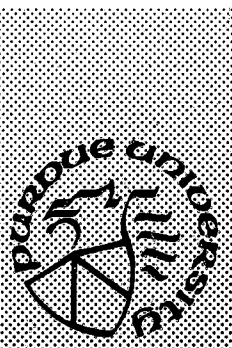
12-1-1989

Robust Techniques for Bearing Estimation in Contaminated Gaussian Noise

R. L. Kashyap
Purdue University

David D. Lee
Purdue University

N. Srinivasa
Purdue University



Robust Techniques for Bearing Estimation in Contaminated Gaussian Noise

R. L. Kashyap
D. D. Lee
N. Srinivasa

TR-EE 89-67
December, 1989

School of Electrical Engineering
Purdue University
West Lafayette, Indiana 47907

Office of Naval Research

Contract N00014-85K-0611

Task R&T Number: s400005 SRB05

Partially supported by

SDIO Innovative Science and Technology(IST) Program

ROBUST TECHNIQUES FOR BEARING ESTIMATION IN CONTAMINATED GAUSSIAN NOISE

by

R. L. Kashyap (Principal Investigator)

David D. Lee

N. Srinivasa

School of Electrical Engineering

Purdue University

West Lafayette, Indiana 47907

Technical Report EE 89-67

December 1989

TABLE OF CONTENTS

Page

Page

LIST OF TABLES

Table Page

LIST OF FIGURES

Figure Page

Figure Page

ABSTRACT

The problem of estimating directions-of-arrival (DOA) of radiating sources from measurements provided by a passive array of sensors is frequently encountered in radar, sonar, radio astronomy and seismology. In this study various *robust* methods for the DOA estimation problem are developed, where the term *robustness* refers to insensitivity against small deviation in the underlying Gaussian noise assumption. The first method utilizes an *eigenvector method* and *robust* reconstruction of the correlation matrix by time series modeling of the array data. Secondly, a *decentralized processing* scheme is considered for geographically distributed array sites. The method provides reliable estimates even when a few of the subarray sites are malfunctioning. The above two techniques are useful for narrow band and incoherent sources. The third *robust* method, which utilizes Radon Transform, is capable of handling both the narrow band and wide band sources as well as the incoherent or coherent sources. The technique is also useful in situations of very low SNR and colored noise with unknown correlation structure. The fourth method is an efficient narrow band *robust maximum likelihood* DOA estimation algorithm which is capable of handling coherent signals as well as the single snapshot cases. Furthermore, relationships between eigenvector methods and a ML DOA estimation, where the source signals are treated as sample functions of Gaussian random processes, are investigated.

CHAPTER 1
INTRODUCTION AND OVERVIEW

1.1. Introduction and Literature Review

Array processing deals with the processing of signals carried by propagating wave phenomena. The received signal is obtained by means of an array of sensors located at different points in space in the field of interest. The aim of array processing is to extract useful characteristics of the received signal field, e.g., direction of arrival (DOA), signature, speed of propagation. The sources of energy responsible for illuminating the array may assume a variety of different forms. They may be narrow band or wide band. Furthermore, they may be incoherent, i.e., independent of each other, or coherently related to each other. Equally, as seen from the location of the array, the radiation may be from diffused media and therefore distributed in nature, or it may be from isolated sources of finite angular extent. The array itself takes on a variety of different geometries depending on the application of interest [13,14]. The most commonly used configuration is the linear array, in which the sensors are uniformly spaced along a straight line. Another common configuration is a planar array, in which the sensors form a rectangular grid or line on the concentric circles.

Different approaches have been followed for solving the direction of arrival (DOA) estimation problem. One of the oldest ideas in array processing for determining the DOA is beamforming [5,8,81]. The idea behind the beamforming is to align the propagation delays of a signal presumed to be propagating in a given direction so as to reinforce it, while signals propagating from other directions and the noise are not reinforced. Directions which exhibit the largest power corresponds to the DOA estimates. Beamforming methods are computationally efficient and yield effective performance in low resolution applications where the incident source spatial separations are sufficiently larger than the inverse of the array aperture [45]. Using this classical approach, increased bearing estimation accuracy can only be obtained by increasing the aperture of the array. In addition, beamforming measures the energy by purely deterministic method, which is liable to be erroneous because of the random variation of sensor outputs caused by noise. For these reasons, modern spectral analysis algorithms have been considered.

Perhaps the most well-known so called high-resolution array processing algorithm is the maximum likelihood method (MLM) first reported by Capon [10,11]. The derivation of this method does not correspond to the standard approach used in maximum likelihood (ML) estimates. Rather, this estimate is derived by finding the steering vector which yields the minimum beam energy subject to a constraint that the processing gain for each direction-of-look to be unity. Minimizing the resulting beam energy reduces the contributions to this energy from sources and noise not propagating in the direction-of-look. The solution of this constrained optimization problem occurs often in the derivation of adaptive array processing algorithms.

The linear-predictive (LP) spectral estimate commonly used in time series problems is also used in array processing problems [40,48,50]. The Fourier transform of the output of a given sensor evaluated at a given frequency is estimated by a weighted

linear combination of those of the other sensors. The LP method is based on finding the weights which minimize the mean-squared prediction error. Another approach for multiple DOA estimation makes use of vector autoregressive moving average (ARMA) modeling of sensor output and combines a special ARMA parameter estimation method with a nonlinear optimization procedure to estimate the relative time delays [51,56].

A class of spectral estimation procedures based on eigenvector-eigenvalue decomposition of the spatial correlation matrix has been developed recently [4,32,64]. The eigenvector method, also called the signal subspace method, makes use of the algebraic property of the spatial covariance matrix that the eigenvetors corresponding to the largest eigenvalues span the same subspace (the signal subspace) as the source direction vectors. Under the condition that the observation period is long and signal to noise ratio (SNR) is not too low, this approach has previously been shown to have substantially higher resolution in estimating DOA's than the conventional beamformer, Capon's MLM [11], and autoregressive (AR) spectral estimators [15]. As in the case of principal factor analysis, an information criterion such as the one developed in [84] can be used to effectively determine the number of sources, thus avoiding a difficult multiple hypothesis testing approach as was done by Press [57].

Eigenvector methods such as MUSIC [65] and ESPRIT [54] have become popular in applications requiring high resolution capability. However, eigenvector methods are usually based on narrow band assumption of signals. One way of solving the wide band DOA estimation problem is to divide the wide frequency band into non-overlapping narrow bands, and then use narrow band signal subspace processing as was proposed by Wax et al. [86]. Alternatively, Wang et al. [82] have considered an eigenvector method where the estimates are obtained by the eigen-decomposition of a frequency domain combination of modified narrow band covariance matrix

estimates. Instead of treating the wide band problem as a multitude of narrow band emitter problems, Su and Morf [77] and Porat and Friedlander [56] have considered using a multivariate rational model for the sensor outputs. Another approach for the DOA estimation problem is to consider it as a 2-D spectral estimation problem by Halpney et al. [19]. An advantage of this approach is that it is applicable when both narrow band and wide band sources are present simultaneously. Jackson and Chien [28], however, have pointed out the severe asymmetry and bias in the estimated spectra using a 2-D quarter plane AR model for bearing estimation.

Although algorithms based on the signal subspace methods claim high resolution capability, they do not perform well at low signal to noise ratio (SNR), and equivalently, when the number of data snapshots available is small. Rapid target movement may also limit the prospective estimation procedure to working with a single snapshot so that the bearing information, along with range and velocity, may be updated continuously. In a low angle radar tracking environment, the estimation problem is complicated by the fact that the signal returning from the target arrives *via* sea or ground reflection within a beamwidth of the direct path echo. A renewed interest in maximum likelihood (ML) estimation procedure, which is equally applicable to single snapshot cases and coherent signals, explains this part of the story. The derivation of this method correspond to the standard approach used in maximum likelihood (ML) estimation [6,7,24]. The ML estimation technique has not been very popular until recently because of the high computational load involved in the multivariate nonlinear maximization. Recently, Ziskind and Wax [87] have presented a computationally attractive method for computing the ML estimate of narrow band sources.

1.2. Robust Estimation

An important issue in array processing is concerning the structure of the noise model. Previously, it was frequently assumed that the noise process was an independent and identically distributed (IID) Gaussian. This assumption has been widely adopted for underlying noise structures and still is used very often in many different applications since it usually reduces the complexity of the problem from both theoretical and empirical standpoints. The assumption of normality is often based on empirical evidence or justified in theory by application of a suitable central limit theorem. But in practical empirical situations, the observed signals contain undesirable imperfections or noise which is inherent to the system under study or which occur because of measurement errors or isolated phenomena.

In many situations the corrupting noise itself can be considered Gaussian with the result that the observations remain Gaussian but with a more complicated structure. However, measurement errors and isolated errors can cause observed data sets to contain small fraction of unusual data points, which are not consistent with a strictly Gaussian assumption. It may not be hard to spot such potentially troublesome data points in the lower dimension, but it becomes exceedingly difficult with higher dimensions, or with multiparameter problems.

An *outlier* in a set of data is defined as an observation which appears to be inconsistent with the remainder of that set of data. The phrase 'appears to be inconsistent' is crucial. It may be a matter of subjective judgement on the part of the observer whether or not he picks out some observation for scrutiny. The important question is whether or not some observations are genuine members of the main population. The next question is how should one react to the outliers, and what methods can be used to support rejecting them, or adjusting their values, prior to processing

the principal mass of the data. The answer depends on the form of the population, i.e., Gaussian ; techniques will be conditioned by the postulated model for that population [2].

Such data in principle can be modeled as having a distribution which is nearly Gaussian in the central region but with heavier tails. For this reason, minor deviations from the Gaussian noise are often modeled by the mixture model for noise [80]. One particular mixture model of interest is the slippage model with the Gaussian distribution as the dominant distribution. If $u(i)$ is a sequence of random variables obeying such a slippage model, then any $u(i)$ is distributed either as a Gaussian distribution of zero mean and variance σ^2 with probability $1-\varepsilon$, or as an unknown distribution of much higher variance with probability ε. In general $\varepsilon < 0.1$, and the mean of the unknown distribution μ, an unknown constant, is of the order of a multiple of σ. This represents a family of distributions characterized by the mixing parameter ε. For $\varepsilon=0$, it reduces to a Gaussian distribution.

In this report, robustness refers to insensitivity against a small deviation in the underlying Gaussian noise assumption. Furthermore, in evaluating DOA estimation methods, the term *resolution* refers to the ability of an algorithm to reveal the presence of two equal-energy sources which have nearly equal bearings. Most previous techniques which claim high resolution capability were developed and tested under the Gaussian assumption. These methods no longer provide high resolution estimates when the underlying noise distribution deviates even slightly from the assumed Gaussian. For an example, even a small deviation from the assumed Gaussian noise model can create havoc with Gaussian maximum likelihood (ML) estimators since the Gaussian ML estimators are extremely sensitive to outliers. Such methods need not and usually do not possess the robust property when the underlying noise distribution is an outlier contaminated Gaussian, which is a mixture of Gaussian distribution and a

small portion of unknown outliers.

1.3. Robust Direction-of-Arrival Estimation by Correlation Matrix Reconstruction

A new narrow band eigenvector method for robust direction-of-arrival (DOA) estimation is considered. The MUltiple SIgnal Classification (MUSIC) algorithm, one of the eigenvector methods, has been shown to yield results which are asymptotically unbiased and efficient by Barabell et al. [1]. An important feature of the above method is the decomposition of an estimate of the received signal correlation matrix onto orthogonal signal and noise subspaces and the formulation of the DOA estimator in the noise subspace. The DOA estimates are given by the positions of the spectral peaks. Thus, sources are resolved if the estimated spectrum contains maxima at or in the immediate neighborhoods of the true DOA's.

When the exact ensemble spatial correlation matrix is used, MUSIC results in unbiased values for the null spectrum of uncorrelated plane waves at the true DOA's irrespective of the SNR and angular separations of the sources. In this category of applications, the noise is usually assumed to be Gaussian, and it is known that a small deviation in the noise distribution from the assumed Gaussian noise model may introduce significant errors into the eigenstructure of the sample correlation matrix estimate, which in turn deteriorates the quality of the DOA estimates.

The focus of this study is to explore an alternative way for estimating the DOA's using eigenvector method in the presence of outlier contaminated Gaussian noise. A multivariate autoregressive (AR) model with proper order is systematically chosen, and the parameters are estimated using a robust technique. Once all the parameters are estimated, the correlation matrix corresponding to the model can be reconstructed.

The number of signal sources and the corresponding DOA's are then estimated using a conventional eigenvector method such as MUSIC.

Simulation results show that the new scheme performs consistently even when the outlier noise is present whereas the performance of the corresponding nonrobust method deteriorates quickly with a slight change of the noise environment. This is especially significant at a low signal to noise ratio (SNR).

1.4. Decentralized Direction-of-Arrival Estimation

There has been an increasing interest in decentralized arrays of sensors, mainly motivated by military requirements. The general scheme of decentralized array processing is as follows. Each subarray is a unit that receives observations and estimates parameters using only its own observations. Estimating parameters at each subarray site is a totally independent process from the estimation process at other subarray sites. Each subarray site then provides its estimates and other necessary information to the fusion center, where the estimates are combined to form a more reliable estimate than the individual estimates from different subarray sites.

If it were possible to transmit all the subarray observations to the central processing unit with trivial delay, the classical theory and the advantages of using the array processing are applicable. However, because of such considerations as cost, reliability, survivality, communication bandwidth, compartmentalization, sensors on platforms under emission control, or even simply the problem of flooding the fusion center with more information than it can process, there is never total centralization of information in practice [78]. Furthermore, the central processing unit has no means of realizing the malfunctioning subarray sites. But with decentralized processing it is possible for the fusion center to recognize the data from malfunctioning subarray sites

or at least minimize the harmful contribution from those subarray sites.

In this study, a robust decentralized scheme for estimating the directions-of-arrival (DOA) will be considered. At each subarray site, a multivariate autoregressive (AR) model with proper order is systematically chosen, and the parameters are estimated using a robust technique. Once all the parameters are estimated, the correlation matrix corresponding to the model can be found. Each subarray site then estimates the number of signal sources, and the estimate is sent to the fusion center along with the statistics for computing the estimate's relative confidence measure. At the fusion center, the estimates of the number of sources are combined based on their relative confidence measures, then the result is sent back to each of the selected subarray sites for their reliability. Each of the chosen subarray sites then provides the determined number of DOA estimates, which are then combined using a robust combining technique at the fusion center.

The algorithm combines the best features of robust parameter estimation technique and the aforementioned advantages of the decentralized processing. One can still obtain reliable estimates when a few of the subarray sites are malfunctioning in addition to the possible deviation of the noise from the assumed Gaussian model. Furthermore, the communication loads between different subarray sites are completely eliminated, while those between each subarray site and the fusion center are minimized.

1.5. Direction-of-Arrival Estimation using Radon Transform

A robust method for the direction-of-arrival (DOA) estimation when there are multiple sources, each of which is either narrow band or wide band, is considered in this study. One importance of this method is that it does not require any information

about the number of received source signals, structure and frequency of the signals, and the correlation structure of sensor noise. The technique is capable of handling narrow band and wide band sources simultaneously at low SNR's, and performs equally well in the presence of colored noise with unknown correlation structure. The proposed DOA estimation scheme which utilizes a 2-D spectral estimation is also useful in outlier contaminated Gaussian noise.

Recently, a new approach of 2-D spectral estimation utilizing 1-D autoregressive (AR) models in the Radon space was investigated by Srinivasa et. al [71,75]. The 2-D PSD is estimated from a finite set of observations of a 2-D stationary random field (SRF) using the Radon transform. In particular, the 2-D PSD estimation problem is converted into a set of 1-D independent problems using the modified central slice theorem for SRF introduced by Jain and Ansari [29].

The 2-D array data is transformed into a set of 1-D sequences, or projections, by the Radon transform. Then an estimate of the 2-D spectrum is obtained on a polar raster by modeling the projection with a 1-D autoregressive (AR) model, where the parameters are estimated by a robust technique, i.e., Huber's M-estimators [27]. The DOA estimates are obtained by locating the peaks in the resulting 2-D spectrum.

Another important aspect of the work presented here is the use of robust 1-D autoregressive (AR) parameter estimation method in the Radon space to obtain a robust 2-D PSD estimate. This reduces the number of parameters to be estimated simultaneously, thus allowing the robust 2-D PSD estimation feasible. Though the DOA estimation method presented in this study is somewhat related to the traditional beamforming, it has a much better resolving capability as we use the spectral density, which is in turn estimated by using a model, to measure the average power. Rough analysis indicates that the resolution of this method is much higher, nearly double, than that of the traditional beamforming method. This algorithm is highly amenable

for parallel processing as well. Furthermore, any particular range of directions of interest can be probed for detecting the presence or absence of sources.

1.6. Robust Maximum Likelihood Direction-of-Arrival Estimation

It is well known that even a small deviation in the noise from the assumed Gaussian can creat havoc with Gaussian maximum likelihood (ML) estimates. Therefore, a robust technique is considered for maximum likelihood (ML) narrow band direction-of-arrival (DOA) estimation problem against outliers and distributional uncertainties. The algorithm employs a robustified Gaussian ML estimator which performs almost as well as a Gaussian ML estimator in pure Gaussian noise, and much better in the presence of outliers. The algorithm is also capable of handling coherent signals as well as single snapshot cases.

The DOA's are estimated by a robust technique based on the so called M-estimators, a generalization of classical ML estimator by Huber [27]. Performances of the estimator in both the Gaussian and outlier contaminated Gaussian noise are evaluated using the Cramer Rao Lower Bound (CRLB) and variance derived from the Influence Function (IF), followed by resolution analysis regarding the ability of the algorithm in resolving two closely spaced sources with equal power.

