# Signal Direction-of-Arrival and Amplitude Estimation for Multiple-Row Bathymetric Sidescan Sonars

by

Wen Xu

B.S., University of Science and Technology of China (1990)
M.S., Institute of Acoustics, Chinese Academy of Sciences (1993)

Submitted in partial fulfillment of the requirements for the degree of

MASTER OF SCIENCE IN OCEANOGRAPHIC ENGINEERING

at the

MASSACHUSETTS INSTITUTE OF TECHNOLOGY

and the

WOODS HOLE OCEANOGRAPHIC INSTITUTION

FEBRUARY, 1998

The author hereby grants to MIT and WHOI permission to reproduce and distribute publicly paper and electronic copies of this thesis document in whole or in part.

Author ............................... ..........................................................................
Joint Program in Applied Ocean Science and Engineering
Massachusetts Institute of Technology/Woods Hole Oceanographic Institution
January 12, 1998

Certified by ..... .......................................
W. Kenneth Stewart
Thesis Supervisor

Accepted by .. ..........................................................................................
Michael S. Triantafyllou
Chairman, Joint Committee for Applied Ocean Science and Engineering
Massachusetts Institute of Technology/Woods Hole Oceanographic Institution

# Signal Direction-of-Arrival and Amplitude Estimation for Multiple-Row Bathymetric Sidescan Sonars

by

Wen Xu

Submitted to the MIT/WHOI Joint Program on Jan. 12, 1998, in
partial fulfillment of the requirements for the degree of Master of
Science in Oceanographic Engineering

## Abstract

In practical applications with bathymetric sidescan sonars, the multipath reflections and other directional interferences are the key limiting factors for a better performance. This thesis proposes a new scheme to deal with the interferences using a multiple-row bathymetric sidescan sonar. Instead of smoothing the measurements over some time or angle intervals, which was previously widely investigated, we resolve the multipath interferences from the direct signal. Two approaches on signal direction-of-arrival (DOA) and amplitude estimation are developed, the correlated signal direction estimate (CSDE) for three-row systems and the ESPRIT-based method. These approaches are compared using different sonar data models, including a stochastic model from the statistical analysis on bottom scattering and a coherent model from the analysis on interference field; the simulations show the ESPRIT-based approach is quite robust at the angular separation of $\geq 10°$ between two sources and at the signal-to-noise ratio above 10dB except for highly coherent or temporally correlated signals, for which CSDE works very well. The computer simulation results and the discussions on practical algorithm implementation indicate the proposed scheme can be applied to a real multiple-row bathymetric sidescan sonar. With the capability to simultaneously resolve two or more directional signals, the new sonar model should work better for a wider variety of practical situations in shallow water without significant increase of the system cost.

Thesis Supervisor: W. Kenneth Stewart
Title: Associate Scientist, Woods Hole Oceanographic Institution

# Acknowledgments

It is with deep appreciation that I express my thanks to my thesis advisor, Dr. Ken Stewart, who provided me this excellent opportunity and led me to the completion of the thesis.

I am especially grateful to Dr. Xiaoou Tang for all kinds of helps, from research to courses, from Matlab to FrameMaker, from driving to cooking.

My thanks also go to Dr. Dezhang Chu and Dr. Dave Custer for reading my thesis and providing me many valuable comments, and Mr. Yanwu Zhang for many times of helpful discussions.

A special thank to Prof. Mark Grossenberg, Prof. Arthur Baggeroer, Prof. Henric Schmidt, Prof. John Leonard and Prof. George Frisk for their academic supports.

I would like to thank Ms. Jean Sucharewicz, Ms. Julia Westwater, and other staffs at the WHOI Education Office and MIT OE Administration Office, who assisted me in many ways during this work.

Finally, I wish to express my love and thanks to my parents and friends in China and here for their constant supports. Particularly, this thesis is dedicated to my coming wife, Wenpeng.

Funding supporting my thesis research project was provided by the Office of Naval Research (ONR).

# Contents

# List of Figures

# List of Tables

# Chapter 1

# Introduction

## 1.1 Bathymetric Sidescan Sonar

Bathymetric sidescan sonar is developed based on the conventional sidescan sonar technique and the phase interferometry technique. Due to its capability to integrate both seafloor bathymetry and imagery and its portability and lower cost compared with the multibeam system, bathymetric sidescan sonar has been used in many different applications, including seafloor mapping, marine geophysical research, harbor and navigation channel surveys for locating and charting underwater hazards, ROV surveys, and some other scientific, commercial and military applications underwater.

The first commercial sidescan sonar appeared in late 1950's. The typical geometry of use associated with sidescan sonar is shown in Fig. 1.1. Generally, the sonar system has two transducers, each per side, and thus two-channel imaging capability. The transducer emits CW or FM pulses in a fan-shape beam, which has a narrow horizontal beampattern ($\theta \approx 1-2°$) and wide vertical beampattern ($\varphi \approx 50-70°$). A scan line of the seafloor is

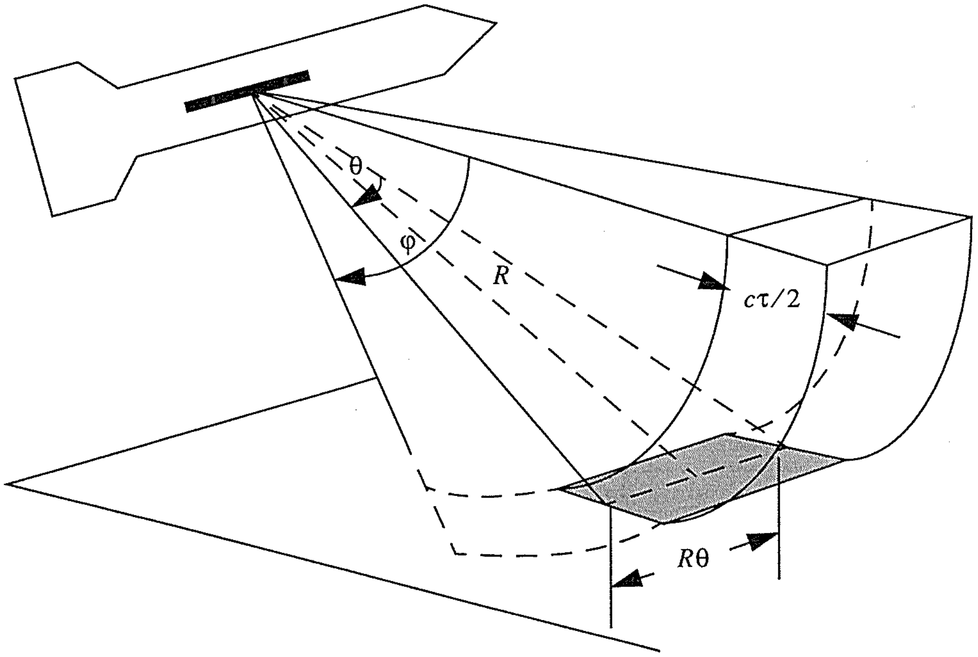


**Figure 1.1:** Geometry of use for conventional sidescan sonar

obtained by displaying the magnitudes of the bottom echoes as a function of across-track distance for a given ping. With the moving of the vehicle (Towfish, ROV or AUV), the scan lines are accumulated ping by ping to form the image. The along-track resolution is determined by the horizontal beam angle, $\theta$, and range, $R$, and the across-track resolution is determined by the transmitted pulse length, $\tau$, (cf. Fig. 1.1, $c$ is the sound speed) and sea bottom geometry. For high frequency, tens or hundreds $kHz$, the narrow horizontal beam angle and short pulse length can be easily obtained; therefore, such kind of system can be used as a high-resolution imaging tool.

One of the major disadvantages for sidescan sonar is that only the range information and no direction information can be given. For a non-flat bottom, it is impossible to relate sidescan to a depth survey, thus making the geometrically corrected interpretation of sidescan images difficult.

The bathymetric sidescan sonar system was proposed in 1970's [1]. Two parallel rows of transducers are used instead of one (Fig. 1.2). Here, each row is long in the fore-aft dimension and short athwartships. $\theta_m$ is the mounting angle of the transducer. By measuring the phase difference, $\Delta\varphi$, between two rows, the signal direction-of-arrival (DOA), $\theta(\theta_t = \theta_m + \theta)$, can be obtained by the relation $\Delta\varphi = \frac{2\pi d \sin\theta}{\lambda}$, where $d$ is the phase center separation between the two rows and $\lambda$ is the wavelength. (In a practical system, generally, a look up table is used instead of above theoretic equation relating $\theta$ and $\Delta\varphi$.) In a

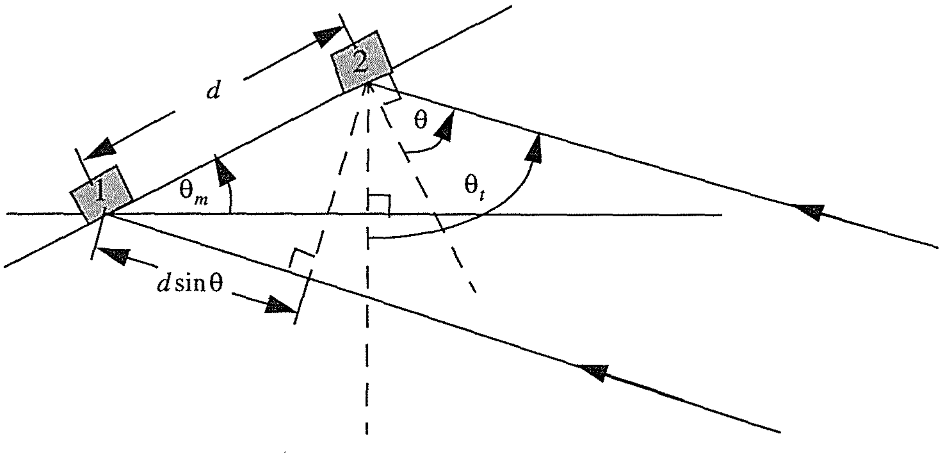


**Figure 1.2:** Geometry for bathymetric sidescan sonar: Sideview

homogeneous medium, knowing the signal angle, $\theta$, and the slant range, $R$, one can simply calculate the bottom depth by $h = R\sin\theta t$ (Fig. 1.1).

To improve resolution, widely spaced rows are preferred. However, the maximum unambiguous phase difference measurement is obtained for half-wavelength row spacing, i.e., $d = \frac{\lambda}{2}$. For $d > \frac{\lambda}{2}$, Denbigh [1] suggested two main techniques to resolve the ambiguities. The first one requires an additional pair of rows having a different spacing and therefore a different ambiguous relationship between phase and angle. Hence, the combination of two phase difference measurements corresponds uniquely with only one signal arriving angle, which is the basis of the "Vernier" technique. A practical example uses several rows with different spacing. The pair with small spacing (less than half wavelength) is used to resolve the ambiguity, while the pair with large spacing is used to obtain high resolution. The second technique of removing ambiguity is tracking the varying phase if the phase can be expected to vary slowly and monotonically along the swath.

In Denbigh's system and some other later developed systems, including SeaMARC II [2] (owned by Hawaii Institute of Geophysics), the phase difference is computed from the individual phase measurements at two rows. However, a common problem with all these bathymetric sidescan sonar systems is that their phase measurements are dispersed by noise and interferences.

The main sources of interference and noise include: 1) multipath reflecting between the ocean surface and bottom, as well as those due to multiple targets within the water column at the same range, 2) volume scattering in the water column, 3) surface scattering, 4) multiple scatters on the bottom within the acoustic footprint, 5) ambient noise. 1), 2), 3) and 5) are discussed in [1], [2] and [3]. 4) is analyzed in [4].

A few signal processing schemes have been proposed to reduce the effects of noise and interference, such as the mean time, mean angle method, and histogram approaches. Their underlying idea is smoothing data over a fixed time interval or angle interval or both. One significant improvement was proposed in [3]. In this method, the complete complex received signal is used instead of only the phase term. The phase difference is computed by estimating the spatial correlation between the received signals at two rows for some time interval (cf. section 3.1 for detail discussions). The basic idea is to identify and reject measurements at instants corresponding to a poor signal spatial correlation. Under the conditions assumed in [3], this kind of estimate is also a maximum likelihood estimate. Furthermore, Zhu, *et al.* [5], incorporated the effects of noise correlation and analytically analyzed the performance of such kind of estimator, which is consistent with the physical intuition.

Woods Hole Oceanographic Institution developed a series of high frequency bathymetric sidescan sonars using a similar method and applied them to obtain the quantitative seafloor characterization. The analysis [6] on the bathymetry power spectral density function and backscattering strength probability density function showed good agreement with previous empirical studies, thus validating the approach to wide area seafloor characterization using calibrated bathymetric sidescan sonar.

It should be pointed out that all above-mentioned methods suffer from another kind of ambiguity resulting from some simultaneously arriving echoes from different directions. This ambiguity is inherent to the sidescan geometry of use, which limits the applications, particularly in shallow water areas with significant multipath propagations.

To solve this ambiguity problem, at least partly, a new system design consideration and its corresponding signal processing scheme using a multiple-row bathymetric sidescan sonar is proposed in this thesis. The echoes from different directions, including the interferences, are to be resolved. The new system should work better for many practical situations in shallow water.

## 1.2 Signal Direction-of-Arrival and Amplitude Estimation

In a strict sense, signal direction-of-arrival (DOA) and amplitude estimation is a parameter estimation problem. It is tightly related to beamforming in that, traditionally, signal DOA and amplitude estimates are obtained on the basis of beamforming, or spatial filtering. They both constitute the very important aspects in array signal processing. Several textbooks [7, 8, 9], tutorial papers [10, 11] have been devoted to these areas. Among them, Van Veen and Buckley [10] used the spatial filtering approach, while Krim and Viberg [11] used the parameter approach. Discussions in this section also refer to [12]. For each method, an applicable data model is assumed unless otherwise indicated.

A beamformer is a processor used in conjunction with an array of sensors (transducers) to receive a signal radiating from a specific location and attenuate signals from other locations. The sensors sample the received signal in space, and the beamformer output is the weighted sum of the sensor outputs. Therefore, the spatial response of a beamformer is determined by the weights used to combine the individual sensor outputs.

In a primitive beamformer, the weight-delay-sum beamformer, the various sensor outputs are weighted and delayed by appropriate amounts relative to the reference sensor to align signal components coming from some target direction and then summed. The weights used are fixed for different steering directions. For a single source, the average power at the beamformer output is maximized when it is steered toward the source, and the beamformer resolution is determined by the array aperture and steered direction. However, this method breaks down completely in the presence of multiple sources.

In order to enable a beamformer to respond to an unknown noise and interference environment, some adaptive weighting methods have been developed. A well-known method, Capon's Maximum Likelihood Method, also called the Minimum Variance Distortionless Response method in a different context, was proposed by Capon [13] in 1969. The basic idea of Capon's method is to choose weights to minimize the power contributed by noise and signals from directions other than the steered direction, while keeping a fixed gain in the steered direction. The power minimization can also be interpreted as placing nulls in the directions of interferences. Many other methods, such as generalized sidelobe canceller (GSC), were developed using the same principle. Capon's method did achieve a significant resolution improvement over conventional beamformer. However, it is not a true ML method (true only for a single source situation), and its performance degrades in the presence of highly correlated signals.

Another important adaptive method is Burg's Maximum Entropy Method (MEM) [14]. Observing that the estimated spatial covariance function has only a finite number of lags due to the finite number of sensors, Burg's method does the covariance extension according to Burg's Entropy criterion, which assumes a maximal random data model. For the deterministic signals or signal with deterministic components, MEM's performance becomes worse. For a uniform linear array, Capon's method and MEM can be related to each other simply [9].

In the context of signal DOA and amplitude estimation, although both Capon's method and MEM are often successfully and widely used, they have certain fundamental limitations, such as the estimation bias and sensitivity, partly because they haven't exploited the underlying data model structure effectively.

Development of subspace-based methods, which explicitly invokes the eigenstructure of the covariance matrix, is a significant contribution for signal parameter estimation problem. One such early attempt was done by Pisarenko [15], who developed a new method to retrieve harmonics from a finite segment of the covariance function. His idea could be extended to the signal DOA estimation problem for a uniform linear sensor array.

The tremendous interests in the subspace approach largely attributed to the MUSIC (Multiple SIgnal Classification) algorithm introduced by Schmidt [16] in later 1970's. One of the important contributions in MUSIC is the geometric approach to signal parameter estimation. In this approach, the signal subspace is estimated by eigendecomposition of the spatial covariance matrix, and the array manifold is obtained by calibrating all possible sensor responses to single rank one signal with varied parameters. For a suitably designed array, the signal parameter space is related to the array manifold without ambiguity, so the signal parameters can be determined uniquely by finding the close regions between the estimated signal subspace and the array manifold according to some optimality criterion. Some common used criteria include the least-squares criterion, which selects the model that minimizes the sum of squared errors between the data and the model output, and the maximum likelihood criteria, which select the parameter vector associated with the most likely measurements. In conventional MUSIC algorithm, a one-dimensional search for parameters is employed, thus reducing the computation load, while producing biased estimates with finite samples. In root-MUSIC, a multi-dimensional search is employed, which, generally, is computationally prohibitive except for the uniform linear array. Many other extensions to conventional MUSIC are developed. In MUSIC, the maximum resolvable source number is the sensor number.

As the first high-resolution algorithm to correctly exploit the underlying data model of narrow-band signals in additive noise, the performance improvement of MUSIC was so significant that it became an alternative to most existing DOA estimation algorithms. However, the prices paid for its excellent performance are the heavy computation load in searching over the parameter space, and a large amount of data storage from the arduous array calibration. Besides, MUSIC is sensitive to calibration errors including the sensor position error and sensor gain/phase error, and cannot give the signal amplitude estimate.

In the middle of 1980's, ESPRIT (Estimation of Signal Parameters Via Rotational Invariance Techniques), a new subspace-based approach to estimation of parameters of cisoids in noise was proposed by Roy, *et al.* [17], and was extended to signal DOA estimation [18]. Following the same geometric approach as MUSIC, ESPRIT exploits an underlying rotational invariance among signal subspaces and obtains the signal DOA estimate without searching parameter space, thus reducing the computation loads and storage costs significantly. Besides, ESPRIT shows more robust performance with respect to array perturbations, and can obtain optimal signal copy thus optimal signal amplitude estimate. On the other hand, ESPRIT is not a general method because it requires the array manifold to possess a displacement invariance; and the maximum resolvable source number is half the sensor number unless an overlapping subarray structure is used (cf. further discussions on ESPRIT at section 3.3).

A number of research papers on ESPRIT have been published since then. Ottersten, *et al.* [19], followed the idea of Su and Morf [20], which models the source signal as the stationary output of a finite dimensional linear system driven by white noise, and developed a wide-band signals DOA estimation method using ESPRIT. Ottersten, *et al.* [21], analyzed the performance of the Total Least Squares (TLS) ESPRIT algorithm and showed the TLS ESPRIT is competitive with the MUSIC, and the performance is close to the calibrated Cramer-Rao bound (CRB) for many practical cases. However, as discussed in later sections, for highly correlated signals, the estimates deviate. Furthermore, Swindlehurst, *et al.* [22], extended the original ESPRIT to exploit arrays with multiple invariances.

In practical application, the ESPRIT algorithm has been used to estimate the angle of arriving signal and then detect the weak targets in SAR (Synthetic-Aperture-Radar) data processing by Curlander [23]. Merwe, *et al.* [24], used the TLS ESPRIT to generate high-resolution two-dimensional microwave images. Wong, *et al.* [25, 26], applied ESPRIT to a velocity-hydrophone array to locate acoustic sources. These results have shown that ESPRIT has great potential to be integrated into a practical real-time processing system.

Above mentioned array signal processing techniques are based on the second-order statistics of the received signals, most of which assume, explicitly or implicitly, a Gaussian signal. In recent years, some DOA estimation algorithms based on high-order statistics, such as fourth-order cumulants, are proposed [27, 28, 29]. For non-Gaussian signals and additive Gaussian noise, such as digital communication signals, some ocean acoustic propagation signals with only a few Gaussian components, cumulant-based algorithms work well. Nevertheless, it is obvious that a cumulant-based algorithm is very sensitive to the signal model, and also the estimation of high-order statistics needs to be further investigated.

Finally, as we have seen from above discussions, estimation of the covariance matrix plays a key role in signal DOA and amplitude estimation. A generalized coherence estimation framework was proposed by Knapp and Carter [30], in which different estimators are chosen to optimize certain different performance criteria. Also, a tutorial overview of the coherence and time delay estimation was given by Carter [31].

## 1.3 Thesis Overview

This thesis proposes a new idea to deal with the interferences using a multiple-row bathymetric sidescan sonar, which is, resolving the interferences from the signals of interests instead of only smoothing the received data (signals + interferences + noise) for given time or angle intervals. Correlated signal direction estimate (CSDE), a new signal DOA and amplitude estimation method for three-row systems, is developed in the absence of noise, and then applied to the situation with noise. ESPRIT-based approaches are also proposed, and they are compared with CSDE using different data model. Results from simulations and a few real data tests are given and analyzed, and some system realization issues are addressed. Because the thesis research is supported by a three-row bathymetric sidescan sonar project from ONR (Office of Naval Research), most of the thesis discussions are in the context of the three-row system, though the basic idea and some of the algorithms can be easily extended to multiple-row systems.

The organization of the rest of this thesis is as follows:

Chapter 2 analyzes the acoustic field in the bathymetric sidescan sonar applications and the stochastic characteristics of the bottom scattering, and then establishes the problem and data model.

Chapter 3 studies the signal processing schemes for DOA and amplitude estimation. Three methods and their applicable data models are discussed in detail.

Chapter 4 presents the simulation and real data test results. The related issues to the data model and algorithm parameters are addressed in detail.

Chapter 5 discusses some issues on the practical implementation of the proposed algorithms in a real sonar system, and summarizes the results of the thesis. Limitation of current work and recommendations for future work are discussed.

# Chapter 2

# Problem and Data Model

## 2.1 Analysis on Interference Field Associated with Sidescan Sonar

As indicated in the first Chapter, one of the main sources of interference associated with sidescan sonar is the multipath reflection between the ocean surface and bottom. To analyze the multipath interference behavior qualitatively, a horizontally stratified ocean acoustic model is chosen as shown in Fig. 2.1.

The ocean surface is assumed to be pressure-release, i.e., pressure $P = 0$, and the surface scattering effect is ignored. In practical, the surface backscattering illuminated by the sidelobe is significant sometimes, but can be reduced by adjusting the beampattern design and system installation. For applications with tow-fish, AUV, or ROV at a depth far away from the surface, this assumption is reasonable.

Sidescan sonar is an active system in that a short pulse of acoustic wave is transmitted, and the backscattering echoes are received. Thus, the bottom can be modeled as an infinite set of discrete point sources (excited by the transmitted acoustic pulse) with strength, $S_i$. Each source is contributed by the whole scattering region of the footprint (Fig. 1.1) ensonified by the acoustic pulses. As horizontal range, $r$, increases, the bottom backscattering strength attenuates. On the other hand, the distance between the contiguous sources, i.e., $r_i - r_{i-1}$, increases to keep consistent with the system resolution issue. Only the first-order scattering is considered, which means we consider only the reflection wave after radiated from the point source. The bottom is penetrable with $|C_R| \le 1$, where $C_R$ is the bottom reflection coefficient determined by the incident angle and medium properties (density and sound velocity). Here the density and velocity are assumed to be constants in the water and bottom, respectively.

The homogeneous, time-independent wave equation with source $f(\vec{R})$ [32] is

$$\nabla^2 P(\vec{R}) + k^2 P(\vec{R}) = -2f(\vec{R}) \tag{2.1}$$

where $k$ is the wavenumber, $\vec{R} = (r, z)$, $f(\vec{R}) = \sum_{i=0}^{\infty} S_i \delta(\vec{R} - \vec{R}_i)$, and $\vec{R}_i$ is the $i$-th source coordinate.

