSIC for DOA estimation

MUSIC, which stands for "Multiple Signal Classification", is a general algorithm for the estimation of the parameters of multiple sources radiating within the vicinity of an array [12]. In the most common MUSIC presentation, it is used to estimate the angles associated with sources radiating plane waves.

Measurements on an array, written as a vector $\boldsymbol{x}$, obey the signal model

$$\boldsymbol{x} = \boldsymbol{A}\boldsymbol{s} + \boldsymbol{w}\,, \tag{3.1}$$

where the N columns of $\boldsymbol{A}$ are the incident plane waves

$$\boldsymbol{A} = [\boldsymbol{a}(\theta_1)\ \ \boldsymbol{a}(\theta_2)\ \ \ldots\ \ \boldsymbol{a}(\theta_N)] \tag{3.2}$$

received on the array from N sources with amplitudes

$$\boldsymbol{s} = [s_1\ \ s_2\ \ \ldots\ \ s_N]^T\ . \tag{3.3}$$

The noise $\boldsymbol{w}$ contains anything uncorrelated with the source signals. The data covariance matrix can be written as

$$\boldsymbol{X} = \boldsymbol{A}\boldsymbol{S}\boldsymbol{A}^\dagger + \lambda\boldsymbol{W}\ , \tag{3.4}$$

where, with $\mathrm{E}(\cdot)$ denoting expectation, $\boldsymbol{S} = \mathrm{E}(\boldsymbol{s}\boldsymbol{s}^\dagger)$ is the source covariance matrix, $\lambda\boldsymbol{W} = \mathrm{E}(\boldsymbol{w}\boldsymbol{w}^\dagger)$ is the covariance matrix of the noise, and λ is a scaling parameter that is relevant for the general derivation of the MUSIC algorithm.

As shown by Schmidt [12], whenever the matrix $\boldsymbol{A}\boldsymbol{S}\boldsymbol{A}^\dagger$ is singular (i.e. there are fewer sources N than there are number of sensors L), the covariance matrix can be written as

$$\boldsymbol{X} = \boldsymbol{A}\boldsymbol{S}\boldsymbol{A}^\dagger + \lambda_{min}\boldsymbol{W}\ , \tag{3.5}$$

where λ_{min} is the smallest solution to $|\boldsymbol{X} - \lambda\boldsymbol{W}| = 0$ (i.e. λ_{min} is the smallest generalized eigenvalue of $\boldsymbol{X}$ and $\boldsymbol{W}$).

As long as the sources are not fully correlated with one another, the Vandermonde structure of $\boldsymbol{A}$ guarantees that the rank of $\boldsymbol{A}\boldsymbol{S}\boldsymbol{A}^\dagger$ is the number of sources N. Schmidt shows that λ_{min} must have multiplicity $L-N$, and the $(L-N)$-dimensional associated eigenspace is orthogonal to the space spanned by the columns of $\boldsymbol{A}\boldsymbol{S}\boldsymbol{A}^\dagger$.

Therefore, given a) the matrix $\boldsymbol{E}_N$ whose columns are the generalized eigenvectors associated with $\lambda_{min}(\boldsymbol{X},\boldsymbol{W})$ and b) a parameterization $\boldsymbol{a}(\theta)$ of all possible columns of $\boldsymbol{A}$, the

MUSIC spectrum can be computed as

$$P_{MU}(\theta) = \left[\boldsymbol{a}^{\dagger}(\theta) \boldsymbol{E}_N \boldsymbol{E}_N^{\dagger} \boldsymbol{a}(\theta) \right]^{-1} \tag{3.6}$$

and will have peaks at the parameter values θ excited by the sources.

With respect to a): in theory, knowledge of $\boldsymbol{W}$ permits one to estimate $\boldsymbol{E}_N$ by computing the generalized eigenvalues $\lambda(\boldsymbol{X}, \boldsymbol{W})$. However, under the stronger assumption that the noise is uncorrelated white noise, the matrix $\boldsymbol{E}_N$ can be estimated directly from the eigenvectors of the covariance matrix $\boldsymbol{X}$ (since the generalized eigenvectors $\lambda(\boldsymbol{X}, a\boldsymbol{I})$ for identity matrix $\boldsymbol{I}$ are simply the eigenvectors of $\boldsymbol{X}$). At high SNR, estimation of the noise subspace can be made by choosing a cutoff eigenvalue from the spectrum of the covariance matrix. A practical way to automate this is to find the peak in the "difference spectrum", defined as $\{\lambda_{i+1} - \lambda_i\}$. Then, the eigenvectors that fall to the right of this peak (assuming eigenvectors sorted into a descending order) span the noise space and form the columns of $\boldsymbol{E}_N$.

With respect to b): the possible columns of $\boldsymbol{A}$ form the "array manifold", which for a plane wave signal model is comprised of the set of all possible arriving plane waves evaluated for the array. In this case it is parameterized by source angle θ and is trivially computed given the positions of the array elements.

3.3 Modal-MUSIC for mode extraction from noise

Here it is shown that by viewing each excited mode in a waveguide as a virtual source (whose waveform on the array is a normal mode rather than a plane wave), the mode estimation problem has the same form as the source estimation problem that MUSIC solves. The field produced by a moving source produces the necessary covariance matrix structure.

We can express the covariance matrix for a VLA, built up from snapshots of a single moving source over a set of ranges $\{r_i\}$, in the familiar matrix form $\boldsymbol{X} = \boldsymbol{A}\boldsymbol{S}\boldsymbol{A}^{\dagger} + \lambda \boldsymbol{W}$, where

now the N columns of $\boldsymbol{A}$ are the normal modes evaluated at the array depths,

$$\boldsymbol{A} = [\boldsymbol{\Phi}_1 \;\; \boldsymbol{\Phi}_2 \;\; \ldots \;\; \boldsymbol{\Phi}_N] \tag{3.7}$$

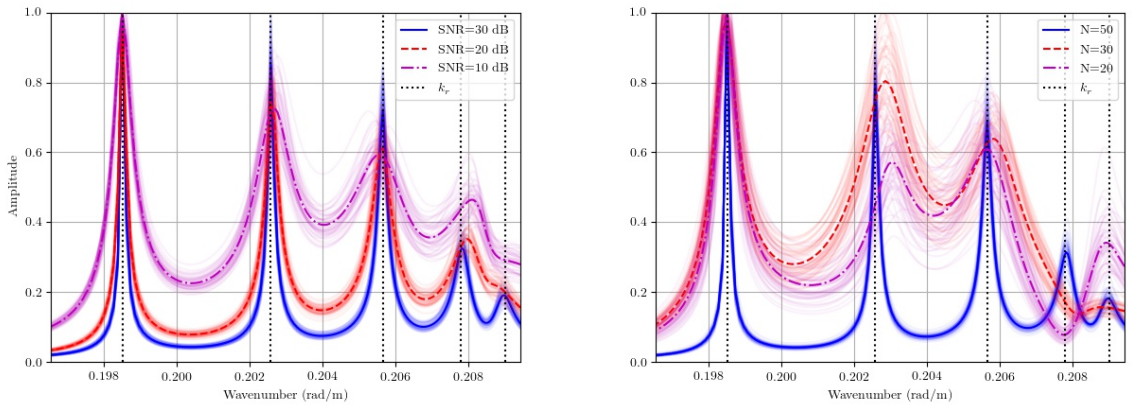
and the source covariance matrix is

$$(\boldsymbol{S})_{mn} = S(\omega)^2 \frac{\Phi_m(z_s)\Phi_n(z_s)}{\sqrt{k_n k_m}} \frac{1}{N_r} \sum_{i=0}^{N_r} \frac{e^{j(k_n - k_m)r_i}}{r_i} \; . \tag{3.8}$$

Here $S(\omega)$ is the source's amplitude at frequency ω, z_s is the depth of the source, $\Phi_i(z)$ is the ith normal mode depth function that solves the depth-separated Helmholtz equation and k_i is the corresponding wavenumber. $\boldsymbol{S}$ is the modal correlation matrix. Although any amount of partial source correlation is theoretically permissible in the derivation of MUSIC, source correlations can significantly degrade the performance of MUSIC in the presence of noise [13]. Therefore, a requirement for the use of modal-MUSIC for a moving source is that the source transits a significant range (see Neilson (2002) for details [10]) in which case $\boldsymbol{S}$ in Eq. 3.8 will be approximately diagonal.

The discussion in Section 3.2 showed that we can express the field due to a moving source in the multiple source matrix form $\boldsymbol{X} = \boldsymbol{A}\boldsymbol{S}\boldsymbol{A}^\dagger + \boldsymbol{W}$. The source parameters determine the columns of $\boldsymbol{A}$: in this case the parameters are the horizontal wavenumbers of the excited modes and the columns of $\boldsymbol{A}$ are the corresponding modes $\{\boldsymbol{\Phi}_i\}$ evaluated at the array depths.

Recall that MUSIC also requires a means of exploring the array manifold. This is trivial for a plane wave signal model, where for each possible source angle the array response is simply the steering vector $\boldsymbol{a}(\theta) = e^{-ik\sin\theta \boldsymbol{z}}$. For the normal mode problem, the waveforms of the "signals" are the normal modes sampled at the array depths. For each candidate wavenumber k_r, the associated normal mode sampled at the array depths, $\boldsymbol{\Phi}(k_r)$, is computed using the shooting method [1]. This requires knowledge of the SSP from the surface down to the bottom of the array. Now that we are equipped with the means of exploring the array manifold, the modal-MUSIC



a) b)

Figure 3.1. (Color online). Simulation results for a) varying SNR and b) array size. For both, all 50 realizations are shown faintly, with the mean spectrum shown darkly. The modeled wavenumbers are shown as dashed lines. a) Normalized modal-MUSIC spectra are shown for SNR of 10, 20, and 30 dB. b) Normalized modal-MUSIC spectra are shown for arrays with 50, 30, and 20 elements spaced at 2 meters and with shallowest element at 100 meter depth. The SNR is fixed at 30 dB for all of the array lengths.

spectrum can be computed as

$$P_{moMU}(k_r) = \left[\boldsymbol{\Phi}^{\dagger}(k_r)\boldsymbol{E}_N\boldsymbol{E}_N^{\dagger}\boldsymbol{\Phi}(k_r)\right]^{-1} \tag{3.9}$$

and will have peaks at the wavenumbers supported by the waveguide.