1.7. Generalization of Eigenspace Methods for Bearing Estimation using Maximum Likelihood

A maximum likelihood (ML) direction-of-arrival (DOA) estimation problem is considered where the source signals are treated as sample functions of random processes instead of unknown deterministic sequences as assumed in most of the

previous approaches. The study reveals a special relationship between this ML DOA estimation scheme and eigenvector methods for estimating DOA's. In particular, the focus is on interconnecting the notions of DOA estimation using eigenvector methods to a more quantitative Gaussian ML approach, i.e., choosing the DOA estimates to be in the directions of the eigenvectors which corresponds to the largest eigenvalues in the signal subspace.

When the number of sources is one, it can be shown that maximizing the likelihood function with respect to the DOA angle is equivalent to choosing the steering vector to be in the direction of the eigenvector which corresponds to the largest eigenvalue in the signal subspace. The equivalence, however, does not hold exactly for multiple sources. The main differences between the eigenvector methods and this ML method for estimating DOA's can be clearly seen for two source cases.

1.8. Layout of the Report

Various aspects of the robust direction-of-arrival (DOA) estimation have been investigated throughout the report. An important aim of this study is to develop robust DOA estimation techniques suitable in many different environments and applications. In evaluating DOA estimation methods, the term *resolution* refers to the ability of an algorithm to reveal the presence of two equal-energy sources which have nearly equal bearings. The robust DOA estimation schemes developed here perform much better than the conventional high resolution methods, which were developed and tested under the Gaussian noise assumption, in the presence of outliers. In the presence of pure Gaussian noise, the robust DOA estimation methods still perform almost as well as the Gaussian based methods.

The organization of the report is as follows. In chapter 2, a robust narrow band DOA estimation technique, which utilizes an eigenvector method and robust reconstruction of correlation matrix by a time series modeling of the array data, is presented . Chapter 3 is an extension of the robust technique developed in chapter 2. The chapter presents a decentralized DOA estimation scheme that can provide much more reliable DOA estimates than those from a similar centralized scheme when a few of the subarray sites are malfunctioning. Chapter 4 then presents a robust wide band DOA estimation method which utilizes a 2-D spectrum estimation approach using Radon Transform. The technique is capable of handling the narrow band and the wide band sources simultaneously, and still performs well in situations of low SNR, and colored noise with unknown correlations. In Chapter 5, a robust maximum likelihood (ML) DOA estimation algorithm, which employs a robustified Gaussian ML estimator, is presented. The technique is equally capable of handling coherent signals as well as the single snapshot cases. Chapter 6 interconnects the notions of DOA estimation using eigenvector methods to a more quantitative Gaussian ML approach, followed by chapter 7 which concludes the report along with topics of the future research.

CHAPTER 2
ROBUST DIRECTION-OF-ARRIVAL ESTIMATION BY CORRELATION MATRIX RECONSTRUCTION

2.1. Introduction

The problem of estimating the direction-of-arrival (DOA) of radiating sources from measurements provided by a passive array of sensors is frequently encountered in radar, sonar, radio astronomy and seismology. In most cases the number of incident plane waves and their DOA's are to be estimated from incident source induced sensor signals. In the case of applications which require high resolution capability and the signal to noise ratio (SNR) is not too low, the eigenspace methods were generally known to perform better than the conventional beam forming, autoregressive (AR) methods, etc.

The MUltiple SIgnal Classification (MUSIC) algorithm, one such method, is shown to yield results which are asymptotically unbiased and efficient by Barabell et al. [1]. An important feature of the above method is the decomposition of an estimate of the received signal correlation matrix onto orthogonal signal and noise subspaces and the formulation of the DOA estimator in the noise subspace. The DOA estimates are given by the positions of the spectral peaks. Thus, sources are "resolved" if the

estimated spectrum contains maxima at or in the immediate neighborhoods of the true DOA's.

When the exact ensemble spatial correlation matrix is used, MUSIC results in unbiased values for the null spectrum of uncorrelated plane waves at the true DOA's irrespective of the SNR and angular separations of the sources. In this category of applications, the noise is usually assumed to be Gaussian, and it is known that a small deviation in the noise distribution from the assumed Gaussian noise model may introduce significant errors into the eigenstructure of the correlation matrix estimate, which in turn deteriorates the quality of the DOA estimates. For this reason, we choose to use the so called "outlier contaminated Gaussian noise model" as it appears to be more realistic than a simple Gaussian model.

This chapter explores an alternative way of estimating the DOA's in the presence of outlier contaminated Gaussian noise. The following scheme is proposed. A multivariate autoregressive (AR) model and its proper order is systematically chosen, and the parameters are estimated using a robust technique from the available array output snapshots, where robustness refers to insensitivity against a small deviation in the underlying noise assumption. Once all the parameters are estimated, the correlation matrix corresponding to the model can be found. The standard MUSIC algorithm is then utilized to estimate the number of sources, and the corresponding DOA's.

The organization of the chapter is as follows. In section 2.2, the basic signal model and the formulation of the problem is presented. Section 2.3 introduces the details of the new scheme. Section 2.4 then presents some of the simulations carried out to compare the performance of the proposed algorithm with that of a similar non-robust algorithm, i.e., MUSIC, followed by the concluding remarks in section 2.5.

2.2. Array Model for Direction-of-Arrival Estimation

The basic problem under consideration is that of the estimation of parameters of finite dimensional signal processes given measurements from a sensor array. In particular, the discussion will be in terms of the problem of multiple incoherent source directions-of-arrival (DOA) estimation from a equispaced linear array. Even though the discussion and results presented here deal only with the single dimensional parameter space, i.e., azimuth only direction finding of far-field point sources, the technique can be easily generalized to higher dimensional parameter spaces. A DOA estimation problem is classified as narrow band if signal bandwidth is small compared to the inverse of the transit time of a wavefront across the array. For simplicity we assume that the incoming signals are narrow band even though the technique can be extended to the wide band cases.

Consider a planar array composed of L identical sensors translationally separated by a known constant displacement δ. Assume that there are $d<L$ narrow band stationary zero mean sources located sufficiently far from the array such that in homogeneous isotropic transmission media, the wavefronts impinging on the array are planar. Additive noise is present at each sensor of the array and is assumed to be a stationary zero mean complex "outlier contaminated Gaussian," which is uncorrelated from sensor to sensor with equal variances.

Frequently, the speckle type noise in signal processing and other patchy disturbances are modeled by the mixture model for noise [80]. One particular mixture model of interest is the slippage model with a Gaussian distribution as the parent distribution. If $w(t)\,;\, t=1,...,N$ is a sequence of random variables obeying such a slippage model, then any $w(t)$ is distributed either as $N(0,\sigma^2)$ with probability $(1-\varepsilon)$ or as an unknown distribution $Q(\mu, a\sigma^2)$ with probability ε, where μ and $a\sigma^2$ are the

mean and the variance of the unknown distribution Q. In general, $a>1$ and $\varepsilon<<1$ and μ, an unknown constant, is of the order of a multiple of σ. The noise distribution in this example can be expressed as

$$p(w) \sim (1-\varepsilon)N(0,\sigma^2)+\varepsilon Q(\mu, a\sigma^2) \tag{2.2.1}$$

and represents a family of distributions characterized by the mixing parameter ε. For $\varepsilon=0$, (2.2.1) reduces to a Gaussian distribution.

The received signal at the lth sensor of the array is denoted as $x_l(t)$, $l=1,...,L$, and given by

$$x_l(t)=\sum_{k=1}^{d} s_k(t)exp(j\,2\pi\,\delta\,l\,\sin\theta_k\,/\,\lambda)+w_l(t) \tag{2.2.2}$$

where $s_k(t)$ is the known complex sinusoidal signal associated with the kth source, λ is the known radar wavelength, δ is the known uniform spacing between the array sensors, and $w_l(t)$ is the additive noise at the lth sensor of the array which is the outlier contaminated Gaussian explained above. Our objective here is to estimate d, the unknown number of signal sources, and θ_k, $k=1,...,d$, the unknown DOA's with respect to the vertical axis stretched above sensor number one as shown in Figure 2.1.

The L-variate signal vector received by the array is denoted by

$$X(t)=\begin{bmatrix} x_1(t) \\ . \\ . \\ x_l(t) \end{bmatrix} = \begin{bmatrix} . & .. & . \\ . & .. & . \\ a(\theta_1) & .. & a(\theta_d) \\ . & .. & . \\ . & .. & . \end{bmatrix} \begin{bmatrix} s_1(t) \\ . \\ . \\ s_d(t) \end{bmatrix} + \begin{bmatrix} w_1(t) \\ . \\ . \\ w_L(t) \end{bmatrix} \tag{2.2.3}$$

where

$$a(\theta_k)=col.\left[1,\exp(j\,2\pi\,\delta\,\sin\theta_k\,/\,\lambda),...,exp(j\,2\pi\,(L-1)\,\delta\,\sin\theta_k\,/\,\lambda)\right],$$

$$k=1,...,d. \tag{2.2.4}$$

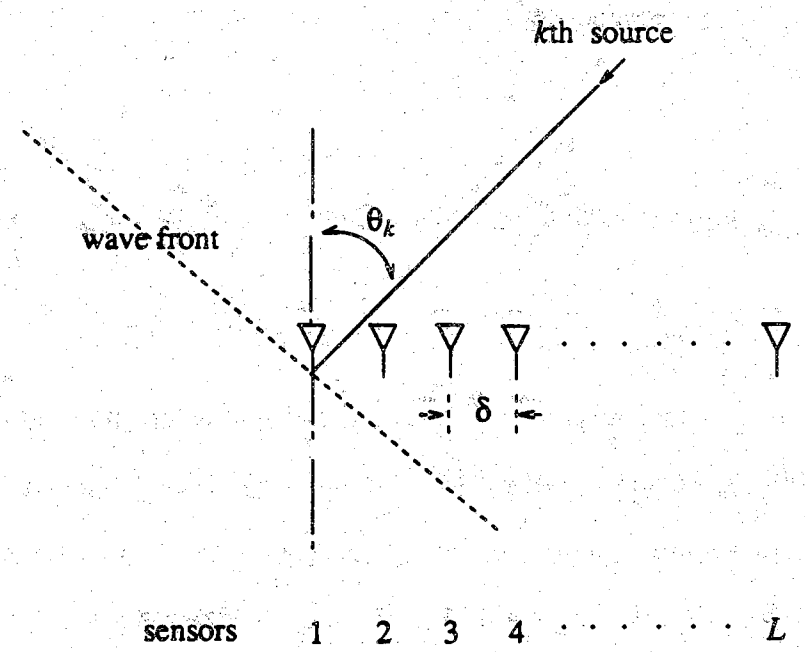


Figure 2.1. A Simple Sketch of a Subarray Site

In other words the waveforms received at the L array elements are linear combinations of d incident wavefronts and noise. The class of eigenspace based algorithms such as MUSIC can achieve high resolution performance only if the quality of the estimated correlation matrix is good. If $X(t_i)$; i=1,...,N, are N independent observations from a complex multivariate normal distribution, then the maximum likelihood estimate of the required correlation matrix is given by

$$\hat{R}_{xx} = \frac{1}{N}\sum_{i=1}^{N} X(t_i)X^H(t_i) \qquad (2.2.5)$$

where H denotes Hermitian transpose.

Unfortunately, $X(t_i)$; $i=1,...,N$ are the N snapshot vectors which may not always be independent from each other, and the Gaussian assumption no longer holds in the presence of even a few outliers in the presumably Gaussian sensor noise. If we insist on using the estimate $\hat{R}_{xx}$ as in (2.2.5), the corresponding DOA estimates will no longer be reliable since the performance of an estimation scheme is critically influenced by the validity of its underlying noise assumption. The common Gaussian assumption, which usually leads to computational and analytical simplicity, is often easily violated, in which case the performance of the estimators based on the assumption may deteriorate seriously. The focus of this chapter is on the remedies for this kind of performance degradation.

2.3. Robust Estimation of the Correlation Matrix

Every model may have a few specific purposes, and the model needs only have just enough significant detail to satisfy these purposes. Thus the basic premise in model building is that complicated systems do not always need complicated models. Our scheme utilizes a multivariate autoregressive (AR) model for computing the robust correlation matrix estimates. In particular, the robust parameter estimates of the multivariate AR model are the M-estimates, a generalization of the Maximum Likelihood estimates by Huber [27]. When all the parameters of the chosen multivariate AR model are estimated from the available set of data, many of its vital statistics such as its correlation matrix can be retrieved from the model.

Three different canonical representations of system equations, which are useful for parameter estimation under different conditions, are discussed by Kashyap et al. [37]. One of the principal reasons for the high degree of computational complexity in the parameter estimation in L-variate AR model is that all the unknowns in the system are estimated simultaneously. One method of reducing the computational complexity is to consider the possibility of separately estimating the unknowns in each of the L individual difference equations, i.e., consider the possibility of replacing one huge estimation problem with L relatively simple estimation problems. For each of the L univariate parameter estimation problems, we consider obtaining the parameter estimates that are robust against outliers and distributional uncertainties.

2.3.1. The Parametric Model

Many deterministic and stochastic discrete time processes encountered in practice are well approximated by a rational transfer function model. The most general linear model is termed an autoregressive moving average (ARMA) model, and the interest in this model stems from its relationship to linear filters with rational transfer functions. It is assumed that the observation, sinusoids plus white noise, obeys a stationary stochastic process whose spectral density has peaks at the relevant frequencies and the series can be represented by a stationary ARMA model.

Typically, if a process obeys an ARMA model, it can be equivalently represented as an infinite autoregressive (AR) process [35]. The predictive ability of the truncated model could be made approximately equal to that of the original ARMA process by choosing a sufficiently large number of terms in the truncated AR model. Since accurate estimation of parameters in a system involving moving average terms is considerably more difficult than the estimation problem in a system without

moving average terms, an AR model of higher order is substituted for the ARMA model.

Suppose that the output samples, $X(t)$, $t=1,\cdots,N$, obey an L-variate AR model with order p, i.e.,

$$X(t) = A_1 X(t-1) + \cdots + A_p X(t-p) + W(t), \tag{2.3.1}$$

where $X(t)$ denotes the L-variate output vector of the array at time t. $W(t)$ denotes the L-variate noise vector whose elements correspond to the noise at each sensor of the array at time t. It is also assumed that the elements of $W(t)$ are uncorrelated from each other and in time t, with zero mean and equal variance. $A_k, k=1,...,p$, is the L by L coefficient matrix of the kth order term. Our immediate objective at this point is to estimate all the components of the A_k's, and the variances of the individual components of $W(t)$. Let us denote

$$X(t) = col.(x_1(t), \ldots, x_L(t))$$

$$A_k = \begin{bmatrix} a_{11}^{(k)} \cdots a_{1L}^{(k)} \\ \cdots \\ \cdots \\ a_{L1}^{(k)} \cdots a_{LL}^{(k)} \end{bmatrix}$$

$$W(t) = col.(w_1(t), \ldots, w_L(t)). \tag{2.3.2}$$

Note that (2.3.1) can be broken into L univariate models.

$$x_j(t) = \Lambda_j^T Z(t-1) + w_j(t), \quad j=1,...,L \tag{2.3.3}$$

where

$$\Lambda_j = col.(a_{j1}^{(1)},..,a_{jL}^{(1)},.....,a_{j1}^{(p)},..,a_{jL}^{(p)}), \quad j=1,...,L. \tag{2.3.4}$$

$$Z(t-1)=col.\,(x_1(t-1),..,x_L(t-1),.....,x_1(t-p),..,x_L(t-p)). \qquad (2.3.5)$$

For each of the L univariate time series model in (2.3.3), we apply a robust parameter estimation technique analogous to the one utilized by Bhargava et al. [3] for estimating parameters of real ARMA model. Since the parameters to be estimated are complex quantities, proper modifications have to be followed. The parameter estimation algorithm presented here involves substantial modifications in the cost function and the gradient finding procedure associated with complex parameters.

2.3.2. Model Order Determination

Robust estimation methods are computationally feasible only when the number of quantities to be estimated is small compared to the number of available observations. Even when the number of parameters to be estimated is small, the minimization of robustified criterion functions often leads to local minima. The situation is very critical if the number of parameters is large. It is also well known that the larger the number of unknown parameters to be estimated for the same number of measurements, the lower is the accuracy of the estimates, the so called principle of parsimony. The choice of orders in the L-variate AR model was done by using the order selection criterion due to Kashyap [34].

2.3.3. Complex Parameter Estimates

In [3], a short review of a robust approach relevant to our problem was presented with an algorithm for implementing the Huber's procedure in parameter estimation. The convergence issues involved in the associated numerical optimization problem

has been also addressed. The following is a modification of the above real parameter estimation algorithm into the complex parameter estimation case.

For each of the L univariate models as shown in (2.3.3), the following robust estimation scheme is proposed. Let the estimate of the true parameter vector, Λ_0, be given by

$$\hat{\Lambda}_0(N;J) = \textit{argument} \min_{\Lambda} J(N) \tag{2.3.6}$$

where

$$J(N) = \sum_{t=p+1}^{N} H(w(t,\Lambda)) . \tag{2.3.7}$$

The function $H(x)$ is given by

$$H(x) = \begin{cases} |x|^2/2 & , \text{ if } |x| \le c \\ c|x| - c^2/2 & , \text{ if } |x| > c, \end{cases} \tag{2.3.8}$$

and $w(t,\Lambda)$ is the residual defined by

$$w(t,\Lambda) = x(t) - \Lambda^T Z(t-1), \tag{2.3.9}$$

which is consistent with (2.3.3). $w(t,\Lambda)$ is also understood as an estimate of $w(t)$ based on the observation set $Z(k)$ up to time $k=t-1$ as if Λ is the correct value, Λ_0.