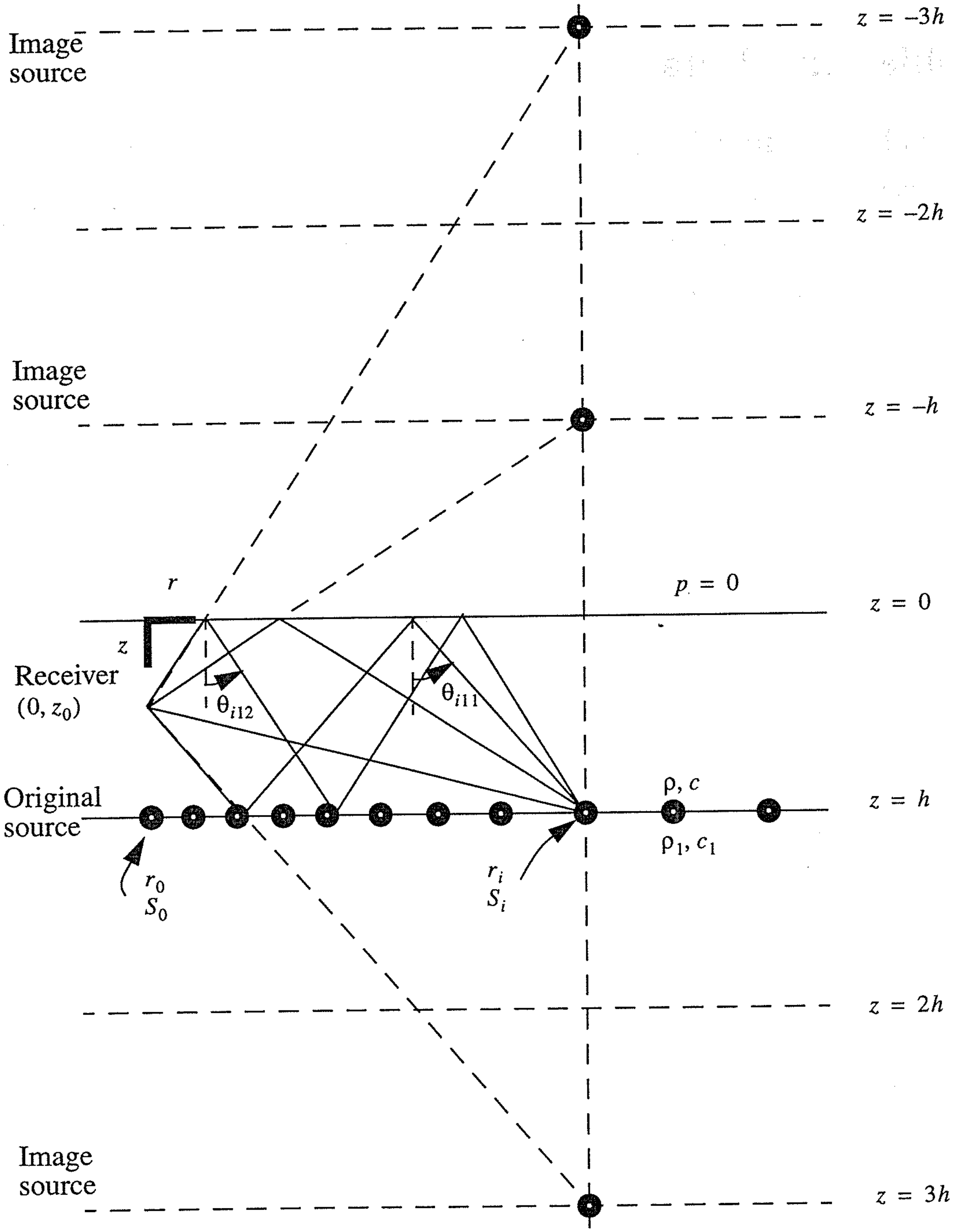


**Figure 2.1:** Ocean acoustic model for sidescan sonar

Even for this simplified model, it is still difficult to obtain the closed-form solution, due to the penetrable bottom. We must resort to some approximation methods. For the typical operating frequency used by a sidescan sonar, tens or hundreds $kHz$, ray theory [32] is a good choice. Combining all possible eigenrays using image method (Fig. 2.1), we obtain the pressure field produced by a source at $\vec{R}_i(r_i, h)$ as

$$P(\vec{R}, \vec{R}_i) = \sum_{n=0}^{\infty} (-1)^n \left( C_R^n(\theta_{in1}) \frac{e^{ikR_{in1}}}{R_{in1}} - C_R^n(\theta_{in2}) \frac{e^{ikR_{in2}}}{R_{in2}} \right) \tag{2.2}$$

$$R_{in1} = \sqrt{(r - r_i)^2 + (z - (2n+1)h)^2} \tag{2.3}$$

$$R_{in2} = \sqrt{(r - r_i)^2 + (z + (2n+1)h)^2} \tag{2.4}$$

where $C_R(\theta_{in1})$ is the bottom reflection coefficient at the angle $\theta_{in1}$ for $n > 0$, and equals to one for $n = 0$. Note, $\theta_{in1}$ is the incident angle to the bottom experienced by the propagation wave from the first $n$-th (image) source to the receiver (Fig. 2.1). Similarly, $C_R(\theta_{in2})$ is associated with the second $n$-th (image) source.

The total wave field at the sonar receiver, $\vec{R}_0 = (0, z_0)$, is

$$P(\vec{R}_0) = \sum_{i=0}^{\infty} S_i \sum_{n=0}^{\infty} (-1)^n \left( C_R^n(\theta_{in1}) \frac{e^{ikR_{in1}}}{R_{in1}} - C_R^n(\theta_{in2}) \frac{e^{ikR_{in2}}}{R_{in2}} \right) \tag{2.5}$$

$$R_{in1} = \sqrt{r_i^2 + (z_0 - (2n+1)h)^2} \tag{2.6}$$

$$R_{in2} = \sqrt{r_i^2 + (z_0 + (2n+1)h)^2} \tag{2.7}$$

Now, we apply above results to the analysis of the interference behavior. Note, the contributions from all the individual (image) sources arrive the receiver at different times. Assume the sonar transmits a pulse at time $t = 0$. The first received signal is from the point source at nadir, $r_0 = 0$, at $t = \frac{2(h - z_0)}{c}$. The first multipath signal occurs at $t = \frac{2h}{c}$. The second multipath signal occurs at $t = \frac{2(2h - z_0)}{c}$ with amplitude attenuated by $C_R(\theta_{011})$. The third multipath signal begins at $t = \frac{4h}{c}$ with amplitude attenuated by $C_R(\theta_{012})$. And so on. As $t$ increases, the propagating distances for both the direct signal and multipath signal, and thus the energy losses for them, increase. Above analysis is consistent with the observation to SeaMARC II data [2].

Here an omnidirectional sensor is implicitly assumed. In a practical system (Fig. 1.1), except for those angular sectors of interests, most other angular sectors are attenuated by the beampattern directivity. Therefore, if a multipath interference appears at these sectors, it will be further attenuated. In sidescan sonar design, a null toward the nadir in beampattern is often preferred to avoid the strong reflection there.

From above discussions, for the time interval, $t = \frac{2(h-z_0)}{c}$ to $t = \frac{2(2h-z_0)}{c}$, one direct signal and at most one multipath propagation signal are received at the same instant. From $t = \frac{2(2h-z_0)}{c}$ to $t = \frac{4h}{c}$, one direct signal and at most two multipath propagation signals are received at the same instant. Higher-order multipath signals are attenuated significantly relative to the direct signal, and thus can be ignored.

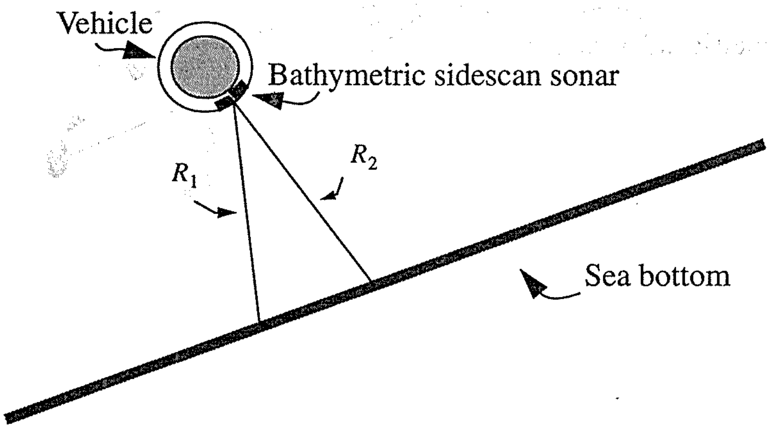
Two examples are given below with $c = 1500m/s$. For $h = 50m$, and $z_0 = 10m$, a typical situation in shallow water, above two time intervals are $0.053s \sim 0.12s$ and $0.12s \sim 0.133s$. Assuming a flat bottom, the corresponding horizontal range intervals are $0 \sim 80m$ and $80 \sim 92m$, and the corresponding grazing angle intervals are $90° \sim 26°$ and $26° \sim 23°$, respectively. For a sidescan sonar with carrier frequency $200kHz$, the typical operating (slant) range is about $200m$, which corresponds to a maximum horizontal range $\sim 200m$ here. In further distance, the direct signals are attenuated significantly and the image data are no longer meaningful.

For $h = 100m$ and $z_0 = 60m$, an example in middle to deep water, the two time intervals are $0.053s \sim 0.187s$ and $0.187s \sim 0.267s$. Assuming a flat bottom, the corresponding horizontal range intervals are $0 \sim 134m$ and $134 \sim 196m$, and the corresponding grazing angle intervals are $90° \sim 17°$ and $17° \sim 12°$, respectively.

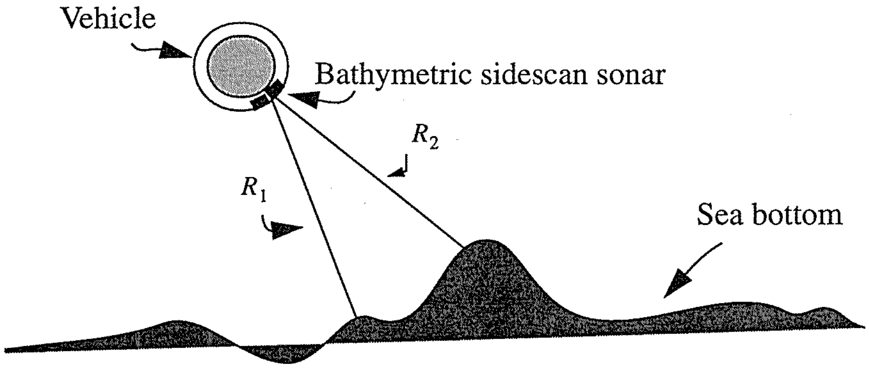
As shown in these two examples, the multipath signals are significant particularly in shallow water, while the direct signals dominate particularly in deep water, which is consistent with the observation to real ocean acoustic field. However, the model used doesn't imply any time-varying behavior, which does exist in the real field.

A key observation is that for grazing angles less than $10°$, the bottom backscattering strength decreases dramatically. So does the signal-to-noise ratio (SNR) at the receiver. Thus the signal DOA estimates and the bathymetry measurements are not reliable. For those regions with larger grazing angles (and thus with reliable bathymetry measurements), a model with two or three simultaneously received signals describes even some shallow water fields very well.

Fig. 2.2 gives some other kinds of interferences, which occur frequently in real applications. In Fig. 2.2 (a), bathymetric sidescan sonar works in a sloping bottom area. For equal $R_1$ and $R_2$, two signals from different directions arrive the receiver at the same time. In Fig. 2.2 (b), the sea bottom is non-flat. Similarly, for equal $R_1$ and $R_2$, two or more signals are received simultaneously at some instants due to the bottom relief.



(a) AUV survey above a sloping bottom

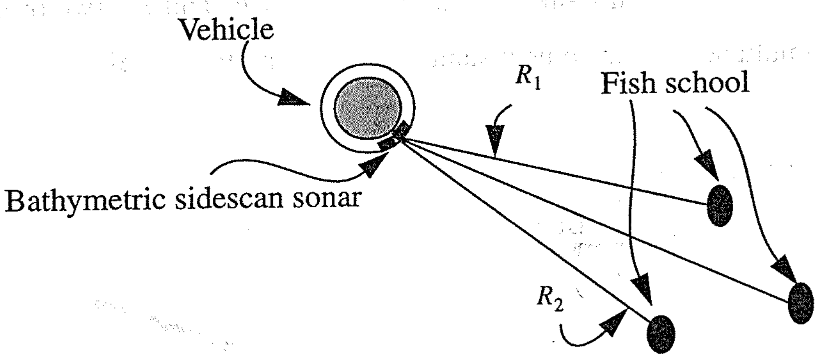


(b) AUV survey above a non-flat bottom

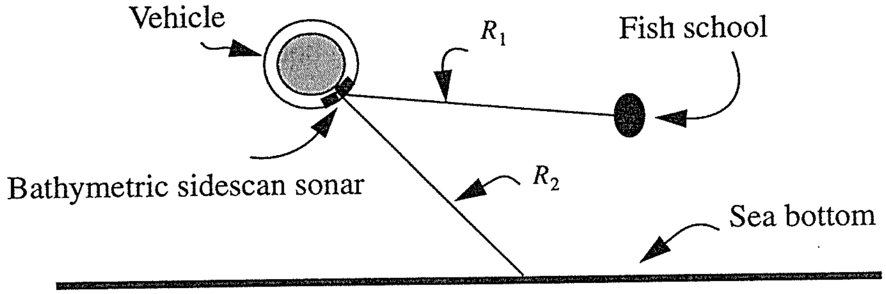
**Figure 2.2:** Two application cases using bathymetric sidescan sonar

## 2.2 Fishery Survey Application

Another application area we may have some interests is the fishery survey. Traditionally, the research fishery surveys are conducted by catching fish with a net and manually counting and measuring the samples. This operation is arduous, time-consuming and expensive. An automatic large-volume high-resolution mapping system is to be developed, which can be realized by combining optical [33] and acoustical methods. Two examples of such an application are given in Fig. 2.3.



(a) Fish schools in the water column



(b) Fish school above the bottom

**Figure 2.3:** Fishery survey using bathymetric sidescan sonar

Hendershot and Jackson [34] proposed an acoustic fish school measurement system of using the phase information, the same approach in the conventional bathymetric sidescan sonar. The fish school is represented by an ellipsoid filled with point scatterers. Depending on range, size and scattering property of the fish school and sonar beampattern, one or more scattering signals from a single fish school are received at the same or different times. In the examples with dense fish schools or fish school close to the bottom (Fig. 2.3 (a) and (b)), when $R_1$ is equal to $R_2$, the system needs to be able to simultaneously resolve two directional signals. The sonar system with this improved capability can provide a coarse estimation about the distribution and density of some fish species for the optical mapping and classification system, and extend the proposed fishery survey to more practical situations.

## 2.3 Statistical Analysis on Bottom Scattering

In section 2.1, we used a point source model for the bottom echo, and mentioned each such source is contributed by the whole scattering region within the footprint. In order to make more clearly about the measurement process, an introduction of the statistical characteristics of the sea bottom scattering (reverberation) is appropriate.

The first systematic presentation on characteristics of sea reverberation using statistical method was given by Ol'shevskii [35] in 1966. In [36], Klepsvik did a similar work with applications to wide-swathe bathymetric mapping. According to [35], the sea bottom reverberation process can be treated as a discrete model

$$S_n(t) = \sum_{i=1}^{n} a_i s(t - t_i, \xi_i) \tag{2.8}$$

where $a_i$ denotes the random amplitude with probability density function $p(a)$, and $t_i$ is the onset time of the $i$-th scattering element, which is often assumed to have a uniform distribution in a small time interval $(-T/2, T/2)$. $s(t)$ describes the transmitted signal, and $\xi_i$ is a stochastic parameter defining the characteristics of the elementary scattering signals with probability density $p(\xi_i)$. For nonstochastic parameter $\xi = \xi_0$, we have $p(\xi) = \delta(\xi - \xi_0)$. The number of scattering element, $n$, arriving at the reception point at the time $t$ is also a random variable, whose distribution satisfies the Poisson Law if the scatterer positions are statistically independent, and the mean scatterer density is constant for a sufficiently large scattering region. In [36], the model used is a little bit different in that the scatterer space distribution is considered.

Generally, the bottom scattering process is non-stationary. For the bathymetric sidescan sonar using the narrow-band signal with the geometry given in Fig. 1.1 and 1.2, we can neglect the relative motion between the transmitter and receivers and doppler effects, and treat the process stationary.

According to the central limit theorem, if the number of elementary scattering signals arriving at the reception position at a given time instant is large, and none of the these signals dominates, then the one-dimensional distribution, $p(S)$, of the scattering process approximates to be Gaussian, i.e.,

$$p(S) = \frac{1}{\sqrt{2\pi}\sigma_S} e^{-\frac{S^2}{2\sigma_S^2}} \tag{2.9}$$

Note in general, $S$ is a complex process. Its instantaneous phase $(\psi)$ distribution is uniform in the interval $(0, 2\pi)$, i.e.,

$$p(\psi) = \frac{1}{2\pi}, 0 \le \psi < 2\pi \tag{2.10}$$

The envelop $(E)$ satisfies the Rayleigh distribution,

$$p(E) = \frac{E}{\sigma_S^2} e^{-\frac{E^2}{2\sigma_S^2}} \tag{2.11}$$

For the sidescan sonar with narrow beam and short pulse, the number of scatterer within the footprint may be reduced significantly, particularly for the bottom where shadowing occurs. Consequently, the resulting distribution departs from a Gaussian law and shows some Poisson characters. Though, the Gaussian model is still a good approximation for off-normal scattering region with larger size of the footprint and random scattering components.

For scattering at normal incidence, a significant contribution from the mean amplitude to the total scattering echo occurs. Hence, the Gaussian model is invalid, and the envelop distribution in this situation approximates to be so-called Rice distribution with

$$p(E) = \frac{E}{\sigma_S^2} e^{-\frac{(E+A_0)^2}{2\sigma_S^2}} I_0\left(\frac{EA_0}{\sigma_S^2}\right) \tag{2.12}$$

where $I_0(\ )$ is a zero-th order modified Bessel function, and $A_0$ is the amplitude of the dominating scattering component.

In a real application, such an analysis is much more complex, and the validity of a statistical model depends on many factors associated with the bottom properties (roughness, relief and material) and system parameters (frequency, signal type and pulse length). Though, analysis of some real bottom scattering data obtained by a bathymetric sidescan sonar [6] has shown a good agreement with the Gaussian model at a narrow grazing-angle band for quite large regions including sediment-pond and axial valley. On the other hand, in a wide range of grazing angles, a multimodal Rayleigh envelop distribution model fits the data well.

The correlation analysis of the scattering process shows that for a rectangular transmitted pulse with a sinusoidal carrier, the correlation interval of the instantaneous value of a scattering process is about one-half of the pulse-length [35]. A measurement correlation interval from a real bathymetric sidescan sonar is given in [36], which approximates twice the specific pulse length. Considering the dispersion of the propagation medium, it is a reasonable result.

The cross-correlation between the signals received at two points is, in general, a function of SNR, wavelength, spacing between the two point receivers, and receiving array orientation to the scattering source [35]. In [36], the cross-correlation is further factorized to include the dependence on both the transmitted pulse and the scattering characteristics. For a two-row high-frequency bathymetric sidescan sonar with half-wavelength row spacing, it can be expected that signals at two rows are highly correlated.

Finally, we give the probability density function for the phase-difference $\Delta\varphi$ obtained on the basis of above analyses [35] as

$$p(\Delta\varphi) = \frac{1-|\gamma|^2}{2\pi(1-\beta^2)^{3/2}}\left(\sqrt{1-\beta^2}+\pi\beta/2+\beta\operatorname{asin}\beta\right) \text{ for } -\pi \le \Delta\varphi-\mu \le \pi \qquad (2.13)$$

where $\mu$ is the mean value, $\gamma$ is the cross-correlation coefficient, and $\beta = |\gamma|\cos(\Delta\varphi-\mu)$. As the cross-correlation coefficient increases, the width of the peaked distribution decreases, and thus the estimated phase difference approaches more to the true value.

## 2.4 Problem Definition and Data Model

The discussions in the first two sections suggest that a new bathymetric sidescan sonar can be developed to obtain more accurate field measurements if two or more simultaneous directional signals can be resolved. This can be done by adding more rows to the conventional sidescan sonar.

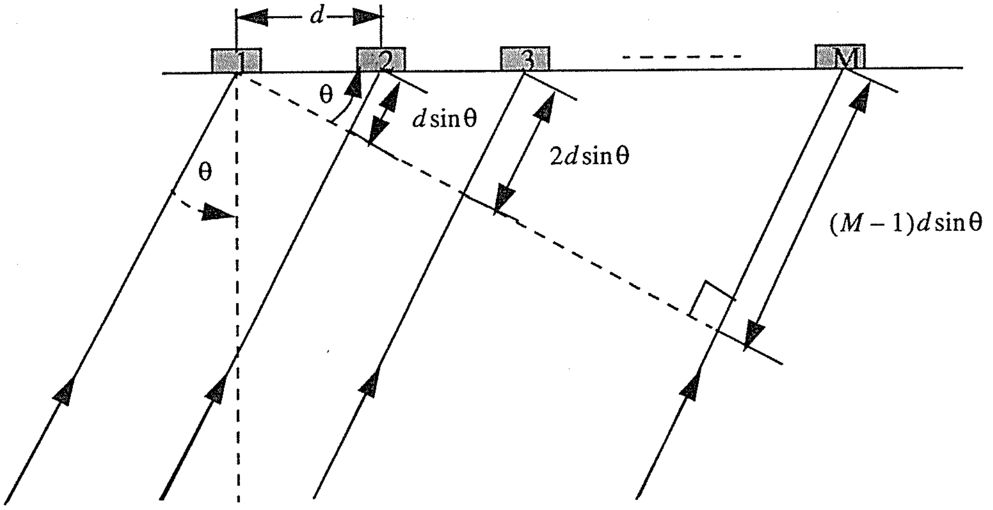


**Figure 2.4:** Multiple-row bathymetric sidescan sonar

A M-row bathymetric sidescan sonar is configured in a uniform linear array (ULA) (Fig. 2.3), in which the mounting angle is ignored for convenience. The row spacing, $d$, is half-wavelength to avoid the direction ambiguity. In the context of bathymetric sidescan sonar, we use "row" instead of "sensor" to emphasize the geometric property of the sensor here, i.e., long in one direction and short in another direction. The choice of row number is almost purely an engineering problem. Taking into account the space available, weight allowable, power limitation and costs, two to five rows may be a reasonable choice.

By collecting data from all the rows, we are to estimate signal DOA and amplitude for multiple directional sources, including the bottom echo, multipath interference, and/or other target return in water column within the sonar operating range.

The sources associated with a bathymetric sidescan sonar are time-varying in the sense that their DOA's change with time. In other words, using the model in section 2.1, for each source, there is a limited data points to be used to estimate its direction and amplitude, if these parameters change slowly. The data length available for a single source is determined by the system parameters including the pulse length and sampling rate, by the application geometry, and by the resolution obtainable from the signal processing algorithm (cf. chapter 4 for further discussion). Here we still treat the continuous time-varying sources as a discrete set of point sources, for each of which a signal DOA estimate can be obtained individually. Therefore, the parameter estimation problem for a whole sonar ping is simplified to the parameter estimation for a discrete set of sources. (Note, because all these discrete sources come from the same sonar ping, their parameters are more or less correlated.)

A few assumptions are invoked in our model and specified further below.

The transmitted medium is assumed to be homogeneous (constant density and velocity), so the propagating rays are in the form of straight lines. Furthermore, the receiver array is in the far-field of the source for $kr \gg 1$, which can be easily satisfied in high-frequency. Under these approximations, the wave field at the receiver array is a sum of plane waves. Note, for the frequency used here, $\sim 200 kHz$, the wavelength, $\lambda$, is about $0.01 m$. With the half-wavelength row spacing, it is reasonable to assume that the received signal at each row is just a delaying version with respect to the reference row, which, generally, is chosen as row 1.

The problem here is a planar problem from the sonar geometry. Assuming a single source, $s_i$, in $\theta_i$ direction, the received data at the $k$-th row can be represented as

$$x_k(t) = a_k(\theta_i) s_{ir}(t - \tau_k(\theta_i)) \tag{2.14}$$

where $s_{ir}(t)$ is the received signal at the reference row, and $\tau_k(\theta_i) = \frac{(k-1)d\sin\theta_i}{c}$, is the propagation delay between the $k$-th row and the reference row for a wavefront from $s_i$. $s_{ir}(t)$ is related to the source by $s_{ir}(t) = s_i / r$, where $r$ is the distance between the source and the reference row. $a_k(\theta_i)$ is the normalized response of the $k$-th row to the wavefront from $\theta_i$ direction.

For typical bathymetric sidescan sonar, the narrow-band signal modulated at center frequency, $w_0$, is transmitted, i.e.,

$$s_{ir}(t) = m_{ir}(t)\cos(w_0 t + \varphi_{ir}(t)) \tag{2.15}$$

where $m_{ir}(t)$, $\varphi_{ir}(t)$ are slowly varying function of time which modulate the amplitude and phase of $s_{ir}(t)$, respectively.

