Spatial Spectrum Estimation with a Maneuvering Sensor Array in a Dynamic Environment

by

Jonathan L. Odom

Department of Electrical and Computer Engineering
Duke University

Date: ______________________

Approved:

__________________________

Jeffrey L. Krolik, Supervisor

__________________________

Loren W. Nolte

__________________________

Matthew S. Reynolds

Thesis submitted in partial fulfillment of the requirements for the degree of Master of Science in the Department of Electrical and Computer Engineering in the Graduate School of Duke University

2011
Abstract

Spatial Spectrum Estimation with a Maneuvering Sensor Array in a Dynamic Environment

by

Jonathan L. Odom

Department of Electrical and Computer Engineering
Duke University

Date: __________________________

Approved:

___________________________
Jeffrey L. Krolik, Supervisor

___________________________
Loren W. Nolte

___________________________
Matthew S. Reynolds

An abstract of a thesis submitted in partial fulfillment of the requirements for the degree of Master of Science in the Department of Electrical and Computer Engineering in the Graduate School of Duke University 2011
Abstract

Estimation of a time-varying field is essential for situational awareness in many subject areas. Adaptive processing often assumes both the field is stationary and the array is fixed for multiple observation windows. For passive sonar, highly dynamic scenarios such as high bearing rate sources or underwater maneuvers severely limit the utilization of multiple snapshots. Several models are considered for time-varying fields, and a broadband maximum-likelihood estimator is introduced that is solved with an expectation maximization algorithm using as few as one snapshot. The number of estimated parameters can be reduced for broadband data when information, such as shape, is known about the source temporal spectrum. Cramér-Rao bound analysis is used to understand the effects of temporal spectrum knowledge on broadband processing. An example is given for the flat spectrum case to compare with conventional processing. Another feature of dynamic environments is array motion. Since underwater arrays are often subject to motion, the estimate must consider arbitrary, dynamic array shapes. Platforms such as autonomous underwater vehicles provide mobility but constrain the number of sensors. Exploiting a maneuverable linear array with the new estimate allows for left-right or front-back disambiguation and suppression of spatial grating lobes. Multi-source simulations are used to demonstrate the ability of a short, maneuvering array to reduce array backlobes as well as operate in the spatial grating lobe region.
**Contents**

Abstract iv

List of Figures vii

Acknowledgements viii

1 Introduction 1

2 Passive Array Processing 4

2.1 Synthetic Aperture Sonar . . . . . . . . . . . . . . . . . . . . . . . . 7

2.2 Adaptive Processing . . . . . . . . . . . . . . . . . . . . . . . . . . . 8

2.3 Estimation Techniques . . . . . . . . . . . . . . . . . . . . . . . . . . 11

2.4 Snapshot Deficient Processing . . . . . . . . . . . . . . . . . . . . . 13

2.5 Other Related Work . . . . . . . . . . . . . . . . . . . . . . . . . . . 16

3 Frequency Domain Data Model 17

4 Narrowband Spatial Spectrum Estimation 20

4.1 Deterministic Dynamic Model . . . . . . . . . . . . . . . . . . . . . 20

4.2 Stochastic Dynamical Model . . . . . . . . . . . . . . . . . . . . . . 22

4.3 Narrowband Simulation . . . . . . . . . . . . . . . . . . . . . . . . . 27

5 Broadband Spatial Spectrum Estimation 31

5.1 Broadband Extension . . . . . . . . . . . . . . . . . . . . . . . . . . 31

5.1.1 Signal Model . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 31

5.1.2 Parameter Reduction . . . . . . . . . . . . . . . . . . . . . . . . . 32
## List of Figures

<table>
<thead>
<tr>
<th>Figure</th>
<th>Description</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>2.1</td>
<td>Overview of processing techniques</td>
<td>5</td>
</tr>
<tr>
<td>2.2</td>
<td>Representation of steering and sensor vector</td>
<td>6</td>
</tr>
<tr>
<td>3.1</td>
<td>Polar coordinate system</td>
<td>18</td>
</tr>
<tr>
<td>4.1</td>
<td>True BTR for simulations</td>
<td>27</td>
</tr>
<tr>
<td>4.2</td>
<td>Narrowband BTR for ideal array</td>
<td>28</td>
</tr>
<tr>
<td>4.3</td>
<td>Various narrowband BTRs using short mobile array</td>
<td>30</td>
</tr>
<tr>
<td>5.1</td>
<td>Cramér-Rao bounds of power estimates</td>
<td>35</td>
</tr>
<tr>
<td>5.2</td>
<td>Fisher information matrix inverse</td>
<td>36</td>
</tr>
<tr>
<td>5.3</td>
<td>Cramér-Rao bound of power estimate with aliasing</td>
<td>37</td>
</tr>
<tr>
<td>5.4</td>
<td>Broadband BTR for circular array</td>
<td>39</td>
</tr>
<tr>
<td>5.5</td>
<td>Various broadband BTRs using short mobile array</td>
<td>40</td>
</tr>
<tr>
<td>5.6</td>
<td>ROC curve with target SNR -25 dB using $M$-of-$N$ detector</td>
<td>41</td>
</tr>
<tr>
<td>5.7</td>
<td>Broadband BTR for circular array with 3m spacing</td>
<td>42</td>
</tr>
<tr>
<td>5.8</td>
<td>Various broadband BTRs using short mobile under-sampled array</td>
<td>44</td>
</tr>
<tr>
<td>5.9</td>
<td>ROC curve with target SNR -25 dB using $M$-of-$N$ detector for under-sampled array</td>
<td>45</td>
</tr>
</tbody>
</table>
Acknowledgements

I would like to thank Dr. Jeffrey Krolik for his support and direction over the last 4 years as I transitioned from undergraduate to graduate work. Also, I would like to thank Dr. Loren Nolte and Dr. Matthew Reynolds for serving on my master’s committee. Special thanks to Dr. Jeffrey Rogers who laid the foundation for the work that I have extended in this thesis and with whom I actively collaborate.

Thanks to the numerous lab mates and other students who have offered generous help my entire time at Duke.

A very special thanks to Ashley, my wife, for supporting me in everything that I do. She has read far too many of my papers and yet continues to be my editor.

This work was supported by ONR.
Spatial spectrum estimation is a common problem in the areas of radar, sonar, astronomy and seismology. This is closely related to signal detection and direction of arrival estimation. The spatial spectrum is an extension of the temporal spectrum into the three physical dimensions. The differences in each of these areas of study is the underlying assumptions for the medium and propagation models. The work described here is situated as a sonar problem but can easily be adapted for other mediums or methods of transmission. For ease in understanding, temporal frequency methods are well-developed and provide an analogy to understand the spatial frequency techniques. Frequency-domain filters are used to shape power spectra, and the rise of Fourier analysis aids in the ability to analyze physical phenomenon. When using a single stationary receive element, a time series signal is often the only data available. However, when multiple receivers are available, the spatial dimensions provide additional exploitable information. A simple approach is to model the medium as free space then combine data across space. When the receive elements have synchronized clocks, the data can be combined in various ways in order to focus a beam and listen to a particular location, known as beamforming or spatial filtering; a com-
prehensive overview was written by Van Veen and Buckley [1]. The spatial spectrum represents the power coming from each direction. In sonar, sea vessels and marine life can be tracked from the emitted acoustic waves, refer to [2] for the earliest experiments. Navel military operations, such as anti-submarine warfare, and marine biologists are interested in the location of underwater sources.

The two general methods of localization considered are sonar (sound navigation and ranging) and radar (radio detection and ranging). The two technologies use similar mathematical frameworks and have complimentary literature. Sonar and radar operate in two broad categories: passive and active. Active methods send a known waveform and then analyze the received return to determine the location (bearing and range) and speed. Passive methods listen for incoming unknown waveforms and determine location, where angle is frequently the parameter of interest. A trade-off exists where active techniques result in more information but require signal transmission and a reflective target. Since passive techniques do not transmit a signal, they do not increase the detectability of a sensing platform and are not limited by transmission requirements. For these reasons, passive processing is an active research area [3].

Underwater arrays are frequently towed, which creates a computationally difficult problem due to platform motion. Iterative methods have been developed to account for motion of the platform holding the sensor array, but little work has been done for the case when sources also move. Motion of the platform can be exploited to improve the performance of the estimate. Towed arrays have reduced usable bearing space due to geometric ambiguities and poor endfire resolution. This paper suggests the use of short maneuvering arrays to regain usable bearing space by combining data from several positions to provide more uniform coverage. Longer arrays are expensive and provide high drag; short arrays can approximate the performance of larger arrays by moving the array. Assuming arbitrary and changing (but known)
array shape, narrowband deterministic and stochastic spatial spectrum models are used to derive adaptive estimates with a sliding window of data. The deterministic model is inspired by radar literature that models beamforming weight vector changes as an array moves [4]. The stochastic approach uses an explicit time-varying model, resulting in a generalization of previous work by Rogers and Krolik [5]. Traditionally, broadband power estimates coherently or incoherently sum the narrowband estimates. However, knowledge of the temporal spectra must be incorporated into the broadband spatial spectrum estimate. Temporal spectrum information is explored for parameter reduction and is shown to provide performance improvements with a simple flat temporal spectrum assumption.