The choice of c is important. Since the approach of Huber [27] is applicable to this case, the constant c has the following expression

$$c = c_0\sigma \tag{2.3.10}$$

where c_0, which depends on ε, the fraction of contamination, is given by

$$2\Phi(c_0) - 1 + 2\phi(c_0)/c_0 = 1/(1-\varepsilon), \tag{2.3.11}$$

and σ^2 is the variance of the dominant Gaussian density. $\Phi(c_0)$ is the standard cumulative Gaussian distribution with zero mean and unit variance, and $\phi(c_0)$ is the

corresponding Gaussian density. Usually c_0 is chosen to lie between 1 and 2, which corresponds to the ε-interval [0.0083, 0.1428] by (2.3.11). In practice, however, both σ and the exact value of ε are unknown. Huber replaced σ by a factor called scale factor, and discusses in detail the choice of this scaling factor for the estimation of the location parameter which depends on scale. However, in the above case the parameter vector Λ is independent of the scale due to the nature of the AR model under consideration, i.e., the scaling affects both sides of (2.3.3) to the same degree, and so these methods are not relevant. Instead, we choose c as follows:

$$c = c_1 \left[\frac{1}{N-p} \sum_{t=p+1}^{N} |w(t,\Lambda)|^2 \right]^{1/2} \tag{2.3.12}$$

where c_1 is a constant between 1 and 2. Note that c changes from iteration to iteration.

2.3.4. Computation Procedure

For computational clarity let us denote

$$\begin{aligned} xr(t) &= Re[x(t)] \\ xi(t) &= Im[x(t)] \\ \Lambda r &= Re[\Lambda] \\ \Lambda i &= Im[\Lambda] \\ Zr(t-1) &= Re[Z(t-1)] \\ Zi(t-1) &= Im[Z(t-1)]. \end{aligned} \tag{2.3.13}$$

Then we may write

$$w(t,\Lambda)w(t,\Lambda)^* = T_1 + T_2, \tag{2.3.14}$$

where

$$T_1 = [\, xr(t) - \Lambda r^T Zr(t-1) + \Lambda i^T Zi(t-1) \,]^2$$

$$T_2 = [\, xi(t) - \Lambda r^T Zi(t-1) + \Lambda i^T Zr(t-1) \,]^2. \tag{2.3.15}$$

Then (2.3.8) becomes

$$H(w(t,\Lambda)) = \begin{cases} (T_1 + T_2)/2 & , \text{ if } |w(t,\Lambda)| \le c \\ c\sqrt{T_1 + T_2} - c^2/2 & , \text{ if } |w(t,\Lambda)| > c. \end{cases} \tag{2.3.16}$$

The usual approach of finding the gradient of $H(w(\Lambda))$ with respect to Λ must be used with caution since the vector Λ is complex. One can evaluate the gradients with respect to the Λ and its conjugate as independent variables, or the real and imaginary parts as independent variables [31]. The gradient vector and the Hessian matrix of the $H(w(t,\Lambda))$ is given by

$$\nabla_\Lambda H(w(t,\Lambda)) = \nabla_{\Lambda r} H(w(t,\Lambda)) + j\nabla_{\Lambda i} H(w(t,\Lambda)) \tag{2.3.17}$$

and

$$\nabla^2_{\Lambda\Lambda} H(w(t,\Lambda)) = \nabla^2_{\Lambda r \Lambda r} H(w(t,\Lambda)) + j\nabla^2_{\Lambda i \Lambda i} H(w(t,\Lambda)), \tag{2.3.18}$$

where the real and imaginary parts of Λ are assumed to be independent.

A Newton-Raphson based algorithm is utilized for the minimization of $J(N)$ in (2.3.7) since its convergence properties are well established [69,70]. The basic step for this method is given by

$$\Lambda^{(i+1)} = \Lambda^{(i)} - a[\sum_{t=p+1}^{N} \nabla^2_{\Lambda\Lambda} H(w(t,\Lambda))]^{-1} [\sum_{t=p+1}^{N} \nabla_\Lambda H(w(t,\Lambda))] \tag{2.3.19}$$

where a is the step size parameter for iteration.

2.3.5. Correlation Matrix Computation

For computational simplicity an equivalent state variable model is utilized in the computation of the correlation matrix. Given the following pth order L-variate AR model as was also shown in (2.3.1),

$$X_m(t) = A_1 X_m(t-1) + \cdots + A_p X_m(t-p) + W_m(t) \tag{2.3.20}$$

where A_k's are the complex coefficient matrices estimated in the previous robust scheme, an equivalent state variable model is found ; i.e.,

$$Y(t) = AY(t-1) + BW(t)$$

$$Y(t) = CX(t). \tag{2.3.21}$$

Given an equivalent state variable model, one may proceed to find $R_Y(0)$, the correlation matrix of $Y(t)$, knowing that one can retrieve all the elements of $R_X(0)$, the correlation matrix of $X(t)$, from $R_Y(0)$.

From (2.3.21), $R_Y(0)$ may be written as

$$R_Y(0) = E[Y(t)Y(t)^H]$$

$$= A\, R_Y(0)\, A^H + B\, \Gamma\, B^H \tag{2.3.22}$$

where $\Gamma = E[W(t)W(t)^H]$ and H denotes Hermitian transpose.

The problem now is to solve for $R_Y(0)$ given all the other terms in (2.3.22). In [33], it was shown that the components of $R_Y(0)$ can be easily obtained by solving

$$(I - A \otimes A^*)\, R = D \tag{2.3.23}$$

where $\otimes$ denotes the kronecker product, and * denote the complex conjugate, R and D are column vectors formed from the rows of $R_Y(0)$ and E respectively; i.e.,

$$R = col.(r_{11},..,r_{1L},....,r_{L1},..,r_{LL}) \tag{2.3.24}$$

$$D = col.(e_{11},..,e_{1L},....,e_{L1},..,e_{LL}) \tag{2.3.25}$$

where

$$E = B\Lambda B^H. \tag{2.3.26}$$

Solving for the vector R from the linear equation (2.3.23) yields all the component of $R_Y(0)$, thus $R_X(0)$ can also be obtained.

2.4. Numerical Simulation

The objective of this simulation study is to investigate and confirm the effectiveness of the new robust DOA estimation technique in many different noise environments. In the simulation, it is assumed that there are eight sensors in the array with identical spacings between them. There are two signal sources with θ_1=0.7854 and θ_2=1.0472, denoting the first and second true DOA's in radians, with respect to the vertical axis stretched above sensor number one in Figure 2.1. The signal sources are chosen as

$$s_1(t) = exp(0.4\pi t) \quad \text{and} \quad s_2(t) = exp(0.8\pi t) \tag{2.4.1}$$

For simplicity, it is also assumed that δ, the spacing between sensors, is exactly one half of λ, the signal wavelength, then (2.4.1) becomes

$$x_l(t)=exp[j\pi(0.4t+l\ \sin\theta_1)] + exp[j\pi(0.8t+l\ \sin\theta_2)] + w_l(t). \tag{2.4.2}$$

from which the data in the simulation is generated.

The order p of the 8-variate complex AR model, i.e. 8 sensor array, is determined using the following decision statistic for multivariate AR model.

$$KIC(p) = N\sum_{j=1}^{8} \ln \rho_j^{(i)} + n_p ln(N/2\pi) \qquad (2.4.3)$$

where N is the number of available 8-variate data, ρ_j is the jth diagonal component of the residual covariance matrix of the fitted AR(p) model, and n_p is the total number of parameters to be estimated for the pth order model. The order, p, is chosen to minimize $KIC(p)$. Details of the decision statistics can be found in [34,35]. With N=100, the decision criterion (2.4.3) is minimized when p=1 or p=2 in most cases, depending on the quality of the additive noise.

At the given array, the complex parameters of the 8-variate AR model are estimated using the technique developed in section 2.3, and the correlation matrix which corresponds to the 8-variate complex AR model is computed using the state variable model method. The MDL criterion by Wax et al. [84] is then utilized to find the estimate of the number of signal sources and the directions-of-arrival estimates at each subarray site.

Table 2.1 shows the performance comparison of the nonrobust and the robust approach for each of the ten experimental runs when the SNR is 13 dB and there exists one percent outlier Gaussian noise, which has five times the variance of the parent Gaussian. For the nonrobust method mentioned above, the correlation matrix estimate is provided by (2.2.5), in which the additive noise is assumed to be a pure Gaussian. * denotes the case where the MUSIC spectrum does not exhibit the any corresponding spectral peaks.

Table 2.2 shows the RMSE's of the available estimates taken from ten independent experimental runs in the presence of one percent outliers with five times the variance of dominant Gaussian noise for different values of signal to noise ratios (SNR). Figure 2.2 shows the average RMSE (average of the two RMSE's which corresponds to the two DOA estimates) *vs.* SNR plot in the pure Gaussian noise for the MUSIC

and for the robust estimates taken from ten independent experiments, while Figure 2.3 shows the similar plot taken in the presence of outliers mentioned above.

The robust method guarantees consistent performance even when the outlier noise is present whereas the performance of the nonrobust method deteriorates very quickly with a slight change of the noise environment. In the following list of tables, d denotes the estimate of the number of source signals, ϵ denotes the percent probability of the outlier noise, and a denotes the ratio of the outlier noise variance vs. the dominant Gaussian noise variance. $\hat{\theta_1}$ and $\hat{\theta_2}$ denote the estimates of the two directions-of-arrival.

Table 2.1. Comparison of nonrobust and robust estimates for ten independent runs when ε=0.01, a=5, and the SNR is equal to 13dB. * denotes the case where the MUSIC spectrum does not exhibit the corresponding spectral peaks. The true DOA's are θ_1=0.7854 and θ_2=1.0472 in radians.

Nonrobust and Robust Results for Each Experimental Run						
	nonrobust			robust		
run no.	d	$\hat{\theta}_1$	$\hat{\theta}_2$	d	$\hat{\theta}_1$	$\hat{\theta}_2$
1	4	0.8388	*	3	0.8242	1.0194
2	2	0.7901	1.0430	2	0.7901	1.0446
3	5	0.7791	*	5	0.8137	0.9959
4	5	0.8624	*	4	0.8184	1.0414
5	3	0.9912	*	4	0.7885	1.0556
6	2	0.7807	1.0414	2	0.7791	1.0430
7	2	0.7854	1.0524	2	0.7885	1.0524
8	3	0.9299	*	4	0.8011	1.0147
9	3	0.7587	1.0509	6	0.7901	1.0540
10	3	0.9739	*	4	0.7854	1.0446

Table 2.2. Averages of the DOA estimates for ten independent runs, when ε=0.01 and a=5 for many different SNR's. Shown in the associated parentheses are the RMSE values of the DOA Estimates taken from the the ten independent runs. The true DOA's are θ_1=0.7854 and θ_2=1.0472 in radians.

Average of the Estimates (RMSE)				
	nonrobust		robust	
SNR(dB)	$\hat{\theta_1}$	$\hat{\theta_2}$	$\hat{\theta_1}$	$\hat{\theta_2}$
13.0	0.7854 (0.0038)	1.0234 (0.0445)	0.7866 (0.0037)	1.0419 (0.0032)
8.5	0.7866 (0.0026)	1.0226 (0.0481)	0.7888 (0.0026)	1.0451 (0.0082)
6.7	0.7825 (0.0084)	1.0051 (0.0688)	0.7876 (0.0047)	1.0506 (0.0049)
5.5	0.7838 (0.0092)	1.0215 (0.0828)	0.7901 (0.0065)	1.0454 (0.0058)
4.5	0.7882 (0.0123)	1.0731 (0.1547)	0.7873 (0.0126)	1.0484 (0.0118)
1.5	*	*	0.7920 (0.0101)	1.0498 (0.0357)
-1.5	*	*	0.7580 (0.0173)	0.9981 (0.0451)

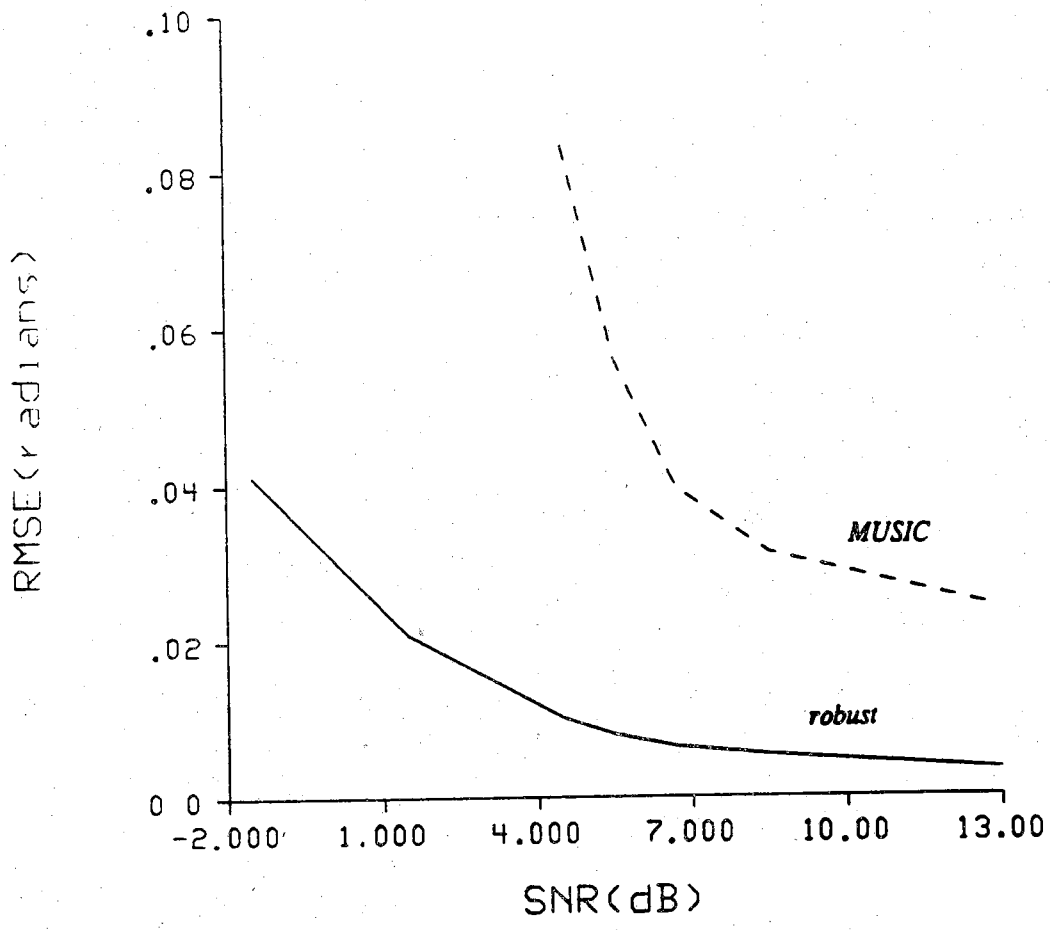


Figure 2.3. *Contaminated Gaussian Noise Case*: Average RMSE vs. SNR for the MUSIC DOA estimates (dotted line) and for the robust DOA estimates (solid line). The contamination is caused by replacing one percent (ϵ=0.01) of the Gaussian data with outliers which has five times the variance (a=5) of the parent Gaussian distribution. The true DOA's are θ_1=0.7854 and θ_2=1.0472 in radians.

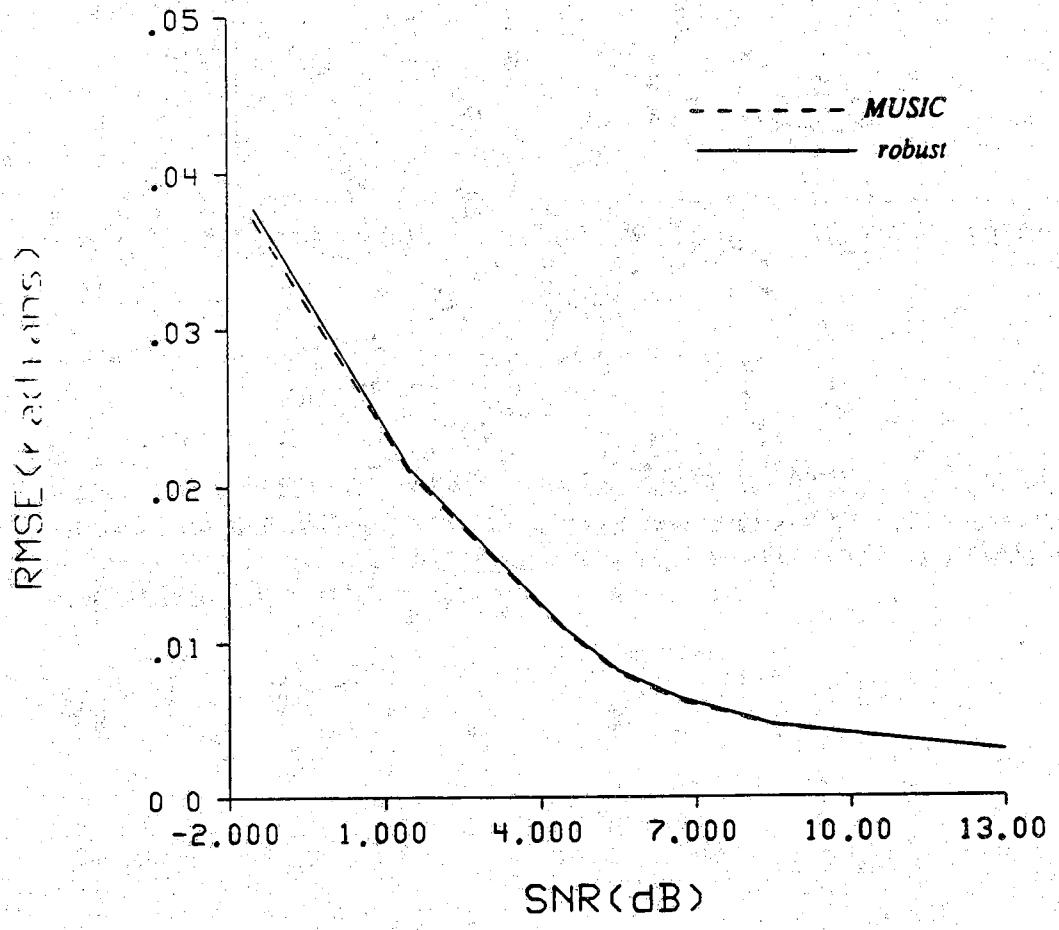


Figure 2.2. *Gaussian Noise Case*: Average RMSE *vs.* SNR plot of the DOA estimates for the MUSIC DOA estimates (dotted line) and for the robust DOA estimates (solid line) from ten independent experiments. The true DOA's are θ_1=0.7854 and θ_2=1.0472 in radians.