Note, the narrow-band signal ensures that the array response is independent of frequency over the signal bandwidth. Using the complex envelope representation, a complex form of the received signal at the reference row is

$$\tilde{s}_{ir}(t) = s_{ir}(t) + j\hat{s}_{ir}(t) = [m_{ir}(t)e^{j\varphi_{ir}(t)}]e^{jw_0 t} \tag{2.16}$$

where $\hat{s}_{ir}(t)$ is the Hilbert transform of $s_{ir}(t)$, given by

$$\hat{s}_{ir}(t) = \frac{1}{\pi}\int_{-\infty}^{\infty} \frac{s_{ir}(\tau)}{t-\tau}d\tau \tag{2.17}$$

Thus, the received data at $k$-th row in a complex form are

$$\tilde{x}_k(t) = a_k(\theta_i)\tilde{s}_{ir}(t-\tau_k(\theta_i)) \approx a_k(\theta_i)\tilde{s}_{ir}(t)e^{-jw_0\tau_k(\theta_i)} \tag{2.18}$$

Here, we used $m_{ir}(t) \approx m_{ir}(t-\tau_k(\theta_i))$ and $\varphi_{ir}(t) \approx \varphi_{ir}(t-\tau_k(\theta_i))$ for all possible delays.

If we have $D$ directional signals at the same instant, the received data including the measurement noise, $\tilde{n}_k(t)$, at the $k$-th row become

$$\tilde{x}_k(t) = \sum_{i=1}^{D} a_k(\theta_i)\tilde{s}_{ir}(t)e^{-jw_0\tau_k(\theta_i)} + \tilde{n}_k(t) \tag{2.19}$$

For $M$ rows, it is more convenient to use the matrix representation.

Defining

$$a_k(\theta) = \left[a_k(\theta_1)e^{-jw_0\tau_k(\theta_1)}, \ldots, a_k(\theta_D)e^{-jw_0\tau_k(\theta_D)}\right] \tag{2.20}$$

$$A(\theta) = \left[a_1(\theta), \ldots, a_M(\theta)\right]^T \tag{2.21}$$

$$\tilde{S}_r(t) = \left[\tilde{s}_{1r}(t), \ldots, \tilde{s}_{Dr}(t)\right]^T \tag{2.22}$$

$$\tilde{X}(t) = \left[\tilde{x}_1(t), \ldots, \tilde{x}_M(t)\right]^T \tag{2.23}$$

$$\tilde{N}(t) = \left[\tilde{n}_1(t), \ldots, \tilde{n}_M(t)\right]^T \tag{2.24}$$

where $[\,]^T$ means the matrix transposition operation, then we have

$$\tilde{X}(t) = A(\theta)\tilde{S}_r(t) + \tilde{N}(t) \tag{2.25}$$

Note, now the direction information is contained in the time delay and thus the phase term of the received signal.

In the problem definition, the sensor output is decomposed into two components, the signal to be observed and the noise due to sources of uncertainty in the process of observing the signal. The noise is due to ambient, sea reverberation, and receiver electronics, and can be treated as additive. Generally speaking, the noise includes both nondeterministic components and deterministic components such as multipath interferences. For the problem of interest here, a multipath interference is treated as a directional signal and thus the noise process is assumed to be a zero mean white Gaussian process. In practical, "white" means the power spectral density of the noise component is flat over the bandwidth of interests. Furthermore, the noise correlation are often ignored in the development of the algorithms. However, for some components of the noise process, there do exist some correlation between the noise and signal and between the noises at two different rows, which will be addressed in the performance evaluation.

Two kinds of signal models are considered. First, the stochastic one from the statistical analysis on bottom scattering in section 2.3, which may also be used for the scattering from fish school under some conditions. Generally, a Gaussian model is assumed for each bottom scattering process from a narrow grazing-angle band. Second, the coherent one from the analysis on interference field in section 2.1. A mathematical model for the received signal $x(t)$ in multipath environment is proposed by Ehrenberg [37] as

$$x(t) = \sum_{i=1}^{N} m_i s(t-\tau_i) + n(t) \tag{2.26}$$

where $m_i$ and $\tau_i$ are the amplitude and arrival time for the $i$-th multipath signal, $N$ is the number of multipaths, $s(t)$ is the propagating waveform of a single path, and $n(t)$ is the additive noise. It is easily to be seen that (2.26) is a particular version of (2.25). From this model, we can expect to see that, for some instants of time, two processes arrive at the reception array simultaneously, one of which is a delayed version of another. In a real environment, the data structure is generally much more complex, and these models might only be two components of the received signals.

# Chapter 3

# Methodologies

## 3.1 Differential Phase Estimate

In this chapter, the solutions for the problem defined in last chapter are developed and studied. First, the cases with a single directional source are considered, where the conventional differential phase estimation (DPE) method [3] works.

Assuming a single source in $\theta$ direction, at the $k$-th row, the received signal in a complex notation (cf. section 2.4) is

$$\tilde{x}_k(t) = a_k(\theta)\tilde{s}_r(t)e^{-jw_0\tau_k(\theta)} + \tilde{n}_k(t) \tag{3.1}$$

We require all the rows have the same beampattern, i.e. $a_1(\theta) = a_2(\theta) = \ldots = a_M(\theta)$ for all $\theta$, and for further convenience, an omnidirectional directivity in vertical plane is assumed for each row, i.e., $a_k(\theta) = 1, \text{for all } \theta$.

For a properly bandlimited received signal, it is possible and sufficient to deal with the base-banded quadrature samples, which can be obtained by a complex demodulator or a real demodulator plus the Hilbert tranformer in digital domain. Thus, the base-banded outputs in digital domain at the $k$-th and $(k+1)$-th row are

$$\tilde{x}_{kb}(n) = \tilde{s}_{rb}(n)e^{-jw_0\tau_k(\theta)} + \tilde{n}_{kb}(n) \tag{3.2}$$

$$\tilde{x}_{(k+1)b}(n) = \tilde{s}_{rb}(n)e^{-jw_0\tau_{k+1}(\theta)} + \tilde{n}_{(k+1)b}(n) \tag{3.3}$$

In the absence of noise, we form a product by multiplying $(3.2)$ by the conjugate of $(3.3)$

$$R(n) = \tilde{x}_{kb}(n)\tilde{x}^*_{(k+1)b}(n) = |\tilde{s}_{rb}(n)|^2 e^{jw_0(\tau_{k+1}(\theta) - \tau_k(\theta))} \tag{3.4}$$

Here $(\ )^*$ is the complex conjugate operation. Then (cf. section 2.4),

$$\phi_R(n) = w_0(\tau_{k+1}(\theta) - \tau_k(\theta)) = \frac{2\pi d \sin\theta}{\lambda} \tag{3.5}$$

where $\phi_R(n)$ is the angle or phase of $R(n)$. Thus,

$$\theta = \operatorname{asin}\frac{\phi_R(n)\lambda}{2\pi d} \tag{3.6}$$

Note, for a linear array, an inherent problem is that it cannot resolve the up-going wave from the down-going wave. Fortunately, a sidescan sonar is so designed that only up-going (or down-going, depending on the manner used) wave with respect to the transducer array plane can be sensed. Therefore, $\theta$ can be chosen in the range of $[-\pi/2, \pi/2]$.

For noisy signal, an estimate of $R(n)$ should be obtained. Given $L$ samples of the process, the least square distance estimate $\hat{R}$ can be obtained by minimizing the cost function

$$C(\hat{R}) = \sum_{n=1}^{L} W^2(n)(R(n) - \hat{R})^2 \tag{3.7}$$

and the resulting $\hat{R}$ is

$$\hat{R} = \frac{\sum_{n=1}^{L} W^2(n)R(n)}{\sum_{n=1}^{L} W^2(n)} \tag{3.8}$$

i.e., the average of the corresponding $L$ samples. Here, $W$ is a window function, and the rectangular window is often used to obtain better resolution [38]. The choice of $L$ is a trade-off between the resolution and smoothness. Due to the time-varying signal DOA in sidescan sonar, large $L$ yields a smoother output, but causes loss in resolution, which is different from the behavior in spectral estimation.

Note,

$$\hat{R}(n) = R(n) + \tilde{s}_{rb}(n)e^{-jw_0\tau_k(\theta)}\tilde{n}_{(k+1)b}(n) + \tilde{s}_{rb}(n)e^{-jw_0\tau_{k+1}(\theta)}\tilde{n}_{kb}(n) + \tilde{n}_{kb}(n)\tilde{n}_{(k+1)b}(n) \tag{3.9}$$

If both the directional signal and noise are zero-mean random processes, and the noises are uncorrelated with the directional signal and with each other, then the estimate $\hat{R}$ approaches to the true $R$ as $L$ increases. A more realistic argument is given in [3] that for highly correlated directional signal and zero-mean Gaussian noise, above estimate is a maximum likelihood estimate at high SNR.

For a $M$-row bathymetric sidescan sonar, better estimate can be obtained by combining all the estimates from arbitrary pairs of rows. Let $\hat{\theta}_{ij} = \operatorname{asin}\frac{\phi_{\hat{R}_{i,j}}\lambda}{2\pi d}$ represents the estimate at the pair of the $i$-th and $j$-th rows, then

$$\hat{\theta} = \frac{2}{M(M-1)} \sum_{i=1,\, j=2,\, i<j}^{M} \hat{\theta}_{ij} \tag{3.10}$$

However, a problem of ambiguity (phase wrapping) arises when the spacing between the rows in a pair is larger than the half-wavelength. For example, for $\hat{\theta}_{13}$, the estimate from row 1 and row 3, $d = \lambda$, and $\phi_R(n) = 2\pi\sin\theta$. The interval of $\theta$, $[-\pi/6, \pi/6]$, is mapped to the interval of $\phi_R(n)$, $[-\pi, \pi]$. Similar, $\theta \in [\pi/6, \pi/2]$ is mapped to $\phi_R \in [\pi, 2\pi]$ and $\theta \in [-\pi/2, -\pi/6]$ to $\phi_R \in [-2\pi, -\pi]$. Therefore, $\phi_R(n)$ ranges from $-2\pi$ to $2\pi$, and we cannot tell in which $2\pi$ circle the estimated $\phi_{\hat{R}}(n)$ should be.

This ambiguity can be resolved using the unambiguous estimation from pairs of rows with half-wavelength spacing. In the same example, if the unambiguous estimation of $\theta$ is in the interval of $[-\pi/6, \pi/6]$, then $\phi_{\hat{R}}(n)$ should be chosen in the interval of $[-\pi, \pi]$; otherwise, chosen in the interval of $[-2\pi, -\pi]$ and $[\pi, 2\pi]$. For a larger row spacing, a similar procedure can be applied.

Note, above DOA estimates, $\hat{\theta}_{ij}$, are not independent. For example, $\hat{\theta}_{13}$ can be obtained from $\hat{\theta}_{12}$ and $\hat{\theta}_{23}$. Averaging all such estimates is not necessary. Moreover, using DPE method, the DOA estimation resolution is inversely proportional to the row spacing. A Vernier technique can be realized by averaging estimates only from those pairs of rows with a larger row spacing.

In the conventional bathymetric sidescan sonar, the amplitude estimation is done by averaging the amplitudes received at each row or simply using any one of the row outputs.

## 3.2 Correlated Signal Direction Estimation (CSDE) for Three-row Systems

As we introduced in last section, the conventional differential phase estimate is obtained in the absence of noise, and then applied to the situation with noise. It is natural to extend this procedure to the multiple-row system. We repeat (2.25) here in the absence of noise using the equivalent base-band representation:

$$\tilde{x}_{1b}(t) = \tilde{s}_{r1b}(t)e^{-jw_0\tau_1(\theta_1)} + \ldots + \tilde{s}_{rDb}(t)e^{-jw_0\tau_1(\theta_D)} \tag{3.11}$$

$$\tilde{x}_{2b}(t) = \tilde{s}_{r1b}(t)e^{-jw_0\tau_2(\theta_1)} + \ldots + \tilde{s}_{rDb}(t)e^{-jw_0\tau_2(\theta_D)} \tag{3.12}$$

$$\vdots$$

$$\tilde{x}_{Mb}(t) = \tilde{s}_{r1b}(t)e^{-jw_0\tau_M(\theta_1)} + \ldots + \tilde{s}_{rDb}(t)e^{-jw_0\tau_M(\theta_D)} \tag{3.13}$$

Again, all the rows are assumed to have the normalized omnidirectional beampattern in vertical plane. As we can see in the consequence, the omnidirectional directivity is not a necessary requirement as long as all the rows have approximately the same beampattern in the angular sector of interests.

For a two-row system, one directional signal can be resolved. From $(3.11-3.13)$, we have three unknowns, $\tilde{s}_{r1b}(t)$ (magnitude and phase) and $\theta_1$, and four equations. Note, each equation in $(3.11-3.13)$ includes two equations by making the real part and imaginary part equal, respectively. This is an overdetermined problem. As a result, the solution $\theta_1$ can be obtained by using only the phase information in $\tilde{x}_{1b}(t)$ and $\tilde{x}_{2b}(t)$, though the magnitude information can be used to obtain the optimal solution in noise.

For a four-row system, signals from three different directions should be able to resolved. (Later we will indicate that for a $M$-row system, the maximum resolvable source number can be $M-1$ by using ESPRIT algorithm.) However, from $(3.11-3.13)$, we have nine unknowns, $\tilde{s}_{r1b}(t)$, $\tilde{s}_{r2b}(t)$, $\tilde{s}_{r3b}(t)$, $\theta_1$, $\theta_2$, and $\theta_3$, and only eight equations. Generally, since three unknowns are associated with each directional source, for a $M$-row system with $M-1$ directional signals to be resolved, we will have $3(M-1)$ unknowns and $2M$ equations. Therefore, the problem for $M \geq 4$ is undetermined. We cannot solve this group of equations analytically and exactly, and must resort to some optimal estimation methods. In next section, such a method will be introduced.

Fortunately, there is a solution for $M = 3$ and $D = 2$, in which case we have 6 unknowns, $\tilde{s}_{r1b}(t)$, $\tilde{s}_{r2b}(t)$, $\theta_1$, and $\theta_2$, and 6 equations. Let's repeat $(3.11-3.13)$ here for $M = 3$ and $D = 2$:

$$\tilde{x}_{1b}(t) = \tilde{s}_{r1b}(t)e^{-jw_0\tau_1(\theta_1)} + \tilde{s}_{r2b}(t)e^{-jw_0\tau_1(\theta_2)} \tag{3.14}$$

$$\tilde{x}_{2b}(t) = \tilde{s}_{r1b}(t)e^{-jw_0\tau_2(\theta_1)} + \tilde{s}_{r2b}(t)e^{-jw_0\tau_2(\theta_2)} \tag{3.15}$$

$$\tilde{x}_{3b}(t) = \tilde{s}_{r1b}(t)e^{-jw_0\tau_3(\theta_1)} + \tilde{s}_{r2b}(t)e^{-jw_0\tau_3(\theta_2)} \tag{3.16}$$

Because the 1st row is chosen as the reference row, we have

$$\tau_1(\theta_1) = \tau_1(\theta_2) = 0 \tag{3.17}$$

$$\tau_3(\theta_1) = 2\tau_2(\theta_1) = \frac{2d\sin\theta_1}{c} \tag{3.18}$$

$$\tau_3(\theta_2) = 2\tau_2(\theta_2) = \frac{2d\sin\theta_2}{c} \tag{3.19}$$

To simplify the equations further, we change variables by

$$X = e^{-jw_0\tau_2(\theta_1)} \tag{3.20}$$

$$Y = e^{-jw_0\tau_2(\theta_2)} \tag{3.21}$$

Then we obtain the following equations

$$\tilde{x}_{1b}(t) = \tilde{s}_{r1b}(t) + \tilde{s}_{r2b}(t) \tag{3.22}$$

$$\tilde{x}_{2b}(t) = \tilde{s}_{r1b}(t)X + \tilde{s}_{r2b}(t)Y \tag{3.23}$$

$$\tilde{x}_{3b}(t) = \tilde{s}_{r1b}(t)X^2 + \tilde{s}_{r2b}(t)Y^2 \tag{3.24}$$

Solving (3.22) and (3.23), we get

$$\tilde{s}_{r1b}(t) = \frac{\tilde{x}_{1b}(t)Y - \tilde{x}_{2b}(t)}{Y - X} \tag{3.25}$$

$$\tilde{s}_{r2b}(t) = \frac{\tilde{x}_{2b}(t) - \tilde{x}_{1b}(t)X}{Y - X} \tag{3.26}$$

Solving (3.22) and (3.24), we get

$$\tilde{s}_{r1b}(t) = \frac{\tilde{x}_{1b}(t)Y^2 - \tilde{x}_{3b}(t)}{Y^2 - X^2} \tag{3.27}$$

$$\tilde{s}_{r2b}(t) = \frac{\tilde{x}_{3b}(t) - \tilde{x}_{1b}(t)X^2}{Y^2 - X^2} \tag{3.28}$$

By equating (3.25), (3.27) and (3.26), (3.28) for $\tilde{s}_{r1b}(t)$ and $\tilde{s}_{r2b}(t)$, respectively, we obtain

$$\tilde{x}_{3b}(t) = \tilde{x}_{2b}(t)Y - \tilde{x}_{1b}(t)XY + \tilde{x}_{2b}(t)X \tag{3.29}$$

and thus

$$Y = \frac{\tilde{x}_{3b}(t) - \tilde{x}_{2b}(t)X}{\tilde{x}_{2b}(t) - \tilde{x}_{1b}(t)X} \tag{3.30}$$

Note, $|X| = 1$ and $|Y| = 1$, i.e., $X \cdot X^* = 1$ and $Y \cdot Y^* = 1$. Using these identities, we obtain a second-order equation of $X$

$$(\tilde{x}_{1b}(t)\tilde{x}_{2b}^*(t) - \tilde{x}_{2b}(t)\tilde{x}_{3b}^*(t))X^2 + (|\tilde{x}_{3b}(t)|^2 - |\tilde{x}_{1b}(t)|^2)X + (\tilde{x}_{1b}^*(t)\tilde{x}_{2b}(t) - \tilde{x}_{2b}^*(t)\tilde{x}_{3b}(t)) = 0 \quad (3.31)$$

Defining

$$A = \tilde{x}_{1b}(t)\tilde{x}_{2b}^*(t) - \tilde{x}_{2b}(t)\tilde{x}_{3b}^*(t) \quad (3.32)$$

$$B = |\tilde{x}_{3b}(t)|^2 - |\tilde{x}_{1b}(t)|^2 \quad (3.33)$$

and noting the symmetry of $X$ and $Y$, we have

$$\begin{matrix} X \\ Y \end{matrix} = \frac{-B \pm \sqrt{B^2 - 4|A|^2}}{2A} \quad (3.34)$$

Finally, the solutions of $(3.14 - 3.16)$ are given by

$$\theta_1 = \operatorname{asin}\frac{\phi_{X(Y)} \cdot \lambda}{2\pi d} \quad (3.35)$$

$$\theta_2 = \operatorname{asin}\frac{\phi_{Y(X)} \cdot \lambda}{2\pi d} \quad (3.36)$$

$$\tilde{s}_{r1b}(t) = \frac{\tilde{x}_{1b}(t)Y - \tilde{x}_{2b}(t)}{Y - X} \quad (3.37)$$

$$\tilde{s}_{r2b}(t) = \frac{\tilde{x}_{2b}(t) - \tilde{x}_{1b}(t)X}{Y - X} \quad (3.38)$$

where $\phi_{X(Y)}$ is the phase of $X(Y)$. The amplitude estimates are just the magnitudes of $\tilde{s}_{r1b}(t)$ and $\tilde{s}_{r2b}(t)$.

Now we apply above solutions for the noisy signals. Again, the estimates of $A$ and $B$ can be obtained over $L$ given samples. In digital domain, using the rectangular window, we have

$$\hat{A} = \hat{R}_{12} - \hat{R}_{23} = \frac{1}{L}\sum_{n=1}^{L}(\tilde{x}_{1b}(n)\tilde{x}_{2b}^*(n) - \tilde{x}_{2b}(n)\tilde{x}_{3b}^*(n)) \quad (3.39)$$

$$\hat{B} = \hat{R}_{33} - \hat{R}_{11} = \frac{1}{L}\sum_{n=1}^{L}(|\tilde{x}_{3b}(n)|^2 - |\tilde{x}_{1b}(n)|^2) \quad (3.40)$$

where $\hat{R}_{ij}$ is the covariance estimate between the $i$-th and $j$-th row. (In the discussions in this section, the measurement process is assumed to be zero-mean or the mean component has been removed, so the covariance estimate is equivalent to the correlation estimate.) However, unlike the differential phase estimate method, the solution here is depended not only the phase but also the magnitude of such covariance estimate. An optimal estimate can be obtained if the two directional sources are highly coherent. To show how this works, let's assume a fully coherent case, i.e., $\tilde{s}_{r2b}(n) = \alpha\tilde{s}_{r1b}(n)$ where $\alpha$ is a complex constant, and represent the receiving data as

$$\tilde{x}_{1b}(n) = \tilde{s}_{r1b}(n) + \alpha\tilde{s}_{r1b}(n) + \tilde{n}_{1b}(n) \tag{3.41}$$

$$\tilde{x}_{2b}(n) = \tilde{s}_{r1b}(n)e^{-jw_0\tau_2(\theta_1)} + \alpha\tilde{s}_{r1b}(n)e^{-jw_0\tau_2(\theta_2)} + \tilde{n}_{2b}(n) \tag{3.42}$$

$$\tilde{x}_{3b}(n) = \tilde{s}_{r1b}(n)e^{-jw_0\tau_3(\theta_1)} + \alpha\tilde{s}_{r1b}(n)e^{-jw_0\tau_3(\theta_2)} + \tilde{n}_{3b}(n) \tag{3.43}$$

then the covariance estimations for $L$ given samples are

$$\hat{R}_{12} = \frac{(1+\alpha)(e^{jw_0\tau_2(\theta_1)} + \alpha^* e^{jw_0\tau_2(\theta_2)})}{L}\sum_{n=1}^{L}|\tilde{s}_{r1b}(n)|^2 + \text{Noise term} \tag{3.44}$$

$$\hat{R}_{23} = \frac{(e^{-jw_0\tau_2(\theta_1)} + \alpha e^{-jw_0\tau_2(\theta_2)})(e^{jw_0\tau_3(\theta_1)} + \alpha^* e^{jw_0\tau_3(\theta_2)})}{L}\sum_{n=1}^{L}|\tilde{s}_{r1b}(n)|^2 + \text{Noise term} \tag{3.45}$$

$$\hat{R}_{11} = \frac{(1+\alpha)(1+\alpha^*)}{L}\sum_{n=1}^{L}|\tilde{s}_{r1b}(n)|^2 + \text{Noise term} \tag{3.46}$$

$$\hat{R}_{33} = \frac{(e^{-jw_0\tau_3(\theta_1)} + \alpha e^{-jw_0\tau_3(\theta_2)})(e^{jw_0\tau_3(\theta_1)} + \alpha^* e^{jw_0\tau_3(\theta_2)})}{L}\sum_{n=1}^{L}|\tilde{s}_{r1b}(n)|^2 + \text{Noise term} \tag{3.47}$$

Note, the signal term in the estimates is the same thing as for a single data sample except that a single $|\tilde{s}_{r1b}(n)|^2$ is replaced by the sum from $L$ samples. Defining

$$\beta_s = \sum_{n=1}^{L}|\tilde{s}_{r1b}(n)|^2 \tag{3.48}$$

as signal weighting factor, we find the signal terms in all the covariance estimates have this same weighting factor, and so do the signal terms in $\hat{A}$ and $\hat{B}$. In the absence of noise, by factorizing $\beta_s$ from $(3.44 - 3.47)$, we can obtain the exact same solutions as by a single data sample. On the other hand, if the noises are zero-mean and uncorrelated with the signal and with each other, then the covariance estimate of $L$ samples adds the noise term incoherently while adding the signal term coherently, thus increasing the array output SNR greatly relative to a single data sample. Therefore, the optimum array gain can be obtained and the covariance-based CSDE is an optimal estimator for above signal and noise models.

To achieve the performance provided by an optimal estimate, the noise samples are generally required to be temporally independent. That means the sampling interval should be significantly larger than the maximum inverse bandwidth of the noise process. This condition is easily satisfied by noting the general noise process has wider, flat spectrum.

A similar issue regarding the signal is quite different. The data model used here (also used by most high-resolution DOA estimate algorithms) requires the signal bandwidth as narrow as possible, while the narrow bandwidth intends to increase the correlation interval of the signal. Though the signals are never pure cisoids in reality, the temporal correlation of a directional signal cannot be ignored. The correlation interval is about twice the transmitted pulse length (cf. section 2.3), for example, which is equivalent to ten data points if a sampling rate of $50kHz$ is used in a $100\mu s$-pulse system.

Let's consider two temporally fully correlated signals, $\tilde{s}_{r1b}(n) = C_1$ and $\tilde{s}_{r2b}(n) = C_2$ both for all $n$, with $C_1, C_2$ constant. Following a similar procedure in coherent issue (actually this is a special coherent case), we can reach the same conclusion. Hence, it is safe to argue that CSDE is also optimal for temporally highly correlated signals. The reason why we call this method correlated signal direction estimation (CSDE) is that the proposed estimator is optimal for highly coherent or temporally correlated signals. Later simulations will verify this point further.