A review of relevant passive array processing schemes is given in Chapter 2. Chapters 3 and 4 introduce two time-varying narrowband spatial spectrum models and estimates. In Section 4.3, the performance of each estimate is compared using simulated data from an interference-dominated environment with a moving source. Simulations are presented with the assumption that a single snapshot is available for each array location in order to ensure few snapshots are required. The broadband model is introduced in Chapter 5, and Cramèr-Rao bounds are used to compare the traditional and proposed broadband approaches. This motivates a new broadband estimator with derivation in Section 5.2 and a simulation demonstrating the estimate working correctly, interestingly, in the spatial grating lobe region of the array.
Spatial spectrum estimation is often used for target tracking and classification in passive sonar analysis. In two dimensions, the estimate corresponds to a map of power as a function of bearing angle, also known as a field directionality map. When the field is dynamic across time, a common representation of the estimate is a bearing-time-record, or BTR. These plots describe the power of sources over the 360° surrounding a platform and provide a history across time. The maps represent the spatial spectra for situational awareness in both human and computer decision processes. Tracking algorithms utilize BTRs and motion models in order to estimate and track target motion [6], but the new work presented here considers field dynamics in the general case. This section provides an overview of the previous work on passive moving arrays.

The simplest processing scheme operates on each observation or dataset separately. Delay-and-sum beamforming processes each set of data independently or averages observations in time, assuming the field does not change. When the method of processing is independent of the received data, it is referred to as conventional processing. For increased performance, synthetic aperture techniques coherently com-
bine observations across space in order to form a virtual aperture. Finally, adaptive techniques frequently assume multiple observations, also known as snapshots, are received at each location. A graphical representation is given in Figure 2.1 where different locations are represented by data \( a, b, \text{ and } c \). The different datasets can represent either the same array at different locations or a different array at each location. The numbers 1 to 3 represent multiple snapshots, which means data is collected multiple times at the same location. Special attention must be given when adaptive techniques are used but few snapshots are available, known as snapshot deficient processing. An example of this case, shown in Figure 2.1, is when only the grayed out squares are available. This represents a situation where only a few (or only one) snapshots are observed at each location. All of the methods in the illustration are covered in this chapter.

![Figure 2.1: Overview of processing techniques](image)

The most straightforward approach is to use conventional beamforming, which is also referred to as delay-and-sum or Bartlett beamforming. The output of each element is multiplied by a complex weight corresponding to the phase delay assumed by propagation. The far-field model assumes that the source distance is much larger than the total length of the array such that the signal propagates across the array in a single line, or plane. For a single source, frequency \( (\omega) \), and location \( (r) \), plane-wave propagation assumes the received signal is given by \( s(r, t) = A \exp(j(\omega t - k^T r(t))) \). Vectors are denoted by bold face variables, \( \mathbf{x} \), and matrices are capitalized, \( \mathbf{X} \). Vector
and matrix transpose is given by $^T$ and conjugate transpose (or Hermitian) is given by $^H$. The spatial frequency is represented by $k$ and defined concretely in Chapter 3 with the data model. The assumed wavenumber vector is referred to as the steering vector and points in the direction of wave propagation, perpendicular to the plane of constant phase. The phase component $k^T r$ is the only component needed for a given frequency. The amplitude, defined by $A$, is used to find source distance when the true source amplitude is known. The spatial dependence of $A$ is ignored in this model. The physical representation is given by Figure 2.2 for the $q$th source and $m$th sensor. Thus a beam can be formed by applying the reverse phase shift to the received signal. When these weights are applied to the received signals, the compact notation is given by (2.1) where $w$ denotes the weights.

$$y = w^H x$$

(2.1)

![Figure 2.2: Representation of steering and sensor vector](image)

This thesis derives algorithms assuming narrowband signals, which can be obtained from demodulated signals so that time, $t$, disappears from the model (but not forgotten). Furthermore, the wavenumber and source location are assumed to be stationary for a given data collection of time $T$, or snapshot, so that $k(t) \approx k(n)$ and $r(t) \approx r(n)$. This approximation is limited by the bandwidth desired, physical size of the array, and source motion. The general rule of thumb is to require the bandwidth-time product, $T \times BW$, to be larger than 16. At broadside, the resolution of a linear
array is $\approx \lambda/D$, or temporal wavelength divided by aperture length. Therefore, the resolution bin size is decreased with larger aperture sizes. The resolution also decreases away from broadside, resulting in less usable bearing space.

2.1 Synthetic Aperture Sonar

Underwater arrays are frequently pulled behind surface or submersible vessels. Synthetic aperture sonar combines data coherently from different array locations as it is towed through the water. A larger, virtual aperture is created in the process. The term coherently refers to the assumption that data observed from different locations have known phase differences. Synthetic aperture sonar was originally developed by Yen and Carey in order to increase resolution performance by opportunistically utilizing tow ship movement [7]. Their proposed processing scheme operates on array beamformer output, $b(\omega, \theta_s)$ as a function of frequency, $\omega$, and steering direction, $\theta_s$. Thus the new beams can be created by applying

$$b(\omega, \theta_s) = b(\omega, \theta_s) \exp \left[ j\omega\tau \left(1 - \frac{v}{c}\sin \theta \right) \right], \quad (2.2)$$

where $\frac{v}{c}\sin \theta$ is the distance travelled by the array toward (or away from) a source and $\tau$ is the time between samples. This knowledge allows beams to be shifted and summed together as long as the signal remains coherent in space and time. Instead of compensating for each steering direction, Stergiopoulos and Urban noted that the FFT could be used to compensate for all directions simultaneously and then simply choose the peak [8]. Beams can only be combined using FFTs with an array that is not turning. They also showed that cross-correlation of received signals can be used to estimate the phase compensation required for different array locations. Carey has continued the work, preforming underwater unmanned vehicle (UUV) towed array experiments with broadband signals [9]. This can alternatively be described in element space, where each physical aperture is used to expand the virtual aperture.
Increasing the number of virtual elements increases the number of dimensions of the space where sources are allowed to exist. The major limitations of this approach are sensitivity to model mismatch and increasing location error with synthetic array size because each array location is assumed to be coherent. This assumption will not generally hold with multiple platforms or distributed arrays. Synthetic aperture approaches allow for increased resolution but do not provide adaptive interference suppression.

2.2 Adaptive Processing

When the environment is dominated by loud interferers, performance is limited by the ability to suppress sources. Interferers are point or distributed sources other than the target and noise. For this discussion, noise is created by individual sensor noise or special cases of isotropic ambient sounds with half-wavelength array element spacing. The ability to suppress sources is based on the structure of the interfering fields. The most common adaptive method in array processing is minimum variance distortionless response (MVDR) beamforming developed by Capon [10]. The beamformer has weights that minimize the power from non-steered directions but require unity gain in the steering direction, $v$.

$$w_{MVDR} = \frac{R^{-1}v}{v^HR^{-1}v}$$ (2.3)

The weights depend on a statistic of the received data, $R = E[xx^H]$. The weights can also be derived assuming knowledge of the signal covariance matrix and will result in an equivalent solution when the steering vector is equal to the source vector. This approach cannot be used directly with a maneuvering array because the received data is a function of time and is therefore difficult to estimate. However, this weighting scheme, shown in (2.3), appears in many processing methods. Estimating
is accomplished by combining snapshots in either frequency or spatial dimensions. While synthetic approaches expand the dimensionality of the received data, many adaptive techniques reduce the space where processing is performed by combining snapshots or constraining the space considered. Historically, a spatial approach was proposed by Wagstaff for combining data incoherently as an array was towed through the water [11]. The method operates on beamformer outputs but takes into consideration the effects of the beampattern. By incoherently combining outputs, the field is reconstructed using knowledge of how the beamformer captures sounds from all directions. The concept is expressed by the following equation,

\[
\hat{r}_{ia} = \frac{1}{2\pi} \int_{0}^{2\pi} \hat{n}(\theta) b_i(\theta - \theta_a)
\]

where \( r \) is the beamformer output, \( n \) is the field power, and \( b \) is the beampattern as a function of steering direction and array orientation respectively. Estimates are denoted with a hat (\( \hat{\cdot} \)). This equation highlights that energy from non-steered directions must be considered as the beams are combined. The estimation process then iterates in an ad-hoc way until it converges to a fit.

Similarly, the data from different array orientations or configurations can be combined by correcting for the different terms. Wagstaff’s work demonstrate that the structure from the beamformer must be accounted for, while Zeira and Friedlander proved that a performance increase is possible using the underlying data from a time-varying array. Using Cramér-Rao bounds, they proved that using a changing array configuration over time improves performance. This is due to diversity of the array from different shapes or orientations [12]. Then they introduced several eigenstructure-based techniques by summing data from multiple array configurations with focusing matrices [13]. A significant body of work exists on frequency domain focusing matrix (see work by Wang and Kaveh [14]) often referred to as coherent signal sub-space methods (CSM). With this method, each signal has a single eigen-
vector across temporal or spatial dimensions. For example, assume there exists a transform matrix $\mathbf{T}$ such that the narrowband covariance matrix can be expressed in terms of a frequency-independent covariance matrix with rank determined by the number of signals.

$$
\mathbf{R}_{\text{CSM}} = \sum_{b=1}^{B} \mathbf{T}(\omega_b) \mathbf{R}(\omega_b) \mathbf{T}^H(\omega_b)
$$

(2.5)

The solution depends on the choice of $\mathbf{T}$ and requires knowledge of the source locations. A more general formulation that does not require known source locations was developed by Krolik and Swingler using a steered covariance matrix (STCM) [15]. This approach steers to each direction instead of assuming all directions can be captured by the same transformation.

$$
\mathbf{R}(\theta)_{\text{STCM}} = \sum_{b=1}^{B} \mathbf{T}(\omega_b, \theta) \mathbf{R}(\omega_b) \mathbf{T}(\omega_b, \theta)
$$

(2.6)

The differences can be seen from CSM in (2.5) to STCM in (2.6) by the dependence on bearing $\theta$. By choosing $\mathbf{T}$ to be a diagonal matrix of the steering vector for direction $\theta$, the STCM provides a better statistic with lower variance for limited snapshots but requires more computations. Calculating a different STCM for each direction increases the computational complexity. Friedlander and Zeira extended the focusing work by transforming data from each array position and summing over time instance instead of frequency. Their work allows the array to vary arbitrarily between snapshots. MVDR weights or data independent weights can then be applied to the steered estimate. Separately, this concept was applied to towed arrays by Greening and Perkins who provided preliminary results with experimental data [16]. The time-changing outputs were not provided in the paper, but time cuts of the power vs bearing show that array curvature can be captured using focusing matrices.