2.5. Conclusion

A new robust narrow band technique has been developed for estimating the number of signal sources and their directions-of-arrival. The scheme utilizes an eigenvector method and the correlation matrix estimate reconstructed from robust time series modeling of the array data. The robust scheme provides estimates that are robust against outliers and distributional uncertainties in the noise environment. Simulation results also confirm that the robust scheme performs almost as well as the nonrobust scheme in the pure Gaussian noise, and much better in the presence of outliers, where the nonrobust method often completely fails to provide any estimates.

CHAPTER 3

DECENTRALIZED DIRECTION-OF-ARRIVAL ESTIMATION

3.1. Introduction

There is also a concern that in many cases the centralized scheme is unattractive, such as in the case of many subarrays at geographically dispersed sites [83]. For an example, the central processing unit has no means of locating the malfunctioning subarray sites. But with decentralized processing, it is possible for the fusion center to recognize data from malfunctioning subarray sites or at least minimize the harmful contribution from such subarray sites.

The general scheme of decentralized processing is as follows. Each subarray is a unit that receives the observation and estimates a set of parameters using only its own observations. Estimating parameters at each subarray site is a totally independent process from those of other subarray sites. Each subarray site then sends its own set of estimates to the fusion center, where the sets of estimates from different subarray sites are combined to form a more reliable set of estimates than each individual set of estimates before the combining.

The following decentralized scheme is proposed. At each subarray site, a multivariate autoregressive (AR) model with proper order is systematically chosen, and

the parameters are estimated using a robust technique. Once all the parameters are estimated, the correlation matrix corresponding to the model can be found. Each subarray site then estimates the number of signal sources, and the estimate is sent to the fusion center along with the statistics for computing the estimate's relative confidence measure. At the fusion center, the estimates of the number of sources are combined using their confidence measures, then the result is sent back to each of the selected subarray sites. Each of the selected subarray sites then provides the same number of DOA estimates, which are then combined using a robust combining technique at the fusion center.

This scheme combines the best features of robust estimation technique and the reliability of decentralized processing. For example, one may still obtain reliable estimates when a few of the subarray sites are malfunctioning in addition to the possible deviation of the noise from the assumed Gaussian model. Furthermore, one can eliminate the communication loads between subarray sites, and minimize those between each subarray site and the fusion center.

The organization of the chapter is as follows. Section 3.2 introduces the new robust decentralized scheme for the estimation of directions-of-arrivals, and Section 3.3 through 3.4 introduce the details of the new scheme. Section 3.5 then presents some of the simulations carried out to compare the performance of the proposed algorithm with that of a similar nonrobust combining algorithm, followed by concluding remarks in section 3.6.

3.2. The Robust Decentralized Scheme

The decentralized scheme at the fusion center involves important integrating steps for the estimates of the number of source signals and for the DOA estimates

from different subarray sites. Wax et al. [83] suggested a method of integrating the estimates of the number of source signals, which takes advantage of the *apriori* information that all the subarrays receive the same number of source signals. However, there is little chance of avoiding catastrophic error on the estimate of the number of source signals with this method if any of the subarray sites are malfunctioning. Hence, we propose the following robust combining scheme which takes advantage of the *apriori* knowledge that the additive noise at each sensor of a given subarray has the same variance.

At each subarray site, the estimate of the number of sources is first obtained from the aforementioned correlation matrix estimate using the MDL criterion by Wax et al. [84], a decision criterion for determining the multiplicity of the smallest eigenvalues of a given correlation matrix. The number of signals is determined as the value for which the MDL criterion is minimized. The estimate from each of the subarray sites is sent to the fusion center along with the statistics for computing the reliability measure of the estimate. The robust estimate of the number of sources is determined with the aid of the reliability measure at the fusion center. The result is sent to back to each of those subarray sites whose initial estimate of the number of sources is equal or very close to the fusion estimate. Using the MUSIC algorithm by Schmidt [64], the selected subarray sites then computes the DOA estimates to be combined at the fusion center by a robust technique.

3.3. Integrating the Estimates of the Number of Signal Sources

Let $\hat{d}_i$ be the estimate of the number of source signals computed at the ith subarray site and d^* be the true value of the number of sources. Then we may write $\hat{d}_i$ as

$$\hat{d}_i = d^* + n_i \quad , \; i=1,...,M \tag{3.3.1}$$

where n_i is considered to be an integer disturbance term to $\hat{d}_i$, and M is the total number of subarray sites. Our objective at this point is to find a robust estimate of d^*. Even though there exists methods for finding such estimates [38], none of them seem to provide the desired performances for this case since the integer disturbance term n_i doesn't necessarily have zero mean; i.e., there may be a tendency of overestimation of the number of source signals. What we need is some type of reliability measure associated with each subarray site on the estimate of the number of sources. Such a reliability measure should indicate how reliable the corresponding estimate of the number of sources is.

The estimate quality of the number of sources is dependent upon the quality of the correlation matrix estimate, which is entirely determined from a multivariate AR model the raw data obey. In particular, the estimate quality of the number of sources is heavily reflected upon the consistency of the variance estimate of the additive noise at each sensor of the given subarray, assuming the *apriori* knowledge that the additive noise has equal variance at all sensors.

We propose a scheme which uses a reliability measure for combining the estimate of the number of signal sources from the subarray sites, each of them consists of L equi-spaced sensors. At each subarray site, the least square (LS) estimate of the Λ, which will be used as a starting value for iteration steps, is computed. Then the robust estimate of the Λ for the L-variate AR model is obtained by the procedure described in the previous section. After the model fitting at each subarray site, we compute the estimate of the sensor noise variance for each sensor by taking the average of the magnitude square of the residuals. The estimate of the jth sensor noise variance at ith subarray is denoted by $\rho_j^{(i)}$, i.e.,

$$\rho_j^{(i)}=\frac{1}{N-p}\sum_{t=p+1}^{N}|w_j^{(i)}(t,\hat{\Lambda}_0)|^2 \quad , \; i=1,...,M \; ; \; j=1,...,L \tag{3.3.2}$$

where M is the total number of subarray sites, L is the number of sensors at each subarray. $w_j^{(i)}(t,\hat{\Lambda}_0)$ is the residual defined by (2.3.9) for the jth univariate model as shown in (2.3.3) for the ith subarray site, and $\hat{\Lambda}_0$ is the estimate of the true parameter vector Λ_0. Then $RS_{(i)}$, the reliability statistic of the ith subarray site which approximately measures the reliability of the estimate of the number of source signals at the subarray site, is defined as

$$RS_{(i)}=\frac{(L-1)S_{(i)}^2}{\hat{\sigma}^2} \quad ,i=1,...,M \tag{3.3.3}$$

where $S_{(i)}^2$ is the sample variance of the $\rho_j^{(i)}$ taken from the ith subarray sensors only, i.e.,

$$S_{(i)}^2=\frac{1}{L-1}\sum_{j=1}^{L}[\rho_j^{(i)}-\bar{\rho}^{(i)}]^2 \tag{3.3.4}$$

$$\bar{\rho}^{(i)}=\frac{1}{L}\sum_{j=1}^{L}\rho_j^{(i)},$$

while $\hat{\sigma}^2$ is the sample variance of the $\rho_j^{(i)}$ from all the subarray sites, i.e.,

$$\hat{\sigma}^2=\frac{1}{ML-1}\sum_{i=1}^{M}\sum_{j=1}^{L}(\rho_j^{(i)}-\hat{\mu})^2 \tag{3.3.5}$$

$$\hat{\mu}=\frac{1}{M}\sum_{i=1}^{M}\bar{\rho}^{(i)}.$$

It can be easily shown that the asymptotic distribution of the test statistic $RS_{(i)}$ will be χ^2 with L-1 degrees of freedom, i.e.,

$$RS_{(i)} \sim \chi^2(L-1). \tag{3.3.6}$$

A risk α and threshold $\beta(\alpha)$ can now be selected such that

$$Prob\,(RS_{(i)} > \beta(\alpha)) = \alpha. \tag{3.3.7}$$

Typically, a value of 0.05 is chosen for α. Then the above criterion rejects the estimate of the number of sources from the ith subarray site if $RS_{(i)} > \beta(\alpha)$. For given L, this is consistent with the idea that the smaller values of $RS_{(i)}$ demonstrate the consistency of the estimates with the *apriori* knowledge that the additive noise at each sensor has the same variance.

The reasoning here is that, as a stronger safeguard against outliers, one does not want to include any obviously unreliable estimates of the number of source signals in the combining scheme. There are many different variants of the robust method which combines the remaining estimates, but the median value of the selected reliable estimates is chosen as the robust estimate of the number of source signals to reduce the complexity of computation. The median is defined by

$$\hat{d}_i = \begin{cases} [\hat{d}_{(\frac{i+1}{2})}] & , \text{ if } i = odd \\ [\frac{1}{2}(\hat{d}_{(\frac{i}{2})} + \hat{d}_{(\frac{i}{2}+1)})] & , \text{ if } i = even \end{cases} \tag{3.3.8}$$

where $\hat{d}_{(j)}$ is the jth order statistic of $\hat{d}_i$, i=1,...,M' and [·] is the closest integer to a real number of any argument. M' is the number of selected subarray sites.

The proposed scheme can be briefly summarized into the following steps using the notation used above. Each (ith) subarray site computes $\rho_j^{(i)}$; j=1,...,L to be sent to the fusion center along with its estimate of the number of sources. The fusion center then computes $\hat{\sigma}^2$, $S_{(i)}{}^2$, and $RS_{(i)}$, i=1,...,M', and eliminates any unreliable estimates of the number of sources using the rejection criterion (3.3.7). After choosing the median of the remaining estimates of the number of sources as the fusion

estimate, the fusion center sends the fusion estimate to those subarray sites whose initial estimate of the number of sources is equal to the fusion estimate.

In cases that only one or very few subarrays have reported this median number of sources, we may also include those subarrays reporting very close results instead of requesting the DOA estimates only from the ones reporting this median number of sources. The degree of closeness required for this inclusion can be subjectively determined according to the accuracy and reliability one pursue from the decentralized scheme. The selected subarray sites then computes the same number of DOA estimates using the MUSIC algorithm. Finally, the DOA estimates from those subarray sites are combined at the fusion center using a robust combining method.

3.4. Integrating the Direction-of-Arrival Estimates

At this point it is assumed that each selected subarray site computes exactly d_f DOA estimates, which correspond to the d_f significant peaks in the corresponding MUSIC spectrum. Here, d_f denotes the fusion estimate of the number of sources. If a selected subarray site did not choose exactly d_f DOA estimates, such DOA estimates are ignored.

Again, the combining scheme at the fusion center may require some sort of reliability measure on the DOA estimates from each subarray site, and a natural choice of such reliability measure seems to be the variances of the corresponding estimates. Barabell et al. [1] derived the expressions for the average deviation of the null spectra, which is the inverse of the MUSIC spectrum utilized in our scheme, at the true DOA's for one and two signal source cases. However, such expressions is not be easily extended to arbitrary number of signal cases. Therefore, we look for other practical alternatives.

In [38], a robust technique is developed for combining the frequency estimates from different sensors, and one may treat the DOA estimates similarly. Since the number of DOA estimates from each subarray site is already known to be identical, we do not consider the case where the number of DOA estimates are different. Let $\theta_1,...,\theta_{d_f}$ be a set of true DOA's in ascending order, and $\hat{\theta}_1^{(i)},..., \hat{\theta}_{d_f}^{(i)}$ be the corresponding set of d_f DOA estimates from the ith selected subarray site.

In general, such a set of d_f DOA estimates is a set of angles with no particular significance to order, and there is the potential for problems in applying vector combining procedures or scalarized procedures based on picking single components from fixed positions in the vector. The probability of such potential problems can be minimized by strictly eliminating the sets of DOA estimates from unreliable subarray sites before the combining. Furthermore, remaining sets of the DOA estimates are combined using a robust technique which is insensitive to the effect of outliers possibly still remaining in the selected DOA estimates. It is assumed that each of the DOA estimate vector are combined is a set of angles in ascending order, i.e., (0.2123, 0.3021, 0.4036).

If the source powers are distinctly different from source to source, one can also rank the estimates according to their corresponding source strengths so that the combination occurs among the estimates of the same ranks. Another possibility is that the estimates exhibiting the strongest source power can be combined first. For the estimates exhibiting the next strongest source power, the differences between these and those of the strongest source power are combined rather than the estimates themselves. The combined differences can be added to the first estimate to get the second estimate exhibiting the second strongest source, etc.

Since each DOA estimate of a given subarray site is a function of observations at each subarray site, it can be represented as

$$\hat{\theta}_j^{(i)} = \theta_j + \gamma_j^{(i)} \quad ; i=1,...,M', \ j=1,...,d_f \tag{3.4.1}$$

where θ_j is the jth true DOA, and $\hat{\theta}_j^{(i)}$ is the jth DOA estimate from the ith selected subarray site. M' is the number of selected subarray sites, and d_f is the fusion estimate of the number of sources. The problem is to estimate θ_j from $\hat{\theta}_j^{(1)},...,\hat{\theta}_j^{(M')}$, $j=1,...d_f$, but the main problem lies in the fact that the distribution of the perturbation $\{\gamma_j^{(i)}\}$ is unknown.

An important issue is whether the fusion estimate is insensitive to a few bad estimates. The bad estimates may be the ones which are sent from a malfunctioning subarray site. A necessary assumption on the perturbation $\{\gamma_i^{(i)}\}$ is that it has a symmetric distribution, and can be approximated as a mixture distribution. For example, γ_i may have the following distribution:

$$\gamma_i \sim (1-\epsilon)\phi + \epsilon\psi \tag{3.4.2}$$

where $\epsilon = 0.05$, $\phi = N(0,\rho)$, and ψ is an unknown outlier distribution with zero mean and the variance of 9ρ. This is an example of the well known mixture distribution [80], and many robust estimation techniques are known to perform well for this kind of mixture distributions. Even in the case where the parent distribution is not Gaussian, one can still use such estimation techniques for combining estimates since the robust estimation methods are not sensitive to distributions.

The robust combining technique is the location parameter estimation problem explored by Huber [27], which is shown in the following. Find $\theta_j^{\dagger}, j=1,...,d_f$ which minimizes

$$\sum_{i=1}^{M} H\left(\frac{\hat{\theta}_j^{(i)} - \theta_j^{\dagger}}{\sqrt{\hat{v}_j}}\right) \tag{3.4.3}$$

where

$$H(x)=\begin{cases}\dfrac{x^2}{2} & , \text{ if } |x|\le c\\ c|x|-\dfrac{c^2}{2} & , \text{ if } |x|>c\end{cases}. \qquad (3.4.4)$$

c is the breakdown point constant obtained by (2.3.12), while $\hat{v}_j$ is the sample variance taken from $\hat{\theta}_j^{(1)}, \hat{\theta}_j^{(2)}, \ldots, \hat{\theta}_j^{(M')}$.

3.5. Numerical Simulation

The objective of this simulation study is to investigate and confirm the effectiveness of the robust decentralized DOA estimation technique in many different contaminated noise situations. For the simulation, It is assumed that there are ten subarray sites, each with eight equispaced identical sensors. There are two signal sources with θ_1=0.7854 and θ_2=1.0472, denoting the first and second true DOA's in radians, with respect to the vertical axis stretched above sensor number one as was shown in Figure 2.1. The source signals are again chosen as

$$s_1(t)=exp(j\ 0.4\pi t) \quad \text{and} \quad s_2(t)=exp(j\ 0.8\pi t). \qquad (3.5.1)$$

For simplicity, it is assumed that δ, the spacing between sensors, is exactly one half of λ, the signal wavelength, then (2.2.2) becomes

$$x_{lm}(t)=exp\,[j\pi(0.4t+l\ \sin\theta_1)]+exp\,[j\pi(0.8t+l\ \sin\theta_2)]+w_{lm}(t), \qquad (3.5.2)$$

from which the data in the simulation is generated. At this point, it is important to point out that different subarray sites have different true DOA's simply because of their geographical location differences. If one knows the locations of all the subarray sites, this problem can be easily overcome by accommodating the subarray location differences at each subarray site. For simplicity, it is assumed that the problem has been already remedied at each subarray site.

The order p_i of the 8-variate complex AR model, i.e. 8 sensor array, at the ith subarray site is determined using the following decision statistic for multivariate AR model.

$$KIC(p_i) = N\sum_{j=1}^{8} \ln \rho_j^{(i)} + n_{p_i} ln\,(N/2\pi) \tag{3.5.3}$$

where N is the number of available 8-variate data, $\rho_j^{(i)}$ is defined by (3.3.2), and n_{p_i} is the total number of parameters to be estimated for the p_ith order model. We choose the order, p_i, which minimizes $KIC(p_i)$. Details of the decision statistics are found in [34,35]. With N=100, the decision criterion (3.5.3) is minimized when p_i=1 or p_i=2 in most of the subarray sites, depending on the quality of the additive noise.

At each of the ten subarray sites, the complex parameters of the 8-variate AR model are estimated using the technique developed in section III, and the correlation matrix which corresponds to the 8-variate AR model is computed. The MDL criterion by Wax et al. [84] is then utilized to find the estimate of the number of signal sources and the DOA estimates at each subarray site are computed.

Estimation of the parameters, which required ten to twenty iterations using the Newton-Raphson algorithm for each of the L univariate model as in (2.3.3), is processed in parallel at each subarray site only once. Incorporating this into a real time procedure would require additional computation, but the trade off is worth while in situations of contaminated noise.

Table 3.1 shows the performance comparison of the nonrobust and the robust approach at each of the ten subarray sites when the SNR is 13 dB and there exists one percent outlier Gaussian noise, which has five times the variance of the parent Gaussian. For the nonrobust method mentioned above, the correlation matrix estimate is provided by (2.2.5), in which the additive noise is assumed to be a pure Gaussian. Note that the correct estimates of the number of sources, which is denoted by +, are

always associated with smaller values of RS values, i.e., 2.17 is used as the threshold for α=0.05. * denotes the case where the MUSIC spectrum does not exhibit the corresponding spectral peaks.