Observations to (3.39) and (3.40) show that the noises at three rows can be correlated with each other as long as they are from the same process and thus have the same spatial correlation characteristics. These noise-noise cross-product terms are cancelled with each other by the minus operation in calculating $\hat{A}$ and $\hat{B}$.

Generally speaking, for incoherent and temporally uncorrelated signals, a covariance-based CSDE is not an optimal estimate. Though, a point-based CSDE can still be used. To reduce the noise effect, a few such estimates are averaged over the given data samples. Specifically, let $\hat{\theta}_1(n), \hat{\theta}_2(n)$ be the solutions from $L$ data samples $\left[\tilde{x}_{1b}(n)\ \tilde{x}_{2b}(n)\ \tilde{x}_{3b}(n)\right]$, then

$$\hat{\theta}_1 = \frac{1}{L}\sum_{n=1}^{L}\hat{\theta}_1(n) \qquad (3.49)$$

and

$$\hat{\theta}_2 = \frac{1}{L}\sum_{n=1}^{L}\hat{\theta}_2(n). \qquad (3.50)$$

We should point out here that the three-row bathymetric sidescan sonar system is of our most interests in real applications since a two-row system can be easily extended to a three-row system, and some such three-row systems have already been available, though they still use the conventional signal processing method. Besides, as we discussed in section 2.1 and 2.2, the new capability, resolving two directional signals at the same instant, can extend the conventional bathymetric sidescan sonar to a wider variety of practical situations.

## 3.3 ESPRIT Approach

ESPRIT is developed on the basis of a particular array structure with displacement invariance. For a uniform linear array used in bathymetric sidescan sonar, two subarrays are formed with maximally overlapping (Fig. 3.1). There are $(M-1)$ sensor doublets (shown by the arrow connections), and the sensors in each doublet have identical beampatterns and are translationally separated by a vector $\vec{\Delta}$, which is the exact requirement by ESPRIT. To see how ESPRIT works, let's inspect the signals received at each doublet first. Denoting $\tilde{x}_k(t)$ as the output of the $k$-th sensor in subarray 1, $\tilde{y}_k(t)$ as the output of the $k$-th sensor in subarray 2, for the $k$-th doublet, (cf. section 2.4)

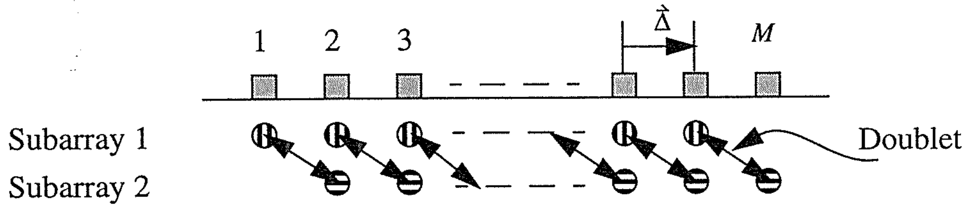


**Figure 3.1:** Array geometry for ESPRIT

$$\tilde{x}_k(t) = \sum_{i=1}^{D} a_{xk}(\theta_i)\tilde{s}_{ir}(t)e^{-jw_0\tau_{xk}(\theta_i)} + \tilde{n}_{xk}(t) \tag{3.51}$$

$$\tilde{y}_k(t) = \sum_{i=1}^{D} a_{yk}(\theta_i)\tilde{s}_{ir}(t)e^{-jw_0\tau_{yk}(\theta_i)} + \tilde{n}_{yk}(t) \tag{3.52}$$

Note, the $k$-th sensor in subarray 2 is the $(k+1)$-th sensor in subarray 1. If $a_k(\theta_i) = a_{k+1}(\theta_i)$, and defining $a'_k(\theta_i) = a_k(\theta_i)e^{-jw_0\tau_k(\theta_i)}$, then we have

$$\tilde{x}_k(t) = \sum_{i=1}^{D} a'_k(\theta_i)\tilde{s}_{ir}(t) + \tilde{n}_{xk}(t) \tag{3.53}$$

$$\tilde{y}_k(t) = \sum_{i=1}^{D} a'_k(\theta_i)\tilde{s}_{ir}(t)e^{-jw_0\Delta\sin\theta_i/c} + \tilde{n}_{yk}(t) \tag{3.54}$$

Combining all the outputs in two subarrays, using the matrix representation (cf. section 2.4), we have

$$\tilde{X}(t) = A(\theta)\tilde{s}_r(t) + \tilde{N}_x(t) \tag{3.55}$$

$$\tilde{Y}(t) = A(\theta)\Phi(\theta)\tilde{s}_r(t) + \tilde{N}_y(t) \tag{3.56}$$

To make the problem more clearly, all the matrices are specified with their sizes as $\tilde{X}_{(M-1)\times 1}$, $\tilde{Y}_{(M-1)\times 1}$, $\tilde{s}_{rD\times 1}$, $A_{(m-1)\times D}$, $\tilde{N}_{x(M-1)\times D}$ and $\tilde{N}_{y(M-1)\times D}$.

$\Phi(\theta)$ is a $D\times D$ diagonal matrix of the phase differences between the sensors in a doublet for $D$ directional sources, i.e.,

$$\Phi(\theta) = diag[e^{-jw_0\Delta\sin\theta_1/c}, \ldots, e^{-jw_0\Delta\sin\theta_D/c}] \tag{3.57}$$

A more compact representation of the data model is given as

$$\tilde{Z}(t) = \begin{bmatrix} \tilde{X}(t) \\ \tilde{Y}(t) \end{bmatrix} = \bar{A}(\theta)\tilde{s}_r(t) + \tilde{N}_z(t) \tag{3.58}$$

with

$$\bar{A}(\theta) = \begin{bmatrix} A(\theta) \\ A(\theta)\Phi(\theta) \end{bmatrix} \tag{3.59}$$

and

$$\tilde{N}_z(t) = \begin{bmatrix} \tilde{N}_x(t) \\ \tilde{N}_y(t) \end{bmatrix} \tag{3.60}$$

So far, the data model has manifest a key factor in ESPRIT that the outputs from the subarray 1 and 2 are related by a complex scaling operator $\Phi(\theta)$, which is equivalent to a real two-dimensional rotation operator (that's why ESPRIT got its name). Thus, the structure of $\bar{A}(\theta)$ can be exploited to obtain estimates of $\Phi(\theta)$ and then the direction information for each source.

The detailed derivation of the ESPRIT algorithm is given in [12]. Here, the basic procedure is summarized simply as following.

From the ideal measurement model, the covariance matrix

$$R_{ZZ} = \bar{A}\tilde{S}_r\bar{A}^* + \sigma^2\Sigma_n \tag{3.61}$$

where $\tilde{S}_r, \sigma^2\Sigma_n$ are signal and noise covariance, respectively. The noise correlation, $\Sigma_n$ is known. For $D \le M-1$ (the sensor number in a subarray), the generalized eigenvalues (GE's) of $(R_{ZZ}, \Sigma_n)$ are $[v_1+\sigma^2, \ldots, v_D+\sigma^2, \sigma^2, \ldots, \sigma^2]$. The $2(M-1)-D$ minimum GE's equal to the noise power, $\sigma^2$. Hence, the noise component is identified and its effect on signal subspace structure can be removed. At the same time, the basis vectors that span the signal subspace is obtained from $D$ generalized eigenvectors (GEV's) corresponding to $D$ largest GE's as $E_s = \Sigma_n[e_1|\ldots|e_D]$. Note, though we assume $D$ is known for the time being, it is obvious that above arguments can be used to estimate $D$.

Here, $E_s$ is $2(M-1)\times D$ matrix, and it can be shown that the space spanned by $E_s$ is equal to that spanned by $\bar{A}$. Thus $E_s$ must be related to $\bar{A}$ by a nonsingular matrix $T$ such that $E_s = \bar{A}T$, and now the invariance structure of $\bar{A}$ suggests a partition of $E_s$ as

$$E_s = \begin{bmatrix} E_x \\ E_y \end{bmatrix} = \begin{bmatrix} AT \\ A\Phi T \end{bmatrix} \tag{3.62}$$

where both $E_x$ and $E_y$ are $(M-1)\times D$ matrices, and they span the same signal subspace spanned by $A$. If we can find a matrix, $\Psi$, which relates $E_x$ to $E_y$ by $E_y = E_x\Psi$, then, from (3.62),

$$AT\Psi T^{-1} = A\Phi \tag{3.63}$$

Assuming $A$ is full rank, we get the key relationship in ESPRIT

$$T\Psi T^{-1} = \Phi \tag{3.64}$$

which means, the diagonal elements of $\Phi$ in (3.57) are equal to the eigenvalues of $\Psi$.

In a real situation, $E_s$ is estimated from a covariance estimation $\hat{R}_{ZZ}$, and generally, $\hat{E}_s$ and $\bar{A}$ don't span the same subspace, and neither do $\hat{E}_x$ and $\hat{E}_y$. To find $\Psi$, the general problem is

$$min_{\Psi}\|\hat{E}_x\Psi - \hat{E}_y\|_F \tag{3.65}$$

where $\| \; \|_F$ is the Frobenius norm [39]. Since both $\hat{E}_x$ and $\hat{E}_y$ are noisy, the problem is further redefined by using the total least square criterion as

$$min_{\Psi_x, \Psi_y}\left\| [\hat{E}_x | -\hat{E}_y] \begin{bmatrix} \Psi_x \\ \Psi_y \end{bmatrix} \right\|_F \tag{3.66}$$

subject to a symmetric constraint on both $\Psi_x$ and $\Psi_y$.

By first forming a new matrix,

$$\begin{bmatrix} \hat{E}_x^{\dagger} \\ \hat{E}_y^{\dagger} \end{bmatrix} \begin{bmatrix} \hat{E}_x & \hat{E}_y \end{bmatrix} \tag{3.67}$$

where $(\;)^{\dagger}$ is the conjugate transpose, and then computing its eigendecomposition and choosing the eigenvectors, $E_{xy}{}^{(D)}$, corresponding to $D$ smallest eigenvalues, and finally, partitioning $E_{xy}{}^{(D)}$ into $D \times D$ matrices, $E_{xy,x}{}^{(D)}$ and $E_{xy,y}{}^{(D)}$, $\Psi$ is obtained as

$$\Psi = -E_{xy,x}{}^{(D)}[E_{xy,y}{}^{(D)}]^{-1} \tag{3.68}$$

The total least square (TLS) ESPRIT algorithm is summarized as following (cf. [12]):

1) Obtain $\hat{R}_{ZZ}$, the estimate of $R_{ZZ}$, from the measurement $\tilde{Z}$. Specifically, for $L$ given data samples at each sensor, i.e., $\tilde{Z} \in C^{2(M-1) \times L}$, the maximum likelihood estimate is

$$\hat{R}_{ZZ} = \frac{1}{2(M-1)L}(\tilde{Z} - \mu_Z) \cdot (\tilde{Z} - \mu_Z)^{\dagger} \tag{3.69}$$

where $\mu_Z$ is the mean of $\tilde{Z}$. The number of snapshots, $L$, is chosen according to the available measurements, SNR at each sensor, and desired DOA estimation accuracy. Generally, $L$ should be at least $(M-1)^2$.

2) Implement the generalized eigendecomposition of $\{\hat{R}_{ZZ}, \; \Sigma_n\}$

$$\hat{R}_{ZZ}\bar{E} = \Sigma_n\bar{E}\Lambda \tag{3.70}$$

where $\Lambda = diag\{\lambda_1, \ldots, \lambda_{2(M-1)}\}, \lambda_1 \geq \ldots \geq \lambda_{2(M-1)}$, and $\bar{E} = [e_1|\ldots|e_{2(M-1)}]$. Note, because we use an overlapping array structure and thus some sensors are members of both subarrays, $\Sigma_n$ should be considered carefully. For example, for $M = 3$, using the noise model proposed, $\Sigma_n$ is given by

$$\begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 1 & 0 \\ 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \tag{3.71}$$

Assuming a decomposition of $\Sigma_n = \Sigma_n^{\frac{1}{2}}\Sigma_n^{\frac{1}{2}}$ exists, the generalized eigendecomposition can be converted a standard eigendecomposition of $\Sigma_n^{-\frac{1}{2}}\hat{R}_{ZZ}\Sigma_n^{-\frac{1}{2}\dagger}$.

3) Estimate the number of sources, $\hat{D}$. For a practical $\hat{R}_{ZZ}$, the $2(M-1)-D$ smallest GE's are clustered around, and not all equal to $\sigma^2$. To obtain $\hat{D}$, some further techniques need to be used [12].

4) Obtain the signal subspace estimate $\hat{S}_Z = \mathcal{R}\{\hat{E}_s\}$, where $\mathcal{R}(\ )$ means spanned by, and decompose it to obtain $\hat{E}_x$ and $\hat{E}_y$,

$$\hat{E}_s = \Sigma_n[\hat{e}_1|\ldots|\hat{e}_{\hat{D}}] = \begin{bmatrix} \hat{E}_x \\ \hat{E}_y \end{bmatrix} \tag{3.72}$$

5) Implement the eigendecomposition,

$$\begin{bmatrix} \hat{E}_x^{\dagger} \\ \hat{E}_y^{\dagger} \end{bmatrix} \begin{bmatrix} \hat{E}_x & \hat{E}_y \end{bmatrix} = E\Lambda E^* \tag{3.73}$$

with the eigenvalues arranged in decreasing order, and partition E into $\hat{D} \times \hat{D}$ submatrices,

$$E = \begin{bmatrix} E_{11} & E_{12} \\ E_{21} & E_{22} \end{bmatrix} \tag{3.74}$$

6) Obtain $\Psi$ by $\Psi = -E_{12}E_{22}^{-1}$, and compute its eigenvalues, denoted by $\lambda_k$, and let $\hat{\phi}_k$ denote the phase of $\lambda_k$, for $k = 1, \ldots, \hat{D}$.

7) Calculate

$$\hat{\theta}_k = \operatorname{asin}\{-c \cdot \hat{\phi}_k/(w_0\Delta)\} \tag{3.75}$$

To avoid squaring the data, which intends to cause some problems on matrix operation, the measurement data can be processed directly using the singular value decomposition (SVD), which theoretically yields the same subspace estimate as the eigendecomposition. In step 5), noting

$$SVD\left\{\begin{bmatrix}\hat{E}_x^{\dagger}\\ \hat{E}_y^{\dagger}\end{bmatrix}\right\} = U\Sigma V^* \tag{3.76}$$

where $\Sigma$ is diagonal and real, and $U$ and $V$ are unitary, then

$$\begin{bmatrix}\hat{E}_x^{\dagger}\\ \hat{E}_y^{\dagger}\end{bmatrix}\begin{bmatrix}\hat{E}_x & \hat{E}_y\end{bmatrix} = U\Sigma^2 U^* \tag{3.77}$$

Thus, the left singular vectors, $U$, of $\begin{bmatrix}\hat{E}_x^{\dagger}\\ \hat{E}_y^{\dagger}\end{bmatrix}$ are the eigenvectors of $\begin{bmatrix}\hat{E}_x^{\dagger}\\ \hat{E}_y^{\dagger}\end{bmatrix}\begin{bmatrix}\hat{E}_x & \hat{E}_y\end{bmatrix}$.

In step 2), a generalized SVD (GSVD) can be applied directly [12]. However, since a standard eigendecomposition problem is obtained there, the GSVD is not necessary. Simply, replacing $\tilde{Z}$ by $\Sigma_n^{-\frac{1}{2}}\tilde{Z}$, the signal subspace estimate can be obtained from the left singular vectors of $\Sigma_n^{-\frac{1}{2}}\tilde{Z}$.

To apply ESPRIT to a real application, several issues are discussed. First, we point out here that the maximum resolvable source number for a $M$-sensor ULA with maximum overlapping is $M-1$, close to that of MUSIC. The performance in terms of estimation bias and variance using overlapping structure need to be further investigated, though a study in [5] supports this usage in the context of a conventional split-aperture system.

Second, the noise may include both nondeterministic and unwanted deterministic components, while knowledge of the noise correlation must be known, which can be estimated by field measurements. Note, the optimum array gain is limited by the logarithm of $M$. For a large $M$, the ESPRIT algorithm may be very sensitive to an incorrect noise correlation model, particularly at low SNR. However, for a small $M$, though the noise model is not a critical issue, SNR will dominate the system performance.

Third, in ESPRIT, the signals are modeled as narrow-band stochastic process and required to be temporally independent. To track the statistics of the signal process, all available measurements should be used. In many practical situations, data length available for estimating the source DOA is limited, and furthermore these data are generally more or less correlated as in the bathymetric sidescan sonar. It can be expected that all these factors will impact the performance of ESPRIT.

Fourth, the correlation between signals from different directions is allowed in ESPRIT. However, for fully coherent signals, ESPRIT cannot resolve them, and gives an incorrect estimate. Though, practically, the fully coherent signals doesn't exist, the performance of ESPRIT will also be degraded depending on the amount of such correlation as for the temporally highly correlated signals. Fortunately, when the highly coherent signals appear together with an incoherent signal, this incoherent signal can still be identified by ESPRIT.

Fifth, for a small $M$, the estimate of source number is not a critical issue and even can be ignored, because, using a maximum possible $D$, ESPRIT can still give correct estimates for the true sources, and the false random sources can be thrown away by further processing. However, a detail inspection on the eigenstructure of $\hat{R}_{ZZ}$ is very helpful to estimate the real wave field structure, and thus gives an evaluation to the quality of the processed data.

Finally, we give the results on optimal signal amplitude estimation, one of the advantages of ESPRIT over MUSIC. Generally, a signal copy is obtained by using a weighted sum of the sensor outputs. Such weight vectors are optimal in the sense that a single output contains only the desired signal while eliminating other $D-1$ signals. Again, using the TLS approach, the optimal weight matrix is given by [12]

$$\hat{W}_{op} = \hat{E}_S[\hat{E}^{\dagger}{}_s \cdot \hat{E}_S]^{-1}\hat{E}_{\Psi}{}^{-\dagger} \tag{3.78}$$

which satisfies $\hat{W}_{op}{}^{\dagger} \cdot \bar{A} = I$ and $I$ is the identity matrix. Therefore, the optimal signal copy is estimated by

$$\tilde{s}_{op}(t) = \hat{W}_{op}{}^{\dagger} \cdot \tilde{Z}(t) \tag{3.79}$$

and the amplitude estimate is the magnitude of $\tilde{s}_{op}(t)$.

# Chapter 4

# Simulations

A new bathymetric sidescan sonar with improved signal processing system is being developed at the Woods Hole Oceanographic Institution (WHOI). In this chapter, computer simulations are carried out to verify the performances of previous algorithms and investigate the possibility of applying these algorithms to the developing system. First, we study a single directional source situation using the DPE and ESPRIT. Then, the performances of ESPRIT and CSDE are evaluated in the cases with two directional sources. They are further compared using different data models with varying degrees of temporal correlation (or coherence). After that, a practical example with two time-varying directional sources and SNR's is simulated. Finally, because the real three-row sonar data are not available at the time of this writing, some two-row data are used to test the algorithms.

The new system is configured as a three-row uniform linear array with half-wavelength row spacing. The operating frequency is $200kHz$, and the transmitted pulse length is $100\mu s$. Because the system base-band bandwidth is $10kHz$, the sampling frequency is chosen as $48KHz$, approximately two and half times of Nyquest rate. Sound velocity, $c$, is assumed to be $1500m/s$. The received signals are passed to the pre-amplifier, demodulator, time-varying and fixed gain amplifier, A/D converter, match-filter, and finally recorded. The raw data are compensated for system parameters according to the sonar equation [40].

In the simulation, a three-sensor array structure is adopted with each sensor being omnidirectional. The sensor spacing is exactly half-wavelength. The signals are constructed as narrow-band stationary stochastic processes on the basis of the statistical characteristics of bottom scattering. Specifically, quadrature samples are generated to form complex samples for each sensor, and the following characteristics of the bottom scattering are considered [35, 3]:

1) The scattering process is Gaussian, and the quadrature components also have Gaussian distributions with zero mean, and variances equal to the variance of the scattering process.

2) The quadrature components are uncorrelated and statistically independent.

3) The envelop of the scattering process has a generalized Rayleigh distribution.

4) The phase has a uniform distribution in the interval $(0, 2\pi)$.

5) The statistical spectrum of bottom scattering in the transmission of determinate-type signals is proportional to the square of their amplitude spectrum.

Under these assumptions, at the reference sensor (1st sensor), for each directional source $s_r$, the envelop $E_r(n)$ can be generated by two zero mean independently Gaussian processes $E_1(n)$ (in-phase component) and $E_2(n)$ (quadrature component) with the same variance $\sigma_r$, as

$$E_r(n)^2 = E_1(n)^2 + E_2(n)^2$$

To include the finite bandwidth effects, the quadrature components are convolved with a window function whose length depends on the pulse length, thus forming the signal component of the simulated data. Therefore, the simulated data are correlated within a finite interval, consistent with the correlation analysis results on the real scattering process. Here, we choose a rectangular window and the window length is five points, the number of samples in a pulse.

In some of the simulations, highly coherent or temporally correlated signals are investigated. For fully temporally correlated signals, $s_{rj}(t)$ is a constant, and we can add a random component to it to adjust its correlation. For fully coherent signals defined as $s_{rj}(t) = \alpha \cdot s_{ri}(t)$ with $\alpha$ a complex constant, the degree of coherence may also be adjusted by letting $\alpha$ have a random component and giving different weights to the deterministic and random components in $\alpha$.

Additive noise is present at all sensors, and assumed to be uncorrelated band-limited zero mean Gaussian process, which is added to the signal quadrature components independently. An FIR low-pass filter corresponding to the system parameters (bandwidth, sampling rate) is used to obtain the band-limited noise.

Two important parameters in simulations are the number of measurements (snapshots) per trial, $L$, and the number of trials, $N$. As previously indicated (cf. section 3.3), the choice of $L$ is dependent on the SNR at array input, desired DOA estimate accuracy (for above two factors, a larger L is preferred), and also, for bathymetric sidescan sonar, estimate resolution (a smaller L is preferred). To establish a relationship between $L$ and the resolution, let's look at the geometry of sonar use. For the bathymetric sidescan sonar being developed, the maximum operating distance is about $200m$. Suppose the operating height above the bottom is about one-third of the horizontal range, i.e., $\sim 60m$. Assuming a flat bottom and $L = 64$ (corresponds to $1m$ in acoustic propagating distance), at normal incidence, the angle difference between two contiguous $64$-point estimates approximates $4°$, while at the maximum operating distance, approximates $0.1°$. Note, the angular resolution of ESPRIT for a three-sensor array is also in the order of $4°$ around the same region. Therefore, $L$ should be smaller at normal incidence region, and larger for far off-normal region. In the simulations, we choose an intermediate value of $L = 64$. In most cases, $N$, the number of trials performed depends on the smoothness of the resulting estimate histogram. Generally, $N$ is large in low SNR, and small in high SNR. For the scenarios investigated here, we choose $N = 512$.

The differential phase estimate is implemented as described in section 3.1. Traditional DPE using only one pair of sensors, averaged DPE using two pairs of sensors, and vernier DPE are tested. In the former two methods, the sensor spacing is half-wavelength. In the vernier method, the DOA estimates are obtained from the pair of sensors with one wavelength spacing, and the associated ambiguity is treated using the algorithm proposed in section 3.1.

In CSDE simulations, the data are synthesized from two directional sources, and in most cases the point-based CSDE (CSDE (1)) is used with all individual estimates averaged for the given samples. For highly temporally correlated or coherent signals, a covariance-based CSDE (CSDE (2)) is applied.

The ESPRIT algorithm employed is the TLS ESPRIT (ESPRIT (1)). One or two directional signals are assumed. In step 5), a singular value decomposition (SVD) is computed instead of an eigendecomposition. Simulations with SVD alternative in step 2) show a tendency to biased estimates in low SNR, and thus are not presented. A modified ESPRIT (ESPRIT (2)) ignoring the effect of noise is also tested. In this modified version, at step 2), a regular eigendecomposition of $\hat{R}_{ZZ}$ is implemented instead of the generalized eigendecomposition of $\{\hat{R}_{ZZ}, \Sigma_n\}$, and thus at step 4), the multiplication factor, $\Sigma_n$, is not necessary.

Simulations are all performed on a SUN SPARC 10 workstation, and the Matlab package is used to perform the actual computations.

First, we consider the situations with just one directional signal, for which DPE and ESPRIT are applicable. The source direction is assumed to be at 5°, 40°, 60° and 75°, respectively. The results for this case are presented in Fig. 4.1 through 4.4. The sample means and standard deviations of the DOA estimates are presented in Table 4.1. As expected, the averaged DPE has approximately half of the estimate variance by the traditional DPE, while, that the vernier DPE performs as well as the averaged DPE is a little bit surprising, which gives vernier technique a few more credits if a pair of sensors with a larger spacing is used with the ambiguity resolved. The results also indicate that the performance similarity between the averaged DPE (or vernier DPE) and ESPRIT at mediate to high SNR. At low SNR, ESPRIT shows better performance in the sense of estimate statistics. As $\theta$ increases, the performances of all these methods degrade, which are further worse considering the lower SNR at larger $\theta$ in a real environment.