Work on towed arrays prior to the experiment by Greening and Perkins generally
ignored time segments when the array was not straight. Adaptive beamforming performs poorly when the true array shape does not match the assumed shape [17]. Gerstoft et al. showed that an estimated array shape can be used to form MVDR weights during a turn [18]. The tow ship GPS location and heading are used in conjunction with a water-pulley model for the array in order to estimate the location of each sensor. The resulting estimate is accurate enough to provide steering vectors for a 128 horizontal ULA. The sensitivity to sensor location error is reduced by diagonal loading, which artificially increases the spatial white noise level [19].

2.3 Estimation Techniques

Directionality mapping can also be directly formulated as an estimation problem instead of using beamforming, but beamforming may appear in the final solution. The weights for the MVDR beamformer can be derived with several criterion: least square error, maximum signal-to-noise ratio, or maximum likelihood; each solution is proportional to the others, differing only by a scalar [20, p 91]. Several common types of estimators in statistical signal processing find the parameter that maximizes a probability density function (pdf). The most complete estimate is based on the prior knowledge of the parameter and the density conditioned on the observed data. This is the maximum a posterior (MAP) estimate. When the prior is assumed to be uniform, the simplified estimate is equivalent to the maximum likelihood (ML) estimate (or MLE). Using large arrays and only a few sources, the dimensionality of the data quickly becomes problematic. Optimization with a large number of dimensions is computationally intractable. Each sensor adds a new dimension, and arrays can have hundreds of sensors. Previous discussions in this paper have considered element-space, which is the space identified with $\mathbb{R}$. More formally, this corresponds to the vector space spanned by the columns of $\mathbf{R}$. The matrix is complex-valued with size $M \times M$, the number of sensors by number of sensors. Sources are generally contained
by a smaller subspace or approximated to reside in a small space. Using physical
knowledge of the array and propagation, a parametric model reduces the number of
dimensions where optimization must occur.

For a given number of sources, the problem transforms into one of estimating the
power level and direction of arrival. When the number of sources is unknown, the
surveillance area of interest is divided into a grid with each grid point corresponding
to a source. In this context, passive sonar methods traditionally model the signals
stochastically and do not generally estimate the time-series signal. Deterministic
signal modeling is used in communications and active sonar where there is more
knowledge of the transmitted signals. Using a grid model, the problem reduces to
covariance matrix estimation over the grid, and the matrix diagonal is source power
[21]. Methods for structured covariance matrix estimation where first described by
Burg et al. [22]. Approaches for direction-of-arrival (DOA) estimation rely heavily on
the expectation maximization (EM) algorithm in order to obtain feasible numerical
solutions. Credit is given to Dempster et al. for unifying EM algorithms into a formal
statistical approach [23]. ML and MAP estimators can be solved using this approach.
The random variables are separated into the incomplete (or observable) data, and
the complete (or unobservable) data. The expectation step calculates the expected
value of an objective function of the complete data given the observed data and a
previous estimate of the parameters. The maximization step then maximizes the
result with respect to the parameters. An ML estimate is derived by using the log-
likelihood function of the parameter as the objective function, and the expectation
step is given as

\[
z = E_C \left[ \ln f(C|\Sigma) \mid x, \Sigma^{old} \right]
\]

where the complete data is \( C \) and the parameter of interest is \( \Sigma \). Feder and Weinstein
were the first to suggest the use of the EM algorithm for parametric signal models.
The work on deterministic signals proposed defining the complete data as the signals prior to mixing, or propagation, and the incomplete data as the received signals [24]. Miller and Fuhrmann extended the work to stochastic narrowband waveforms by jointly estimating the direction-of-arrival and signal [25]. The work was then applied to broadband signals by Chung and Böhme but only considers direction-of-arrival [26]. The field directionality or mapping problem has been solved with an EM algorithm in the context of astronomy by various authors, most recently Lanterman [27–29], which corresponds closely to the problem in passive sonar. Research in astronomy is focused on creating images of the field intensity. Using a complex normal model in a stationary field with uncorrelated sources, the iterative solution is written,

\[ \Sigma_{\text{new}} = \frac{1}{N} \left( \Sigma_{\text{old}}^2 \right) \sum_{n=1}^{N} \left( \mathbf{D}^H(n) \mathbf{K}^{-1}(n) \mathbf{D}(n) - \mathbf{D}^H(n) \mathbf{K}^{-1}(n) \hat{\mathbf{R}}(n) \mathbf{K}^{-1}(n) \mathbf{D}(n) \right) \]

(2.8)

where \( \mathbf{K}(n) = \mathbf{D}(n) \Sigma_{\text{old}} \mathbf{D}^H(n) + \sigma_n^2 \mathbf{I} \) and \( \hat{\mathbf{R}} = \frac{1}{N} \sum_{n=1}^{N} \mathbf{x}(n) \mathbf{x}^H(n) \). The variance of the noise is given by \( \sigma_n^2 \). Each source is assumed to be uncorrelated so that the update is uncoupled given the previous iteration’s estimate; however, estimates are not uncoupled between iterations.

2.4 Snapshot Deficient Processing

The data-dependent techniques discussed thus far assume the field is stationary and many snapshots are available. Consider a fast-moving target crossing an angular resolution bin between snapshots. In the limit, the field is only stationary for a single snapshot. The problem for many adaptive methods is the estimation of the covariance matrix, which must be inverted for optimal beamforming. When the field is assumed stationary but a source crosses a resolution bin, the eigenvalues spread and result in null broadening. The nulling performance of adaptive beamformers is
reduced and determined by the number of resolution bins crossed during all snapshots considered [30]. One approach is to limit the number of snapshots used to estimate the covariance matrix. The eigenstructure can be used to estimate the inverse of the covariance without requiring it to be full rank. For direct inversion, the matrix must be full rank: the rule of thumb is that the minimum number of snapshots for numerical stability is two times the number of elements (2M). Abraham and Owsley introduced a method that does not constrain the eigenvector space but assumes they are formed from unknown signal vectors. The eigenvector decomposition of the covariance matrix is given by (2.9). The largest eigenvalues, \( \lambda \), correspond to the largest signals, and the eigenvectors, \( e \), correspond to linear combinations of signal vectors. With this assumption, the covariance matrix can be approximated with the \( D \) largest eigenvectors as given by (2.10).

\[
R = \sum_{q=1}^{Q} \lambda_q e_q e_q^H
\]  

(2.9)

\[
\hat{R} = e \sum_{d=1}^{D} (\lambda_d - \sigma^2) e_d e_d^H + \sigma^2 I
\]  

(2.10)

Using properties of eigenvectors, the inverse of the covariance matrix can be obtained by

\[
R^{-1} = \sum_{q=1}^{Q} \lambda_q^{-1} e_q e_q^H
\]  

(2.11)

This can be used to form an estimate of the covariance matrix by using estimates of the \( D \) largest eigenvectors instead of a full rank matrix. Beamforming with this estimate is called Dominant Mode Rejection (DMR) [31]. However, this method is limited by eigenvalue decomposition of the sample covariance matrix.

Recently, methods have been introduced to reduce the number of snapshots required by utilizing special knowledge of the eigenstructure. This amounts to com-
bining the methods discussed in Section 2.2 and Section 2.3. Kraay and Baggeroer used the term “physically constrained maximum likelihood” to describe estimates which only consider signal vectors that are physically possible. The basis of the update equation, shown in (2.12), considers the possible spaces summed over \( q \), where weights are included with \( W \). The power from a particular direction is given by \( P \), which is estimated using an MVDR beamformer with \( \mathbf{R}_{n-1} \). The concept has also been applied to a least squares solution and extensions for sparse spectra [32, 33].

\[
[R_{n}]_{i,j} = \sigma_{n-1}^2 \delta_{i,j} + \sum_{q} P_{n-1}(\omega, \mathbf{k}_q) e^{-jk_{q}^{T}(r_i-r_j)} \tag{2.12}
\]

The physical constraint is already incorporated in previous estimation techniques as a direct result of the model, which can be seen by considering the \( K \) term in (2.8). The estimation technique considers a \( Q \) point grid and enforces the physical constraints of (2.12) by forming \( K \) from \( D \). Also, the estimate from (2.8) does not require sample matrix inversion. Thus, the estimation techniques can be applied to snapshot deficient scenarios under modified stationary assumptions. Rogers and Krolik developed this approach for long towed arrays. The EM requirements for field stationarity are relaxed by approximating the field as slowly changing between snapshots. Using Markov assumptions with a state transition pdf

\[
f(\Sigma(n)|\Sigma(n-1)) = \delta(\Sigma(n) - \Sigma(n - 1)), \tag{2.13}
\]

the time-varying estimate of the field can be approximated by using a small sliding window of snapshots [5]. The estimate can also be accomplished online or recursively by setting the size of the window to one and using the previous estimate to initialize the EM algorithm [34]. The solution is of the same form as (2.8). Broadband estimates are then formed by incoherently summing the estimates from narrowband data. The method reduces ambiguous backlobes and increases endfire resolution in the same way as the methods for stationary fields. This thesis proposes a similar
approach then applies it to short array scenarios. The new work presented in Section 4.2 is a generalization of the work by Rogers and Krolik.