3.4 Simulations

Simulations demonstrate the performance of modal-MUSIC under conditions of varying signal-to-noise ratio (SNR) and array length. It is seen that the quality of the estimates is always better for the higher order modes, which are better sampled by the array. This effect is accentuated because we use a surface source that emphasizes the high order modes.

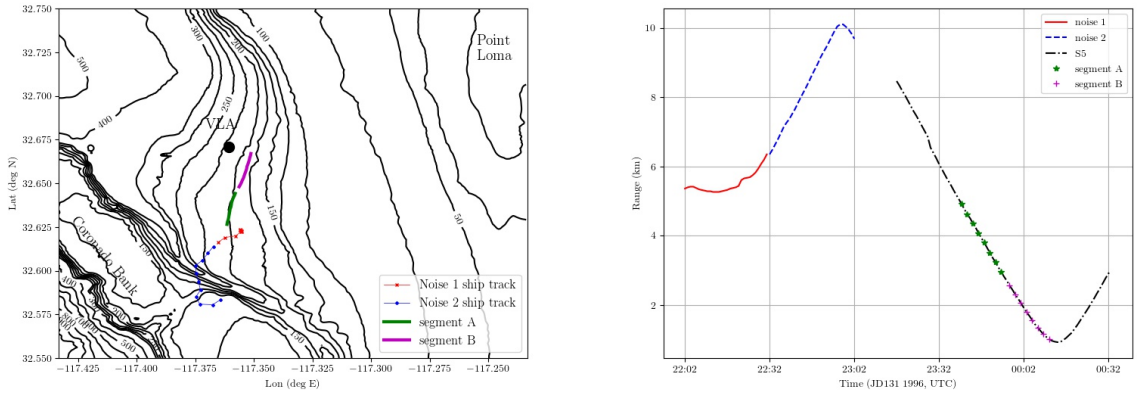
We simulate an 8 meter deep, 50 Hz source moving through a 216.5 meter deep Pekeris waveguide with water column sound speed $c = 1500$ m/s, bottom sound speed $c_b = 1600$ m/s, and bottom density $\rho_b = 2.0$ g/cm^3. The source has 5 excited modes in this waveguide. For each

simulation setup, we generate a sequence of pressure field values $[\boldsymbol{p}_1, \boldsymbol{p}_2, \ldots, \boldsymbol{p}_{Nr}]$ by sampling the field received on the array from the source as it moves from 5 km range to 8 km range every 5 meters (for a total of 301 snapshots). The normal modes and wavenumbers are computed using the Sturm sequence eigenvalue method [14].

An ensemble of data is generated by adding realizations of white gaussian noise to the sequence of pressure field values. The noise variance σ_N^2 required to produce a desired SNR in dB is given by $\sigma_N^2 = \sigma_s^2 / 10^{SNR/10}$, where σ_s^2 is the average received power from the source over all ranges and array elements. For each realization, uncorrelated complex Gaussian noise with real and imaginary variance of $\sigma_N^2/2$ is added to each element in the pressure field sequence. Then the sample covariance matrix is formed from the 301 noise-corrupted pressure field snapshots. The signal subspace rank is estimated from the eigenvalue difference spectrum, and that estimate is used to compute the modal-MUSIC spectrum.

To look at the effect of SNR on modal-MUSIC performance, we use a half-spanning VLA similar to the VLA deployed in the SWellEx-96 experiment. It consists of 50 elements spaced at 2 meters, and starts 100 meters deep in the 216.5 meter deep waveguide. For SNR values from 10 dB to 30 dB, 50 realizations of data are generated and the modal-MUSIC spectrum is computed (shown in Figure 3.1). At 30 dB SNR, modal-MUSIC resolves all five modes of the waveguide. At 10 dB and 20 dB SNR, modal-MUSIC can no longer resolve mode 1 (which has a vertical wavenumber of X), and the estimate of mode 2 becomes biased.

We also demonstrate the effect of shortening the array from 50 elements (100 meters) to 20 elements (40 meters), keeping the SNR at 30 dB for all of the data realizations. Note that the wavelength of the 50 Hz source is around 30 meters. The ability to estimate the low-angle modes quickly drops off with array length, though the ability to estimate the highest mode is consistent even as the array length drops close to the wavelength of the sound.



a) b)

Figure 3.2. (Color online). a) SWellEx-96 experiment region. The two tracks of the ship during the half-hour segment of data used for mode estimation are shown to the right of the Coronado Bank and south of the VLA. Segments A and B, used for localization are also shown. b) Ranges from the ship to the VLA during noise 1 and noise 2 (segments used for mode extraction) and segment A and segment B (used for localization). The full ship track during the S5 event is also shown.

3.5 Matched Field Processing

MFP [15] is a source localization method used in underwater acoustics. For each potential source position, which in a range-independent environment consists of range and depth (r,z), a model is used to compute a replica $\boldsymbol{r}(r,z)$ which is compared with the received field $\boldsymbol{x}$. Using a simple linear processor (the Bartlett), the ambiguity surface is computed as

$$P(r,z) = |\boldsymbol{w}^{\dagger}\boldsymbol{x}|^2 = \boldsymbol{w}^{\dagger}\boldsymbol{X}\boldsymbol{w}\,, \tag{3.10}$$

where the weight vector $\boldsymbol{w} = \boldsymbol{r}/||\boldsymbol{r}||$ is the normalized replica vector and the sample covariance matrix of the data $\boldsymbol{X}$ is used when multiple snapshots are available. Assuming the model is correct, the ambiguity surface will show a peak at the true source location. When a source transmits at multiple frequencies, a single ambiguity surface is formed by summing the normalized ambiguity surfaces from each frequency in dB[15].

As mentioned, a full acoustic model (water column SSP plus geoacoustic model) is

typically used to compute the replica $\boldsymbol{r}(r,z)$. Here, we will use the normal modes extracted from noise to calculate the replica vectors, thus sidestepping the use of a geoacoustic model.

3.6 Data details

Noise from the R/V Sproul during the SWellEx-96 experiment is used to extract normal modes using data from a VLA. Then, matched field processing using the extracted normal modes is used to demonstrate the source localization capability on the S5 event.

SWellEx-96 is a well-known shallow-water MFP experiment performed off the coast of San Diego [16]. This work uses recordings from the experiment made by a 64-element VLA with element spacing 1.875 meters that spans the bottom half of the water column. The array is cut for 400 Hz, and the sampling rate of the sensors is 1500 Hz.

For normal mode extraction, two thirty-minute segments, labeled 'noise 1' and 'noise 2', of data were selected from the recording during which sound from the ship driving in the experiment region is received on the VLA. For localization, two ten-minute segments, labeled 'segment A' and 'segment B' were selected from the S5 event, which was a 75-minute source tow. During the source tow, both a shallow and a deep source were towed, although we focus on the deep source which was towed at a depth of 60 meters and transmitted 13 tonals between 49 Hz and 388 Hz. The ship's position in the experiment during these four segments is shown in Figure 3.2a, and the source-receiver range is shown in Figure 3.2b.

To obtain covariance matrices for modal-MUSIC, a short time Fourier transform (STFT) with a 2048 point Hamming-windowed FFT and no overlap was applied to both noise 1 and noise 2 to obtain snapshots with a time step of 1.36 seconds and a frequency spacing around 0.73 Hz. The roughly 1300 snapshots at each frequency between 30 and 400 Hz were then used to build the roughly 500 sample covariance matrices (one for each frequency) used in modal-MUSIC.

Similarly for the MFP application, tones were extracted from a 2048 point, Hamming windowed STFT with a 50% overlap between snapshots, resulting in snapshots spaced at 0.68