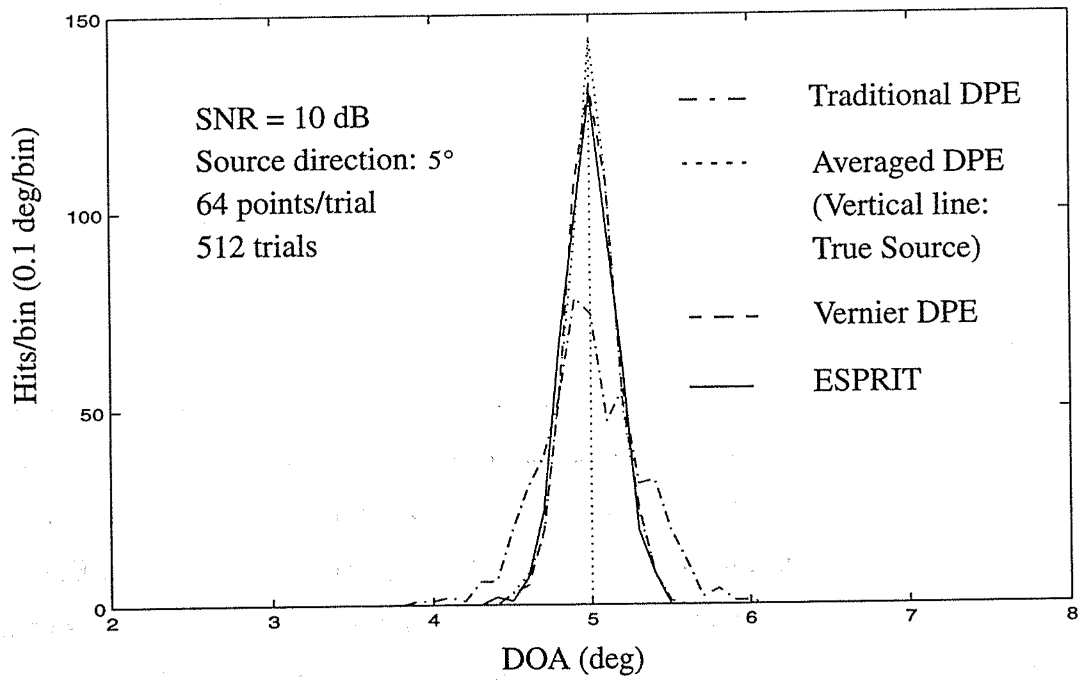


**Figure 4.1:** Histogram of DPE and ESPRIT simulation results for a single directional source (1)

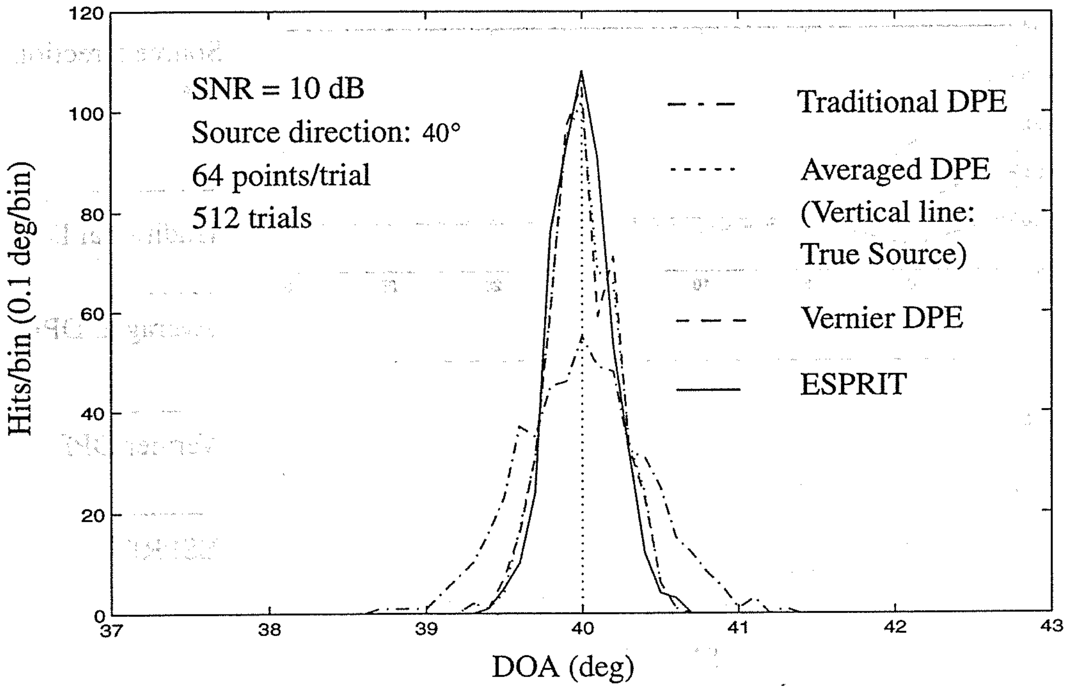


**Figure 4.2:** Histogram of DPE and ESPRIT simulation results for a single directional source (2)

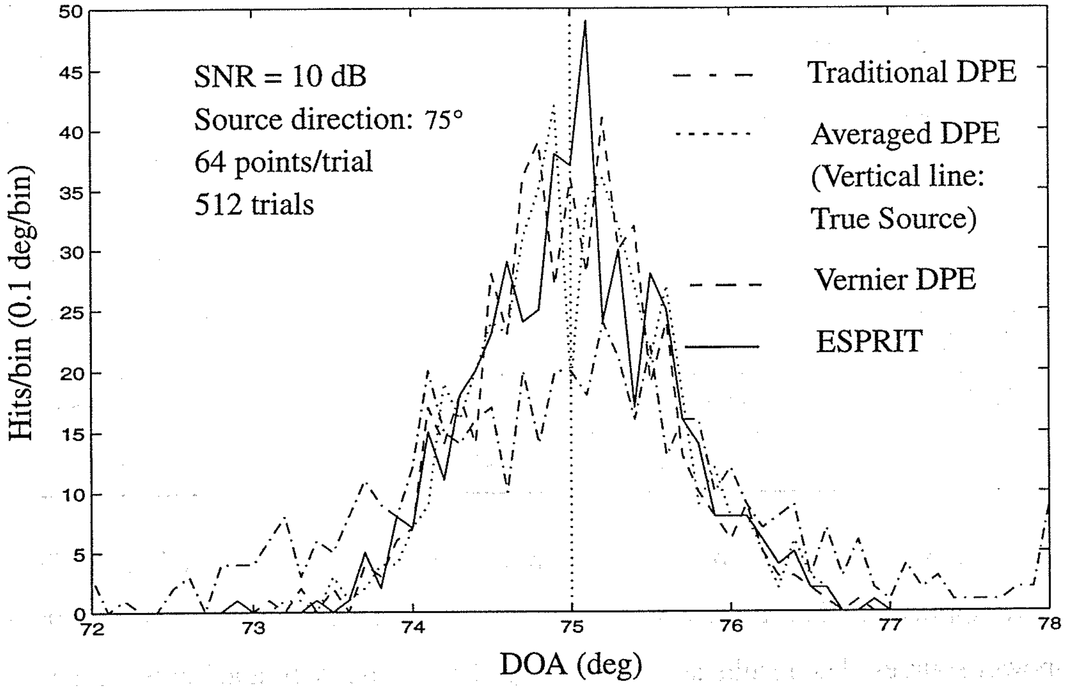


**Figure 4.3:** Histogram of DPE and ESPRIT simulation results for a single directional source (3)

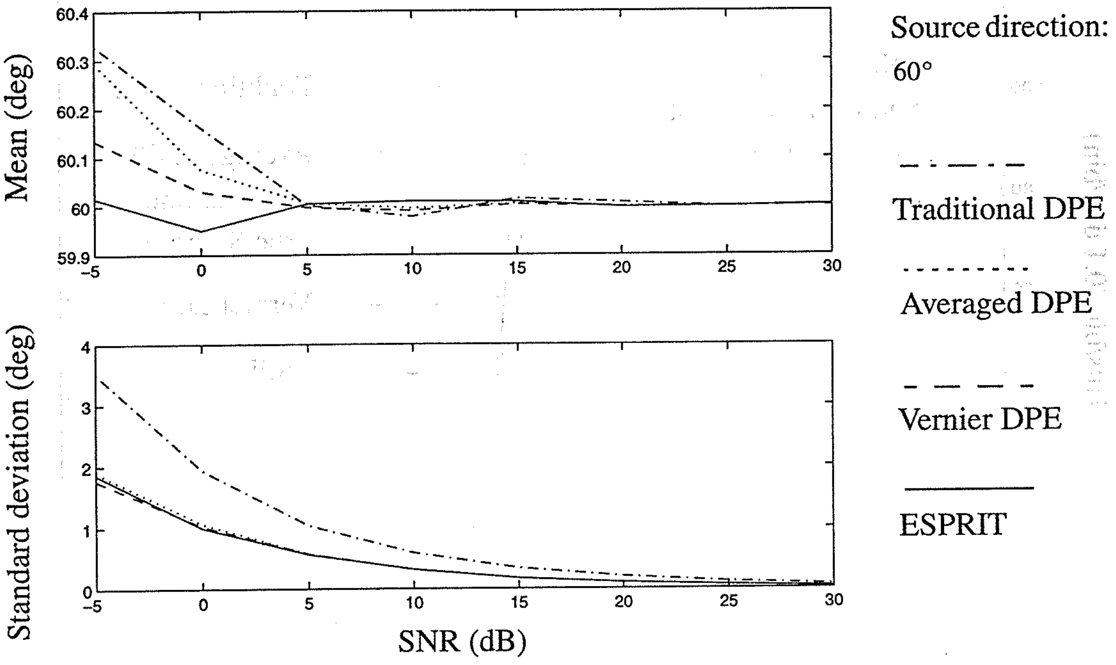


**Figure 4.4:** DPE and ESPRIT simulation results for a single directional source: mean and standard deviation v.s. SNR

**Table 4.1: DPE and ESPRIT simulation results for a single directional source at SNR = 10dB**

| Estimator | Source DOA Estimates | | |
|---|---|---|---|
| | $\hat{\theta}(\theta = 5^\circ)$ | $\hat{\theta}(\theta = 40^\circ)$ | $\hat{\theta}(\theta = 75^\circ)$ |
| Traditional DPE | 4.99° ± 0.31° | 40.00° ± 0.41° | 75.01° ± 1.19° |
| Averaged DPE | 5.00° ± 0.16° | 40.00° ± 0.21° | 75.03° ± 0.60° |
| Vernier DPE | 5.00° ± 0.16° | 40.00° ± 0.21° | 75.00° ± 0.59° |
| ESPRIT | 4.99° ± 0.17° | 40.00° ± 0.19° | 75.02° ± 0.61° |

The results with two directional sources are also investigated, for which ESPRIT and CSDE are applicable. The SNR and angular resolution issues are addressed with two equal-power sources. The results are given in Fig. 4.5 through 4.10 and Table 4.2 and 4.3. ESPRIT (1), ESPRIT (2), and CSDE (1) are tested. Note here, a simple angle-ordering scheme is used. The larger of the estimated DOA's is chosen as the larger of the true DOA's. So does the smaller. This is reasonable for large SNR and source separation. The results indicate that the performance of DOA estimation is the function of SNR for all methods. At SNR=5dB, two sources at 40° and 50° can still be resolved using both ESPRIT (1) and ESPRIT (2), while for CSDE (1), a SNR=10dB is needed, and the estimate bias is still obvious. The estimate performance also depends on the angular separation between two sources. Using ESPRIT algorithms, two sources with angular separation of 4° in the direction around 50° can be resolved at SNR=20dB, while for CSDE (1), the resolvable angular separation is about 5 – 10° at the same SNR. For all the situations here, ESPRIT's perform better than CSDE (1). However, the results show no significant difference between ESPRIT (1) and ESPRIT (2). It should be pointed out that all the estimate performances are also the function of the source direction (see Fig. 4.11, 4.12, and Table 4.4). As the sources move from the near-normal region to far off-normal region, the ESPRIT estimate bias and standard deviation become larger, while CSDE estimate is better in the intermediate region.

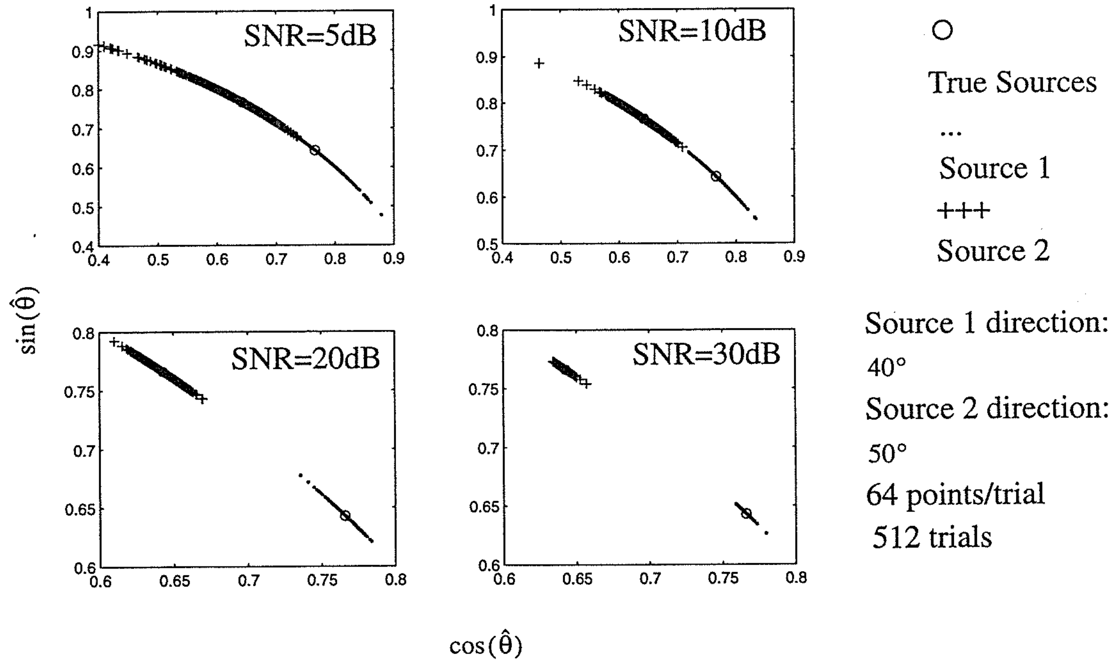


**Figure 4.5:** ESPRIT (1) simulation results for two directional sources: estimated DOA distribution v.s. SNR

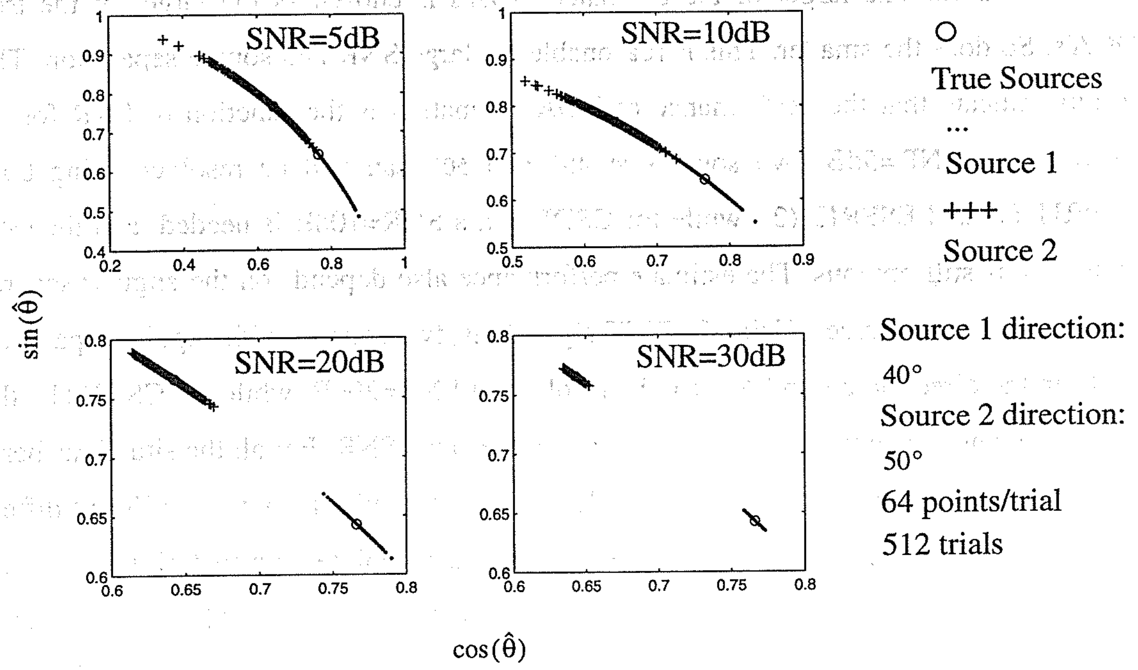


**Figure 4.6:** ESPRIT (2) simulation results for two directional sources: estimated DOA distribution v.s: SNR

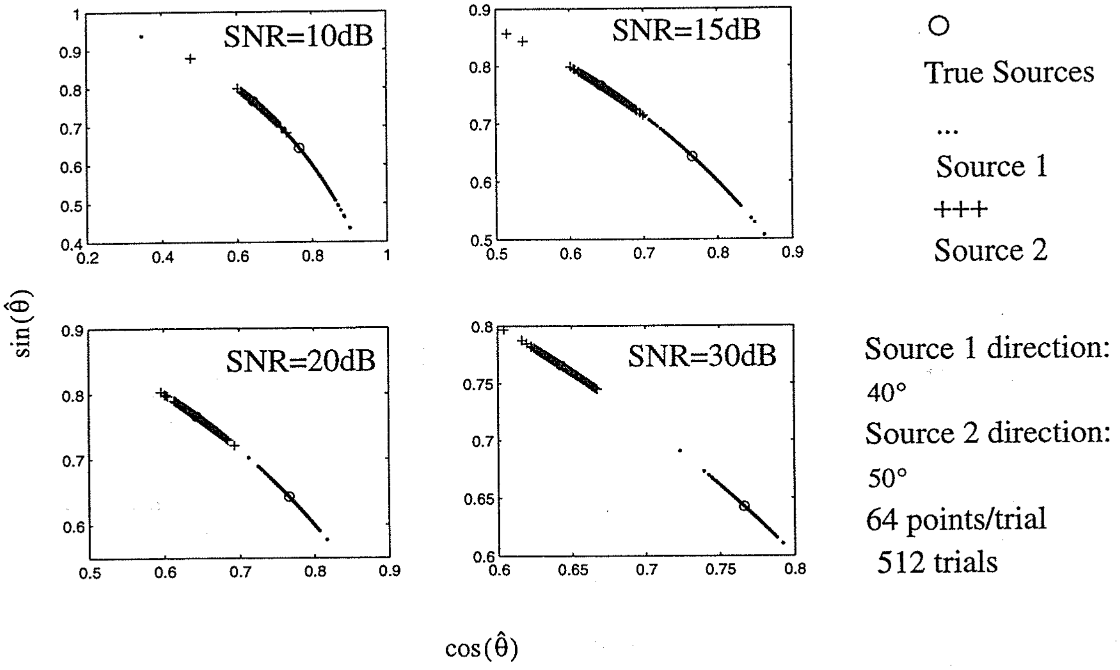


**Figure 4.7:** CSDE (1) simulation results for two directional sources: estimated DOA distribution v.s. SNR

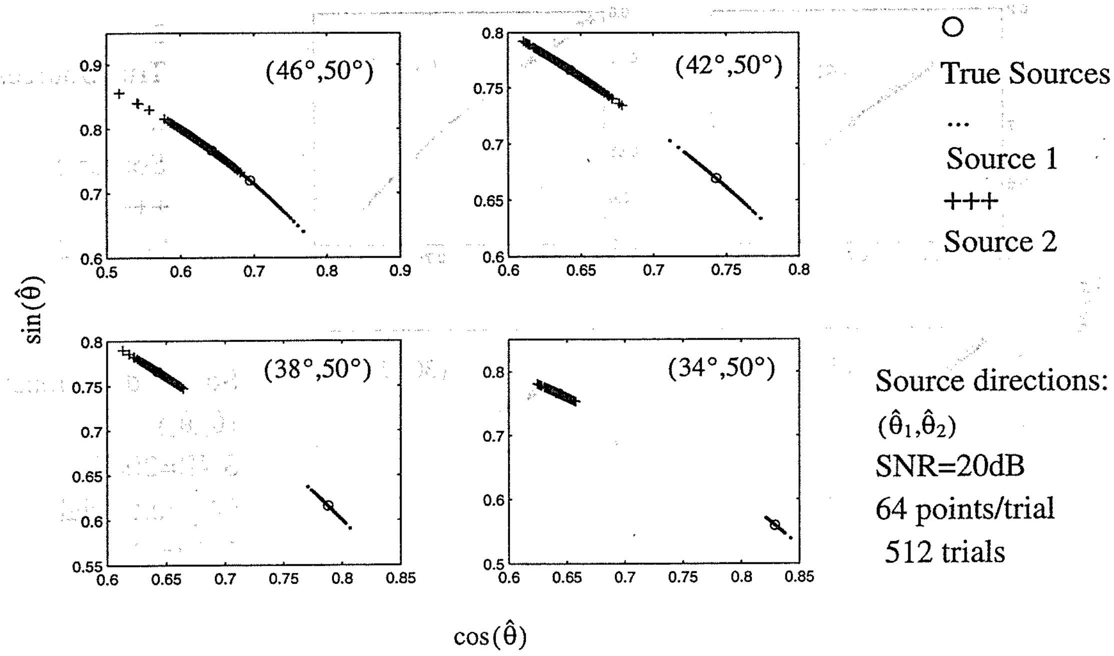


**Figure 4.8:** ESPRIT (1) simulation results for two directional sources: estimated DOA distribution v.s. angular separation

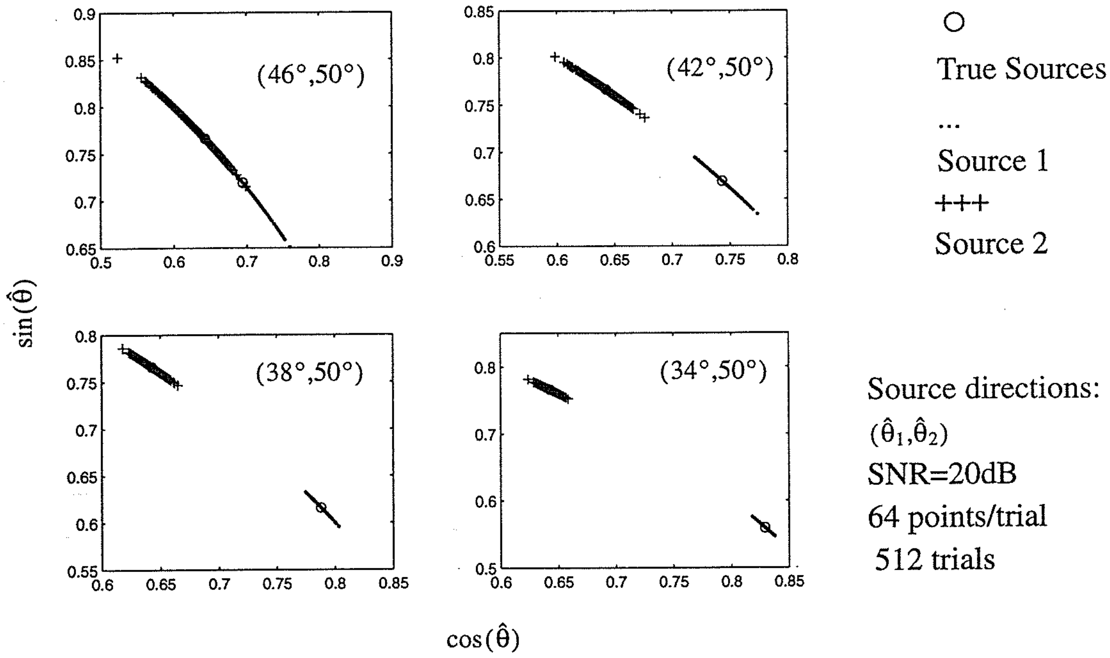


**Figure 4.9:** ESPRIT (2) simulation results for two directional sources: estimated DOA distribution v.s. angular separation