Table 3.2 shows the final DOA estimates for the nonrobust and the robust approach after the combining, and the root mean square error (RMSE) of the estimates from subarray sites before the combining, for many contaminated noise situations with the SNR of 13 dB. Shown in the associated parentheses are the RMSE values of the DOA estimates from different subarray sites before the combining but after eliminating any unreliable estimates. Here, * indicate the case where none of the ten subarray sites detected the MUSIC spectrum peak which are at the vicinity of the desired estimate. Table 3.3 shows similar comparisons as in Table 3.2, but with many different SNR's in a fixed contaminated noise environment, i.e., there exists one percent outlier Gaussian noise which has five times the variance of the parent Gaussian.

In combining the nonrobust estimates, the algebraic mean of all the available estimates are taken as the final DOA estimates. The combining of the DOA estimates only involves those from the subarray sites where the subarray's estimate of the number of sources coincides with the robust estimate from the fusion center. In other words, if a given subarray site did not choose exactly two DOA estimates, they were not included in the combining scheme.

The change in the RMSE values caused by the variation in the noise environment is evident in Table 3.2 and Table 3.3. The decentralized processing scheme indeed packs more reliability into the robust algorithm which already provides much more consistent performance than the corresponding nonrobust method which deteriorates very quickly with a slight change of the noise environment.

In the following tables, d denotes the estimate of the number of source signals. ε denotes the percent probability of the outlier noise, and a denotes the ratio of the outlier noise variance vs. the dominant Gaussian noise variance. $\hat{\theta_1}$ and $\hat{\theta_2}$ denote the two DOA estimates. The "reliability statistic" is denoted by RS.

Table 3.1. Comparison of nonrobust and robust estimates at ten different subarray sites when ε=0.01, a=5, and the SNR is equal to 13dB. Note that the correct estimates of the number of sources, which is denoted by +, are always associated with smaller values of RS values (2.17 is used as the threshold for α=0.05). * denotes the case where the MUSIC spectrum does not exhibit the corresponding spectral peaks. The true DOA's are θ_1=0.7854 and θ_2=1.0472 in radians.

Nonrobust and Robust Results at Each Subarray Site							
	nonrobust			robust			
site	d	$\hat{\theta}_1$	$\hat{\theta}_2$	d	$\hat{\theta}_1$	$\hat{\theta}_2$	RS
1	4	0.8388	*	3	0.8242	1.0194	2.76
2	2	0.7901	1.0430	2^+	0.7901	1.0446	0.0017
3	5	0.7791	*	5	0.8137	0.9959	22.82
4	5	0.8624	*	4	0.8184	1.0414	3.81
5	3	0.9912	*	4	0.7885	1.0556	2.29
6	2	0.7807	1.0414	2^+	0.7791	1.0430	0.0059
7	2	0.7854	1.0524	2^+	0.7885	1.0524	0.0047
8	3	0.9299	*	4	0.8011	1.0147	5.98
9	3	0.7587	1.0509	6	0.7901	1.0540	1.03
10	3	0.9739	*	4	0.7854	1.0446	5.99

Table 3.2. Final DOA Estimates at the fusion center after the combining when the SNR is equal to 13 dB. Shown in the associated parentheses are the RMSE values of the DOA estimates taken from the subarray sites before the combining but after rejecting unreliable estimates. * indicate the case where none of the ten subarray sites detected the MUSIC spectrum peak which corresponds to the desired estimate. The true DOA's are θ_1=0.7854, and θ_2=1.0472 in radians.

Combined Estimates (& RMSE before combining)					
		nonrobust		robust	
ε	a	$\hat{\theta}_1$	$\hat{\theta}_2$	$\hat{\theta}_1$	$\hat{\theta}_2$
0		0.7863 (0.0038)	1.0482 (0.0073)	0.7860 (0.0028)	1.0495 (0.0051)
0.01	5	0.7854 (0.0039)	1.0234 (0.0445)	0.7880 (0.0037)	1.0470 (0.0032)
0.05	5	0.8438 (0.0870)	1.0487 (0.0093)	0.7864 (0.0034)	1.0488 (0.0085)
0.01	10	0.8323 (0.1630)	0.9843 (0.0972)	0.7876 (0.0141)	1.0389 (0.0163)
0.05	10	0.9475 (0.3945)	*	0.8222 (0.0628)	1.0642 (0.0226)

Table 3.3. Final DOA estimates at the fusion center after the combining when ε=0.01 and a=5 for many different SNR's. Shown in the associated parentheses are the RMSE values of the DOA estimates taken from the subarray sites before combining but after rejecting unreliable estimates. * indicate the case where none of the ten subarray sites detected the MUSIC spectrum peak which corresponds to the desired estimate. The true DOA's are θ_1=0.7854, and θ_2=1.0472 in radians.

Combined Estimates (& RMSE before combining)				
	nonrobust		robust	
SNR(dB)	$\hat{\theta}_1$	$\hat{\theta}_2$	$\hat{\theta}_1$	$\hat{\theta}_2$
13.0	0.7854 (0.0038)	1.0234 (0.0445)	0.7880 (0.0037)	1.0470 (0.0032)
8.5	0.7866 (0.0026)	1.0226 (0.0481)	0.7866 (0.0026)	1.0419 (0.0082)
6.7	0.7825 (0.0084)	1.0051 (0.0688)	0.7858 (0.0047)	1.0485 (0.0049)
5.5	0.7838 (0.0092)	1.0215 (0.0828)	0.7849 (0.0065)	1.0454 (0.0058)
4.5	0.7882 (0.0123)	1.0731 (0.1547)	0.7872 (0.0126)	1.0484 (0.0118)
1.5	*	*	0.7901 (0.0101)	1.0496 (0.0357)
-1.5	*	*	0.7681 (0.0173)	1.0021 (0.0451)

3.6. Conclusion

We have considered a new decentralized processing scheme to estimate the number of signal sources and their directions-of-arrival. It employs a decentralized processing scheme such that each subarray site provides a robust estimate of the number of sources accompanied by its corresponding reliability statistic such that only the reliable estimates of the number of sources are combined at the fusion center. A robust combination technique is used to combine the corresponding DOA estimates from the subarray sites. Simulation results show that the new decentralized procedure provides much more reliable estimates that are also robust against outliers and distributional uncertainties in the noise environment. Simulation results also confirm that the new scheme performs especially well at low values of the SNR.

CHAPTER 4

DIRECTION-OF-ARRIVAL ESTIMATION USING RADON TRANSFORM

4.1. Introduction

The problem of estimating the direction of arrival (DOA) of radiating sources from measurements provided by a passive array of sensors is frequently encountered in radar, sonar, radio astronomy and seismology. This chapter specifically consider the DOA estimation problem when there are many sources, each of which is either narrow band or wide band, in situations of low SNR, outlier contaminated Gaussian noise, and colored noise with unknown correlations.

Different approaches have been followed for solving the DOA estimation problem: beamforming, maximum likelihood, eigenspace methods, etc. Beamforming methods are computationally efficient and yield effective performance in low resolution applications where the incident source spatial separations are sufficiently larger than the inverse of the array aperture [45]. The ML technique has not been popular because of the high computational load involved in the multivariate nonlinear maximization. Recently, Ziskind and Wax [87] have presented a computationally attractive method for computing the ML estimate of narrow band sources. Eigenspace methods such as MUSIC [65] and ESPRIT [54] have become popular in applications requiring

high resolution capability. However, eigenspace methods are usually based on narrow band assumption of signals. One way of solving the wide band DOA estimation problem is to divide the wide frequency band into non-overlapping narrow bands, and then use narrow band signal subspace processing [86]. Alternatively, Wang and Kaveh [82] have considered an eigenspace method where the estimates are obtained by the eigen-decomposition of a frequency domain combination of modified narrow band covariance matrix estimates.

Instead of treating the wide band problem as a multitude of narrow band emitter problems, Su and Morf [77] and Porat and Friedlander [56] have considered using a multivariate rational model for the sensor outputs. Another approach for the DOA estimation problem is to consider it as a 2-D spectral estimation problem [19]. An advantage of this approach is that it is applicable when both narrow band and wide band sources are present simultaneously. Jackson and Chien [28], however, have pointed out the severe asymmetry and bias in the estimated spectra using a 2-D quarter plane AR model for bearing estimation.

There has been a growing interest in the development of theory and applications of *robust* methods, where the term "robustness" refers to insensitivity against small deviation in the underlying Gaussian noise assumption. Previous schemes, which were developed and tested under the Gaussian assumption, usually fail to resolve close DOA's when the underlying noise distribution deviates even slightly from the assumed Gaussian since they are very sensitive to minor deviations from the underlying assumptions. Therefore, the importance of robust methods need not be overemphasized.

Hansen and Chellappa [23] have recently considered 2-D robust spectral estimation, and have found that it requires very extensive computation. The necessity of a large order non-causal model for resolving fine details in a 2-D PSD has been also

pointed out in [67]. It is well known, however, by the principle of parsimony that the accuracy of the parameter estimates decreases when the number of unknown parameters to be estimated increases. Hence, robust spectral estimation methods are computationally feasible only when the number of parameters to be estimated is small.

Recently, a new approach of 2-D spectral estimation utilizing 1-D autoregressive (AR) models in the Radon space was investigated by Srinivasa et al. [71,75]. The 2-D PSD is estimated from a finite set of observations of a 2-D stationary random field (SRF) using the Radon transform. In particular, the 2-D PSD estimation problem is converted into a set of 1-D independent problems using the modified central slice theorem for SRF introduced by Jain and Ansari [29]. The projections of the array data are computed, and then 1-D models are utilized for each projection to obtain an estimate of the 2-D PSD. Since the number of parameters to be estimated in the 1-D model is small, robust methods of parameter estimation are feasible.

The contribution of this chapter is an application of the Radon transform approach of 2-D PSD estimation to the DOA estimation problem. The importance of the method presented here is that it does not require any information about the number of received signals, structure of the signals, and the correlation structure of sensor noise. The technique is capable of handling narrow band and wide band sources simultaneously at low SNR's, and performs equally well in the presence of colored noise with unknown correlation structure.

The quality of the estimates obtained by this method may not be as good as that of model based methods such as maximum likelihood (ML) estimation if the number of signals, signal type, frequency of signals, and the sensor noise structure is already known exactly. In practice, however, no such information is given beforehand. Furthermore, the expression for the variance of the estimates cannot be obtained using this method. Though the DOA estimation method presented in this chapter is

related to the traditional beamforming, it has much better resolving capability as we use the spectral density, which is in turn estimated by using a model, to measure the average power.

An additional important aspect of the work presented here is the use of robust 1-D AR parameter estimation method in the Radon space to obtain a robust 2-D PSD estimate. This reduces the number of parameters to be estimated simultaneously, thus allowing the robust 2-D PSD estimation feasible. This algorithm is highly amenable for parallel processing, and any particular range of directions of interest can be probed for detecting the presence or absence of sources.

The organization of the chapter is as follows. In Section 4.2 we introduce the signal and the noise model, and Section 4.3 outlines the DOA estimation scheme using Radon transform. Section 4.4 then brings out the important similarities and differences between this approach and traditional beamforming. Section 4.5 brings out the significance of this method in wide band signals and correlated noise with unknown structure. Section 4.6 then discusses some of the simulation results carried out to demonstrate the performance of the proposed algorithm, followed by Section 4.7 which concludes the chapter.

4.2. Problem Formulation

Consider a uniform linear array with M identical sensors with interelement spacing d. Let us assume that ν stationary zero-mean sources, with directions of arrival (bearings) θ_i , $i=1,...\nu$, impinge on the array. The sources are located sufficiently far from the array such that in homogeneous isotropic transmission media, the wavefronts impinging on the array are planar. If we treat the sampled outputs from the sensors as a 2-D data sequence, the received 2-D signal is given by

$$y(n,m) = \sum_{i=1}^{\nu} \alpha_i s_i(nT - mD_i) + u(n,m), \tag{4.2.1}$$

where T is the sampling period, α_i and D_i are the unknown amplitude and time delay between elements of the array associated with the ith source, $s_i(t)$. The DOA for the ith source, θ_i, is measured with respect to the array normal, i.e., $D_i = d \sin\theta_i$.

The additive noise $u(n,m)$ present at each sensor of the array is assumed to be a stationary outlier contaminated Gaussian process which may also be correlated from sensor to sensor, but is statistically independent of the signal. For reasons explained earlier, the noise is assumed to obey a slippage model, i.e., $u(n,m)$ is distributed either as a Gaussian distribution $N(0,\sigma^2)$ with probability $(1-\varepsilon)$, or as an unknown distribution $Q(\mu, \beta\sigma^2)$ with probability ε, where μ and $\beta\sigma^2$ are the mean and the variance of the unknown distribution Q. In general, $\beta>1$, $\varepsilon < 0.1$ and, μ, an unknown constant, is of the order of a multiple of σ. The noise distribution can then be expressed as

$$P[u(n,m)] \sim (1-\varepsilon)N(0,\sigma^2)+\varepsilon Q(\mu, \beta\sigma^2), \tag{4.2.2}$$

which represents a family of distributions characterized by the mixing parameter ε. Note that for $\varepsilon=0$, equation (4.2.2) reduces to a Gaussian distribution.

The objective here is to estimate the unknown DOA's, θ_i; $i=1,...,\nu$, with respect to the vertical axis stretched above sensor number one as shown in Figure 4.1, from the observations $\{y(n,m), n=1,2,...,N; m=1,2,...,M\}$ which are obtained by sampling the array output at the Nyquist rate.

4.3. Robust Direction-of-Arrival Estimation using Radon Transform

The simultaneous estimation of the DOA and spectral densities of radiating sources is equivalent to a 2-D spectral estimation problem. This is a general problem which not only arises in passive sonar but also in other application areas as well. Recently, a Radon transform approach of 2-D spectral estimation utilizing 1-D autoregressive (AR) models in the Radon space was investigated [75]. In the following we consider the application of this new approach of 2-D spectral estimation to the DOA estimation problem.

4.3.1. 2-D Spectral Estimation using AR Modeling in the Radon Space

The basic idea here is to use the Radon transform to convert the 2-D spectral estimation problem into a set of 1-D independent spectral estimation problems. The 2-D sequence is transformed to a set of 1-D sequences by forming projections. A projection at an angle ψ_k is a weighted summation of the observations and is given by

$$p_{\psi_k}(j) = \sum_{n=1}^{N} \sum_{m=1}^{M} w_{kj}(n,m)\, y(n,m), \tag{4.3.1}$$

where ψ_k denotes the kth projection angle, and $w_{kj}(n,m)$ is a weighting factor determined by some geometrical considerations. For example $w_{kj}(n,m)$ can be made proportional to the length of intersection of the jth ray in the kth projection with the element (n,m), (see Figure 4.1). Here, we have assumed a grid structure and each data sample is located at the center of the cell. The 1-D sequence $\{p_{\psi_k}(j),\ j = 1,2,...,L\}$ is thus obtained by summing up samples which fall along a set of parallel lines which are normal to the projection angle ψ_k.

By the Central Slice Theorem for random fields [29], $P_{\psi_k}(\Omega)$, the 1-D power spectral density of the projection taken at the projection angle ψ_k, is related to the slice of the 2-D power spectral density $S(\Omega_1,\Omega_2)$ of the infinite array data,

$$S(\Omega\cos\psi_k,\Omega\sin\psi_k) = |\Omega|\, P_{\psi_k}(\Omega). \tag{4.3.2}$$

Hence, an estimate of the 2-D PSD of the array data can be approximated slice by slice from the estimates of 1-D PSD of the projections on a polar raster. Any 1-D modeling technique can be utilized for the projection data. In the present study we have utilized an Autoregressive (AR) model, and a robust method of estimating the AR parameters is described next. For the sake of notational simplicity only subscript k is retained in the following.

4.3.2. Robust Estimation of Parameters of the AR Model for 1-D Projections

Each of the projection sequences is modeled by a Γth order AR model, i.e.,

$$p_k(j) = {\Lambda^0}_k{}^T Z_k(j-1) + \eta_k(j) \tag{4.3.3}$$

where

$${\Lambda^0}_k = col.\,(\lambda_k(1),\ldots,\lambda_k(\Gamma)), \tag{4.3.4}$$

is a column vector containing the AR parameters,

$$Z_k(j-1) = col.\,(p_k(j-1),..,p_k(j-\Gamma)), \tag{4.3.5}$$

is the lag sequence of the projection samples, and $\eta_k(j)$ is a white noise sequence. The AR parameters are estimated by a robust technique based on the so called M-estimators, a generalization of classical maximum likelihood (ML) estimator by Huber [27].