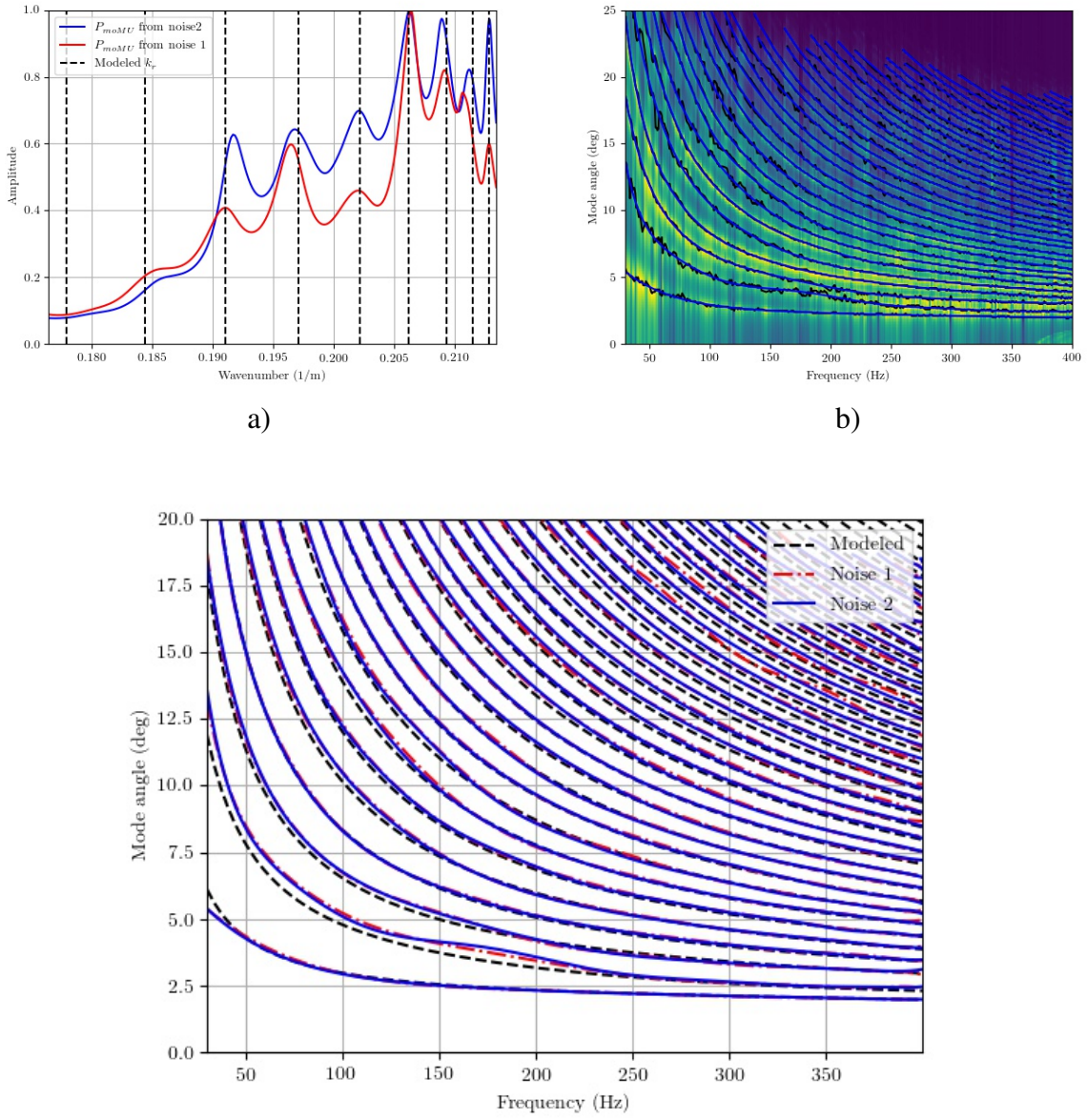


Figure 3.3. (Color online). a) Normalized modal-MUSIC spectrum at 50.54 Hz, shown for both noise 1 and noise 2 and compared to modeled wavenumbers. b) Normalized modal-MUSIC spectra and mode curves extracted from noise 2 data segment. The peaks are shown as black lines, and their polynomial fits are shown in blue. c) Smoothed wavenumber curves extracted from both segments noise 1 and noise 2, shown in blue and red. The wavenumbers computed from the model are shown in black. The model and data agree fairly well up to mode 15.

seconds. The signal bins were selected as those closest to the Doppler-shifted frequency, as estimated from mean range-rates for the source over the segments (-2.45 m/s for segment A and -2.1 m/s for segment B) and a nominal soundspeed of 1500 m/s. Then, covariance matrices were formed using sets of 100 snapshots, corresponding to a covariance integration time of 68 seconds. By overlapping each covariance integration interval by 50 snapshots, the final sequence of covariance matrices have a time step of 34 seconds. These covariance matrix sequences are then used for multi-tone MFP.

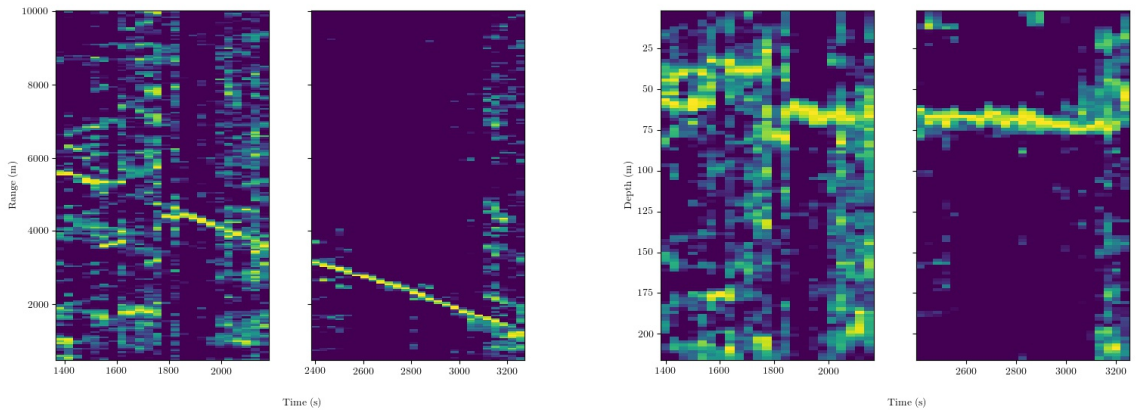
3.7 Modal-MUSIC mode extraction and localization

3.7.1 Modal-MUSIC results

Estimation of the noise subspace $\boldsymbol{E}_N$ required by modal-MUSIC requires estimates of the rank $\hat{q}(f)$ of the signal subspace for each frequency. This was estimated using the peak in the eigenvalue difference spectrum. The resulting rank estimates $\hat{q}(f)$ are then fit to a curve over frequency to remove outliers and noise from the subspace dimension estimates. These smoothed rank estimates are used to build up $\hat{\boldsymbol{E}}_N(f)$ at each frequency by taking the $64-\hat{q}(f)$ smallest eigenvectors (64 is the number of array elements). Once the noise subspace is estimated, a conductivity-temperature-depth (CTD) cast from the experiment is used to generate an SSP for shooting candidate modes, and the modal-MUSIC spectrum is computed using Eq. 3.9. The single modal-MUSIC spectrum from ship noise at 50.54 Hz during both noise 1 and noise 2 is compared to wavenumbers computed using an established geoacoustic model [17] in Figure 3.3b.

In order to compare the spectra across frequency, the modes are parameterized by an "angle" $\theta(k_r) = \arccos(k_r/k)$, where $k = \omega/\min_z c(z)$ is the maximum wavenumber for the waveguide. The normalized spectra for every processed frequency are shown in Figure 3.3b for noise 2.

The mode estimates at each frequency are obtained by picking peaks from each fre-



a) b)

Figure 3.4. (Color online). Results for localization on segments A and B from S5 data use data-derived normal modes. In each of the two plots, segment A is shown on the left and segment B is shown on the right. The range and depth are well-estimated out to around 5 km. No geoacoustic model was used for these results.

quency's modal-MUSIC spectrum. The spectra are smooth enough that picking peaks based on local optima is sufficient. Peaks are thresholded at 10% of the peak-maximum to limit the selection to well-excited normal modes.

As seen in Figure 3.3b, the wavenumber estimates are noisy. To obtain denoised estimates of the modes, the estimates from all 506 frequencies can be combined using a smoothness constraint. To enforce smoothness of the normal mode curves in frequency, a polynomial fit is made for each measured mode. Polynomials of increasing degree are fit through each mode curve until the change in the residuals with increasing degree falls below a threshold value (we selected 0.1), or the maximum degree of 14 is reached. These polynomial curves provide the final product of normal modes as a function of frequency, and are compared to modes computed with an established geoacoustic model for the SWellEx environment [17] in Figure 3.3c for both noise 1 and noise 2.

3.7.2 MFP localization results

Typically MFP uses a geoacoustic model to compute the replica vectors. Here, we rely on the wavenumbers extracted from data to compute the replica vectors for each candidate source position (r, z). These replicas are combined with the covariance matrix sequences (described in Section 3.6) to compute a sequence of ambiguity surfaces using Eq. 3.10. By taking the maximum over depth, a range-time ambiguity surface can be formed (and similarly taking the maximum over range yields a depth-time ambiguity surface), which is useful for visualizing the three-dimensional range-depth-time ambiguity surfaced produced by MFP. These two two-d surfaces are shown in Figure 3.4. It can be seen that the extracted normal modes allow for accurate source range and depth estimation out to a range of around 5 kilometers, although the quality of estimation decreases with range.

3.8 Conclusion

Knowledge of the water column SSP is sufficient to explore the array manifold of an array measuring a normal mode acoustic field. This opens up the use of the high resolution estimator MUSIC to then find the excited modes in the waveguide, provided a suitable source of acoustic energy. A moving source transiting a sufficient (unknown) range is a suitable source for normal mode extraction.

The estimated excited modes can be related to the environment properties for a geoacoustic inversion (not explored here), or be used directly for performing source localization without any knowledge of the bottom properties of the environment. It was shown that normal mode wavenumbers extracted from ship noise and smoothed over frequency had sufficient accuracy to perform localization using broadband MFP.

3.9 Acknowledgments

This work was funded by the TFO program of the United States Office of Naval Research. Bill Hodgkiss and Dave Ensberg helped with data access for the SWellEx-96 experiment.

This chapter, in full, is a reprint of the material as it appears in Akins, F. H. and Kuperman, W. A. (2022). Modal-MUSIC: A passive mode estimation algorithm for partially spanning arrays. *JASA Express Letters*, 2(7), 074802. The dissertation author was the primary investigator and author of this paper. The coauthor listed in this publication directed and supervised the research.

Bibliography

[1] M. B. Porter and E. L. Reiss, "A note on the relationship between finite-difference and shooting methods for ode eigenvalue problems," *SIAM Journal on Numerical Analysis*, vol. 23, no. 5, pp. 1034–1039, 1986.