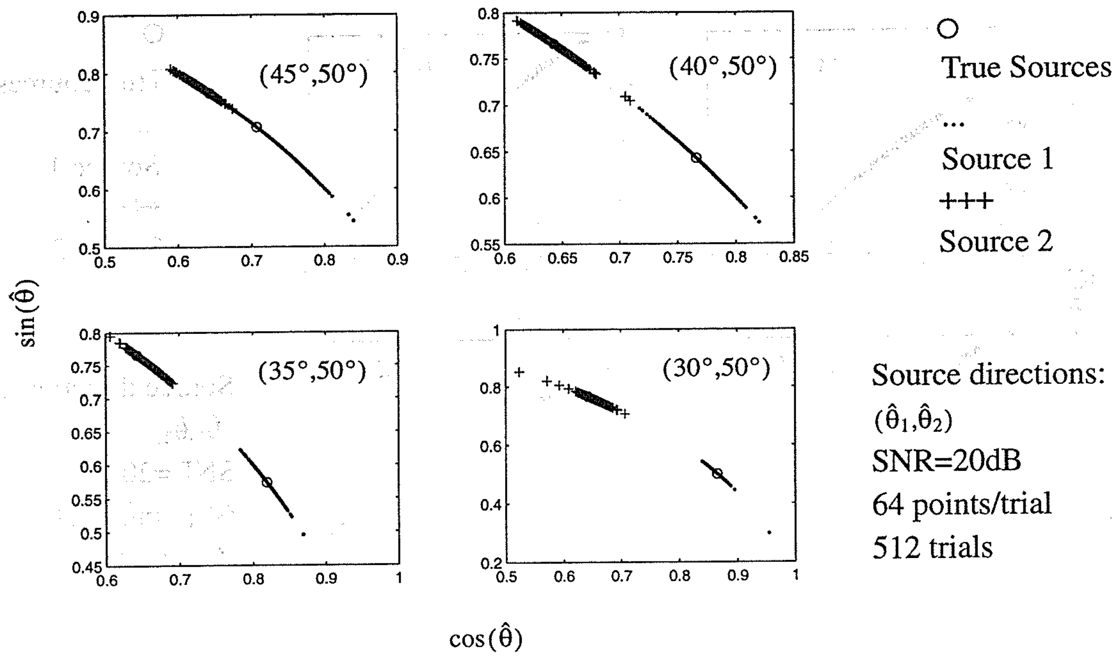


**Figure 4.10:** CSDE (1) simulation results for two directional sources: estimated DOA distribution v.s. angular separation

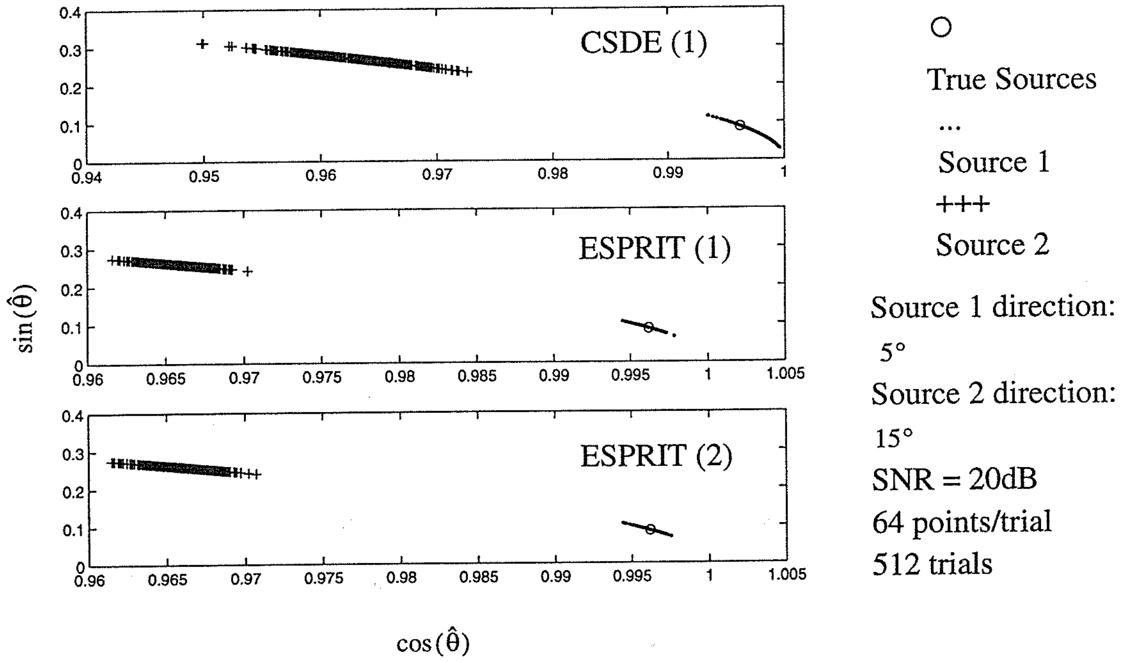


**Figure 4.11:** Simulation results for two directional sources: estimated DOA distribution at near-normal region

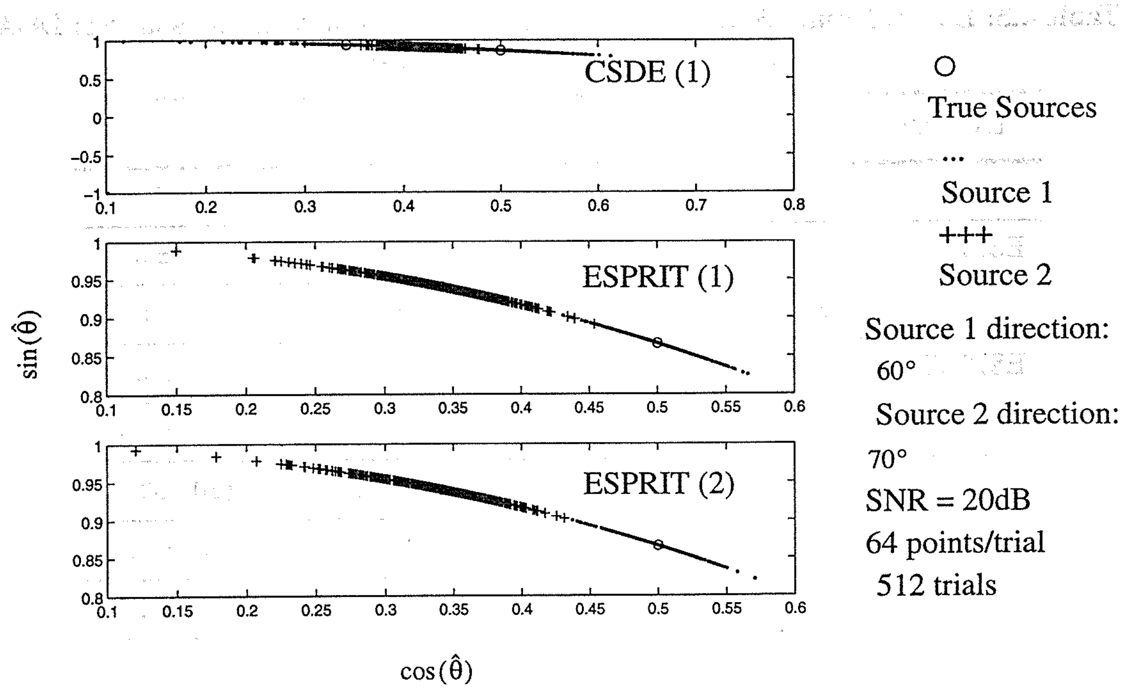


**Figure 4.12:** Simulation results for two directional sources: estimated DOA distribution at far off-normal region

**Table 4.2: ESPRIT and CSDE simulation results for two directional sources: DOA estimates v.s. SNR**

| Estimator | Source DOA Estimates ($40°, 50°$) | | | | |
|---|---|---|---|---|---|
| | SNR=5dB | SNR=10dB | SNR=15dB | SNR=20dB | SNR=30dB |
| ESPRIT (1) | 39.82° ± 3.12° | 39.83° ± 1.84° | | 40.01° ± 0.60° | 40.00° ± 0.20° |
| | 50.90° ± 4.03° | 50.31° ± 2.16° | | 50.00° ± 0.72° | 50.01° ± 0.22° |
| ESPRIT (2) | 39.51° ± 3.18° | 39.83° ± 1.78° | | 39.98° ± 0.61° | 40.00° ± 0.19° |
| | 50.75° ± 4.03° | 50.20° ± 2.14° | | 50.00° ± 0.73° | 49.99° ± 0.21° |
| CSDE (1) | | 38.76° ± 4.20° | 39.11° ± 2.40° | 39.51° ± 1.46° | 39.84° ± 0.76° |
| | | 48.53° ± 1.56° | 49.42° ± 1.36° | 49.73° ± 1.06° | 49.89° ± 0.62° |

**Table 4.3: ESPRIT and CSDE simulation results for two directional sources: DOA estimates v.s. angular separation**

| Estimator | Source DOA Estimates (SNR=20dB) | | | |
|---|---|---|---|---|
| | (46°, 50°) | (42°, 50°) | (38°, 50°) | (34°, 50°) |
| ESPRIT (1) | 45.84° ± 1.52° | 42.02° ± 0.74° | 38.00° ± 0.47° | 33.99° ± 0.34° |
| | 50.34° ± 1.76° | 50.10° ± 0.88° | 49.98° ± 0.56° | 49.96° ± 0.42° |
| ESPRIT (2) | 45.80° ± 1.56° | 42.03° ± 0.81° | 38.00° ± 0.46° | 33.99° ± 0.31° |
| | 50.38° ± 1.72° | 50.02° ± 0.87° | 49.94° ± 0.58° | 50.00° ± 0.42° |
| | (45°, 50°) | (40°, 50°) | (35°, 50°) | (30°, 50°) |
| CSDE (1) | 42.98° ± 2.98° | 39.43° ± 1.51° | 34.91° ± 1.21° | 30.03° ± 1.10° |
| | 50.70° ± 0.97° | 49.75° ± 1.03° | 49.23° ± 0.93° | 49.18° ± 1.11° |

**Table 4.4: ESPRIT and CSDE simulation results for two directional sources: DOA estimates v.s. source direction**

| Estimator | Source DOA Estimates (SNR=20dB) | | |
|---|---|---|---|
| | ( 5°, 15° ) | (40°, 50°) | (60°, 70°) |
| ESPRIT (1) | 4.98° ± 0.32° | 40.01° ± 0.60° | 60.07° ± 1.49° |
| | 15.11° ± 0.31° | 50.00° ± 0.72° | 70.26° ± 2.41° |
| ESPRIT (2) | 4.99° ± 0.33° | 39.98° ± 0.61° | 60.09° ± 1.48° |
| | 15.00° ± 0.33° | 50.00° ± 0.73° | 70.17° ± 2.32° |
| CSDE (1) | 4.22° ± 0.82° | 39.51° ± 1.46° | 63.40° ± 8.20° |
| | 15.63° ± 0.81° | 49.73° ± 1.06° | 65.42° ± 1.22° |

The amplitude estimate is investigated using both ESPRIT and CSDE. The results are shown in Fig. 4.13 through 4.18. Note, the ESPRIT solution is the estimate of source amplitude within a scalar of constant. Hence, we can observe the displacement between the true amplitude and ESPRIT estimate. However, the shape of the source amplitude is well recovered. The scalar factor must be determined by calibrating at least one sensor. On the other hand, the CSDE solution is the exact amplitude estimate. The example estimate standard deviation is given in Table 4.5. It is not strange that the performance of the amplitude estimation depends on that of the DOA estimation, and thus the SNR available. As shown in simulations, for SNR=10 - 20dB, the amplitude estimate is quite good as long as the DOA estimate is accurate.

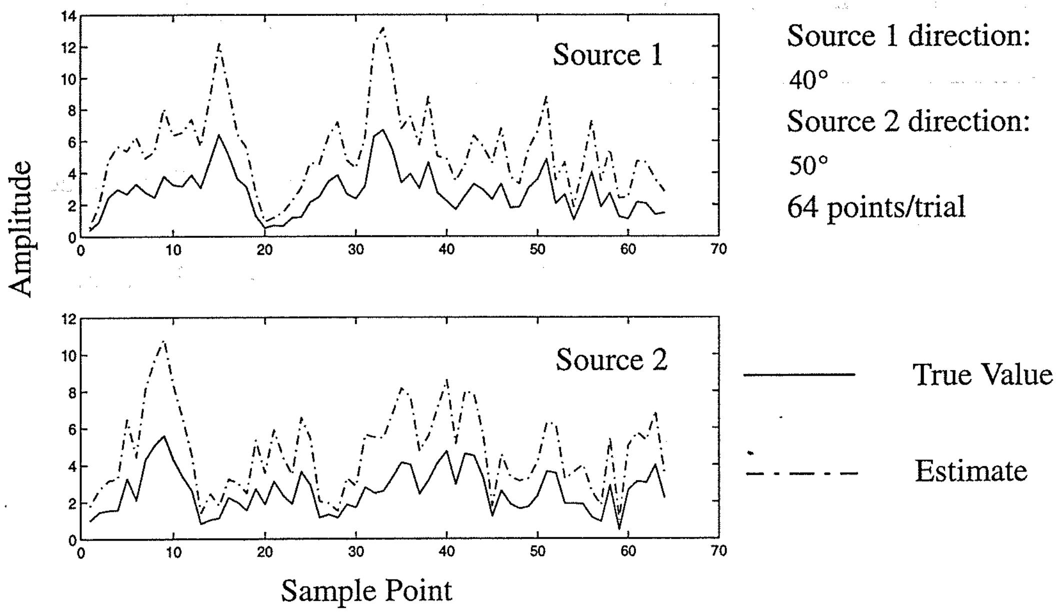


**Figure 4.13:** ESPRIT (1) simulation results for two directional sources: amplitude estimate example at SNR=20dB

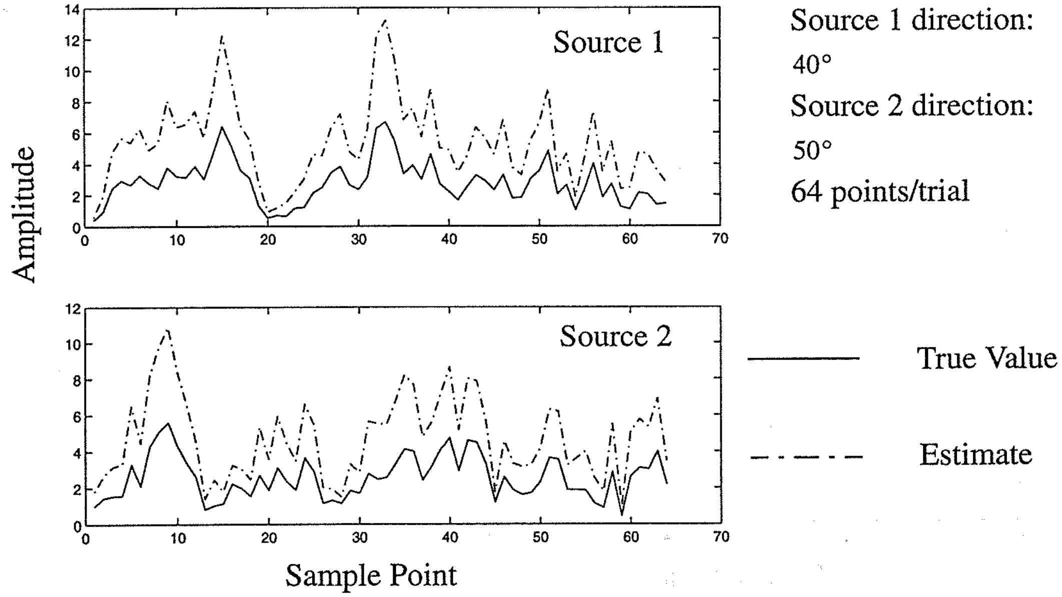


**Figure 4.14:** ESPRIT (2) simulation results for two directional sources: amplitude estimate example at SNR=20dB

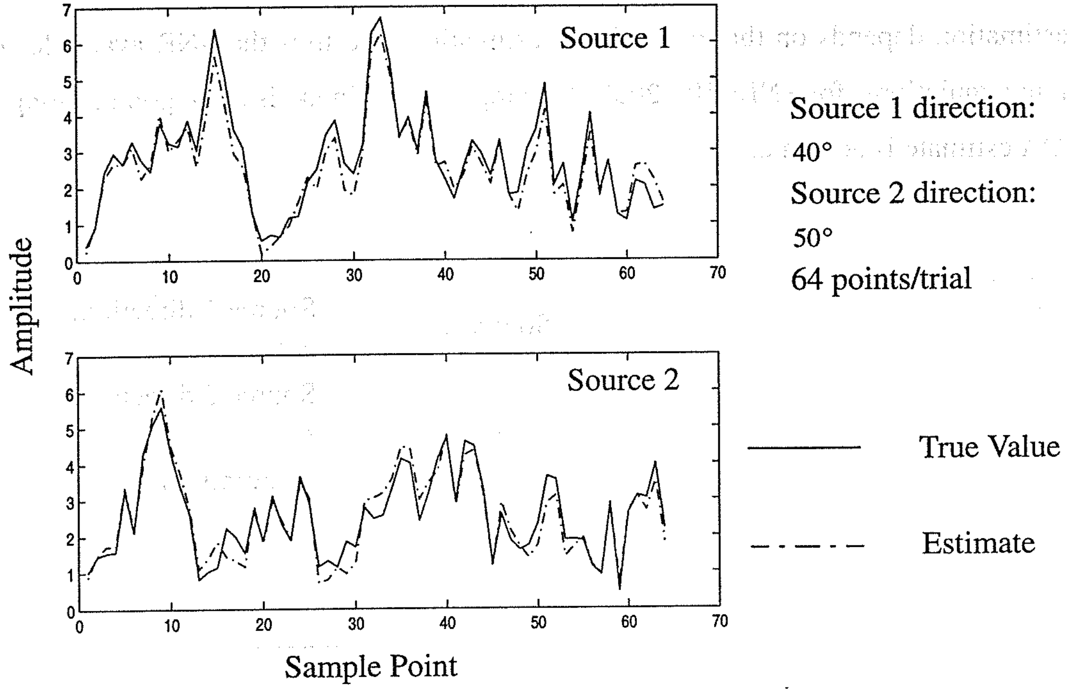


**Figure 4.15:** CSDE (1) simulation results for two directional sources: amplitude estimate example at SNR=20dB

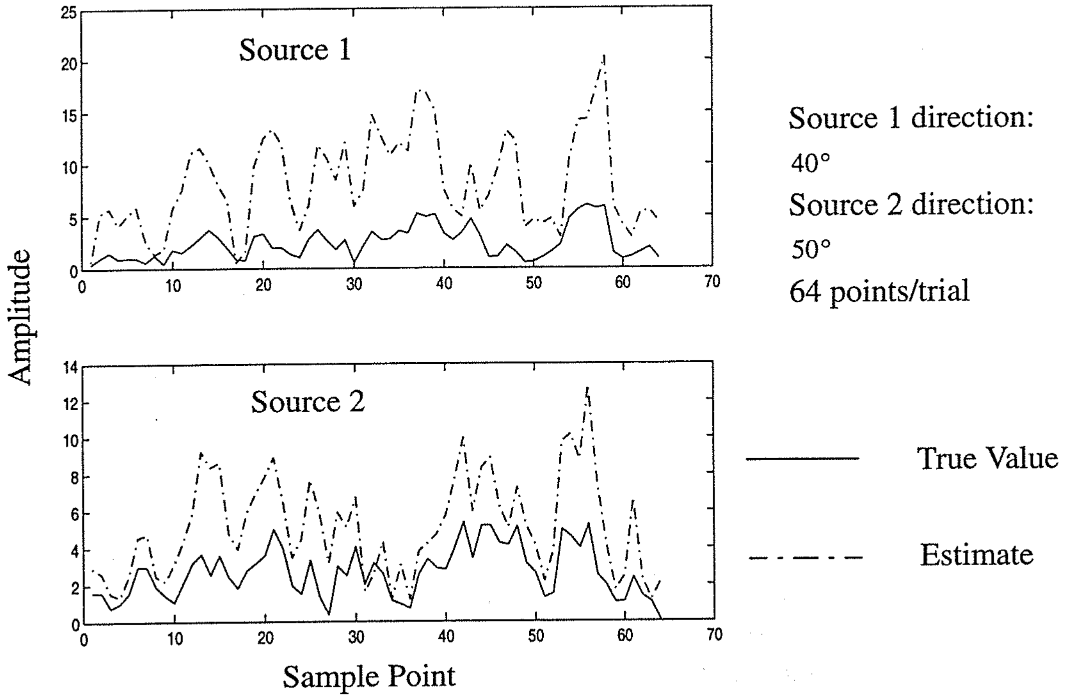


**Figure 4.16:** ESPRIT (1) simulation results for two directional sources: amplitude estimate example at SNR=10dB

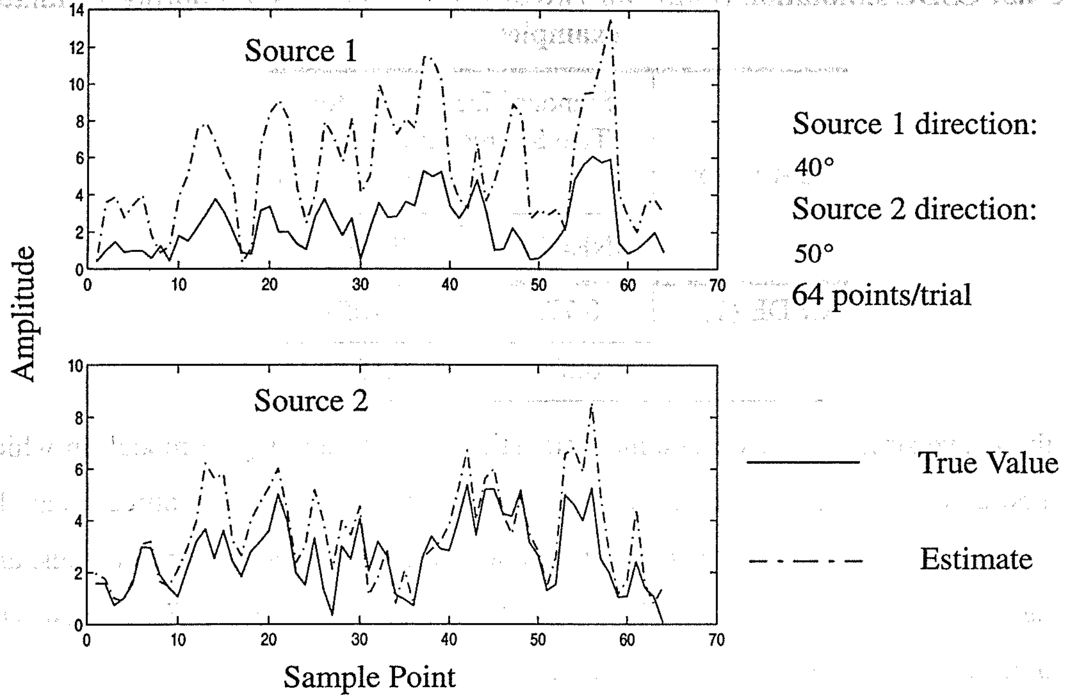


**Figure 4.17:** ESPRIT (2) simulation results for two directional sources: amplitude estimate example at SNR=10dB

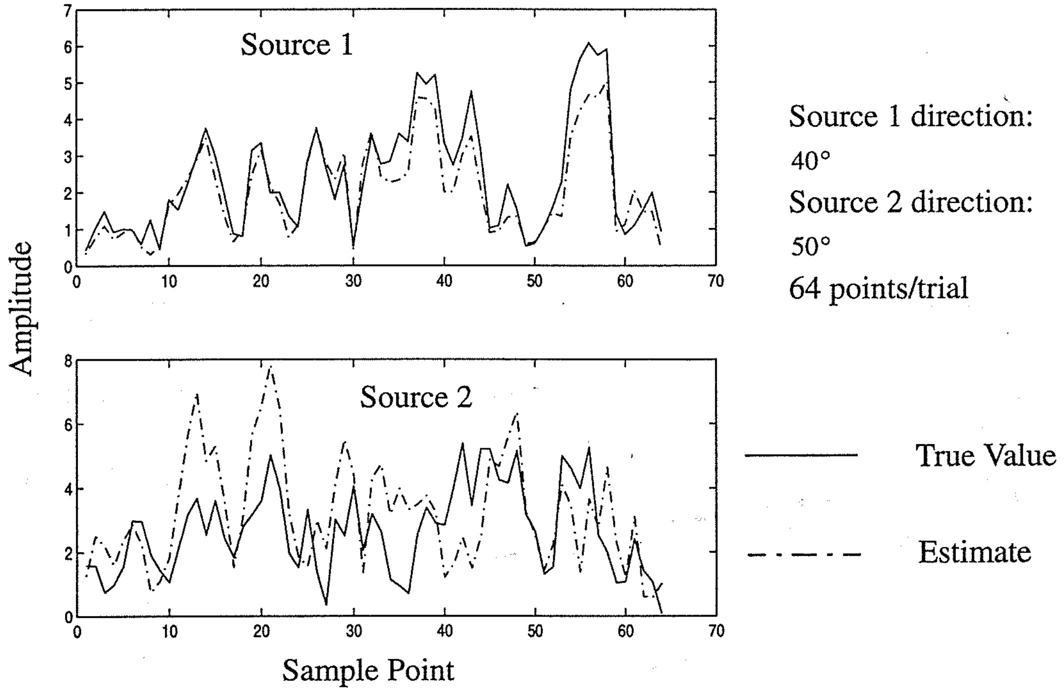


**Figure 4.18:** CSDE (1) simulation results for two directional sources: amplitude estimate example at SNR=10dB

**Table 4.5: CSDE simulation results for two directional sources: amplitude estimate examples**

| Estimator | Standard Deviations for Two Source Amplitude Estimates | |
|---|---|---|
| | SNR=10dB | SNR=20dB |
| CSDE (1) | 0.7734 | 0.4051 |
| | 0.8717 | 0.3575 |

In the above simulations, we assume a statistical independent signal model, in which the temporal correlation, or the coherence between two directional sources, can be ignored. Now, we will consider the coherent and temporally correlated signal models, and the covariance-based CSDE (CSDE (2)) is tested together with ESPRIT. First, sensitivity to temporal correlation is investigated. The results are shown in Fig. 4.19 through 4.21. The sample mean and standard deviation are given in Table 4.6. The averaged correlation coefficient, $c_{av}$, for any time lag within the sample data length is chosen as 0.5, 0.85, and 0.97, respectively. The results indicate that, though its performance decreases a little bit, ESPRIT is quite robust to the signal correlation except for very highly (close to 1.0) correlated model. On the contrary, for CSDE (2), a fully temporally correlated signal model is preferred.