The following robust estimation scheme is utilized for each of the 1-D AR parameter estimation problem. The subscript k has been dropped for notational simplicity in the following discussion. In order to enforce scale invariance, the parameter vector Λ and a scale σ are estimated simultaneously. In [27] this was done by minimizing

$$J(\Lambda,\sigma) = \sum_{j=\Gamma+1}^{L} [\rho(\Delta_j/\sigma)\,\sigma] + a\sigma\,, \qquad \sigma > 0 \tag{4.3.6}$$

with respect to Λ and σ, and Δ_j and $\rho(\cdot)$ are defined in (4.3.8) and (4.3.9) respectively. The constant a, given by (4.3.12), is chosen to make the estimates consistent at the nominal distribution. Unless the minimum of $J(\Lambda,\sigma)$ occurs on the boundary σ=0, it can be equivalently characterized by the Γ+1 equations

$$\sum_{j=\Gamma+1}^{L} \psi(\Delta_j/\sigma) \left\{ \frac{\partial \Delta_j}{\partial \lambda(i)} \right\} = 0; \qquad i=1,\ldots,\Gamma$$

and

$$\sum_{j=\Gamma+1}^{L} \chi(\Delta_j/\sigma) = a. \tag{4.3.7}$$

Here,

$$\Delta_j = p(j) - \Lambda^T Z(j-1), \tag{4.3.8}$$

$$\rho(x) = \begin{cases} x^2/2 & , \text{ if } |x| \le c \\ c|x| - c^2/2 & , \text{ if } |x| > c \end{cases} \tag{4.3.9}$$

$$\psi(x) = \partial/\partial x\ \rho(x), \tag{4.3.10}$$

$$\chi(x) = x\ \psi(x) - \rho(x), \tag{4.3.11}$$

and

$$a = (L-\Gamma) \int_{-\infty}^{\infty} \chi(x)\,(1/\sqrt{2\pi})\,\exp\{-x^2/2\}\,dx. \qquad (4.3.12)$$

The constant c is related to the fraction of contamination ε by

$$2\Phi(c) - 1 + 2\phi(c)/c = 1/(1-\varepsilon), \qquad (4.3.13)$$

where, $\Phi(c)$ is the standard cumulative Gaussian distribution with zero mean and unit variance, and $\phi(c)$ is the corresponding Gaussian density. Usually c is chosen to lie between 1 and 2, which corresponds to the ε-interval [0.0083, 0.1428] by (4.3.13). We use the following algorithm since its convergence properties are well established [27].

4.3.3. Robust Estimation Procedure

(1). Choose starting values of $\Lambda^{(m)}$, $\sigma^{(m)}$; $m=0$, and and a tolerance value $\varepsilon > 0$.

(2). Compute residuals $\Delta_j^{(m)}$, $j=\Gamma+1,...,L$ by

$$\Delta_j^{(m)} = p(j) - \Lambda^{(m)T} Z(j-1).$$

(3). Compute a new value of $\sigma^{(m)}$ by

$$(\sigma^{(m+1)})^2 = (1/a) \sum_{j=\Gamma+1}^{L} \chi(\Delta_j^{(m)}/\sigma^{(m)})\,(\sigma^{(m)})^2.$$

(4). "Winsorize" the residuals

$$q_j = \psi(\Delta_j^{(m)}/\sigma^{(m+1)})\,\sigma^{(m+1)} \quad ; \; j=\Gamma+1,...,L.$$

(5). Compute the partial derivatives

$$b_{ji} = (\partial/\partial\lambda(j))\,[\Lambda^{(m)T}\, Z(j-1)] = q(\Gamma+j-i) \quad ; \; j=\Gamma+1,...,L. \;\; i=1,...,\Gamma.$$

Note that one needs to compute b_{ji}'s only once at the beginning of the iteration since $\Lambda^{(m)T}\, Z(j-1)$ is a linear function of $\Lambda^{(m)}$.

(6). Solve for τ

$$B^T B\tau = B^T Q$$

where $Q = col.\,(q_1, \ldots, q_L)$ and B is a $(L-\Gamma)$ by Γ matrix whose elements are b_{ji}'s.

(7). Update $\Lambda^{(m)}$

$$\Lambda^{(m+1)} = \Lambda^{(m)} + \xi\tau,$$

where $0 < \xi < 2$ is an arbitrary (fixed) relaxation factor.

(8). Stop iterating and go to step (9) if the parameters change by less than ε times their standard deviation, where ε is an arbitrarily chosen small value, i.e., if for all j

$$|\tau_j| < \varepsilon \sqrt{\bar{b}_{jj}}\;\; \sigma^{(m+1)}$$

where $\bar{b}_{jj}$ is the jth diagonal element of the matrix $\bar{B} = (B^T B)^{-1}$; otherwise $m = m + 1$ and go to step (2).

(9). The final estimate of Λ^0 is given by $\Lambda^{(m+1)}$, and the variance of the residual sequence by $(\sigma^{(m+1)})^2$.

Once the AR parameters are estimated, a slice of the 2-D PSD is estimated by

$$\hat{S}(\Omega\cos\psi_k, \Omega\sin\psi_k) = \frac{|\Omega|\ \hat{\nu}_k}{\left|1 - \sum_{j=1}^{\Gamma} \hat{\lambda}_k(j)\ exp\,[-i(j\Omega)]\right|^2}, \qquad (4.3.14)$$

where $\hat{\lambda}_k(j)$ for $1 \le j \le \Gamma$ are the estimated coefficients of the Γth order AR model and $\hat{\nu}_k$ is the variance of the residual sequence for the projection at angle ψ_k. The 2-D PSD estimate is obtained on a polar raster by repeating this procedure by taking projections over the angular range $[0^\circ, 180^\circ)$.

4.3.4. Estimation of the Directions-of-Arrival

The bearing and the spectral densities of radiating sources can be estimated from the 2-D spectrum $S(\Omega_1, \Omega_2)$ of the spatio-temporal array data [19,28]. For simplicity consider the case of a single source with DOA θ and center frequency ω received by a uniform linear array, and assume unit sampling for both the spatial and temporal domain. The temporal frequency variable $\Omega_1 = \omega$ while the spatial frequency variable $\Omega_2 = \omega \sin\theta$. Hence a peak in the spectrum at (Ω_1, Ω_2) corresponds to a signal with frequency ω and direction of arrival θ measured with respect to the array normal. Since the projection taken in the plane perpendicular to the DOA captures the energy distribution of the source, the corresponding slice angle ψ in the polar raster is the complimentary angle to the DOA. Using Cartesian-to-polar conversion,

$$\Omega_1 = \Omega\ cos\,(\psi - 90^\circ) = \omega$$

$$\Omega_2 = \Omega\ sin\,(\psi - 90^\circ) = \omega \sin\theta \qquad (4.3.15)$$

Hence, the DOA θ is estimated by

$$\hat{\theta} = \sin^{-1}[tan(\hat{\psi} - 90^{\circ})] \tag{4.3.16}$$

where $\hat{\psi}$ is the slice angle which contains the peak in the polar raster.

In the case of several sources, we use the principle of superposition to detect the DOA's. Thus for ν sources coming from directions $\theta_1, \theta_2, \ldots, \theta_\nu$, the 2-D spectrum of the array data exhibits a corresponding number of distinct peaks, with each peak being located at a point determined by the direction of the corresponding source. The DOA θ_i is thus estimated by

$$\hat{\theta}_i = \sin^{-1}[tan(\hat{\psi}_i - 90^{\circ})]$$

where $\hat{\psi}_i$ denotes the slice angle of the ith peak in the polar array. We would like to point out that only those slices of the 2-D PSD over any desired range of angles, dictated by the range of DOA's of interest, can be estimated by forming the corresponding projections. This is particularly useful in tracking applications. Further, note that the frequency of the sources need not be known *apriori* to estimate the DOA's. In fact, the source frequencies can be estimated simultaneously using (4.3.15). In the next section we will derive the relation (4.3.16) alternatively from a purely spatio-temporal analysis.

4.4. Relation between Beamforming and Radon Transform

Classical beamforming method utilizes a delay-and-sum processor. The idea behind beamforming is to align the propagation delays of a signal, presumed to be propagating in some particular direction, so as to reinforce it. Signal propagating from other directions and the noise are not reinforced. The energy in the beam is computed for many directions of look which is in turn achieved by manipulating the delays. The DOA of signals correspond to the location of the maxima of this energy

plotted against the direction of look. A major drawback well known about this method is the poor angular resolution which is directly related to and limited by the physical length of the array. Robinson [60], and Scheibner and Parks [63] have pointed out that time domain beamforming, which involves shifting and summing the receiver outputs, is equivalent to performing a discrete Radon transform given by (4.3.1). The beamformer output is formed by summing the array data along lines of constant slowness and time intercept. The Radon transform consists of the integration (summing) of a function of two dimensions along straight lines each given by its slope and intercept. We first present an alternative derivation of (4.3.16) and then discuss why we wish to estimate the DOA from the location of the spectral peaks which are obtained by modeling the projection data. In the following continuous functions are assumed and summations are replaced by integrals.

Consider the case where a linear array receives a narrow band signal and the array output is given by

$$y(x,t) = a\ exp\,[j\omega(x\sin\theta + t)] + u(x,t) \tag{4.4.1}$$

where the signal and noise are assumed to be statistically independent. Integrating $s(x,t)$ along the line

$$x\cos\psi + t\sin\psi = r, \tag{4.4.2}$$

the projection at an angle ψ is given by

$$p_{\psi}(r) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} y(x,t)\,\delta(x\cos\psi + t\sin\psi - r)\,dx\,dt. \tag{4.4.3}$$

From (4.4.2)

$$t = r\,cosec\,\psi - x\cot\psi.$$

Hence,

$$p_{\psi}(r) = \int_{-\infty}^{\infty} a\, exp\,[j\omega(x\sin\theta + r cosec\psi - x\ cot\ \psi)]\ dx$$

$$+ \int_{-\infty}^{\infty} n(x,\ r cosec\psi - x cot\ \psi)\ dx. \qquad (4.4.4)$$

Upon simplifying

$$p_{\psi}(r) = \frac{a\ \exp[j\omega(x\sin\theta + r cosec\psi - x cot\psi)]}{j\omega(\sin\theta - cot\psi)} + h(\psi, r), \qquad (4.4.5)$$

where the last term denotes the second integral in (4.4.4). The power in the projection at angle ψ is given by

$$J(\psi) = \int_{-\infty}^{\infty} p_{\psi}(r)\, p_{\psi}^{*}(r)\ dr. \qquad (4.4.6)$$

Substituting for $p_{\psi}(r)$ from (4.4.5),

$$J(\psi) = \frac{a^2}{[\omega(\sin\theta - cot\psi)]^2} + h_1(\psi). \qquad (4.4.7)$$

Assuming the noise component to be small, the maximum of $J(\psi)$ occurs when

$$\sin\theta = cot\psi.$$

Hence,

$$\theta = \sin^{-1}[\ \tan(\psi - 90^{\circ})] \qquad (4.4.8)$$

However, instead of computing the power in each projection and finding the maximum of this power among various projections taken at different angles, we compute the spectrum of each projection. Traditional beamforming measures the energy by purely deterministic method, i.e., by summing up the squares of the amplitude. In principle, it is liable to be erroneous because of the random variation of sensor

outputs caused by noise. This also explains why beamforming fails to perform well at low SNRs. In our case we use the spectral density to measure the average power, and it is well known that the spectral density is a robust measure of the energy in stochastic signals. Further, the poor resolution of the beamforming method is overcome by modeling the projection data and then computing the spectrum. These are the important differences between the classical beamforming approach and the method presented in this chapter.

4.5. Importance of the Method in Wide Band Signals and Correlated Noise

The method presented here does not require any information about the number of received signals, type of the source signals, frequencies of the signals, and structure of the sensor noise, since it uses the spectral density to measure the average power in each projection and finding the maximum of this power among various projections taken at different angles. For the same reason, the technique is capable of handling narrow band and wide band sources simultaneously, and performs equally well in the presence of colored noise with unknown correlation structure.

One drawback of this method is that the quality of the estimates may not be as good as those of model based methods such as maximum likelihood (ML) estimation if the structure of the signal is known exactly and sensor noise obeys a Gaussian distribution. Secondly, explicit expression for the variance of the estimates obtained by this method cannot be derived.

4.6. Simulation Results

A number of experiments with synthetic data are carried out to study the performance of our DOA estimation procedure in low SNR, outlier contaminated Gaussian noise, correlated noise situations, and the combination of narrow band and wide band sources. The data set in experiment 1 and 2 is ($32\,X\,32$), while in the rest of the experiments it is ($16\,X\,16$). Unit sampling will be assumed for the sake of simplicity. The 2-D spectrum is obtained on a polar raster from the 1-D PSD estimated using an AR model for each of the 180 equi-spaced projections. Since the location of the peak in the 2-D spectrum is sufficient to deduce the DOA, the conversion from polar to Cartesian co-ordinates is not required. In the following, by non-robust method we mean that the spectrum is computed from AR parameters estimated using Marple's least squares algorithm [49], while robust method implies that the spectrum is computed from AR parameters estimated using the robust method outlined earlier. The order of the 1-D AR model is chosen according to the order selection criterion due to Kashyap [34]. The DOA is estimated by locating the peaks in the 2-D spectrum and using (4.3.16).

4.6.1. Experiment 1

In this experiment the DOA estimation of a single source was studied in the case of pure Gaussian noise using the nonrobust method with SNR set at -9 dB. In order to get a feel for the statistical performance, 20 different data sets were generated when the signal source is injected to arrive at an angle $\theta°$. The mean and RMSE of the estimated DOA obtained for different values of $\theta°$ are given in Table 4.1. In practical situations, the DOA estimates can be obtained more accurately by using the following

strategy : First compute the 2-D spectrum using a larger angular spacing between projections (say 5°) and locate the peaks. In the neighborhood of the located peaks, more projections can be computed by decreasing the angular spacing between projections. This procedure can be repeated two or three times to get more accurate DOA estimates.

4.6.2. Experiment 2

In this experiment the DOA estimation of a single signal source arriving at 30.0 degrees is considered in various outlier contaminated noise environments. Table 4.2. gives the details of the results obtained using the nonrobust technique and the robust technique. Note in the case of pure Gaussian noise, the results obtained from both methods are accurate. However, it shows that the robust method always perform better in the presence of outliers. Computational experience has however shown that the cost function may have various local minima, especially when the SNR is low. Thus for low SNR's, in spite of using sophisticated optimization techniques, the minimization algorithms do not always converge to the global minimum. For modest values of SNR, a systematic method of avoiding convergence to a local minima by a "reduction of poles technique" has been presented [3]. But for very low SNR's this problem cannot be avoided completely and the results from the local minima have to be accepted. One can always obtain the global minimum by repeating the algorithm with several different starting points and choosing the one with the least value, which may be very time consuming.

4.6.3. Experiment 3

In order to demonstrate the effectiveness of the proposed DOA estimation scheme under colored noise, the problem of resolving two DOA's under spatially correlated array sensor noise is considered. Specifically, two sinusoidal sources with normalized frequencies 5/16 and 9/16 arriving at 30.0° and 33.0°, respectively, having individual signal to noise ratio (SNR) of 0 dB, are considered. For comparison, the estimated radial slices of the 2-D PSD for spatially uncorrelated array sensor noise case are shown in Figure 4.2, while the result obtained for spatially correlated situation is shown in Figure 4.3. For the result shown in Figure 4.3, the noise at each sensor is correlated with those of two adjacent sensors on either side with correlation coefficient 0.5. There are no significant differences between the two results. Note the good resolving capability in the colored noise situation as well.

4.6.4. Experiment 4

This experiment was conducted to investigate the capability of the proposed technique in handling wide band sources as well as narrow band sources. Figure 4.4 shows the radial slices of the 2-D PSD for a narrow band source with normalized frequency of 11/16 arriving at 30.0° with SNR of -3dB. The DOA estimate obtained is 30.6°. Figure 4.5 shows an example of a wide band source generated using an AR(2) model (coefficients 1.096 and -0.87), and arriving at 30.0° with SNR of 0dB. The DOA estimate obtained is 29.7°.

Finally, Figure 4.6 shows an example of two narrow band sources and a wide band source. The narrow band sources arrive at 14.5° and 15.8° with SNR of 0 dB, while the wide band source is generated using the same AR(2) model as in the

previous experiment, and arrives at 30.0°. The DOA estimates obtained are 14.7°, 16.1° for the narrow band sources and 30.0° for the wide band. Figure 4.7 shows the side view of the radial slices shown in Figure 4.6. Note that the projection angle of three peaks can be clearly seen from this view, and the DOA estimation technique requires locating only this projection angle ψ_k at which the peak occurred.

Table 4.1. The DOA estimates and their RMSE of a single source under the pure Gaussian noise using the Marple's algorithm when the SNR is -9 dB. The array data size is (32x32). Twenty independent experiments were performed.

True θ	Mean	RMSE
15.00°	14.99	0.69
30.00°	29.71	0.83
45.00°	45.76	1.25
60.00°	59.15	1.82

Table 4.2. Comparison of the DOA estimates under different noise environments in a single source (true DOA = 30.0°) case. Also shown are the results from a non-robust method (Marple's algorithm) and a robust method. The array data size is (32x32). Ten independent experiments were performed. a is the ratio of the outlier noise variance *vs.* the dominant Gaussian noise variance, and ε is the fraction of outliers in noise.