[2] G. V. Frisk, K. M. Becker, S. D. Rajan, C. J. Sellers, K. von der Heydt, C. M. Smith, and M. S. Ballard, "Modal mapping experiment and geoacoustic inversion using sonobuoys," *IEEE Journal of Oceanic Engineering*, vol. 40, no. 3, pp. 607–620, 2015.

[3] S. C. Walker, P. Roux, and W. A. Kuperman, "Modal doppler theory of an arbitrarily accelerating continuous-wave source applied to mode extraction in the oceanic waveguide," *The Journal of the Acoustical Society of America*, vol. 122, no. 3, pp. 1426–1439, 2007.

[4] Y. Park, P. Gerstoft, and W. Seong, "Grid-free compressive mode extraction," *The Journal of the Acoustical Society of America*, vol. 145, no. 3, pp. 1427–1442, 2019.

[5] T. C. Yang, "Data-based matched-mode source localization for a moving source," *The Journal of the Acoustical Society of America*, vol. 135, no. 3, pp. 1218–1230, 2014.

[6] T. C. Yang, "Source depth estimation based on synthetic aperture beamfoming for a moving source," *The Journal of the Acoustical Society of America*, vol. 138, no. 3, pp. 1678–1686, 2015.

[7] J. Bonnel, A. Thode, D. Wright, and R. Chapman, "Nonlinear time-warping made simple: A step-by-step tutorial on underwater acoustic modal separation with a single hydrophone," *The Journal of the Acoustical Society of America*, vol. 147, no. 3, pp. 1897–1926, 2020.

[8] H. Niu, P. Gerstoft, E. Ozanich, Z. Li, R. Zhang, Z. Gong, and H. Wang, "Block sparse bayesian learning for broadband mode extraction in shallow water from a vertical array," *The Journal of the Acoustical Society of America*, vol. 147, no. 6, pp. 3729–3739, 2020.

[9] H. Niu, P. Gerstoft, R. Zhang, Z. Li, Z. Gong, and H. Wang, "Mode separation with one hydrophone in shallow water: A sparse bayesian learning approach based on phase speed," *The Journal of the Acoustical Society of America*, vol. 149, no. 6, pp. 4366–4376, 2021.

[10] T. B. Neilsen and E. K. Westwood, "Extraction of acoustic normal mode depth functions using vertical line array data," *The Journal of the Acoustical Society of America*, vol. 111, no. 2, pp. 748–756, 2002.

[11] P. Hursky, W. S. Hodgkiss, and W. A. Kuperman, "Matched field processing with data-derived modes," *The Journal of the Acoustical Society of America*, vol. 109, no. 4, pp. 1355–1366, 2001.

[12] R. Schmidt, "Multiple emitter location and signal parameter estimation," *IEEE Transactions on Antennas and Propagation*, vol. 34, no. 3, pp. 276–280, 1986.

[13] T.-J. Shan, M. Wax, and T. Kailath, "On spatial smoothing for direction-of-arrival estimation of coherent signals," *IEEE Transactions on Acoustics, Speech, and Signal Processing*, vol. 33, no. 4, pp. 806–811, 1985.

[14] M. Porter and E. L. Reiss, "A numerical method for oceanacoustic normal modes," *The Journal of the Acoustical Society of America*, vol. 76, no. 1, pp. 244–252, 1984.

[15] A. B. Baggeroer, W. A. Kuperman, and H. Schmidt, "Matched field processing: Source localization in correlated noise as an optimum parameter estimation problem," *The Journal of the Acoustical Society of America*, vol. 83, no. 2, pp. 571–587, 1988.

[16] J. Murray and D. Ensberg, "The swellex-96 experiment," *URL: http://www. mpl. ucsd. edu/swellex96*, 1996.

[17] P. A. Baxley, N. O. Booth, and W. S. Hodgkiss, "Matched-field replica model optimization and bottom property inversion in shallow water," *The Journal of the Acoustical Society of America*, vol. 107, no. 3, pp. 1301–1323, 2000.

Chapter 4

Experimental demonstration of low signal-to-noise ratio matched field processing without prior geoacoustic information

Passive localization of a low signal-to-noise ratio (SNR) source in a shallow water waveguide without prior geoacoustic information is accomplished by combining the mode-extraction method modal-MUSIC with range-coherent matched field processing (MFP). Range-coherent MFP [J. Acoust. Soc. Am. **150(1)**, 270-280] coherently combines snapshots from different resolution cells to obtain gain over noise. Modal-MUSIC [JASA EL **2(7)**, 074802] uses knowledge of the water column sound speed profile (no bottom information) to extract noisy estimates of modal wavenumbers from ship noise recorded on a partially-spanning vertical line array (VLA). A geoacoustic model is then fit to the wavenumber estimates extracted from noise with modal-MUSIC and used to compute replicas for range-coherent MFP. The combination of these two methods applied to a 21-element VLA achieves successful source localization at SNR levels as low as -20 dB using ten tonals transmitted during the SWellEx96 experiment.

4.1 Introduction

Matched field processing (MFP) passively localizes a source by comparing the acoustic field measured on an array to modeled fields for various candidate source positions. To successfully localize a source, MFP requires an accurate environmental model, which, at low frequency

in shallow water, includes information about propagation characteristics in the seabed. The challenges of obtaining such a model limit the applicability of MFP. By passively extracting the geoacoustic portion of the environmental model from ambient noise, this work successfully applies MFP in the absence of comprehensive *a priori* environmental knowledge.

Geoacoustic parameters can be extracted from various acoustic observables, such as: modal group speed [1], noise correlations between vertically-separated sensors [2], the power reflection coefficient [3], complex pressure field measured at known ranges [4], echosounder measurements, and so on. This work is concerned with passive localization in a (partially) unknown environment. Therefore, estimation of an unknown environment from acoustic data must be carried out without the benefit of an active source. For the experimental scenario of interest, the real parts of the wavenumbers of the modal field are extracted from the covariance matrix of noise in the band from 50 Hz to 250 Hz on a partially-spanning vertical line array (VLA) using modal-MUSIC [5]. The estimated wavenumbers yield information on the sediment compressional sound speed, density, and layer thickness, but do not contain direct information on sediment attenuation. However, modal-MUSIC also estimates the number of modes present in the field, which can be used for accurate replica calculation in the absence of sediment attenuation [6]. Geoacoustic parameters could also be included in the MFP search space such as in focalization [7, 8, 9]. However, focalization greatly increases the computational demands of MFP and has never been experimentally demonstrated at low-SNR. Therefore, passive acquisition of a geoacoustic model from shipping noise is a useful step towards successful localization of a quiet source with MFP.

Once equipped with a sufficiently accurate model to perform MFP, localization of a quiet moving signal is still limited by coherent integration time and array gain. A moving source remains within a resolution cell for a finite time, which limits covariance matrix estimation. A recent method, "range-coherent MFP" [10], coherently compares snapshots from separate resolution cells by modeling the effect of candidate source motions on the snapshots and comparing them with the data, with the caveat that the source frequency is known. Range-

coherent MFP essentially allows one to add additional synthetic array staves to obtain array gain at the price of additional computation.

In this work, range-coherent MFP is used to localize a quiet source (frequency averaged SNR of -20 dB over ten tonals) at ranges out to 8 kilometers using a 21-element vertical line array (VLA). The replicas used for MFP are computed using a geoacoustic model extracted from shipping noise using modal-MUSIC. Therefore, no prior geoacoustic information was necessary for the localization.

4.2 Modal-MUSIC

Modal-MUSIC is an array processing method that combines the MUltiple SIgnal Classification (MUSIC) algorithm [11] with the shooting method [12] to extract normal modes from shipping noise using a VLA [5].

Under certain conditions [5], 1) the eigenvectors of the covariance matrix of the array can be partitioned into signal and noise subspaces and 2) candidate normal modes that are orthogonal to the noise subspace are the true modes of the waveguide. The modal-MUSIC spectrum is then defined as

$$P_{moMU}(k_r) = \left[\boldsymbol{\Phi}(k_r)\boldsymbol{E}_N\boldsymbol{E}_N^{\dagger}\boldsymbol{\Phi}(k_r)\right]^{-1} \tag{4.1}$$

where $\boldsymbol{\Phi}(k_r)$ is the mode shape associated with candidate wavenumber k_r, sampled at the array elements, obtained by the shooting method using the water column sound speed profile (SSP). The columns of the matrix $\boldsymbol{E}_N$ are the eigenvectors of the sample covariance matrix (SCM) associated with the noise subspace.

In order to compare the spectra at different frequencies, it is convenient to introduce a "mode angle", which is defined as

$$\theta \equiv \arccos(k_r/(\omega/c_{min})) \,, \tag{4.2}$$

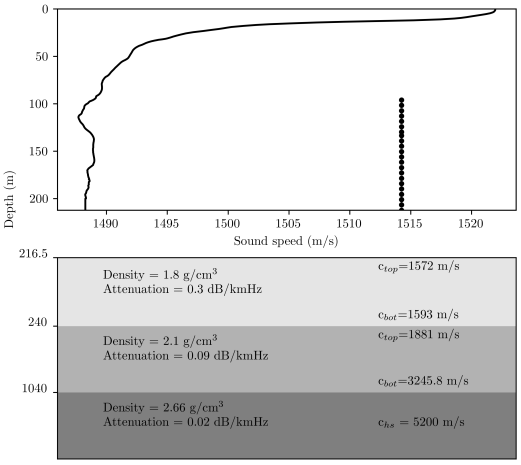


Figure 4.1. Array and environmental model used for Monte Carlo simulations. The SSP comes from a CTD cast during the SWellEx96 experiment, and is also used for modal-MUSIC in data processing. The array geometry is also representative of that used during the SWellEx96-experiment, consisting of a 21 element array spanning approximately 120 meters. The geoacoustic model is due to Bachman [13] and is used for conventional MFP in this work. The CTD shown is also used for the shooting method for modal-MUSIC.

where $\omega = 2\pi f$ is the frequency under consideration, and c_{min} is the minimum value of the SSP.