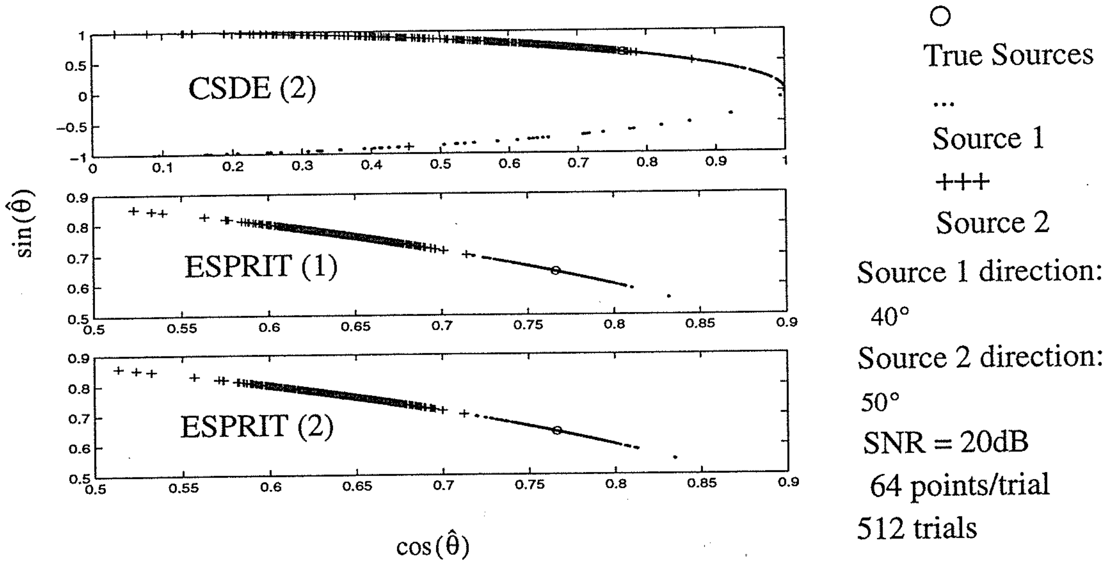


**Figure 4.19:** Simulation results for two directional sources: temporally correlated signal model with $c_{av}$=0.5

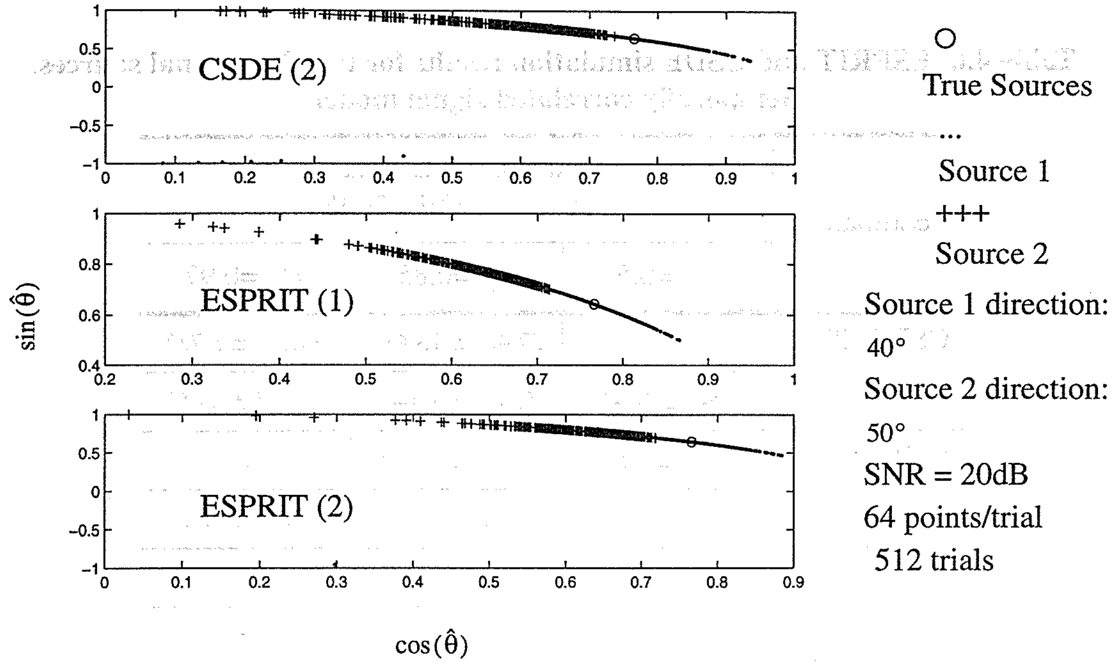


**Figure 4.20:** Simulation results for two directional sources: temporally correlated signal model with $C_{av}$=0.85

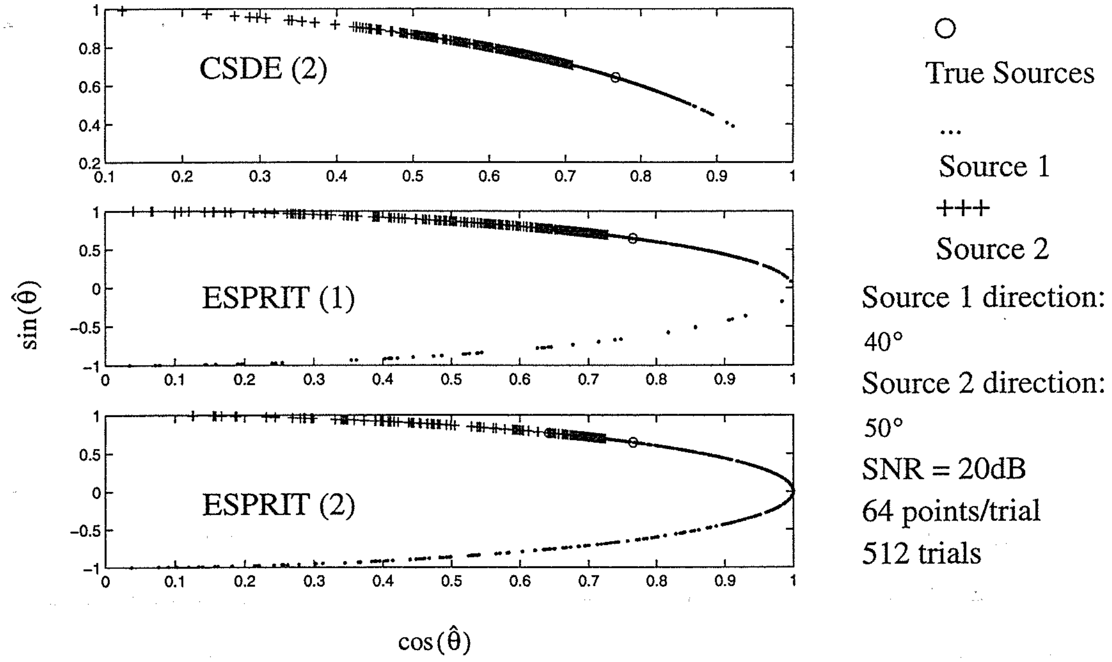


**Figure 4.21:** Simulation results for two directional sources: temporally correlated signal model with $C_{av}$=0.97

**Table 4.6: ESPRIT and CSDE simulation results for two directional sources: temporally correlated signal model**

| Estimator | Source DOA Estimates (40°, 50°, SNR=20dB) | | |
|---|---|---|---|
| | $C_{av}$=0.5 | $C_{av}$=0.85 | $C_{av}$=0.97 |
| CSDE (2) | / | 37.48° ± 13.43° | 39.18° ± 3.73° |
| | 51.54° ± 10.49° | 52.02° ± 6.64° | 51.91° ± 5.65° |
| ESPRIT (1) | 39.99° ± 1.52° | 39.94° ± 2.96° | / |
| | 50.02° ± 1.90° | 50.12° ± 3.93° | 51.20° ± 9.32° |
| ESPRIT (2) | 39.91° ± 1.55° | 39.53° ± 6.06° | / |
| | 50.10° ± 1.95° | 50.26° ± 4.71° | 48.32° ± 8.01° |

In reality, 100-percent correlated signals in time domain seldom occur, if not impossible. A much more meaningful model is the coherent signals model, which is investigated here. The simulation results are shown in Fig. 4.22 through 4.24. The sample mean and standard deviation are given in table 4.7. The degree of coherent, $C_d$, between two directional sources is chosen as 0.68, 0.86, and 0.98, respectively. Though ESPRIT is still not very sensitive to the coherent signal model except at highly coherent signals with degree of coherent larger than ~0.95, CSDE (2) shows much better performance than ESPRIT even at degree of coherent 0.70. This suggests the coherent signal be the best model for the covariance-based CSDE method.

The simulation for a real bathymetric sidescan sonar is implemented. The statistical signal model is the same as previous simulations for non-temporally-correlated, non-coherent signals. Two source directions are assumed to change continuously from 10° to 60° and from 20° to 70°, respectively, with a constant angular separation of 10° between them. Note, in a real system, the signal strength decreases as range increases, while the noise level at the receiver is almost a constant, independent of the signal range. Thus, the SNR decreases as range increases. Here, the SNR is assumed to change continuously from 40dB at near-normal region to 10dB at far off-normal region. All the system parameters are consistent with the sonar being developed at WHOI. For a sonar system $50m$ above the flat bottom, the chosen angular sector corresponds to the signal range of $60 - 200m$. Both ESPRIT and point-based CSDE (CSDE (1)) are applied. The results are shown in Fig. 4.25 through 4.27. The standard deviations from the true value for each source DOA estimate are given in Table 4.8. Except at far off-normal region, where CSDE (1) produces some biased estimates for one of the sources, all the methods present a good estimate for the source DOA's. At far off-normal region, the performance degradation can be expected, because the SNR's there are lower, and the algorithms perform worse at that region. The simulations here show that these algorithms has the capability to trace two time-varying source DOA's with a small angular separation between them for a wide angular region.

Finally, let's look at some real data from two-row bathymetric sidescan sonars. Fig. 4.28 shows an example of the recorded real received data at each of two rows. The estimated averaged SNR's at row 1 and row 2 are about 12.9dB and 10.2dB, respectively, and the noise correlation between two rows is about 0.12. The complex data can be obtained by using the Hilbert transform. Traditional DPE and ESPRIT for two-row structure are applied to the bottom echo DOA and amplitude estimate. Note, in this configuration, both ESPRIT (1) and ESPRIT (2) are equivalent. The covariance estimation is obtained at every data block with 64 samples. The results are shown in Fig. 4.29. Because the data were collected from a relatively flat bottom, we can observe an approximately linear source DOA relation with the increasing range. The ESPRIT-based approach achieves exactly the same performance for echo DOA estimates as the traditional DPE. Moreover, the amplitude estimate is obtained by the ESPRIT-based approach as well. Because the traditional DPE has been widely used in current bathymetric sidescan sonar techniques, the results here indicate a great application potential using the ESPRIT-based approach.

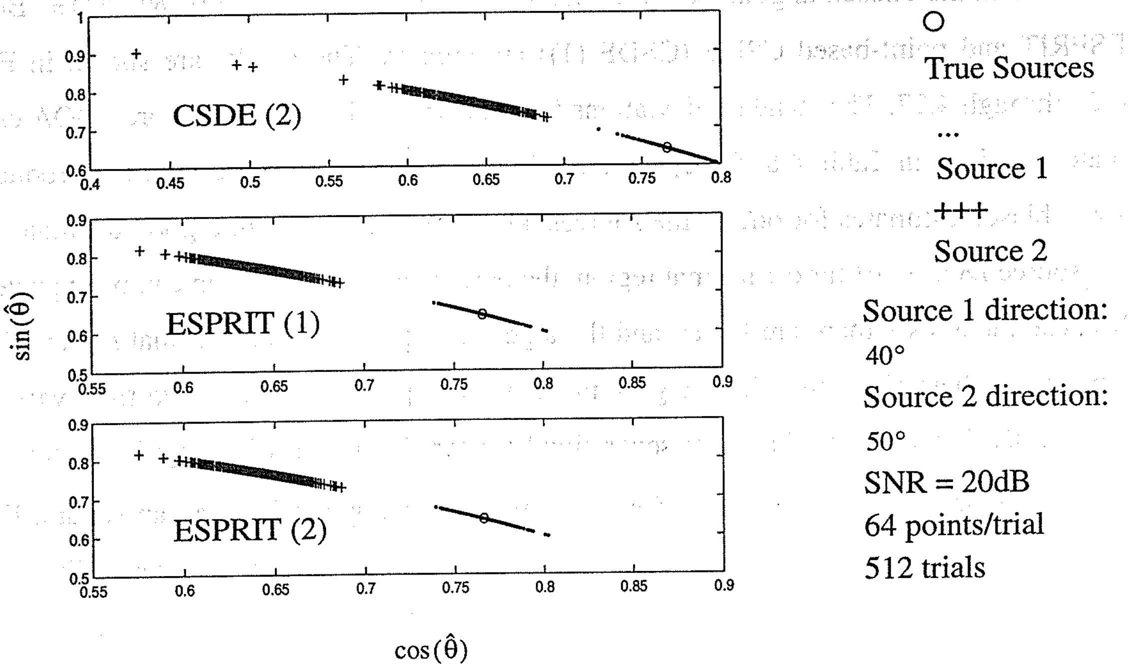


**Figure 4.22:** Simulation results for two directional sources: coherent signal model with $C_d$ =0:68

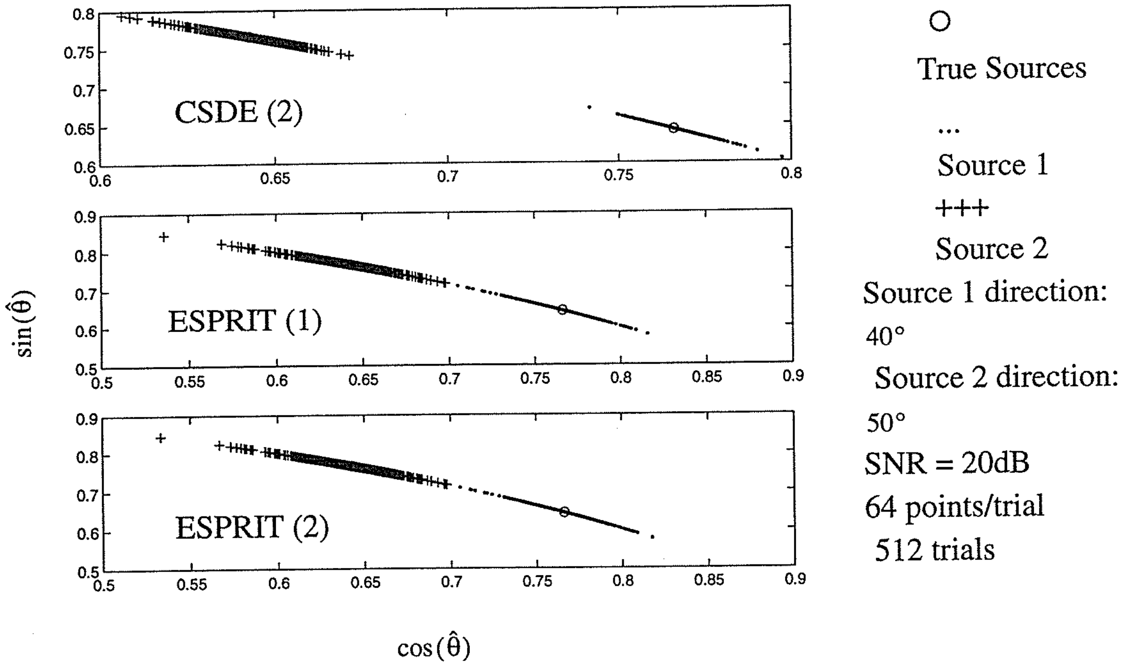


**Figure 4.23:** Simulation results for two directional sources: coherent signal model with $C_d$ =0.86

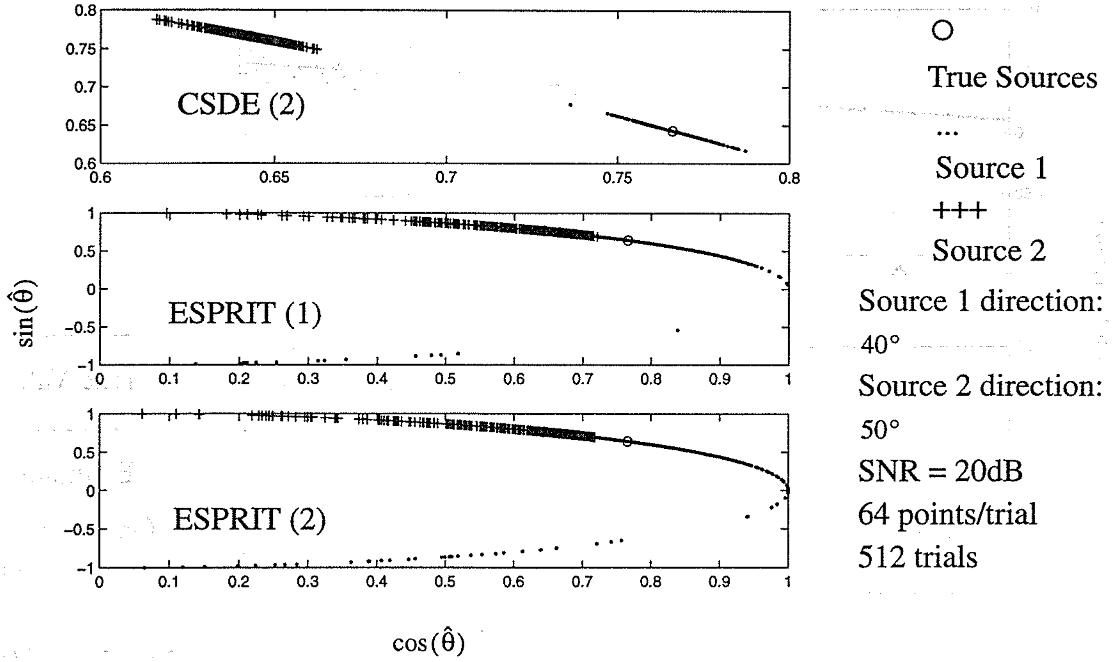


**Figure 4.24:** Simulation results for two directional sources: coherent signal model with $C_d$ =0.98

**Table 4.7: ESPRIT and CSDE simulation results for two directional sources: coherent signal model**

| Estimator | Source DOA Estimate ($40°$, $50°$, SNR=20dB) | | |
|---|---|---|---|
| | $C_d$ = 0.68 | $C_d$ = 0.86 | $C_d$ = 0.98 |
| CSDE (2) | 40.05° ± 0.92° | 39.97° ± 0.62° | 40.02° ± 0.66° |
| | 50.14° ± 1.60° | 50.02° ± 0.71° | 50.03° ± 0.57° |
| ESPRIT (1) | 40.07° ± 0.90° | 40.03° ± 1.45° | / |
| | 50.07° ± 1.20° | 50.10° ± 1.60° | 51.18° ± 6.60° |
| ESPRIT (2) | 40.06° ± 0.90° | 39.98° ± 1.46° | / |
| | 50.09° ± 1.21° | 50.14° ± 1.62° | 51.11° ± 7.55° |