2.5 Other Related Work

*Derivative Based Updating (DBU)*

The problem of moving arrays also occurs in radar with a few solutions that have not been considered in-depth for underwater applications. Hayward suggests the use of deterministic modeling of beamforming weights as an array moves [4, 35]. Using a linear model, the weights are expressed

\[ w \approx w_o + n\dot{w} \quad (2.14) \]

where a dot (′) is used to express the derivative. The simple use of Taylor series expansion is shown to help maintain nulls for rotating arrays.

*Subband Energy Detection*

For sonar, an ad-hoc method for creating BTRs was developed by Bono et al. called Subband Energy Detection (SED). The algorithm used in real systems is Subband Peak Energy Detection (SPED) [36]. The interesting feature of this algorithm is that it combines data across frequencies in a non-traditional manner. Knowledge of temporal spectral shape can increase the performance of target detection and localization, but few algorithms have been developed to exploit this [37]. SPED considers the beamformed output at each frequency and locates the peaks in azimuth/bearing. This is then weighted and summed over frequencies to form the BTR for a given time. SPED is currently being implemented as the modern approach in real sonar systems [3].
Assume an array of $M$ sensors collects band-limited discrete-time data. For a finite time $T$, the array and field are assumed to be stationary. The observation time, $T$, is longer than the maximum time required for the phase of a wavefront to travel across the length of the array. Each period of time is referred to as a time instance indexed by $n$. Time series data is transformed to the frequency domain through a Discrete Fourier Transform. Initially, a narrowband model is considered and then expanded to the broadband case as a set of narrowband frequencies. The coordinate system is defined by Figure 3.1, where bearing and elevation are $\theta$ and $\phi$ respectively. Bearing is held to a constant direction (e.g. North) such that array motion does not rotate the coordinate system. The origin is the phase center of the array and translates over time. Only a single elevation angle is considered, 90°, without loss of generality.

Narrowband Model

The narrowband frequency domain model considers a single frequency, $\omega$ in radians, such that data received at the $m$th sensor during the $n$th time instance is represented...
by \( x_m(n, \omega) \). The field is modeled as a collection of point sources in three dimensional space as a function of time. Each source is described by a wavenumber vector, \( \mathbf{k} \), with length corresponding to the spatial wavelength of the transmitted signal and pointing from the source position to the array

\[
\mathbf{k}(n, \omega) = \frac{\mathbf{\varepsilon}}{c} \begin{bmatrix}
\cos \theta(n) \sin \phi(n) \\
\sin \theta(n) \sin \phi(n) \\
\cos \phi(n)
\end{bmatrix},
\]  

(3.1)

where \( c \) is the phase speed of the medium in meters/sec. Note the definition of (3.1) actually points from array to the source using a vector similar the one in Figure 3.1. However, the derivations utilize \(-\mathbf{k}\) to denote a direction pointing from the source to the array. The medium is assumed to be non-dispersive. Sources are assumed to be in the far-field such that the wavefront arriving at the array is linear in phase and amplitude loss across the array is negligible compared to the loss from total distance traveled. The phase of the wavefront from the \( q \)th source to the \( m \)th element is given by an \( M \times 1 \) vector.

\[
[\mathbf{d}_q(n, \omega)]_m = \exp \left( -j \mathbf{k}_q^T(n, \omega) \mathbf{r}_m(n) \right),
\]  

(3.2)

where \( \mathbf{r}_m \) is the vector to the \( m \)th array element. Stacking the received data elements as \( \mathbf{x}(n, \omega) = [x_1(n, \omega), x_2(n, \omega) \ldots x_M(n, \omega)]^T \) creates an \( M \times 1 \) received data vector. Under this plane wave model, the received data for a field containing \( Q \) sources

**Figure 3.1: Polar coordinate system**
is expressed

\[ \mathbf{x}(n, \omega) = \sum_{q=1}^{Q} \mathbf{d}_q(n, \omega)s_q(n, \omega) + \mathbf{\eta}(n, \omega). \]  \hfill (3.3)

Array or source motion is captured by the time-dependent steering vector. Many passive applications employ a stochastic model for signals since the originating signal is unknown. In this case, the frequency domain signals are zero-mean circular symmetric complex normal distributed where the variance of the signal denotes the signal power. Thus, the \( Q \times 1 \) signal vector \( \mathbf{s}(n, \omega) = [s_1(n, \omega), s_2(n, \omega) \cdots s_Q(n, \omega)]^T \) is a random process distributed according to a circular symmetric complex normal distribution with zero mean and covariance matrix \( \Sigma(n, \omega) \). Noise is assumed to be uncorrelated between sensors and time instances with constant variance \( \sigma^2_\eta \) resulting in an additive white Gaussian model, but can be extended to include colored noise. The covariance of the resulting received signal vector is denote by \( \mathbf{R}(n, \omega) \). The statistical assumptions are summarized below.

\[ \mathbf{s}(n, \omega) \sim CN(\mathbf{0}, \Sigma(n, \omega)) \]  \hfill (3.4)

\[ \mathbf{\eta}(n, \omega) \sim CN(\mathbf{0}, \mathbf{I}\sigma^2_\eta(n, \omega)) \]  \hfill (3.5)

\[ \mathbf{x}(n, \omega) \sim CN(\mathbf{0}, \mathbf{R}(n, \omega)) \]  \hfill (3.6)

A steering matrix matrix is formed by augmenting columns of steering vectors for each source, \( \mathbf{D}(n, \omega) = [\mathbf{d}_1(n, \omega), \mathbf{d}_2(n, \omega) \cdots \mathbf{d}_Q(n, \omega)] \), with size \( M \times Q \). Using (3.3), the received data covariance matrix is expressed

\[ \mathbf{R}(n, \omega) = \mathbf{D}(n, \omega)\Sigma(n, \omega)\mathbf{D}^H(n, \omega) + \sigma^2_\eta \mathbf{I}. \]  \hfill (3.7)
The mapping problem is derived as a covariance matrix estimation problem in this thesis, where the source signal covariance is the unknown parameter of interest. Assume the angular space around the array platform is a grid of $Q$ points. Discretization provides the bearing angles for a BTR and assumes a fixed number of sources at each location. Thus, the power originating from each direction (or grid point) is represented by the source covariance matrix, $\Sigma$. Assume all of the sources are uncorrelated so that the source covariance matrix is diagonal. Source and platform movement are physically constrained and can be captured by a deterministic or stochastic model of the underlying covariance matrix. Applying target motion models is often performed post-detection and differs from the following approach where the spectrum dynamics are modeled pre-detection.

4.1 Deterministic Dynamic Model

The first approach models the spectrum deterministically through a set of known basis functions. Dependence on frequency is dropped for notational convenience throughout the narrowband derivations. In general, the estimate is a function of the
joint probability density function (pdf) of the received data and covariance matrix, written as

\[ \hat{\Sigma}_n = \arg \max_{\Sigma(n)} f(x(n), \ldots, x(n - N + 1), \Sigma(n)) . \] (4.1)

The simplest case is to assume a linear model written as

\[ \Sigma(n) = \Sigma_o + n \hat{\Sigma} \] (4.2)

where the model considers a sliding window from \( n - N + 1 \) to \( n \) and \( \Sigma_o \) represents the spatial spectrum at the start of the window; the notation in (4.2) assumes a sliding reference frame for \( n \) always beginning at 0. By substituting (4.2) into the covariance matrix definition (3.7), the approximation for the received signal covariance matrix is

\[ R_x(n) \approx D(n)(\Sigma_o + n \hat{\Sigma}) D(n)^H + \sigma_n^2 \mathbf{I}, \] (4.3)

and an uniform prior distribution on the spatial spectrum results in the equivalent maximum likelihood estimator,

\[ \hat{\Sigma}_n, \hat{\Sigma}_n = \arg \max_{\Sigma_o, \hat{\Sigma}} f \left( x \mid \Sigma_o, \hat{\Sigma}, \hat{\Sigma}_{n-1} \right) \] (4.4)

where \( x = \{x(n), x(n - 1), \ldots, x(n - N + 1)\} \). The inclusion of the previous time estimate is used for a recursive, online formulation and is an abusive notation since the statistics of the previous estimate are not considered. Maximization of the likelihood function is computationally infeasible, therefore an expectation maximization iterative technique is used to approximate the solution. For the EM derivation, the complete data is the hidden source and noise signals, \( C_N(k) = \{s(k), \eta(k), \ldots, s(k - N + 1), \eta(k - N + 1)\} \). The likelihood function with the complete data is given,

\[ f(C_N(n) | \Sigma_o, \hat{\Sigma}) = f(C_N(n) | \Sigma_o, \hat{\Sigma}) \]

\[ = \prod_{k=0}^{N-1} G(0, \Sigma_o + k \hat{\Sigma}) G(0, \sigma_n^2 \mathbf{I}) \] (4.5)
where summation limits are shifted for notational convenience, and \( G(\mu, \Sigma) \) represents the circular symmetric complex normal density function. Assuming each source is independent, the expectation of the log-likelihood given a previous iteration with respect to \( s, \eta \) is written

\[
E_{CN(n)} \left[ \ln f(C_N(n)|\Sigma_o, \hat{\Sigma}) \mid x, \Sigma_{o}^{old}, \hat{\Sigma}^{old} \right] =
\sum_{k=0}^{N-1} \sum_{q=1}^{Q} - \ln([\Sigma]_{qq} + k[\hat{\Sigma}]_{qq}) - \frac{h_q(k)}{[\Sigma]_{qq} + k[\hat{\Sigma}]_{qq}}
\]