4.3 Geoacoustic information in the modal-MUSIC spectrum

This section discusses the extraction of a geoacoustic model from modal-MUSIC wavenumber estimates and presents relevant uncertainty analysis for the model fit. In previous work [5], wavenumber estimates were de-noised by fitting polynomials to the identified modes in the dispersion curve. These polynomials produced sufficiently accurate wavenumber estimates to localize a loud source to around five kilometers, but failed to localize at greater range or to localize a quiet source. The failure is believed to be due to poor estimation of low-order modes, which will be improved by fitting a geoacoustic model to high order modes. In particular, simulations demonstrate that the high-order modes will be estimated with lower variance than low order modes. This is fortuitous for geoacoustic model fit, as the high-order modes contain more geoacoustic information than low-order modes.

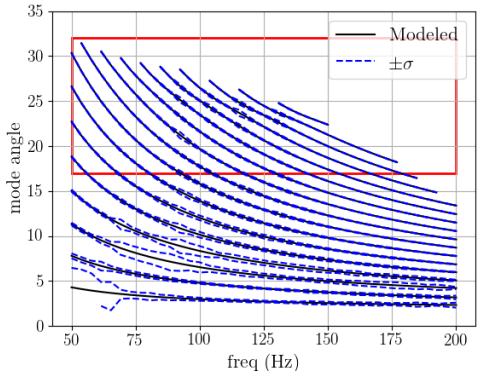


Figure 4.2. (Color online) Monte Carlo simulation results demonstrating uncertainty in the modal-MUSIC spectrum. One standard deviation is shown as blue dashed lines bracketing the real modes shown as black lines from 100 realizations of the SCM for a modal matrix with a source at 10 meters depth and no correlation between modes. The red box emphasizes a region of the dispersion curve that a) is well-estimated by modal-MUSIC and b) is sensitive to the geoacoustic properties.

Once modes have been identified in the modal-MUSIC dispersion curve, a misfit function between model and data may be defined as follows:

$$J(\boldsymbol{x}) = \sum_{i=1}^{N_f} \sum_{j \in M_i} (\hat{k}_j(f_i) - k_j(f_i;\boldsymbol{x}))^2 / \sigma_{ij}^2 \,, \tag{4.3}$$

where $\hat{k}_j(f_i)$ is the estimate of the jth mode wavenumber at frequency f_i with error variance σ_{ij}^2, M_i is the set of well-estimated modes at frequency f_i, N_f is the number of frequencies at which modal-MUSIC was performed, and $k_j(f_i;\boldsymbol{x})$ is the forward model calculation of the jth mode at frequency f_i for bottom parameters $\boldsymbol{x}$. The minimum of this function $\boldsymbol{x}_*$ corresponds to the best model fit.

In addition to mode-dependent estimate error variance σ_{ij}^2, there is also a mode-specific model sensitivity. Modes with a phase speed less than the sediment sound speed have an exponentially decaying tail in the bottom, and the horizontal wavenumber is relatively insensitive to bottom parameters [14] compared to modes that are oscillating in the sediment. In the next section, Monte Carlo simulations will show that *the modes which are most sensitive to the*

properties of the sediment are also the modes that can be estimated with the least variance.

4.3.1 Monte-Carlo simulations for uncertainty demonstration

Monte Carlo simulations are carried out to demonstrate modal-MUSIC uncertainty under ideal conditions over a range of frequencies. The simulations use the environmental model and array geometry shown in Figure 4.1, which parallel the SWellEx-96 experiment setup.

Each snapshot realization, expressed as a vector, is

$$\boldsymbol{p}_i = \sum_{m=1}^{M} x_m^i \Phi_m(\boldsymbol{z}) + \boldsymbol{n}_i \ , \tag{4.4}$$

where $x_m^i \sim \mathcal{N}(0, \Phi_m(z_s)^2)$ is a random Gaussian amplitude with variance equal to the mode excitation strength for a source depth at 10 meters (chosen to be characteristic of a ship), $\Phi_m(\boldsymbol{z})$ is the mth mode evaluated at the array depths in vector $\boldsymbol{z}$, M is the number of modes, and $n_i \sim \mathcal{N}(0, \sigma_n^2 I)$ is zero-mean Gaussian noise. The diagonal value σ_n^2 of the noise covariance is selected to give an average single-sensor "SNR" of 20 dB. Here, SNR refers to the ratio of the variance of the modal component of the noise to the variance of the white noise component.

Then, an ensemble of 1000 snapshot realizations was used to form a single realization of the array sample covariance matrix (SCM)

$$\hat{K} = \frac{1}{1000} \sum_{i=1}^{1000} (\boldsymbol{p}_i - \hat{\boldsymbol{\mu}})(\boldsymbol{p}_i - \hat{\boldsymbol{\mu}})^\dagger \tag{4.5}$$

where $\hat{\boldsymbol{\mu}}$ is the sample mean and $(\cdot)^\dagger$ denotes the conjugate transpose. For simulation, the number of modes is assumed to be known and is used to partition the eigenvectors of the SCM into the signal and noise subspaces. Finally, the noise subspace estimated this way is used to compute a modal-MUSIC spectrum using Equation (3.9).

100 SCM realizations are used to obtain 100 realizations of the modal-MUSIC spectrum. From this ensemble, means and variances are estimated for each mode at each frequency. The

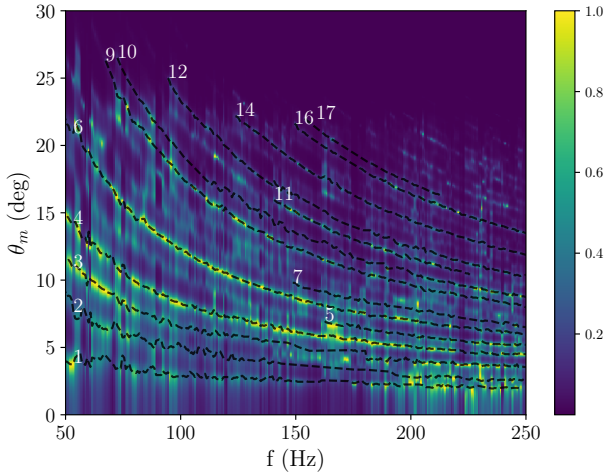


Figure 4.3. (Color online) Normalized modal dispersion curve from modal-MUSIC with identified modes superimposed as black dashed lines and labeled with white text.

results are shown in Figure 4.2 and demonstrate that, for the SWellEx-96 experiment at 20 dB SNR with 1000 snapshots, modal-MUSIC will estimate the low-order modes with high error variance but will estimate the higher-order modes accurately. Figure 4.2 also includes a red box indicating the modes that are both sensitive to bottom properties and estimated with low variance.

4.3.2 Number of propagating modes: cutoff modal angle

Attenuation is an important geoacoustic parameter for low-frequency shallow water MFP replicas. Baxley et al. [6] show that, within a certain source radius, the main effect of bottom attenuation on MFP replicas is the determination of the effective mode cutoff in the modal sum. That is, the attenuation determines the *number* of propagating modes, while the effect of attenuation on relative modal amplitudes within the set of propagating modes is relatively small.

For example, consider the Bachman geoacoustic model (Figure 4.1) for the SWellEx experimental region [13]. It has a halfspace compressional speed of 5200 m/s, meaning that, in the absence of attenuation, there are *propagating* modes with a modal angle up to 73 degrees. The presence of sediment attenuation adds a significant imaginary part to the modal wavenumbers

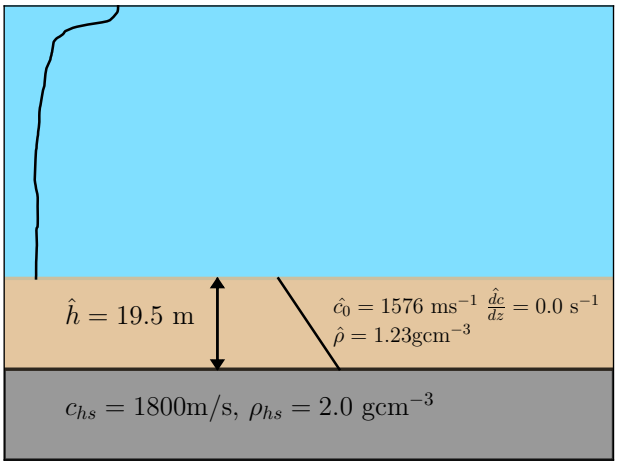


Figure 4.4. (Color online) Single layer model (not to scale) fit to the modal-MUSIC dispersion curve. This model, along with the modal cutoff angle shown in Figure 4.5, is used for range-coherent MFP results.

above a certain frequency-dependent cutoff mode angle $\theta_c(f)$ (an effective "modal critical angle"), so the modes with $\theta(k_r) > \theta_c(f)$ are stripped from the waveguide after a short range. However, the sediment attenuation is still low enough to allow modes below this modal critical angle to significantly affect the replicas out to a range of ten kilometers. An approximation to the replicas that uses the real part of the wavenumbers and an accurate, frequency-dependent modal critical angle will provide good agreement with data from within a certain radius.

Modal-MUSIC only produces estimates of the real part of the wavenumbers, and therefore carries no precise information on the sediment attenuation. Examination of the modal-MUSIC spectrum on data (for example, see Figure 4.3) suggests that modal-MUSIC may provide a good estimate $\hat{\theta}_c(f)$ of the modal cutoff angle. The compressional speed, density, and layer properties of the geoacoustic model can be fit to the modal-MUSIC spectrum, and then this model can be augmented with $\hat{\theta}_c(f)$. Together, these provide a reasonably accurate description of sound propagation in the shallow water waveguide without the need to explicitly include model attenuation.