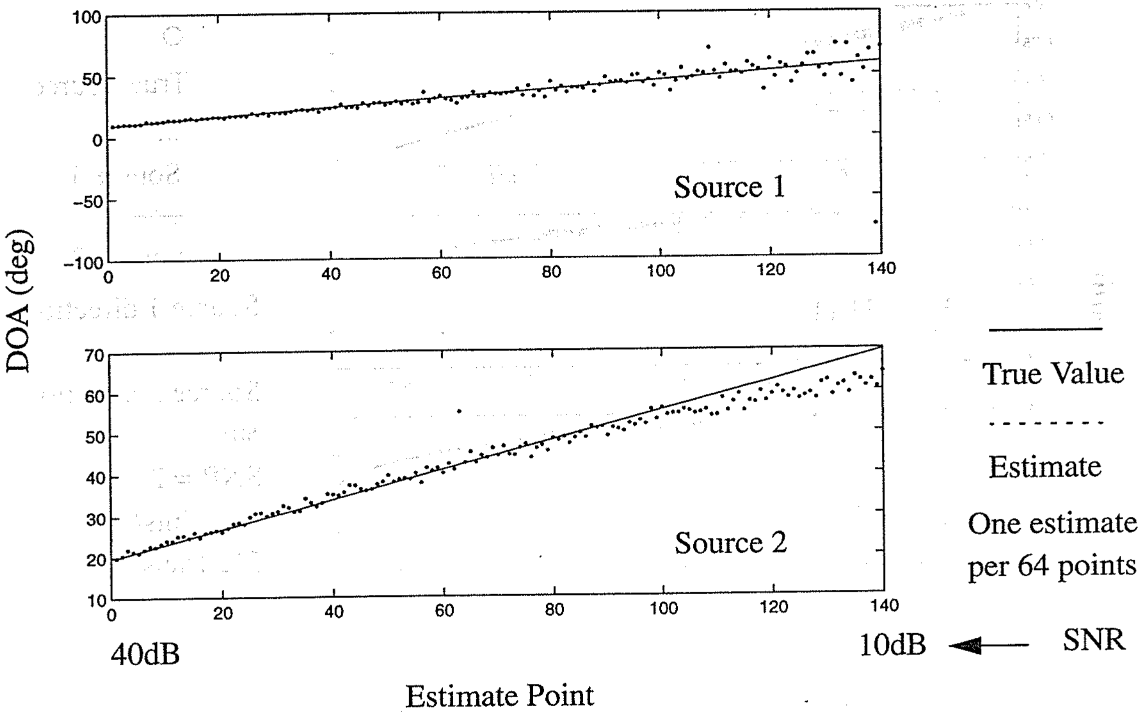


**Figure 4.25:** CSDE (1) simulation results for two time-varying directional sources with time-varying SNR.

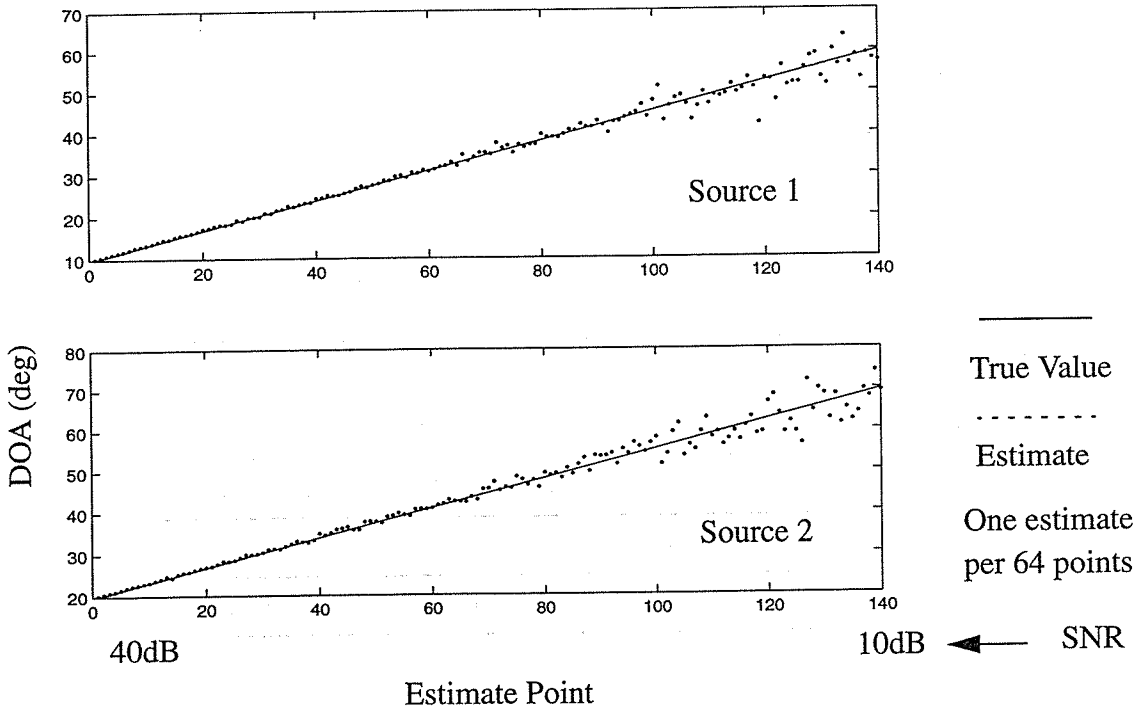


**Figure 4.26:** ESPRIT (1) simulation results for two time-varying directional sources with time-varying SNR.

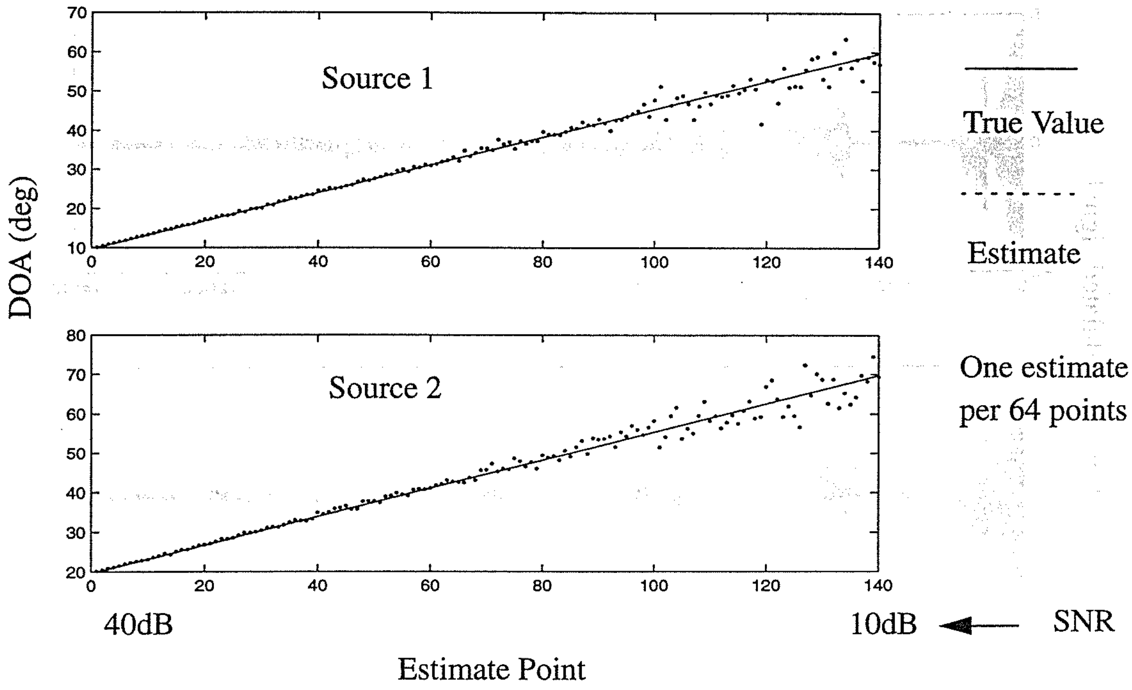


**Figure 4.27:** ESPRIT (2) simulation results for two time-varying directional sources with time-varying SNR.

**Table 4.8: ESPRIT and CSDE simulation results for two time-varying directional sources with time-varying SNR's**

| Estimator | CSDE (1) | | ESPRIT (1) | | ESPRIT (2) | |
|---|---|---|---|---|---|---|
| Source DOA estimates: Standard deviation | 12.27° | 2.99° | 1.75° | 2.07° | 1.83° | 2.08° |

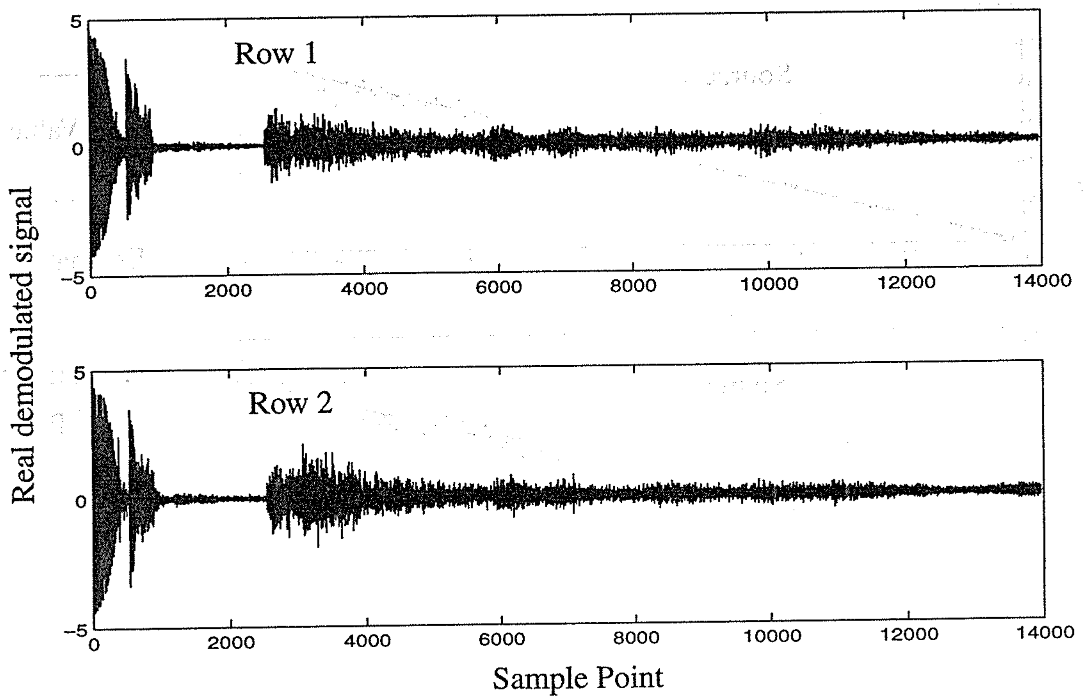


**Figure 4.28:** Real data from a flat bottom using a two-row bathymetric sidescan sonar

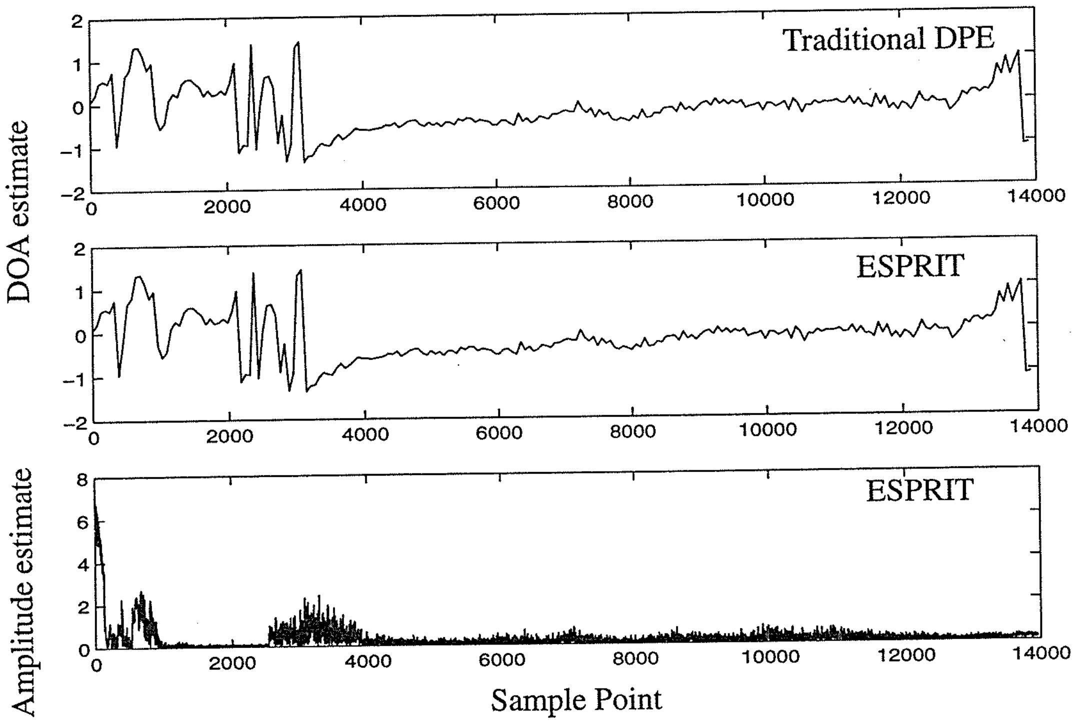


**Figure 4.29:** Real data test results

# Chapter 5

# Discussions and Conclusions

## 5.1 Discussions for Real System Application

It is well known that the computation load and the sensitivity to signal and noise field are the two key factors, which make it difficult to use a high-resolution adaptive DOA estimation method in a real commercial sonar system. In this section, these two issues are discussed simply.

First, the computation loads are evaluated for ESPRIT and CSDE, which are widely tested in Chapter 4. Each estimate is obtained by using 64 data samples, and the sensor number is assumed to be three. Obviously, a significant computational effort is expended in forming the measurement covariance matrix, which requires on the order of $64 \times 3^2 = 576$ complex operations. In ESPRIT (1), eigendecompositions of three matrices with the sizes of $4 \times 4$, $4 \times 4$ and $2 \times 2$, respectively, are computed. Because an eigendecomposition of a $m \times m$ complex matrix requires on the order of $10m^3$ complex operations, the total computation load for eigendecompositions is on the order of $1360$ complex operations. (It can be shown that the computation requirement is the same order for both eigendecomposition and SVD if the covariance matrix forming is included and only singular values and left singular vector are computed [39].) Besides, to reduce the generalized eigendecomposition to the standard eigendecomposition, a Mahalanobis transformation is implemented, which needs $16 \times 7 = 112$ complex operations. To obtain the signal subspace estimate from the first eigendecomposition, $8 \times 7 = 56$ complex operations are needed, and to estimate the rotation operator $\Psi$, about $4 \times 3 + 7 = 19$ complex operations are needed. Therefore, the computation load for DOA estimation is on the order of $2123$ complex operations. If a modified ESPRIT (ESPRIT (2)) is used with the noise covariance ignored, this number is on the order of $1955$. Furthermore, the amplitude estimate includes the weights computation, which needs $78$ complex operations, and the output computation, which needs $128 \times 7 = 896$ complex operations. Finally, accounting for the eigenvalue ordering operation and memory operation, the total computation load is on the order of $3300$ complex operations for ESPRIT (1), and $3100$ for ESPRIT (2). Assuming a sampling rate of $48KHz$, the number of the complex operations per second is on the order of $2.475M$ for ESPRIT (1), and $2.325M$ for ESPRIT (2).

A similar analysis can be applied to CSDE. Actually, the covariance-based CSDE is most computationally efficient, whose total computation load per second is on the order of $750K$ complex operations, while the point-based CSDE requires more computations, which are on the order of $2M$ complex operations per second.

Note, in average, a complex operation (multiply or add/subtract) requires four real operations. Assuming floating-point operation, the DSP system must be able to implement $10M$ such operations per second, which is not a critical requirement for a modern DSP chip. For example, TMS320C30 produced by Texas Instruments has the capability of $33.3M$ FLOPS (floating-point operations per second). Though some other facts may increase (e.g., other input/output, procedure management tasks) or decrease (e.g., parallel instruction, matrix-orient instruction) the computation load, it is reasonable to declaim that the computation load is no longer a critical issue in a real bathymetric sidescan sonar system realization. This is due to a relatively simple array configuration and also the significant improvement on computational efficiency by ESPRIT over MUSIC and some other high-resolution adaptive methods.

Second, the real data imperfection should be taken into account because it will affect the performance of the DOA and amplitude estimate by varying degrees. In general, most adaptive algorithms require a known noise correlation estimate, which can be obtained either by real measuring or by theoretically modeling. If a wrong model is used, the mismatching will cause the performance degradation, even totally failure operation. In ESPRIT, a noise correlation estimate is also needed. However, in the simulations in chapter 4, we have observed no significant difference between the standard ESPRIT and the modified ESPRIT, which doesn't use the noise correlation information. This is not strange because in the context of bathymetric sidescan sonar, only several sensors are used, and the optimum array gain is ten times the logarithm of the sensor number. Consequently, though a significant improvement in SNR won't happen, the noise model mismatching is not a serious problem here.

However, some other data imperfections may be important. As we have discussed before, the degree of temporal correlation for one source, spatial correlation among sensors for one source, coherence between different sources, affect more or less the estimate performance. The Gaussian assumption of the bottom scattering doesn't hold as the resolution increases. Furthermore, the strongly dispersive medium makes the plane wave assumption invalid, and for a complicated bottom structure, the elementary scattering signal loses its similarity with the impinging signal. In the fishery survey application, the rapidly moving of the fish schools causes the Doppler effects no longer ignorable.

Besides, the system design imperfections, such as non-matching beampatterns, perturbed sensor gain and phase response, and sensor position error, should be considered. In system implementation, an ill-conditioning matrix operation should also be avoided.

It is impossible to deal with all these imperfections well. However, if some informations about them are available, a basis can be established to evaluate quantitatively the performance obtained by the system algorithms. For example, as we know, the performance of all the proposed algorithms depends strongly on the input SNR, so a better estimate can be expected at the high SNR region, and vice versa. Therefore, the measurement data quality is indicated by an estimated SNR map associated with a bathymetric sidescan output. Also, as we have discussed in section 3.3, the eigendecomposition of the measurement covariance matrix describes the signal field structure to some extent.

## 5.2 Conclusions and Future Works

In the real application with bathymetric sidescan sonar, the multipath reflection and other directional interferences are the key limiting factors for a better performance. A new scheme to deal with these interferences using a multiple-row bathymetric sidescan sonar is proposed. Instead of smoothing the measurements over some time or angle intervals as previously widely investigated, we resolve the interferences from the signals of interests. The proposed scheme is supported by the analysis on the interference field associated with bathymetric sidescan sonar, which shows that, for those regions with necessary signal-to-noise ratio for reliable bathymetry measurements, a model with only one to three dominating directional sources (including the interferences) at the same instant is a reasonable approximation to the wave field.

A few approaches are studied or developed to obtain the signal DOA and amplitude estimate, and because the real three or more-row sonar data are not available at the time of this writing, these approaches are tested using the simulation data based on both the statistical and coherent signal models. For one directional source situation, the traditional differential phase estimate (DPE), two of its modified versions, and ESPRIT are investigated. The results show that, at mediate to high SNR, the averaged DPE and vernier DPE perform as well as ESPRIT. However, at lower SNR, ESPRIT shows better performance in the sense of estimate statistics, possibly because it exploits the underlying noise correlation knowledge.

For the cases with two directional sources, a correlated signal direction estimate (CSDE) for three-row systems is proposed, and compared with the ESPRIT-based approaches using different data models. The simulation results show ESPRIT-based approaches are quite robust at the angular separation of $\geq 10°$ between two sources and at the signal-to-noise ratio above 10dB except for highly coherent or temporally correlated signals, for which the covariance-based CSDE works very well. Besides, both ESPRIT and CSDE can give the amplitude estimate for each directional source on the basis of the DOA estimate.

Basically, ESPRIT algorithm exploits the phase information embedded in the signal covariance estimate as DPE does. Therefore, for a two-row system and one directional source, both ESPRIT and DPE are equivalent. However, because ESPRIT can be applied to a multiple-row system, more directional sources may be resolved. On the other hand, CSDE depends on not only the received phase information but also the exact amplitude measurement at each row, thus limiting its performance in the non-coherent, non-temporally-correlated signal field. However, a point-based CSDE doesn't assume any signal or noise model and can always be used, though its performance depends. Furthermore, CSDE is most computationally efficient.

Under the array configuration with only three or a few more sensors, a high-resolution adaptive method such as ESPRIT cannot fully show its advantages in DOA and amplitude estimation. On the other hand, the sensitivity to signal and noise field inherent to such a method is reduced. The simulation results and real-data test using two-row sonar data, as well as the discussions on the computation load and data imperfection, have shown that

ESPRIT has a great potential to be used in a real multiple-row bathymetric sidescan sonar system.

There is certainly more research to be done into the practical aspects of the above algorithms. First, all these algorithms need to be tested using the real three or more-row sonar data. Second, the imperfections in the real data should be further investigated. Based on the real data test and analysis results, the optimum row number should be able to be determined. With the capability of simultaneously resolving two or more directional sources, the new multiple-row bathymetric sidescan sonar should work better for a wider variety of practical situations, particularly in shallow water; this improvement can be obtained without significant increase of the system cost.

So far, our discussions concentrate on the signal DOA and amplitude estimation. A logic step for further data processing is to interpret and smooth these estimates. Therefore, by combining the smoothed estimates with other sensor data, the measurements of the bottom bathymetry, backscattering properties and target imagery can be obtained.

# Bibliography

[1] P. N. Denbigh. Swath bathymetry: Principles of operation and analysis of errors. *IEEE Journal of Oceanic Engineering*, 14(4):289-298, October 1989.

[2] H. Matsumoto. Characteristics of SeaMARC II phase data. *IEEE Journal of Oceanic Engineering,* 15(4):350-360, October 1990.

[3] M. A. Masnadi-Shirazi, C. de Moustier, P. Cervenko, and S. H. Zisk. Differential phase estimation with the SeaMARC II bathymetric sidescan sonar system. *IEEE Journal of Oceanic Engineering,* 17(3):239-251, July 1992.

[4] G. Jin and D. Tang. Uncertainties of differential phase estimation associated with interferometric sonars. *IEEE Journal of Oceanic Engineering,* 21(1):1-11, January 1996.

[5] W. Zhu, W. Xu, and J. Yu. Error estimation of pulse pair correlation differential phase estimator of sonar array. In *Proc. of IEEE OCEANS'96*, pages 1071-1075, Fort Lauderdale, FL., September 1996.

[6] W. K. Stewart, D. Chu, S. Malik, S. Lerner and H. Singh. Quantitative seafloor characterization using a bathymetric sidescan sonar. *IEEE Journal of Oceanic Engineering,* 19(4):599-610, October 1994.

[7] S. Haykin and A. Steinhardt, editors. *Adaptive Radar Detection and Estimation.* John Wiley & Sons, New York, 1992.

[8] S. Haykin. *Adaptive Filter Theory, 3rd Edition.* Prentice Hall, New Jersey, 1996.

[9] A. B. Baggeroer. *Array Signal Processing.* 13.742J course notes, Massachusetts Institute of Technology, Cambridge, MA, 1997.

[10] B. D. Van Veen and K. M. Buckley. Beamforming: A versatile approach to spatial filtering. *IEEE ASSP Magazine*, pages 4-24, April 1988.

[11] H. Krim and M. Viberg. Two decades of array signal processing research. *IEEE Signal Processing Magazine*, pages 67-94, July 1996.

[12] R. H. Roy. *ESPRIT: Estimation of Signal Parameters via Rotational Invariance Techniques*. Ph.D. dissertation, Stanford University, Stanford, CA, August 1987.

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

[14] J. P. Burg. Maximum entropy spectral analysis. In *Proc. of the 37th Annual International SEG Meeting*, Oklahoma City, OK, October 1967.

[15] V. F. Pisarenko. The retrieval of harmonics from a covariance function. *Geophys. J. R. astr. Soc.*, 33:347-366, 1973.

[16] R. O. Schmidt. *A Signal Subspace Approach to Multiple Emitter Location and Spectral Estimation*. Ph.D. dissertation, Stanford University, Stanford, CA, 1981.

[17] R. Roy, A. Paulraj, and T. Kailath. ESPRIT-A subspace rotation approach to

estimation of parameters of cisoids in noise. *IEEE Transaction on Acoustics, Speech, and Signal Processing*, 34(5): 1340-1342, October 1986.

[18] R. Roy and T. Kailath. ESPRIT-Estimation of signal parameters via rotational invariance techniques. *IEEE Transaction on Acoustics, Speech, and Signal Processing*, 37(7): 984-995, July 1989.

[19] B. Ottersten and T. Kailath. Direction-of-arrival estimation for wide-band signals using the ESPRIT algorithm. *IEEE Transaction on Acoustics, Speech, and Signal Processing*, 38(2): 317-327, February 1990.

[20] G. Su and M. Morf. Modal decomposition signal subspace algorithms. *IEEE Transaction on Acoustics, Speech, and Signal Processing*, 34(3): 585-602, June 1986.

[21] B. Ottersten, M. Viberg, and T. Kailath. Performance analysis of the total least squares ESPRIT algorithm. *IEEE Transaction on Acoustics, Speech, and Signal Processing*, 39(5): 1122-1135, May 1991.

[22] A. L. Swindlehurst, B. Ottersten, R. Roy, and T. Kailath. Multiple invariance ESPRIT. *IEEE Transaction on Acoustics, Speech, and Signal Processing*, 40(4): 867-881, April 1992.

[23] J. C. Curlander. Data pre-processing of SAR data to enhance target to clutter ratio. Vexcal Corporation, Boulder, CO.

[24] A. van der Merwe, J. W. Odendaal, and E. C. Botha. 2D microwave imaging using TLS-ESPRIT with matrix enhancement. *Microwave and Optical Technology Letters,* 8(4):218-222, March 1995.

[25] K. T. Wong and M. D. Zoltowski. Orthogonal velocity-hydrophone ESPRIT for sonar source localization. In *Proc. of IEEE OCEANS'96*, pages 1307-1312, Fort Lauderdale, FL, September 1996.

[26] K. T. Wong and M. D. Zoltowski. ESPRIT-based extended-aperture source localization using velocity-hydrophones. In *Proc. of IEEE OCEANS'96*, pages 1427-1432, Fort Lauderdale, FL, September 1996.

[27] B. Porat and B. Friedlander. Direction finding algorithms based on high-order statistics. *IEEE Transaction on Signal Processing*, 39(9): 2016-2024, September 1991.

[28] M. C. Dogan and J. M. Mendel. Applications of cumulants to array processing-Part I: Aperture extension and array calibration. *IEEE Transaction on Signal Processing*, 43(5): 1200-1216, May 1995.

[29] N. Yuen and B. Friedlander. DOA estimation in multipath: An approach using fourth-order cumulants. *IEEE Transaction on Signal Processing*, 45(5): 1253-1263, May 1997.

[30] C. H. Knapp and G. C. Carter. The generalized correlation method for estimation of time delay. *IEEE Transaction on Acoustics, Speech, and Signal Processing*, 24(4): 320-327, August 1976.

[31] G. C. Carter, editor. *Coherence and time delay estimation.* Piscataway, New Jersey, 1993.

[32] G. V. Frisk. *Ocean and Seabed Acoustics.* Prentice Hall, Englewood Cliffs, New Jersey, 1994.

[33] X. Tang and W. K. Stewart. Plankton image classification using novel parallel-training learning vector quantization network. In *Proc. of IEEE OCEANS'96*, pages 1227-1236, Fort Lauderdale, FL., September 1996.

[34] R. G. Hendershot and D. R. Jackson. Size and angular location estimates of fish schools using spatial correlation. In *Proc. of IEEE OCEANS'80*, pages 386-391, September 1980.

[35] V. V. Ol'shevskii. *Characteristics of Sea Reverberation.* Consultants Bureau, New York, 1967.

[36] J. O. Klepsvik. *Statistical characteristics of the sea-bed reverberation process.* Continental Shelf Institute, Norway, August 1983.

[37] J. E. Ehrenberg, T. E. Ewart, and R. D. Morris. Signal-processing techniques for resolving individual pulses in a multipath signal. *Journal of Acoustical Society of America*, 63(6):1861-1865, June 1978.

[38] Fredric J. Harris. On the use of windows for harmonic analysis with the discrete fourier transform. *Proceedings of The IEEE*, 66(1):51-83, January 1978.

[39] G. H. Golub and C. F. Van Loan. *Matrix Computations, 3rd edition.* The Johns Hopkins University Press, 1996

[40] R. J. Urick. *Principles of Underwater Sound, 3rd edition.* McGraw-Hill, New York, 1983.