(4.6)

where

\[
h_q(k) = E \left[ |s_q(k)|^2 \mid \Sigma^{old}, \hat{\Sigma}, x \right] .
\]

(4.7)

The maximization step must be accomplished numerically and the solution is separable between sources.

4.2 Stochastic Dynamical Model

Using a more Bayesian approach and stochastic model, assume the spectrum is a Markov process of the form

\[
\Sigma(n + 1) = \Sigma(n) + \Delta,
\]

(4.8)

where \( \Delta \) represents zero-mean state noise on the power from each angle. A more general model can be used, but this simple model provides a random walk approach to the time-varying nature of the frequency domain spatial spectrum. Knowledge, or conversely uncertainty, of the change in the spectrum is described by the state noise. This results in a probabilistic relation between sequential spectra that corresponds to a transition pdf

\[
f (\Sigma(k) \mid \Sigma(k - 1)) .
\]

(4.9)
The estimation problem is the same as in (4.1), condensed as

\[ \hat{\Sigma}_n = \arg \max_{\Sigma(n)} f(x, \Sigma(n)). \]  

(4.10)

However, the pdf is now separable due to the Markov property, where the conditional pdf of the received data is expressed as

\[ f(x|\Sigma(n - N - 1), \cdots, \Sigma(n)) = \prod_{k=n-N+1}^{n} f(x(k)|\Sigma(k)). \]

While this greatly simplifies the problem, the optimization still remains difficult due to the “hidden” nature of the source signals, \( s(k) \). Again using the approach from Section 4.1, an EM algorithm is derived using the signal and noise as the complete data. The resulting slightly modified conditional density function

\[ f(C_{N}(n) | \Sigma(n - N + 1) \cdots \Sigma(n)) = \prod_{k=n-N+1}^{n} f(s(k), \eta(k) | \Sigma(k)) \]  

(4.11)

provides the likelihood function for an ML estimator assuming the spectrum is independent for each time instance. The full pdf is expressed in terms of the Markov relation (4.11) and the transition pdf (4.9),

\[ f(C_{N}(n), \Sigma(n - N + 1) \cdots \Sigma(n)) = \prod_{k=n-N+1}^{n} f(s(k), \eta(k)|\Sigma(k)) \cdot f(\Sigma(k)|\Sigma(k-1)). \]  

(4.12)

The forward method for solving hidden Markov chains is applied to express the joint pdf of the complete data and the spatial spectrum, [38, pp. 112-113]

\[ f(C_{N}(n), \Sigma(k)) = \left\{ \int_{\Sigma(n-1)} f(C_{N-1}(k-1), \Sigma(k-1)) \int_{\Sigma(k-1)} f(\Sigma(k)|\Sigma(k-1))d\Sigma(k-1) \right\} f(s(n), \eta(n)|\Sigma(k)). \]  

(4.13)
Now consider the solution one step backwards in time generically as a function \( g(x) = f(C_{N-1}(k-1), \Sigma(k-1)) \). The first few steps in the Taylors series expansion around \( \mu \) is written

\[
g(x) = g(\mu) + g'(\mu)(x - \mu) + \frac{1}{2} g''(\mu)(x - \mu)^2
\]

which offers a simple approximation \( E[g(x)] \approx g(\mu) + \frac{x^2}{2} g''(\mu) \). This is essentially a local approximation of \( g(x) \) assuming the pdf is tight around \( x \) near \( \mu \). This approximation applied to the term in brackets from (4.13) is analogous to assuming the field does not vary quickly between time instances. Using a uniform prior on the spatial spectrum and ignoring noise constants, the joint pdf of the spatial spectrum and the complete data is written

\[
f(C_N(n), \Sigma(n)) = f(s(n), \eta(n) | \Sigma(n))
\]

\[
= \prod_{k=n-N+1}^{n-1} G(s(k); 0, \Sigma(n)) + \frac{1}{2} \text{Tr} \left[ S H \Sigma_n (G(s_k; 0, \Sigma(n))) \right], \tag{4.14}
\]

where \( H \) represents the Hessian operator (with respect to the diagonal elements of \( \Sigma_n \)) and \( \text{Tr} \) is the trace operator. The covariance matrix of \( \Delta \) is represented by \( S \). Let the diagonal of the covariance \( S \) be \( d^2_q \) for the \( q \)th element. Simplification with the natural log and discarding constants results in

\[
\ln f(C_N(n), \Sigma(n)) = \sum_{k=n-N+1}^{n} \ln G(s(k); 0, \Sigma(n)) + \sum_{k=n-N+1}^{n-1} l_k \tag{4.15}
\]

The term in the second sum from (4.15) is \( l_k \) defined in (4.16) and will be small for slowly changing spectra. Assuming state noise in uncorrelated between grid points, \( S \) will be diagonal and can be approximated by

\[
l_k = \ln \left[ 1 + \frac{1}{2} \text{Tr} \left[ S (\Sigma^{-1}(n)s(k)s(k)^H\Sigma^{-1}(n) - \Sigma^{-1}(n)) \right] \right] \tag{4.16}
\]

\[
\approx \frac{1}{2} \sum_{q=1}^{Q} \frac{d^2_q|s_q(k)|^2}{(\Sigma_{qq})^2(n)} - \frac{d^2_q}{\Sigma_{qq}} \tag{4.17}
\]
An EM algorithm is derived by taking the expectation of the joint pdf of the complete data and the spectrum given the incomplete data, \( \mathbf{x} \), and the previous estimate \( \Sigma^{old} \).

With terms independent of \( \Sigma \) ignored, the natural logarithm of the joint pdf is written

\[
\ln f(C_N(n), \Sigma(n)) = \sum_{k=n-N+1}^{n} -\ln \det [\pi \Sigma(n)] - s^H(k) \Sigma^{-1}(n) s(k) + \sum_{k=n-N+1}^{n-1} l_k.
\]

(4.18)

Let the source signal terms of the following expectation step be defined \( h_q(k) = \mathbb{E} [s_q(k)^2 | \Sigma^{old}, \mathbf{x}] \). Now taking the expectation with respect to the complete data results in

\[
\mathbb{E}_C [f(C_N(n), \Sigma(n)) | \Sigma^{old}, \mathbf{x}] = \sum_{q=1}^{Q} \left\{ -N \ln [\Sigma(n)]_{qq} - \frac{(N-1)d_q^2}{2[\Sigma(n)]_{qq}} - \frac{1}{n-N+1} \sum_{k=n-N+1}^{n} \frac{h_q(k)}{[\Sigma(n)]_{qq}} + \frac{1}{n-1} \sum_{k=n-N+1}^{n-1} \frac{d_q^2 h_q(k)}{[\Sigma(n)]_{qq}^2} \right\}.
\]

(4.19)

The separable sum over the \( Q \) sources is a result of the sources being statistically independent. Let each source be represented across time by defining \( h_q = \sum_{k=n-N+1}^{n} h_q(k) \). The maximization step of the EM algorithm is solved by setting the derivative of the expectation step to zero and solving. Since each source is separable, consider the \( q \)th source and refer to the result from (4.19) as \( z \) where the derivative is expressed

\[
\frac{dz}{d[\Sigma(n)]_{qq}} = -N \frac{[\Sigma(n)]_{qq}}{[\Sigma(n)]_{qq}} + \frac{0.5(N-1)d_q^2}{[\Sigma(n)]_{qq}^2} \frac{\sum_{k=n-N+1}^{n} h_q(k)}{[\Sigma(n)]_{qq}} - \frac{d_q^2}{[\Sigma(n)]_{qq}^2} \sum_{k=n-N+1}^{n-1} h_q(k).
\]

(4.20)

The cubic equation has 2 non-trivial solutions with only one solution assuming that
the spatial spectrum variance is small, given by (4.21)

\[
[\tilde{\Sigma}(n)]_{qq} = \frac{h_q}{2N} + \frac{N - 1}{4N} d_q^2 + \frac{1}{2N} \left( \frac{N - 1}{2} d_q^2 + h_q \right)^2 - 4Nd_q^2 \sum_{k=n-N+1}^{n-1} h_q(k) \right) ^{\frac{1}{2}}
\]

(4.21)