4.4 Range-coherent matched field processing

By modeling source motion, range-coherent MFP [10] coherently combines snapshots from a moving source crossing multiple resolution cells. This section introduces notation necessary to present results in the following sections.

Consider a sequence of L snapshots taken at a frequency f. At each time t_i, the source range and depth are r_i and z_i. Then the snapshot at time t_i is written as

$$\boldsymbol{d}_i = S_i(f)\boldsymbol{p}(r_i, z_i) + \boldsymbol{n}_i \,, \tag{4.6}$$

where $\boldsymbol{n}_i$ is noise from the sensors, interferers, etc., $\boldsymbol{p}$ is the pressure field due to the source and $S_i(f)$ is the complex source spectrum at frequency f. For a perfectly narrowband source, $S_i(f) = Ae^{j2\pi f t_i}$, with A equal to the complex value $S_0(f)$ at $t = 0$.

Range-coherent MFP first removes the source phase from each snapshot to get a new sequence of measurements

$$\boldsymbol{d}'_i = \boldsymbol{d}_i e^{-j2\pi f t_i} = A\boldsymbol{p}(r_i, z_i) + \boldsymbol{n}'_i \,. \tag{4.7}$$

If the source is truly coherent and the frequency f is correctly specified, the effect of this correction is to a) add a phase shift to the noise and b) remove the effect of time on the phase of the pressure field due to the source. The leftover change in the pressure field is due only to change in source position, and can therefore be modeled.

For a candidate source track $\{r_i, z_i\}$, the model produces a sequence $\{\boldsymbol{p}(r_i, z_i)\}$. Both the snapshot set $\{\boldsymbol{d}'_i\}$ and modeled set $\{\boldsymbol{p}(r_i, z_i)\}$ are stacked into "supervectors"

$$\boldsymbol{d}_{sup} = [\boldsymbol{d}'^T_0 \;\; \boldsymbol{d}'^T_1 \;\; \ldots \;\; \boldsymbol{d}'^T_{L-1}]^T \tag{4.8}$$

and

$$\boldsymbol{p}_{sup}(\{r_i, z_i\}) = [\boldsymbol{p}_0^T(r_0, z_0) \ \ \cdots \ \ \boldsymbol{p}_{L-1}^T(r_{L-1}, z_{L-1})]^T \ , \tag{4.9}$$

with L denoting the number of snapshots to be coherently combined.

In the case of a source moving at a constant depth and a constant range-rate, source position can be modeled using three parameters: constant depth z, initial range r_0 at initial time $t = 0$, and constant range-rate $\dot{r}$. Multiple realizations of $\boldsymbol{d}_{sup}$ can be used to form a "super-covariance matrix" :

$$\hat{\boldsymbol{K}} = \frac{1}{N_s} \sum_{i=1}^{N_s} (\boldsymbol{d}_{sup}^{(i)} - \hat{\boldsymbol{\mu}})(\boldsymbol{d}_{sup}^{(i)} - \hat{\boldsymbol{\mu}})^\dagger \tag{4.10}$$

($\hat{\boldsymbol{\mu}}$ is the sample mean). The Bartlett ambiguity surface is then computed as

$$B(r_0, z, \dot{r}, f) = \frac{\boldsymbol{p}_{sup}(r_0, z, \dot{r}, f)^\dagger \hat{\boldsymbol{K}} \boldsymbol{p}_{sup}(r_0, z, \dot{r}, f)}{\left\| \boldsymbol{p}_{sup}(r_0, z, \dot{r}, f) \right\|^2 \mathrm{tr}(\hat{\boldsymbol{K}})} \tag{4.11}$$

where $\mathrm{tr}(\cdot)$ is the trace operation.

When multiple tonals are available, the ambiguity surfaces from each frequency are averaged in decibels [15]

$$B(r_0, z, \dot{r}) = \frac{1}{N_f} \sum_{k=1}^{N_f} 10 \log_{10} B(r_0, z, \dot{r}, f_k) \ . \tag{4.12}$$

Finally, the source depth, range and range-rate are estimated as the values r_0, z, and $\dot{r}$ that maximize the function (4.12).

For purposes of comparison, the range-coherent processor in Equation (4.12) can be computed without applying the source phase correction in Equation (4.7), yielding a range-"incoherent" matched field processor which performs similarly to matched field tracking [16]. The improvement of range-coherent over range-"incoherent" MFP is then attributable to coherent gain from combining snapshots from separate resolution cells, as opposed to tracking gain from

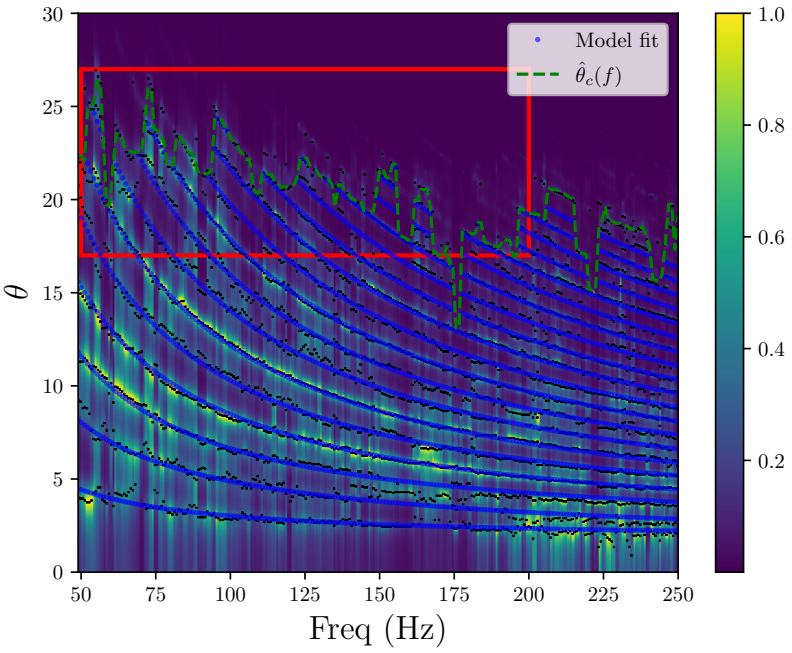


Figure 4.5. (Color online) Results of performing the geoacoustic fit to modal-MUSIC spectrum. Wavenumbers used for geoacoustic model fit are shown in the red box. The blue lines show the forward model with the optimized dispersion curve. Data points from modal-MUSIC are shown as black dots superimposed over the normalized modal-MUSIC spectrum. The green dashed line shows the mode angle cutoff used in replica calculation for MFP as a function of frequency. The color scale shows the value of the modal-MUSIC spectrum, normalized to the maximum value at each frequency.

incoherently combining conventional MFP surfaces.

4.5 Experimental results: Modal-MUSIC and model fit

This section details the application of modal-MUSIC and geoacoustic model fitting to experimental data collected during the SWellEx96 experiment [17]. These results are shown in Figure 4.5.

4.5.1 Modal-MUSIC dispersion curve

Modal-MUSIC is applied to a 30 minute segment of data from JD131 2232 GMT to 2302 GMT, collected on 21 elements with a 5.625 meter spacing. The first element is at a depth of 96

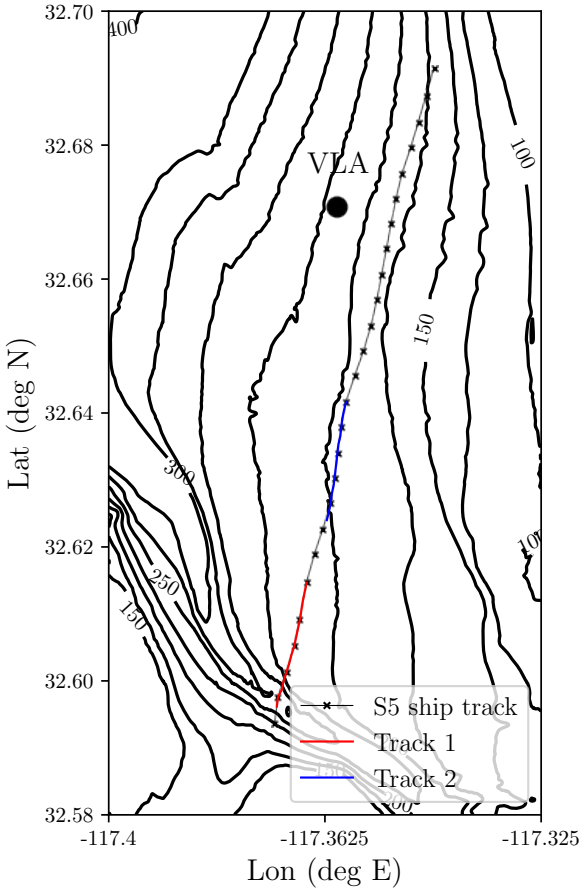


Figure 4.6. (Color online). Bathymetry and source position during the S5 event of the SWellEx96 experiment.

meters, and the total array length is around 120 meters. The dominant source of noise during this portion of data is thought to be from the research vessel transiting at ranges from approximately eight to ten kilometers.

Each element samples the pressure field at a 1500 Hz sampling rate. To extract the modal-MUSIC dispersion curve, snapshots at frequencies from 50 to 250 Hz are formed using a 2048 point FFT (1.36 second length) with a Hamming window, with no overlap between snapshots. This results in 1318 snapshots used to form each frequency SCM (frequency spacing of 0.73 Hz).