			DOA Estimates (RMSE)	
SNR(dB)	a	ε	*marple*	*robust*
9.0	-	0.	29.34 (1.0118)	29.92 (0.9598)
9.0	5.	1.	30.52 (1.9141)	30.50 (1.3146)
6.0	-	0.	29.05 (1.4118)	29.92 (1.0349)
6.0	5.	1.	29.02 (1.6697)	30.46 (1.2323)
3.0	-	0.	29.91 (1.7348)	30.48 (1.4252)
3.0	5.	1.	30.10 (1.6339)	30.89 (1.1260)
0.0	-	0.	30.37 (1.7512)	31.38 (1.7781)
0.0	5.	1.	29.50 (1.8598)	30.07 (1.1675)

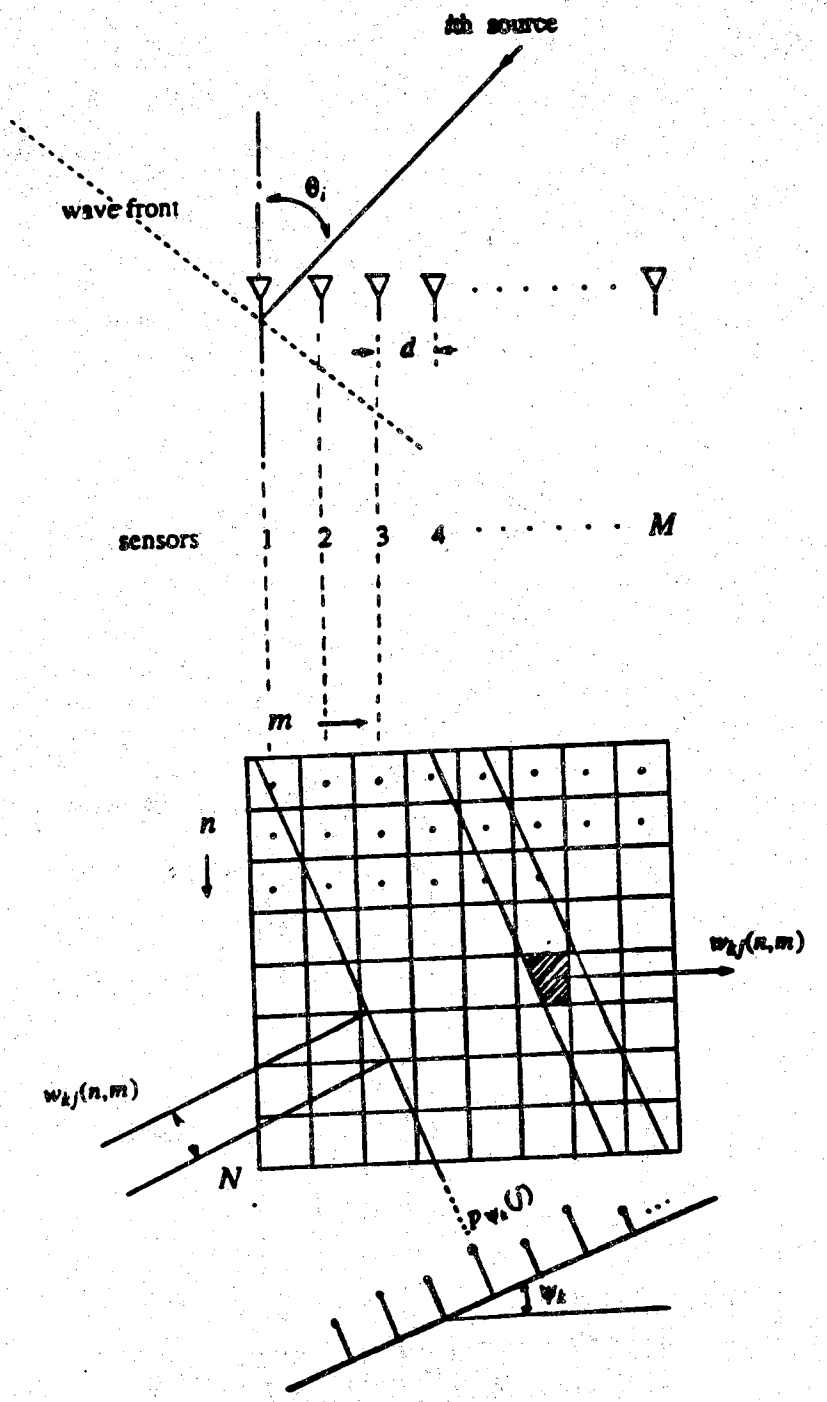


Figure 4.1. A simple sketch of a linear array with uniform spacing d between the sensors. The sensor outputs sampled in time form a 2-D data set $\{y(n,m), n=1,...,N\ ;\ m=1,...,M\}$. Different weights $w_{kl}(n,m)$ can be used to compute the Radon transform of a discrete data set.

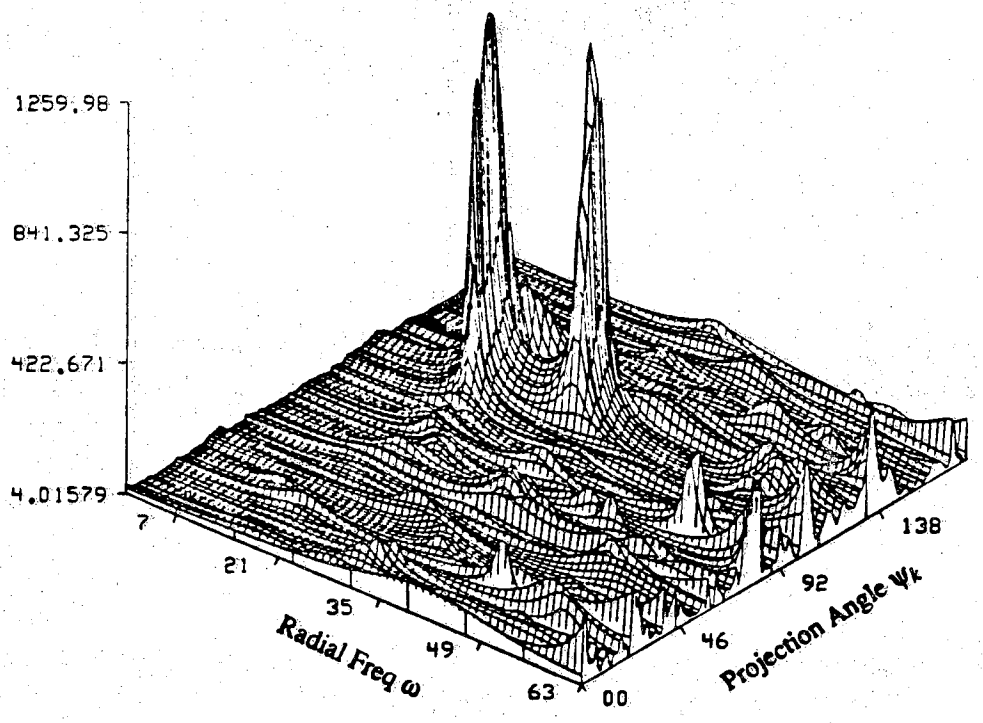


Figure 4.2. *Spatially Uncorrelated Noise Case*: Radial slices of the estimated 2-D PSD, where the AR parameters (order 6) are estimated by the Marple algorithm. The (16X16) 2-D data consists of two sources with normalized frequencies 5/16 and 9/16 arriving at 30.0° and 33.0° with individual signal-to-noise ratio (SNR) of 0 dB. The estimated DOA's are 30.6° and 33.7°.

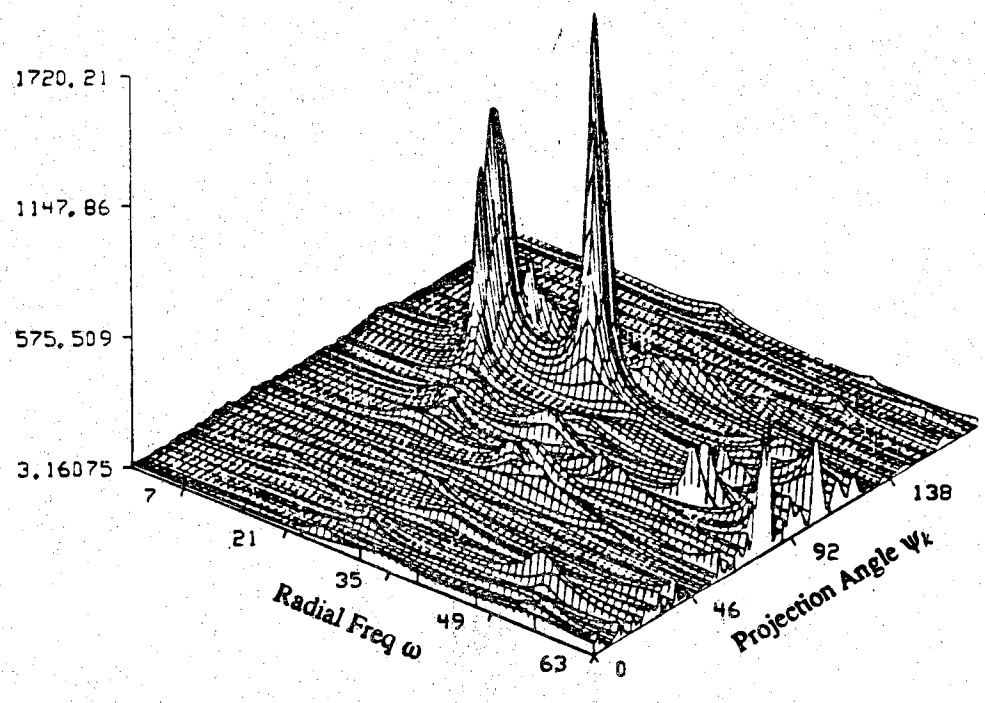


Figure 4.3. *Spatially Correlated Noise Case*: Radial slices of the estimated 2-D PSD. The noise at each sensor is correlated with those of two neighboring sensors with correlation coefficient of 0.5. The AR parameters (order 6) are estimated by the Marple algorithm. The (16X16) 2-D data consists of two sources with normalized frequencies 5/16 and 9/16 and arriving at 30.0° and 33.0° with individual signal-to-noise ratio (SNR) of 0 dB. The estimated DOA's are 30.6° and 33.7° even in this case.

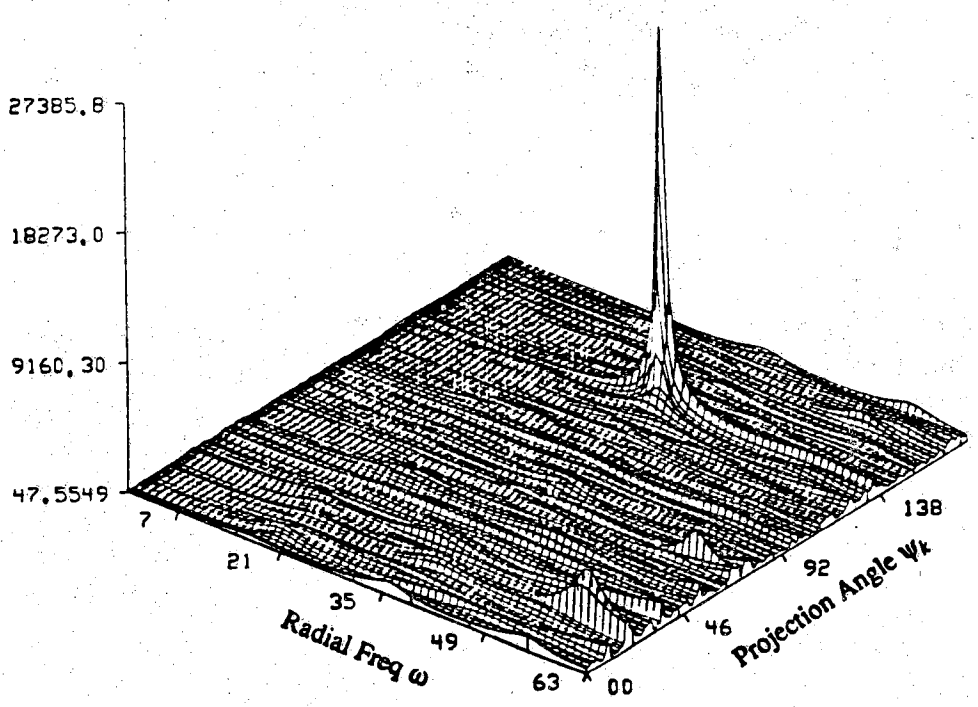


Figure 4.4. Plot of the 2-D PSD estimate obtained using AR (6) for each of the 180 projections displayed slice by slice. The AR parameters are estimated by the Marple algorithm. The (16X16) 2-D data consists of a *narrow band* source with normalized frequency of 11/16 arriving at 30.0° with SNR of -3 dB. The estimated DOA is 29.2°.

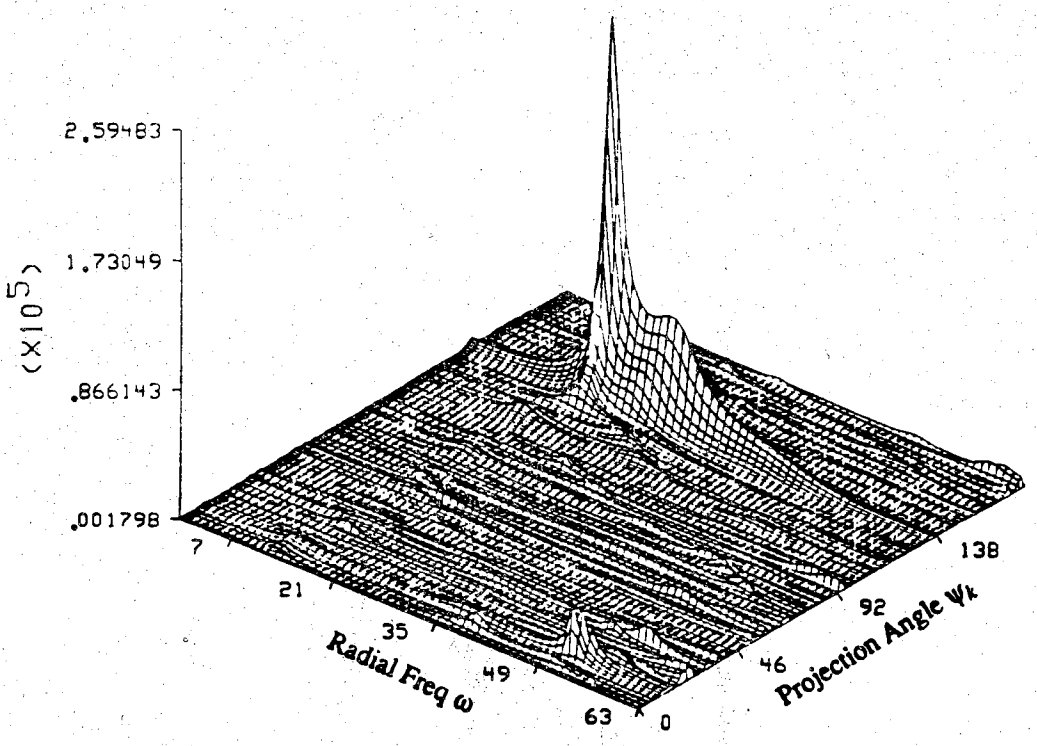


Figure 4.5. Plot of the 2-D PSD estimate obtained using AR (6) for each of the 180 projections displayed slice by slice. The AR parameters are estimated by the Marple algorithm. The (16X16) data consists of a *wide band* source arriving at 30.0°. The DOA estimate obtained is 30.6°.

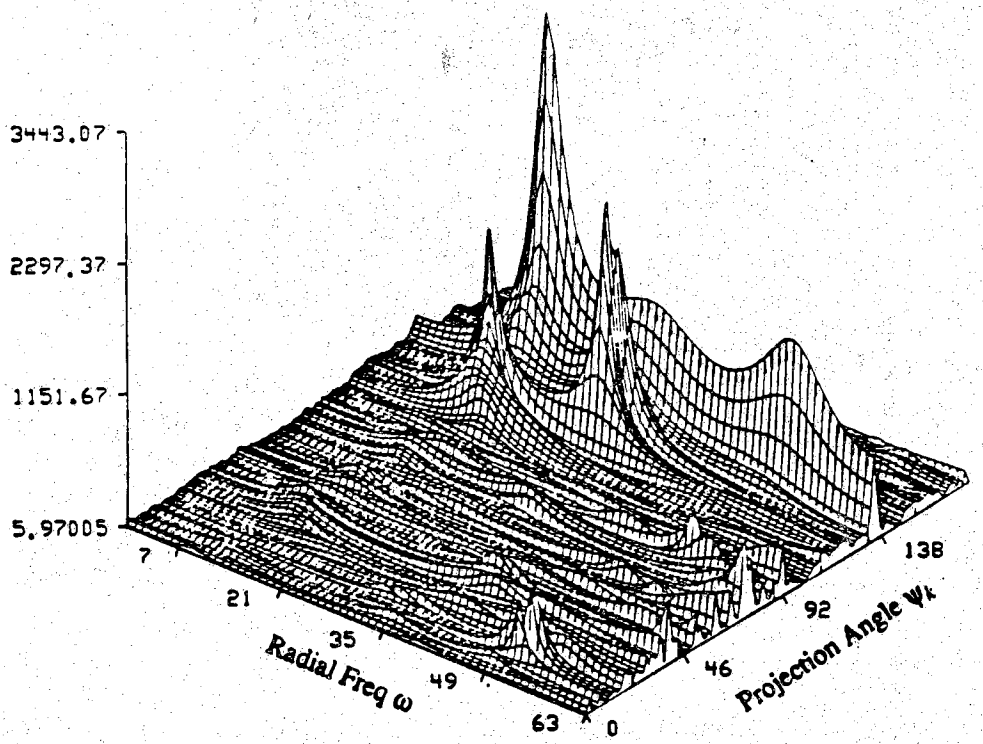


Figure 4.6. Plot of the 2-D PSD estimate obtained using AR (6) for each of the 180 projections displayed slice by slice. The (16X16) 2-D data consists of two *narrow band* sources and a *wide band* source. The *narrow band* sources arrive at 14.5° and 15.8° with SNR of 0 dB. The *wide band* source arrives at 30.0°. The DOA estimates obtained are 14.7°, 16.1° for the narrow band sources and 30.0° for the wide band.

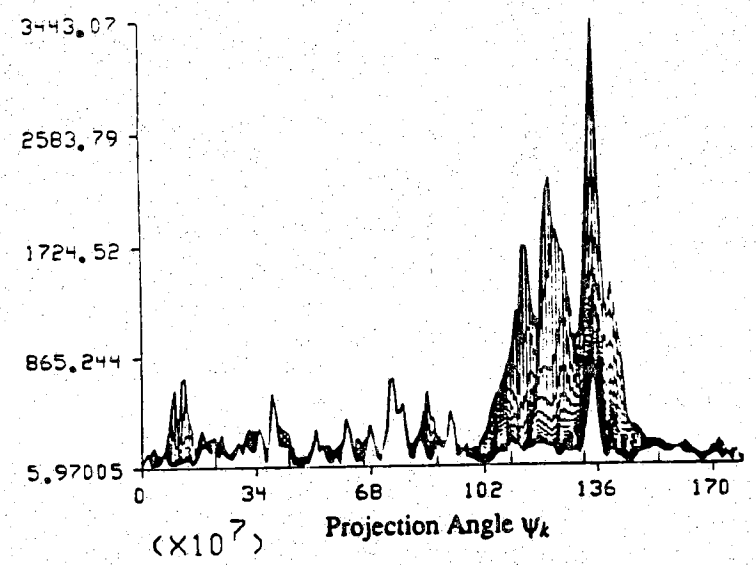


Figure 4.7. Side view of the radial slices shown in Figure 4.6.