This provides a generalization of the stationary model assumed by Rogers and Krolik [5, 34]. The result from their previous work can be obtained by setting \(d = 0\) in the solution (4.21), which calculates a time-average of the expected values of the signal covariance. The simplification results in a solution of \(h_q/N\). It is useful to consider the relation of \(h_q(k)\) with the previous terms. The term \(h_q(k)\) results from the conditional pdf for a linear model, given by

\[
E \left[ |s(k)|^2 \mid \Sigma^{\text{old}}, x \right] = \text{Var}[s(k)] + |E[s(k)]|^2 = \Sigma^{\text{old}} + \Sigma^{\text{old}} G(k) \Sigma^{\text{old}}
\]

where term from the conditional mean of the signal estimate is

\[
G(k) = D^H(k) \left[ K^{-1}(k) - K^{-1}(k) \hat{R}(k) K^{-1}(k) \right] D(k)
\]

(4.22)

and the expected received signal covariance is

\[
K(k) = D(k) \Sigma^{\text{old}} D^H(k) + \sigma^2 \eta I.
\]

(4.23)

Lanterman considers estimates using a combination of a likelihood and prior (often improper and referred to as regularization) [28], but it frequently results in numerical optimizations with high computational cost, as in Section 4.1. The solution in (4.21) is derived with approximate prior knowledge and provides a quantitative framework for pseudo-stationary approximations. The limiting case where \(d = 0\) provides a different modeling approach than the one described by Rogers, seen in (2.13) by relating it to a simple model with white state noise described by (4.8).
4.3 Narrowband Simulation

A simulation is used to demonstrate the improvements of the new spatial spectrum estimators over conventional processing. Capturing array dynamics allows a much smaller maneuvering uniform linear array (ULA) to provide maps comparable to much larger arrays by enhancing endfire directions and suppressing backlobes. Interference-to-noise ratio (INR) and signal-to-noise ratio (SNR) are defined as \( \sigma^2/\sigma_n^2 \).

Consider a scenario with 4 interferers at positions relative to North at -90°, -14°, 135° each with 10 dB INR and -135° with 3 dB INR. The target begins at 60° and transitions to 15° with a maximum instantaneous bearing rate of -0.26° per second with SNR of 3 dB. The cartoon bearing-time-record (BTR) illustrates the scenario in Figure 4.1. BTRs provide a method of viewing the angular map as it changes over time.

![Figure 4.1: Illustrative BTR for all simulations](image)

For the narrowband case, each BTR is computed from -180° to 180° with 4° spacing and snapshot sampling at 1/2 Hz. The simulated data is generated assuming a narrowband frequency of 750 Hz, and the speed of sound is assumed to be 1560 m/s. Conventional beamforming is accomplished by computing \( \hat{\Sigma}(k) = \mathbf{D}(k)\mathbf{R}(k)\mathbf{D}(k) \) at each time instance independently using a rectangular window. The oracle in this
example is a 660 element circular array with inter-element spacing of $\lambda/2$ with diameter $\approx 110\lambda$. Circular arrays provide uniform angular resolution. This is compared against an ULA maneuvering around the same circle at 2 m/s for 5 minutes with heading increments of $+1^\circ$ per second. The circle is not fully completed in the simulation, and the circular array is matched to the path covered. The SNR and INR are reduced for the clairvoyant array such that the post-array gain SNR is the same for subsequent arrays, which is equivalent to placing shorter mobile arrays at $1/10$ the distance to all sources under spherical spreading. This *ideal* circular array is assumed to have a perfect estimate of the received data covariance matrix. The resulting BTR, shown in Figure 4.2, displays each source clearly with a noise floor prescribed by the uniform weights. An alternative approach would be to compare the circular array with several short arrays at scattered locations using the same total number of elements, and optimum placement has been considered by Bilik *et al.* [39].

![Figure 4.2: BTR showing power estimates, in dB, from narrowband conventional beamformer output for clairvoyant circular array.](image)

Conventional beamforming is data independent and only optimal in white-noise limited environments with performance determined by number of sensors and weight-
ing schemes. When using few sensors with many sources, the interference dominates the targets. This occurs in the second array configuration, which utilizes 7 sensors in a rigid ULA at $\lambda/2$ spacing while maneuvering over the same shape as the circular array. The symmetric geometry of a ULA causes ambiguities that appear as backlobes on the BTR. Conventional beamforming at each time instance is unable to distinguish true sources from backlobes, which appear as diagonal stripes or extremely high bearing rate sources with circular platform maneuvers, shown in Figure 4.3(a). Poor endfire resolution of the ULA is visible as thick lines progressing from $-180^\circ$ to $180^\circ$ and also from $0^\circ$ wrapping around to $0^\circ$ over time. These problems can be mitigated with time-varying models. The EM algorithms perform 10 iterations after each data collection (using only a single snapshot for each location) with the deterministic method using the previous $N = 9$ snapshots. The deterministic method, shown in Figure 4.3(b), displays the interferer and target tracks clearly without backlobes. For the first 50 seconds, the target is transitioning across the backlobe from the source at $135^\circ$ and across the end-fire of the array creating uncertainty in the target power estimate, also occurring after the 200 second mark when the target passes through endfire again. The stochastic method uses $N = 1$ with $d = 0$, reflecting a recent online technique originally developed for maneuvering long towed arrays [34]. The maneuverability of the short array and time-varying models allow significant suppression of the array ambiguities and increase endfire resolution. Comparing Figure 4.2 with Figures 4.3(b) and 4.3(c), the results show that short mobile arrays can provide performance the of much larger arrays. Differences in array gain can be addressed by increasing the number of short arrays or by reducing transmission loss through location or maneuverability of the much smaller array.
Figure 4.3: BTR showing narrowband power estimates, in dB, with a maneuvering short array using (a) conventional beamforming, (b) deterministic model, and (c) stochastic model.
5.1 Broadband Extension

While previous sections only consider narrowband sources, this section introduces a broadband model and analyzes the effect of temporal spectrum knowledge. Broadband signals are modeled as a set of narrowband signals, using a statistical model first published by Bangs [40].

5.1.1 Signal Model

The signal vector is formed by stacking $B$ narrowband vectors.

$$\bar{s}(n) = \left[ s^T(n, \omega_1), s^T(n, \omega_2), \ldots, s^T(n, \omega_B) \right]^T. \quad (5.1)$$

A bar ($\bar{\cdot}$) is used to refer to vectors and matrices resulting from stacking narrowband variables. The number of frequency bins cannot be larger than the number of time-domain samples. Assuming a long enough observation window (or large time-bandwidth product), each frequency is independent with $\bar{s} \sim \mathcal{CN}(0, \bar{\Sigma}(n))$. In this
case, the covariance matrix is block diagonal expressed as

\[
\Sigma(n) = \begin{bmatrix}
\Sigma(n, \omega_1) & 0 & \cdots & 0 \\
0 & \Sigma(n, \omega_B) & \cdots & 0 \\
\vdots & \vdots & \ddots & \vdots \\
0 & 0 & \cdots & \Sigma(n, \omega_B)
\end{bmatrix}.
\] (5.2)

Similarly, the received signal and steering vectors are stacked into a broadband form.

\[
\tilde{x}(n) = [x^T(n, \omega_1), x^T(n, \omega_2), \cdots, x^T(n, \omega_B)]^T
\] (5.3)

\[
\tilde{D}(n) = [D^T(n, \omega_1), D^T(n, \omega_2), \cdots, D^T(n, \omega_B)]^T
\] (5.4)

The broadband received signal vector has a form similar to the narrowband vector and has a circularly symmetric complex normal distribution, \( \tilde{x} \sim CN(\mathbf{0}, \tilde{\mathbf{R}}) \). The block diagonal structure from (5.2) creates a block diagonal structure in the received signal vector expressed by \( \tilde{\mathbf{R}}(n) = \tilde{\mathbf{D}}(n)\tilde{\Sigma}(n)\tilde{\mathbf{D}}^H(n) + \sigma_q^2\mathbf{I} \) or

\[
\tilde{\mathbf{R}}(n) = \begin{bmatrix}
\mathbf{R}(n, \omega_1) & 0 & \cdots & 0 \\
0 & \mathbf{R}(n, \omega_B) & \cdots & 0 \\
\vdots & \vdots & \ddots & \vdots \\
0 & 0 & \cdots & \mathbf{R}(n, \omega_B)
\end{bmatrix}.
\] (5.5)

The broadband model assumes frequency bins are independent and resemble multiple observations of the same random variables when the signals have a flat temporal spectrum. For this case, the underlying signal covariance matrices are approximately the same or \( \Sigma(n, \omega_b) \approx \Sigma(n, \omega_a) \forall a \in [1, B] \). This approximation is used for small bandwidths by Gerstoft et al. [18]. However, each frequency creates a slightly different steering vector and results in interesting geometries, which will motivate a new broadband method.