To compute the modal-MUSIC spectrum (Equation 4.1), the eigenvector decomposition of each SCM is computed, and the signal subspace rank at each frequency is estimated by

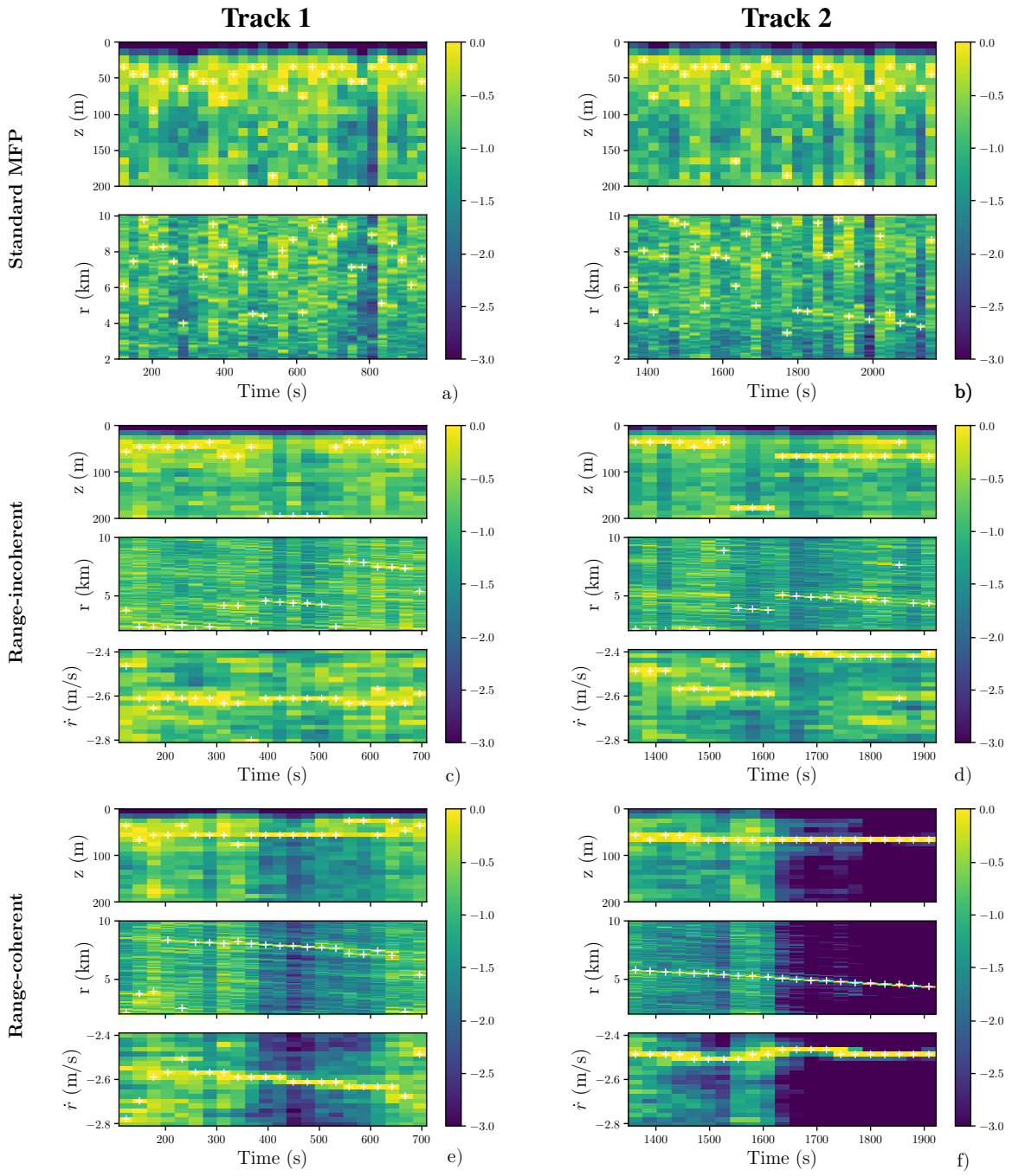


Figure 4.7. (Color online) Comparison between a), b) conventional MFP c), d) range-'incoherent' MFP and e), f) range-coherent MFP. White crosses denote maximum of the marginal ambiguity surface. The left column a), c) and e) correspond to Track 1. The right column b), d) and f) correspond to Track 2. Color scale is the value of Bartlett processor (Equation 4.12) in decibels, normalized to maximum beamformer output at each covariance integration period. Color range is from 0 dB to -3 dB. Range-"incoherent" and range-coherent MFP also include a range-rate vs. time ambiguity surface.

looking for the maximum change in the eigenvalue difference. Specifically, for eigenvalues sorted in decreasing order, the signal subspace dimension is defined as $\text{argmax}\{(\lambda_i - \lambda_{i+1})\}$ for $i = 1, 2, \ldots, 20$. Outliers from these subspace estimates are removed, and then the subspace estimates are smoothed over frequency by fitting them to a quadratic function. The smoothed subspace estimates are used to estimate the noise subspace matrix $\boldsymbol{E}_N$. A CTD from a nearby time in the experiment (depicted in Figure 4.1) is used for shooting candidate modes over the array.

Modes are then manually identified as smooth lines over frequency connecting the peaks in the dispersion curve. Figure 4.3 shows the modal-MUSIC spectrum as a function of frequency and mode angle θ_m with the identified modes, labeled. As seen in the figure, modes up to mode number 17 are identified in the dispersion curve.

4.5.2 Geoacoustic model fit

The emphasis here is to find a model that fits the modal-MUSIC curve well enough to perform MFP, rather than to find accurate estimates of the geoacoustic parameters with uncertainty quantification.

Precise mode estimate uncertainty cannot be known *a priori* since it depends on the (unknown) geoacoustic properties of the waveguide. However, the uncertainty analysis detailed in Section 4.3.1 shows that the estimate error variance is roughly constant for estimated modes with a phase speed greater than or equal to 1550 m/s ($\approx$ 16 degree mode angle), independent of frequency and mode number (see red box in Fig. 4.5). Furthermore, the objective function in Eq. 4.3 is insensitive to modes below this phase speed cutoff. Therefore the geoacoustic fit is carried out using the wavenumber estimates of modes above this phase speed cutoff. The presented results are insensitive to the precise choice of this cutoff phase speed.

The geoacoustic model consists of a single layer with constant density and linear SSP overlying a halfspace, as shown in Figure 4.4. The four parameters in the layer include the layer width h, the compressional speed at the top of the layer c_0, the sound speed gradient in the layer

$s = \frac{\mathrm{d}c}{\mathrm{d}z}$, and the layer density ρ assumed to be constant. The dispersion curve is insensitive to the halfspace parameters, as long as the compressional speed in the halfspace is high enough to support all the estimated modes. The maximum observed phase speed is 1750 m/s in Figure 4.3, so the halfspace speed is fixed arbitrarily at 1800 m/s and the density at 2 g·cm^{-3} for the optimization.

Equipped with this model parameterization, the optimal parameter vector $\boldsymbol{x} = [h \; c_0 \; s \; \rho]^T$ is found by minimizing the objective function in Equation 4.3. The layer thickness is constrained to lie between 5 and 100 meters, the compressional speed between 1540 and 1800 m/s, the gradient between 0 and 4 s^{-1}, and density between 1.2 and 2.0 g cm^{-3}. Wavenumbers for each candidate parameter vector $\boldsymbol{x}$ were computed using the Sturm sequence eigenvalue search [18] (the method used in KRAKEN [19]). The objective function is minimized by performing a sequential 1-d global optimization over each parameter, holding the others fixed at their latest optimized value. Six iterations were seen to provide good convergence. The optimal value is $\boldsymbol{x}_* = [19.5 \text{ m}, 1576 \text{ m/s}, 0 \text{ s}^{-1}, 1.23 \text{ g cm}^{-3}]^T$.

The wavenumbers from the optimized model are shown superimposed over the modal-MUSIC spectrum in Figure 4.5. The phase speed of the highest angle estimated mode is used as a phase speed cutoff for each frequency when calculating replicas.

4.6 Experimental results: Range-coherent matched field processing

This section details the localization of a quiet towed source using range-coherent MFP with the modal-MUSIC-derived model replicas described in the previous section. These results are compared to traditional MFP using replicas computed from the geoacoustic model shown in Figure 4.1. An additional comparison is made to range-"incoherent" MFP, where range-coherent MFP is carried out without correcting the source phase as in Equation 4.7. The incoherent approach will average down sidelobes that move at a different speed from the candidate search

speed, but will not get the coherent gain over noise from coherent snapshot comparison.

4.6.1 S5 event

The S5 event of the SWellEx96-experiment consists of a roughly straight source tow during a 75 minute period. During this time, a source towed at roughly 60 meters depth played five tonal sets each consisting of 13 tonals between 49 Hz and 400 Hz. Ten tones from the second tonal set, from 52 Hz to 238 Hz, are used for testing localization results. From the 75 minute event, two segments labeled "Track 1" and "Track 2" are selected for analysis. The source track during the event and these segments is shown in Figure 4.6, and includes source ranges from eight kilometers to four kilometers.

Snapshots are formed by taking a 16384-point Hamming-windowed FFT with a 50% overlap (5.46 second spacing). To simplify the presentation, prior knowledge of source range-rate was used to calculate the approximate Doppler-shifted signal bin from nominal frequency. Note that bin selection can be easily incorporated into range-coherent MFP since it searches over range-rate. The snapshot length results in an estimated mean SNR of -20 dB over the first ten quiet source tones at a source range of 8 km and -12 dB at 5 km source range.

Resolution cell size is roughly 60 meters for a 400 Hz source, corresponding to a 25 second integration time for a source moving 2.5 m/s. For a traditional MFP benchmark, five snapshots are used to form covariance matrices at the ten frequencies. Supervectors for range-"incoherent" and range-coherent MFP are formed by stacking ten snapshots, each separated by 27 seconds. This draws each snapshot in the supervector from a separate resolution cell. Then, five realizations of the supervector are formed by using five consecutive snapshots for the first snapshot in the supervector. These five realizations are averaged to form a "super" covariance matrix of dimension 210 by 210 (21 real vertical array elements times 10 synthetic horizontal elements).

The covariance matrices are then used to evaluate the Bartlett ambiguity surface on a grid of source range, depth and, for range-coherent and range-"incoherent" MFP, range-

rate. For traditional MFP, this yields a 3-d surface over time, range, and depth. For range-coherent/"incoherent" MFP, it yields a 4-d surface over time, range, depth, and range-rate. To visualize these results, 2-d marginal surfaces are formed by holding two axes fixed and taking the maximum over the other axis/axes. For example, a depth-time ambiguity surface for range-coherent MFP results can be formed by taking the maximum over all ranges and range-rates for each fixed covariance time and depth.

The results are shown in Figure 4.7. For the quiet tonal set, conventional MFP is unable to localize the source for either of the two segments. Range-"incoherent" MFP produces some weak estimates out to a range of around 5 kilometers before breaking down. Range-coherent MFP with the modal-MUSIC derived replicas successfully localizes the source out to a range of 8 kilometers (estimated SNR of -20 dB).

4.7 Conclusion

This work presents an approach for localizing a quiet coherent source without prior geoacoustic knowledge. This method requires knowledge of the water-column SSP down to the bottom of the array. Relevant geoacoustic information is extracted from acoustic measurements of a ship of opportunity using modal-MUSIC. This information is used to compute replicas for range-coherent MFP, which successfully localizes a multi-tone source at -20 dB SNR.

In this work, range-coherent MFP manages to extract gain by coherently comparing data from resolution cells that are separated by five minutes, or more than 800 meters in range. This is perhaps surprising, given that the source motion is not perfectly uniform during this interval and the replicas are not perfectly matched to the environment. This suggests that the phase evolution in range of the modal field is a quantity robust to both non-uniform motion and replica mismatch. One consequence of this is that range-rate estimates using range-coherent MFP should be successful, even in the event of insufficient SNR/replica accuracy to perform full source position localization. Importantly, this range-rate estimation takes advantage of vertical

array gain as well as coherent gain by comparing phase across many resolution cells.

The contribution of modal-MUSIC is to provide both parameter estimates for a simple geoacoustic model, as well as estimates of a frequency-dependent, modal cutoff angle $\hat{\theta}_c(f)$. The use of a model improves localization results over the purely data-based approach used in previous work [5]. This improvement is likely due to the reduction of error in the low-order modes, which should be well constrained by the SSP (which is presumed known) but are poorly estimated by modal-MUSIC. Therefore, the model fit ensures a physically consistent set of modes, which improves the replica accuracy.

4.8 Acknowledgements

This work was funded by the TFO program of the United States Office of Naval Research. Bill Hodgkiss and Dave Ensberg helped with data access for the SWellEx-96 experiment.

The text of this chapter, in full, has been submitted for publication of the material as it may appear in Akins and Kuperman (2023), *The Journal of the Acoustical Society of America*. The dissertation author was the primary investigator and author of this paper. The coauthor listed in this publication directed and supervised the research.

Bibliography

[1] J. Bonnel, Y.-T. Lin, D. Eleftherakis, J. A. Goff, S. Dosso, R. Chapman, J. H. Miller, and G. R. Potty, "Geoacoustic inversion on the new england mud patch using warping and dispersion curves of high-order modes," *The Journal of the Acoustical Society of America*, vol. 143, no. 5, pp. EL405–EL411, 2018.

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

[3] C. H. Harrison and D. G. Simons, "Geoacoustic inversion of ambient noise: A simple method," *The Journal of the Acoustical Society of America*, vol. 112, no. 4, pp. 1377–1389, 2002.

[4] G. V. Frisk, K. M. Becker, S. D. Rajan, C. J. Sellers, K. von der Heydt, C. M. Smith, and M. S. Ballard, "Modal mapping experiment and geoacoustic inversion using sonobuoys," *IEEE Journal of Oceanic Engineering*, vol. 40, no. 3, pp. 607–620, 2015.

[5] F. H. Akins and W. A. Kuperman, "Modal-MUSIC: A passive mode estimation algorithm for partially spanning arrays," *JASA Express Letters*, vol. 2, no. 7, p. 074802, 2022.

[6] P. A. Baxley, N. O. Booth, and W. S. Hodgkiss, "Matched-field replica model optimization and bottom property inversion in shallow water," *The Journal of the Acoustical Society of America*, vol. 107, no. 3, pp. 1301–1323, 2000.

[7] M. D. Collins and W. A. Kuperman, "Focalization: Environmental focusing and source localization," *The Journal of the Acoustical Society of America*, vol. 90, no. 3, pp. 1410–1422, 1991.

[8] S. E. Dosso and M. J. Wilmut, "Bayesian focalization: Quantifying source localization with environmental uncertainty," *The Journal of the Acoustical Society of America*, vol. 121, no. 5, pp. 2567–2574, 2007.

[9] A. M. Thode, G. L. DSpain, and W. A. Kuperman, "Matched-field processing, geoacoustic inversion, and source signature recovery of blue whale vocalizations," *The Journal of the Acoustical Society of America*, vol. 107, no. 3, pp. 1286–1300, 2000.

[10] F. Hunter Akins and W. A. Kuperman, "Range-coherent matched field processing for low signal-to-noise ratio localization," *The Journal of the Acoustical Society of America*, vol. 150, no. 1, pp. 270–280, 2021.

[11] R. Schmidt, "Multiple emitter location and signal parameter estimation," *IEEE Transactions on Antennas and Propagation*, vol. 34, no. 3, pp. 276–280, 1986.

[12] M. B. Porter and E. L. Reiss, "A note on the relationship between finite-difference and shooting methods for ode eigenvalue problems," *SIAM Journal on Numerical Analysis*, vol. 23, no. 5, pp. 1034–1039, 1986.

[13] R. T. Bachman, P. W. Schey, N. O. Booth, and F. J. Ryan, "Geoacoustic databases for matched-field processing: Preliminary results in shallow water off San Diego, California," *The Journal of the Acoustical Society of America*, vol. 99, no. 4, pp. 2077–2085, 1996.

[14] S. D. Rajan, J. F. Lynch, and G. V. Frisk, "Perturbative inversion methods for obtaining bottom geoacoustic parameters in shallow water," *The Journal of the Acoustical Society of America*, vol. 82, no. 3, pp. 998–1017, 1987.

[15] A. B. Baggeroer, W. A. Kuperman, and H. Schmidt, "Matched field processing: Source localization in correlated noise as an optimum parameter estimation problem," *The Journal of the Acoustical Society of America*, vol. 83, no. 2, pp. 571–587, 1988.

[16] L. T. Fialkowski, J. S. Perkins, M. D. Collins, M. Nicholas, J. A. Fawcett, and W. A. Kuperman, "Matched-field source tracking by ambiguity surface averaging," *The Journal of the Acoustical Society of America*, vol. 110, no. 2, pp. 739–746, 2001.

[17] J. Murray and D. Ensberg, "The swellex-96 experiment," *URL: http://www. mpl. ucsd. edu/swellex96*, 1996.

[18] M. Porter and E. L. Reiss, "A numerical method for oceanacoustic normal modes," *The Journal of the Acoustical Society of America*, vol. 76, no. 1, pp. 244–252, 1984.

[19] M. B. Porter, "The KRAKEN Normal Mode Program," tech. rep., SACLANT Undersea Research Centre, La Spezia, Italy Rep. SM-245, 1991.

Chapter 5

Conclusion

This dissertation presents two new passive acoustic methods to extend MFP to the localization of a low SNR source with limited prior environmental information. MFP has previously been applied to the localization of high SNR sources in the presence of limited environmental information [1, 2, 3]. Chapter two of this dissertation demonstrated successful passive localization of a low SNR source by exploiting signal coherence between multiple resolution cells. Chapter three of this dissertation presented a method called modal-MUSIC for estimating modal wavenumbers from noise on a partially-spanning array. The estimated modal wavenumbers and mode shapes are used in place of a full environmental model to perform localization at high SNR. Chapter four fit a seabed model to the modal-MUSIC estimated wavenumbers of chapter three to obtain a sufficiently accurate model for low-SNR MFP using the range-coherent MFP method of chapter two.

Taken together, the studies in this dissertation indicate that passive acoustic measurements with a drifting VLA may compose a sufficient standalone system for detection and ranging in shallow water. To fully realize this goal, further work must address the need for a water column SSP to estimate bottom parameters with modal-MUSIC.

Bibliography

[1] M. D. Collins and W. A. Kuperman, "Focalization: Environmental focusing and source localization," *The Journal of the Acoustical Society of America*, vol. 90, no. 3, pp. 1410–1422, 1991.

[2] S. E. Dosso and M. J. Wilmut, "Bayesian focalization: Quantifying source localization with environmental uncertainty," *The Journal of the Acoustical Society of America*, vol. 121, no. 5, pp. 2567–2574, 2007.

[3] A. M. Thode, G. L. DSpain, and W. A. Kuperman, "Matched-field processing, geoacoustic inversion, and source signature recovery of blue whale vocalizations," *The Journal of the Acoustical Society of America*, vol. 107, no. 3, pp. 1286–1300, 2000.

[4] J. Murray and D. Ensberg, "The swellex-96 experiment," *URL: http://www. mpl. ucsd. edu/swellex96*, 1996.