5.1.2 Parameter Reduction

Temporal spectrum knowledge must be tied to the model of the broadband signal covariance matrix, \( \tilde{\Sigma} \), and can reduce the number of parameters. The Cramér-Rao bound is derived to demonstrate the effect of multiple frequency bins on the
performance of power estimates. The problem of snapshot deficiency is ignored in CRB derivations because the purpose is to understand the impact of broadband data. The multivariate circular symmetric complex normal probability density, with covariance $\mathbf{R}$ and mean $\mathbf{\mu}$, for a single snapshot is

$$f(x) = \frac{1}{\det(\pi \mathbf{R})} \exp \left( - (\mathbf{x} - \mathbf{\mu})^H \mathbf{R}^{-1} (\mathbf{x} - \mathbf{\mu}) \right). \quad (5.6)$$

Consecutive snapshots are assumed independent and identically distributed (i.i.d.). The Fisher information matrix, $\mathbf{J}$, is well known for the single snapshot case

$$[\mathbf{J}]_{pq} = \text{Tr} \left\{ \mathbf{R}^{-1} \frac{\partial \mathbf{R}}{\partial \sigma_p} \mathbf{R}^{-1} \frac{\partial \mathbf{R}}{\partial \sigma_q} \right\}, \quad (5.7)$$

where $\text{Tr}$ is the trace operator. The bound on the covariance of an unbiased estimator is given by $\text{cov}(\hat{\Sigma}) \geq \mathbf{J}^{-1}$ or $\text{var}(\hat{\sigma}^2) \geq J^{-1}$, for the scalar case. Considering a single frequency at a time then expanding to $B$ frequencies, the derivative term and inverse are block diagonal resulting in a simple expanded form of the Fisher information matrix,

$$[\mathbf{J}]_{pq} = \sum_{b=1}^{B} \text{Tr} \left\{ \mathbf{R}^{-1}(\omega_b) \frac{\partial \mathbf{R}(\omega_b)}{\partial \sigma_p^2} \mathbf{R}^{-1}(\omega_b) \frac{\partial \mathbf{R}(\omega_b)}{\partial \sigma_q^2} \right\}. \quad (5.8)$$

The key to this simplification is independence across frequency bins. For a single source, the resulting bound is trivial and independent of source location. The definitions of $p$ and $q$ in (5.8) have yet to be defined and are intentionally vague, indexing the unknown parameters. Traditional frequency-domain techniques estimate power at each frequency then use weighted summation to combine estimates. This requires estimating the full set of parameters instead of directly estimating the final parameter. The parameter space reduces from $Q \times B$ to as few as $Q$ when the source temporal spectrum is known or a single parameter describes each source, such as unknown source level with known shape. Consider the case where sources are assumed
to have a flat spectrum. This and several other cases were considered for detection
and localization bounds by Messer, who provides one of the few comparisons for
the bounds on various temporal spectrum assumptions [37]. Typically, processing
schemes incoherently average the estimate across frequencies, resulting in a bound

\[
J_c^{-1} = \frac{1}{B^2} \sum_{b=1}^{B} J^{-1}(\omega_b),
\]

(5.9)

while parameter reduction has a bound with the form

\[
J_B^{-1} = \left[ \sum_{b=1}^{B} J(\omega_b) \right]^{-1}.
\]

(5.10)

Refer to Appendix A for bound derivations.

The bound formed in (5.9) will be referred to as the full ML approach since it
estimates each parameter then averages across frequency. The bound from (5.10) will
be referred to as the reduced ML approach since it performs parameter reduction
before estimation. Comparing the full ML approach with the reduced ML case, it
can be seen that there will only be a significant difference when \( J(\omega_b) \) varies
between frequencies. For example, consider a 10 element ULA with static source at
10° and 10 dB SNR without spatial aliasing. The CRB of the power estimate of a
second source is computed as a function of bearing in order to explore the effects
of angular separation and bandwidth. Increasing separation quickly decreases the
bound, as shown in Figure 5.1. The full and reduced bounds do not show significant
differences for 550-750 Hz, however separation of approximately 2 dB occurs with
250-750 Hz. The full method reaches the reduced bound near 30°. It is clear that
angular separation and not bandwidth impacts the bound the most here. This is
because the bound does not have large variations between frequencies in this simple
case.
In contrast, spatial aliasing causes large variations in the bound since grating lobes create incorrect power estimates in ambiguous directions. However, grating lobes occur at known frequency dependant locations and can be exploited to reduce uncertainty. Comparing the forms of (5.9) and (5.10), very uncertain parameters will dominate the full estimate while the reduced estimate is less sensitive to individual parameter uncertainty. Consider the previous example with a 10 element ULA but now with minimum 3\(\lambda\)/2 spacing over the 550-750 Hz range. The large spacing results in grating lobes as well as geometric ambiguities from the linear array. The inverse of the first element of the FIM increases where grating lobes occur, as shown in Figure 5.2. The static source at 10° increases the bound as in the previous example. However, the additional peaks crossing 80° and from 160° to 120° result from aliasing and are coupled with frequency unlike backlobes or left-right ambiguities due to the linear arrays. This allows source power to be estimated by combining information across frequency for a single bearing, using knowledge of the aliasing.

The two ambiguous regions over 180° can be seen as jumps in the bound shown.

**Figure 5.1:** CRB of power estimate with 10 dB source at 10° as function of second source (0 dB) using 10 element ULA array with half wavelength spacing for various bandwidth extending below spatial sampling limit.
Figure 5.2: Inverse of FIM element corresponding to parameterized source (0 dB) with 10 dB static source at 10° using 10 element ULA array with 3λ/2 spacing in Figure 5.3. The large jumps in the bound are a result of summing the bounds across frequency from Figure 5.2. The full parameter bound in Figure 5.3 uses an approximation discussed in Appendix A for numerical stability. The reduced parameter bound remains significantly lower than the full method, reducing the impact of spatial aliasing. This motives incorporating parameter reduction in broadband extensions of the spatial spectrum estimates.

5.2 Reduced Broadband Spatial Spectrum Estimate

This section extends the stochastic narrowband spatial spectrum estimate to broadband sources. In order to focus on the broadband effects, the derivations presented consider \( d = 0 \), but any dynamic model can be incorporated resulting in different likelihood functions and priors. Following the same procedure given in Chapter 4,
Figure 5.3: CRB of power estimate in aliasing region with 10 dB source at 10° as function of second source (0 dB) using 10 element ULA array with 3λ/2 spacing

An EM algorithm is again used with expectation step (5.11) defined

\[
E_{c} \left[ \ln f(\hat{C}, \Sigma(n)|\hat{\Sigma}^{old}, \bar{x}) \right] = -N \sum_{b=1}^{B} \sum_{q=1}^{Q} \ln[\Sigma(n, \omega_{b})]_{qq} - \sum_{k=n-N+1}^{n} \left( N \sum_{b=1}^{B} \sum_{q=1}^{Q} \frac{h_{q}(k)}{[\Sigma(n, \omega_{b})]_{qq}} \right) \tag{5.11}
\]

and maximization step

\[
\frac{\partial z_{B}}{\partial \Sigma_{qq}} = \frac{-NB}{[\Sigma(n)]_{qq}} + \sum_{k=n-N+1}^{n} \sum_{b=1}^{B} \frac{h_{q}(k)}{[\hat{\Sigma}]_{qq}} \tag{5.12}
\]

The broadband estimate results in an estimate over the previous data and across all frequencies, giving an intuitive solution,

\[
[\hat{\Sigma}]_{qq} = \frac{1}{NB} \sum_{k=n-N+1}^{n} \sum_{b=1}^{B} E \left[ |s_{q}(k, \omega_{b})|^{2} | \hat{\Sigma}^{old}, \bar{x} \right] \tag{5.12}
\]
From this form it appears that multiple frequencies can be combined by simply averaging the final result, which is the traditional solution. However, this is an iterative solution where each estimate is uncoupled across angle but not time or frequency. The iterative form is expressed

\[
\hat{\Sigma}_n^{\text{new}} = \hat{\Sigma}_n^{\text{old}} - \frac{1}{NB} \sum_{k=n-N+1}^{n} \sum_{b=1}^{B} (G(n, \omega_b)) \hat{\Sigma}_n^{\text{old}}
\]  

(5.13)

where the frequency and time dependent terms from the conditional pdf are

\[
G(n, \omega_b) = \left( \mathbf{D}^H (K^{-1} - K^{-1} \hat{R} K^{-1}) \mathbf{D} \right)
\]  

(5.14)

\[
K(n, \omega_b) = D \hat{\Sigma}_n^{\text{old}} D^H + \sigma_n^2 \mathbf{I}.
\]  

(5.15)

It is clear from (5.13) that simply averaging across frequencies as a post-processing method does give the same result as (5.12) appears to suggest. Incorporating the temporal spectrum knowledge into the estimate directly forces convergence to a joint solution. In contrast, the typical estimate that moves the average outside of this iteration, given by (5.16), expands the space where the solution is allowed to exist then shrinks the space after convergence, thus increasing the possible number of incorrect local maxima.

\[
\hat{\Sigma}_n = \frac{1}{B} \sum_{b=1}^{B} \hat{\Sigma}_n(\omega_b)
\]  

(5.16)

5.3 Broadband Simulation

The differences between the full ML approach and reduced ML approach is simulated first without spatial aliasing. This is presented for completeness and to compare broadband expansion versus spatial aliasing. Using the previous simulation from Section 4.3 with extended frequency range from 550-750 Hz with 2 Hz bins, 1 Hz snapshot sampling, and 181 angle bins. The inter-element spacing is limited
to \approx 1\text{m} at \lambda/2 spacing for the smallest wavelength. The circular array containing the path of the UUV contains a total of 660 elements and has a diameter of 110 wavelengths as before. The SNR continues to be normalized for all simulations in order to maintain constant post-array gain. The conventional beamformer output, incoherently summed across frequency, is shown in Figure 5.4. The color scale of the plot is normalized in order to provide a noise floor that corresponds to the possible sidelobe levels. A rectangular window is used for the beamformer, and the sources can be clearly seen throughout the run.

![Figure 5.4: Broadband BTR for circular array](image)

BTR showing power estimates, in dB, from broadband conventional beamformer output for clairvoyant circular array.

The 7 element mobile platform is the same size as before and maneuvers over the shape of the circular array while maintaining the rigid ULA structure. The beams for the conventional beamformer are widened slightly as the wider lower frequency beams are combined with the high frequency beams. The ambiguities have the same structure as before, appearing as high bearing rate sources, shown in Figure 5.5(a). The full ML technique provides improvement by combining the frequency domain data to reduce noise. As derived by the CRB, the additional data reduces the variance in the BTR shown in Figure 5.5(b). The reduced ML approach shown
in Figure 5.5(c) shows insignificant improvement over the full ML approach. This is expected because each frequency bin uniformly provides consistent information. The overall improvement from narrowband ML to broadband ML is due to the multiple frequency bins available for each time instance.

![BTRs showing broadband power estimates](image)

**Figure 5.5**: BTRs showing broadband power estimates, in dB, with an maneuvering short array using (a) conventional beamforming, (b) full ML, and (c) reduced ML.

A detection algorithm is used in order to provide a metric for comparison between the three different methods. This is also used to demonstrate low SNR target detection performance with the small UUV. An $M$-of-$N$ detector is used by performing a test at $N$ observations then determining if a target exists when $M$ tests show a target. This detector is used because ambiguities create false targets [6]. Even
the true target can create false targets. Using $N = 100$ observations of 1 second intervals, a detection occurs when $M = 80$ or more power (BTR) estimates exceed a given threshold along a hypothesized target track. The true target track is used in this case. The threshold is a function of the desired probability of detection, $P_d$, or probability of false alarm, $P_f$. When there is a single target in white noise, conventional beamforming corresponds to the likelihood ratio and the optimal detector for a single snapshot. The curves in Figure 5.6 use 200 realizations for each method under target present ($H_1$) and target free ($H_0$) conditions in the interference dominated environment. The levels of each of the interferers is kept the same as before, but the target SNR is dropped to -25dB. This shows the ability of each method for low SNR target detection. The conventional delay-and-sum beamformer is unable to distinguish the target from the noise and interferers. The full and reduced ML methods perform significantly better, but neither ML method is clearly superior to the other, as is also shown in the BTRs.

![ROC curve, M-80 of N-100 detector](image)

**Figure 5.6**: ROC curve with target SNR -25 dB using $M$-of-$N$ detector
5.4 Broadband Simulation with Under-sampled Array

A simulation is used to demonstrate performance differences between the full ML approach, (5.16), and the reduced ML approach, (5.13) with under-sampled arrays resulting in spatial aliasing. The simulation from Section 5.3 is continued with true BTR given in Figure 4.1. The array inter-element spacing is $3\lambda/2$ for the largest frequency, resulting in spacing of $\approx 3m$. For a uniform linear array, it is possible to have as many as five additional incorrect ambiguities from two grating lobes and a total of three backlobes. The circular array following the path of the maneuvering array is reduced to 220 elements due to the larger element spacing. The SNR is normalized for each array such that post-array gain is held constant as before. Conventional beamforming is used on the circular array with incoherent frequency averaging resulting in high sidelobes, shown in Figure 5.7. The circular array does not suffer from the distinct lines from spatial grating lobes. A 7 element rigid ULA array maneuvers over the path of the circular array. Note the length of the mobile platform is increased by a factor of 3 due to the larger spacing. Conventional beamforming results

![Figure 5.7: Broadband BTR for circular array with 3m spacing](image)

BTR showing power estimates, in dB, from broadband conventional beamformer output for under-sampled clairvoyant circular array.
in ambiguities from spatial aliasing and array geometry, which vary as a function of platform orientation. This creates structured patterns appearing as maneuvering sources or swirls in the background of the BTR, shown in Figure 5.8(a). The full ML technique estimates the covariance matrix at each frequency across all time and then averages the estimates across frequencies, which contain spatial grating lobes. The averaging reduces the effects of spatial aliasing but results in sporadic sources appearing in ambiguous regions, shown in Figure 5.8(b). It is possible to introduce regularization, for example by setting $d$ to a non-zero value, but this will also reduce the estimated power of high bearing rate sources. A trade-off exists in this case but is not explored in this thesis. The reduced ML method provides significant suppression of spatial grating lobes, shown in Figure 5.8(c). The performance improvement from full ML to reduced ML is significant, with only a small ambiguity remaining near 100°. There is also improvement from the narrowband case in Figure 4.3(c) to the broadband cases, although an increase in resolution is due to the larger length of the array in the broadband simulation. This resolution increase is possible using the same number of sensors because the estimate is extended into the spatially aliased region.

As a quantitative metric, the ROC analysis from Section 5.3 is continued for the under-sampled array case. Again the target SNR is -25 dB, and the true trajectory is used for the hypothesized locations. Using 100 realizations with the same conditions as before, the ROC curve is given in Figure 5.9. This demonstrates that the new method provides a detection performance increase in the spatial grating lobe regions. For the full ML method, the ambiguities introduced by the spatial aliasing increases the false alarm rate for a given probability of detection. However, the reduced ML method is able to maintain a similar level of performance.
Figure 5.8: BTRs showing broadband power estimates, in dB, with an undersampled maneuvering short array using (a) conventional beamforming, (b) full ML, and (c) reduced ML.
Figure 5.9: ROC curve with target SNR -25 dB using $M$-of-$N$ detector for undersampled array
This thesis introduces several approaches for spatial spectrum estimation in dynamic environments for passive array processing. A maneuvering array is considered because combining different array orientations or configurations can increase the performance over a stationary array for the full 360° bearing space. Using a frequency domain model, a narrowband deterministic model is derived to incorporate source motion across observations without source tracking. A narrowband stochastic model is derived using a basic Markov model and, for a limiting case, shown to simplify to recent work in the sonar literature. The new formulation presented here provides an alternative and intuitive model for the previous work. As an extension, broadband data is shown to increase the performance of an estimate by incorporating varying information across frequency bins. A framework for broadband estimates is suggested, and an estimate is derived assuming a flat temporal spectrum. This temporal knowledge allows the estimate to operate in the spatial grating lobe region, which is generally ignored in the literature because it violates traditional spatial sampling requirements. The simulations in this thesis assume a single snapshot is available at each location in order to show multiple snapshots are not required.
Appendix A

Cramér-Rao Bounds

The Cramér-Rao bound on the spatial spectrum estimate given the model from Chapter 5 is in general expressed by (5.7). Defining a form of the parameter vector $\bar{\sigma} = [\sigma_1^2(\omega_1) \cdots \sigma_Q^2(\omega_1), \sigma_1^2(\omega_2) \cdots \sigma_Q^2(\omega_B)]$, with order sources then frequencies, creates a structured FIM

$$J = \begin{bmatrix} J(\omega_1) & 0 & \cdots & 0 \\ 0 & \ddots & \ddots & \ddots \\ \vdots & \ddots & \ddots & \ddots \\ 0 & \cdots & 0 & J(\omega_B) \end{bmatrix}.$$ 

The full ML method estimates each frequency and then averages the estimates, thus the bound must introduce an additional parameter $\sigma_q = \frac{1}{B} \sum_{b=1}^{B} \sigma_q^2(\omega_b)$. The resulting Fisher information matrix is denoted $J_c$.

$$[J_c] = \begin{bmatrix} J & J_{\sigma q} \\ J_{\sigma q} & J_{\sigma} \end{bmatrix}$$

The inverse can be written in the form $[J_c]^{-1} = H^T J^{-1} H$ where $[H]_{ab} = \frac{\delta_{a b}}{\sigma_a^2}$ [41, p. 230]. The effect of $H$ in matrix form is to perform the addition on each estimate,
and the bound can be rewritten

\[ J_e^{-1} = \frac{1}{B^2} \sum_{b=1}^{B} J^{-1}(\omega_b). \]  
(A.1)

The reduced ML method assumes a single parameter is used for each source and results in a parameter vector \( \bar{\sigma} = [\sigma_1^2 \cdots \sigma_Q^2] \). This broadband derivation uses only \( Q \) parameters with FIM denoted \( J_B \), resulting directly from (5.8) and inverse expressed as (A.2).

\[ J_B = \sum_{b=1}^{B} J(\omega_b) \]

\[ J_B^{-1} = \left[ \sum_{b=1}^{B} J(\omega_b) \right]^{-1} \]  
(A.2)

It is useful to consider the partitioned case for a single variable without loss of generality

\[ J(\omega_b) = \begin{bmatrix} J_{11} & J_{12}^T \\ J_{21} & J_{22} \end{bmatrix} \]

where all elements of \( J \) are non-negative and \( J_{22} \) is positive semidefinite. Note that the bound on the parameter is

\[ J_{11} = [J_{11} - J_{21}^T J_{22}^{-1} J_{21}]^{-1} \]

\[ J_{11} \geq J_{11}^{-1} \]

\[ [J^{-1}]_{qq} \geq \{[J]_{qq}\}^{-1} \]  
(A.3)

where the new bound is *looser* than the CRB but provides an approximation at least as low as the CRB.

48
Bibliography