4.7. Conclusions

In this chapter a Radon transform approach of robust DOA estimation has been presented. An important aspect of the work presented here is the use of robust 1-D AR parameter estimation method in the Radon space to obtain a 2-D robust spectral estimate. The technique requires no information about the number, type, and frequency of the received signals, and the structure of sensor noise. It is capable of handling narrow band and wide band sources simultaneously at low SNR's, and performs equally well in the presence of colored noise with unknown correlation structure. Though the DOA estimation method presented in this chapter is related to the traditional beamforming, it has much better resolving capability as we use the spectral density, which is in turn estimated by using a model, to measure the average power.

The total number of parameters estimated while computing the 2-D spectrum on a polar raster is quite large. However, all these parameters are not estimated "simultaneously" from the array data. Instead, the Radon transform is used to convert the basic 2-D problem into a set of independent 1-D problems, which can be processed concurrently. Another advantageous feature of this method is that any particular range of directions of interest can be probed. This is particularly useful in tracking applications.

Computer simulation studies demonstrates the performance of the new procedure in accurately estimating DOA in various situations. Rough analysis indicates that the resolution of our method is much higher, nearly double, than that of the traditional beamforming method. Though this is confirmed by simulations, a more detailed theoretical analysis is however required. We have considered a uniform linear array in the present study.

The method can be generalized to the case of a linear array with known non-uniform sensor spacing. In this case the weighting factor $w_{kj}(n,m)$ needs to be chosen accordingly. The discussions and results presented here deal only with single dimensional parameter space, i.e., azimuth only direction finding of far field sources. However, it can be easily generalized to higher dimensional parameter spaces.

CHAPTER 5
ROBUST MAXIMUM LIKELIHOOD DIRECTION-OF-ARRIVAL ESTIMATION

5.1. Introduction

The maximum likelihood (ML) technique was one of the first to be investigated [45]. Because of the high computational load of the multivariate nonlinear maximization problem involved, however, it did not become popular until recently. There are many suboptimal techniques with reduced computational load, but the performance of these techniques is usually inferior to that of the ML technique. As was also pointed out by Ziskind et al. [87], the inferiority is especially conspicuous in the threshold region, namely, when the signal to noise ratio (SNR) is small, or alternatively, when the number of snapshots is small. Moreover, these techniques cannot handle the case of coherent signals. This case appears, for example, in specular multipath propagation problems and, therefore, it is of great practical importance. The preprocessing spatial smoothing techniques proposed to cope with this problem remedy the situation only partially [66].

This chapter specifically consider a robust maximum likelihood (ML) direction-of-arrival (DOA) estimation problem in situations of outlier contaminated Gaussian

noise. Again, the term "robustness" refers to insensitivity against small unknown deviation in the underlying Gaussian noise assumption. Even a small deviation from the assumed Gaussian noise model can create havoc with Gaussian ML estimates since the Gaussian ML estimators are extremely sensitive to outliers. The Gaussian ML estimation scheme, which were developed and tested under the Gaussian assumption, usually fail to resolve close DOA's when there are just one or two *outliers* out of one hundred observed sensor array snapshots.

The DOA's are estimated by a robust technique based on the so called M-estimators, a generalization of classical ML estimator by Huber [27]. Performances of the estimator in both the Gaussian and outlier contaminated Gaussian noise are evaluated using the Cramer Rao Lower Bound (CRLB) and the variance derived from the Influence Function (IF). The organization of the chapter is as follows. Section 5.2 formulates the Gaussian ML DOA estimation problem, followed by Section 5.3 which shows the formulation of the robust estimation problem. Section 5.4 then shows the details of the robust ML DOA estimation scheme. Section 5.5 then presents some of the analysis carried out to compare the performance of the robust algorithm with that of the Gaussian ML estimation algorithm, followed by concluding remarks in section 5.6.

5.2. Problem Formulation

Consider a linear array composed of M identical equi-spaced sensors. It is assumed that there are q coherent or incoherent narrow band sources, centered around a known frequency with wavelength λ, impinge on the array from directions $\theta_1,...,\theta_q$. Since narrow-bandness in the sensor array means that the propagation delays of the signals along the array are much smaller than the reciprocal of the bandwidth of the

signals, the envelopes of the signals received by the array can be expressed by the following.

Let $x_j(t_i)$ and $u_j(t_i)$ denote the ith snapshots of the jth sensor output and noise, respectively. Then $x_j(t_i)$ can be expressed as

$$x_j(t_i) = \sum_{k=1}^{q} \alpha_k \sin[\omega_k t_i + (2\pi/\lambda)(j-1)\, d\, \sin\theta_k] + u_j(t_i)$$

$$i=1,...,N\ ;\ j=1,...,M, \tag{5.2.1}$$

where N is the total number of snapshots, M is the number of sensors in the array, d is the spacing between sensors, and θ_k's are the unknown DOA's.

It is well known by the principle of parsimony that the accuracy of the parameter estimates decreases when the number of unknown parameters to be estimated increases. In order to minimize the number of unknown parameters to be estimated, we assume that α_k's and ω_k's, the amplitudes and frequencies of the envelopes associated with the sources in equation (5.2.1), are known quantities even though they can be estimated simultaneously using this technique. Note that α_k's and ω_k's can also be estimated using a robust technique proposed by Oh et al. [53].

Furthermore, in coherent multipath problems, one can rewrite equation (5.2.1) as

$$x_j(t_i) = \sum_{k=1}^{q} \sum_{l=1}^{L_k} \alpha_{kl} \sin[\omega_k t_i + (2\pi/\lambda)(j-1)\, d\, \sin\theta_{kl} + \nu_{kl}] + u_j(t_i)\,,$$

where α_{kl}, θ_{kl}, and ν_{kl} are the amplitude, DOA, and phase of the signal envelope arriving from the kth source via lth path, respectively. L_k is the number of different paths the kth source signal take.

The vector of the received signals $X(t_i)$ can be expressed as

$$X(t_i) = S(t_i) + U(t_i) \qquad i=1,...,N \tag{5.2.2}$$

where

$$X(t_i) = col.[\, x_1(t_i), x_2(t_i), \ldots, x_M(t_i) \,],$$

$$S(t_i) = col.[\, s_1(t_i), s_2(t_i), \ldots, s_M(t_i) \,],$$

and

$$U(t_i) = col.[\, u_1(t_i), u_2(t_i), \ldots, u_M(t_i) \,],$$

$$s_j(t_i) = \sum_{k=1}^{q} \alpha_k \sin\,[\omega_k t_i + (2\pi/\lambda)(j-1)\, d\, \sin\theta_k],$$

$$i = 1, \ldots, N \;;\; j = 1, \ldots, M,$$

and

$$U(t_i) \sim N(0, \rho I)\,. \tag{5.2.3}$$

The joint density function of the sampled data is given by

$$f[X(t_1), \ldots, X(t_N); \Theta]$$

$$= \frac{1}{(2\pi)^{MN/2} (\det[\rho I])^{N/2}} \exp\left\{ -\frac{1}{2\rho} \sum_{i=1}^{N} Y^T(t_i;\Theta) Y(t_i;\Theta) \right\} \tag{5.2.4}$$

where

$$\Theta = col.[\theta_1, \ldots, \theta_q],$$

$$Y(t_i;\Theta) = col.[y_1(t_i;\Theta), \ldots, y_M(t_i;\Theta)],$$

and

$$y_j(t_i;\Theta) = x_j(t_i) - \sum_{k=1}^{q} \alpha_k \sin\,[\omega_k t_i + \pi\,(j-1)\, \sin\theta_k]. \tag{5.2.5}$$

Note that the spacing d has been chosen to be exactly half of wavelength λ for simplicity. Thus, the log likelihood, ignoring constant terms, is given by

$$L(\Theta) = -\frac{MN}{2} \log \rho - \frac{1}{2\rho} \sum_{i=1}^{N} Y^T(t_i;\Theta) Y(t_i;\Theta). \tag{5.2.6}$$

To compute the maximum likelihood (ML) estimator we have to maximize the log likelihood with respect to Θ and ρ. Fixing the Θ, and maximizing with respect to ρ, we get

$$\hat{\rho} = \frac{1}{MN} \sum_{i=1}^{N} Y^T(t_i;\Theta) Y(t_i;\Theta). \tag{5.2.7}$$

Substituting this back to equation (5.2.6), ignoring constant terms, we can obtain the ML estimates of θ_k's by maximizing

$$-MN \log \left[\frac{1}{MN} \sum_{i=1}^{N} Y^T(t_i;\Theta) Y(t_i;\Theta) \right].$$

Since the logarithm is a monotonic function, the estimates can be obtained by

$$\min_{\Theta} \sum_{i=1}^{N} Y^T(t_i;\Theta) Y(t_i;\Theta).$$

This can be rewritten as

$$\min_{\Theta} \sum_{i=1}^{N} \sum_{j=1}^{M} \left\{ x_j(t_i) - \sum_{k=1}^{q} \alpha_k \sin [\omega_k t_i + \pi (j-1) \sin\theta_k] \right\}^2. \tag{5.2.8}$$

5.3. Robust Direction-of-Arrival Estimation

In the above formulation, the estimates obtained by equation (5.2.8) are ML estimates if the underlying noise distribution is exactly Gaussian. In practical situations, however, the commonly made assumption in statistics is at most approximation to reality. An *outlier* in a set of data is defined to be an observation which appears to be inconsistent with the remainder of that set of data. It may not be hard to spot the

potentially troublesome data point in the lower dimension, but it becomes exceedingly difficult with higher dimensions, or with multiparameter problems.

For this reason, minor deviations from the Gaussian noise are often modeled by the mixture model for noise [80]. The mixture model of interest as introduced in the previous chapter is the slippage model with the Gaussian distribution as the dominant distribution. If $U(t_i)$ is a sequence of random variables obeying such a slippage model, then any $U(t_i)$ is distributed either as a Gaussian distribution $N(0,\sigma^2)$ with probability $(1-\varepsilon)$, or as an unknown distribution $Q(\mu, \beta\sigma^2)$ with probability ε, where μ and $\beta\sigma^2$ are the mean and the variance of the unknown distribution Q. In general, $\beta>1$, $0<\varepsilon< 0.1$ and, μ, an unknown constant, is of the order of a multiple of σ. The noise distribution can then be expressed as

$$P\{U(t_i)\} \sim (1-\varepsilon)N(0,\sigma^2)+\varepsilon Q(\mu, \beta\sigma^2), \qquad (5.3.1)$$

which represents a family of distributions characterized by the mixing parameter ε. For $\varepsilon=0$, (5.3.1) reduces to a Gaussian distribution.

The parameter vector, Θ, is estimated by a robust technique based on the so called M-estimators, a generalization of classical maximum likelihood (ML) estimator by Huber [27]. In order to enforce scale invariance, the Θ and a scale parameter σ are estimated simultaneously, and this can be done by minimizing

$$J(\Theta,\sigma) = \sum_{i=1}^{N}\sum_{j=1}^{M} [\rho(y_j(t_i;\Theta)/\sigma)\,\sigma] + a\sigma\,, \qquad \sigma > 0 \qquad (5.3.2)$$

with respect to Θ and σ, where $y_j(t_i;\Theta)$ and $\rho(\cdot)$ are defined by (5.2.5) and (5.3.4) respectively. The constant a, given by (5.3.7), is chosen to make the estimates consistent at the nominal distribution. Unless the minimum of $J(\Theta,\sigma)$ occurs on the boundary $\sigma=0$, it can be equivalently characterized by the $q+1$ equations

$$\sum_{i=1}^{N}\sum_{j=1}^{M} \psi(y_j(t_i)/\sigma)\,[\partial y_j(t_i;\Theta)/\partial\theta_k] = 0; \qquad k=1,\ldots,q$$

$$\sum_{i=1}^{N}\sum_{j=1}^{M} \chi(y_j(t_i;\Theta)/\sigma) = a\ , \tag{5.3.3}$$

where

$$\rho(x) = \begin{cases} x^2/2 & ,\ \text{if } |x| \le c \\ c|x| - c^2/2 & ,\ \text{if } |x| > c \end{cases} \tag{5.3.4}$$

$$\psi(x) = (\partial\rho(x)/\partial x) \tag{5.3.5}$$

$$\chi(x) = x\,\psi(x) - \rho(x), \tag{5.3.6}$$

$$a = MN \int_{-\infty}^{\infty} \chi(x)\,(1/\sqrt{2\pi})\exp[-x^2/2]\,dx. \tag{5.3.7}$$

and $y_j(t_i;\Theta)$ is as defined in equation (5.2.5). The constant c is related to the fraction of contamination ε by

$$2\Phi(c) - 1 + 2\phi(c)/c = 1/(1-\varepsilon), \tag{5.3.8}$$

where $\Phi(c)$ is the standard cumulative Gaussian distribution with zero mean and unit variance, and $\phi(c)$ is the corresponding Gaussian density. Usually c is chosen to lie between 1 and 2, which corresponds to the ε-interval [0.0083, 0.1428] by (5.3.8). We use the following algorithm to solve the $q+1$ equations of (5.3.3) since its convergence properties are well established [27].

5.4. Robust Estimation Procedure

Given $x_j(t_i)'s$,$i=1,..,N$; $j=1,..,M$, solve equation (5.3.3) for Θ and σ.

step 1. Choose starting values $m=0$, $\Theta^{(m)}$, $\sigma^{(m)}$, and a tolerance value $\zeta > 0$.

step 2. Compute residuals.

$$y_j(t_i;\Theta)^{(m)} = x_j(t_i) - \sum_{k=1}^{q} \alpha_k \sin[\omega_k t_i + \pi\,(j-1)\,\sin\theta_k{}^{(m)}] \quad , \quad i=1,..,N\ ;\ j=1,..,M.$$

step 3. Compute a new value of $\sigma^{(m)}$.

$$[\sigma^{(m+1)}]^2 = (1/a) \sum_{i=1}^{N} \sum_{j=1}^{M} \chi(y_j(t_i;\Theta)^{(m)}/\sigma^{(m)})\,[\sigma^{(m)}]^2.$$

step 4. "Winsorize" the residuals.

$$Z = col.[z_{11},..,z_{1M},.....,z_{N1},..,z_{NM}]$$

where

$$z_{ij} = \psi(y_j(t_i;\Theta)^{(m)}/\sigma^{(m+1)})\,\sigma^{(m+1)} \quad ;\quad i=1,..,N\ ;\ j=1,..,M.$$

step 5. Compute the partial derivatives.

$$b_{mk} = (\frac{\partial}{\partial\theta_k}) \left\{ \sum_{k=1}^{q} \alpha_k \sin\,[\omega_k t_i + \pi(j-1)\,\sin\theta_k] \right\} \quad , m=1,...,NM\ ;\ k=1,...,q\ .$$

step 6. Solve $[B^T B]\tau = B^T Z$ for τ, where B is a NM by q matrix whose elements are b_{mk}'s.

step 7. Update $\Theta^{(m)}$ by $\Theta^{(m+1)} = \Theta^{(m)} + \xi\tau$, where $0 < \xi < 2$ is an arbitrary relaxation factor.

step 8. Stop iterating and go to step 9 if the parameters change by less than ζ times their standard deviation, where ζ is an arbitrarily chosen small value, i.e., if for all j,

$$|t_j| < \zeta\,[\bar{b}_{jj}]^{1/2}\,\sigma^{(m+1)}$$

$\bar{b}_{jj}$ is the jth diagonal element of the matrix $\bar{B} = (B^T B)^{-1}$; otherwise $m = m + 1$ and go to step 2.

step 9. The final estimate of Θ_0 is given by $\Theta^{(m+1)}$, and the variance of the residual sequence by $[\sigma^{(m+1)}]^2$.

5.5. Performance Analysis

Let us first compute the Cramer Rao Lower Bound (CRLB) for a Gaussian noise case, and then for a contaminated Gaussian noise case to observe the changes in the CRLB with small changes in the Gaussian noise distribution. The variances of the robust estimates are then obtained from Influence Function (IF), which was introduced by Hampel [22] for investigating the infinitesimal behavior of real-valued functionals on a heuristic basis. Details of the derivation can be found in [22,52]. Finally, the resolution capability of the robust algorithm is investigated.

Throughout the experiment two incoherent or coherent emitters impinging from 30° and 33° are simulated with snapshots taken from eight equispaced linear array of sensors. For the Gaussian noise case the noise is generated from a zero mean Gaussian density, and for the contaminated noise setting, the unknown distribution Q in equation (5.3.1) is assumed to be a zero mean Gaussian with ten times the variance of the dominant Gaussian. The percentage of outliers, ε, is chosen to be five.

The amplitudes of the envelopes associated with the sources are chosen to be α_1=1, α_2=1. For the uncorrelated sources the frequencies of the envelopes associated with the sources are chosen to be $\omega_1 = 10\pi$ and $\omega_2 = 20\pi$. On the other hand, $\omega_1 = 10\pi$ and $\omega_2 = 10\pi$ are chosen for the case of coherent sources. One hundred

snapshots were taken unless specified otherwise. Less than ten iterations were required for either the Gaussian ML estimation or the robust estimation before convergence was achieved.

5.5.1. Cramer Rao Lower Bound (CRLB) for Gaussian Noise Case

Let us denote

$$A_1 = \frac{1}{(2\pi)^{M/2}(det[\rho I])^{N/2}}$$

and

$$A_2(t_i;\Theta) = -\frac{1}{2\rho}\sum_{j=1}^{M}\left\{x_j(t_i) - \sum_{k=1}^{q}\alpha_k \sin[\omega_k t_i + \pi(j-1)\sin\theta_k]\right\}^2. \qquad (5.5.1)$$

Then the density function can be written as

$$f[X(t_1),...,X(t_N);\Theta] = \prod_{i=1}^{N} f[X(t_i);\Theta], \qquad (5.5.2)$$

where

$$f[X(t_i);\Theta] = A_1 \exp[A_2(t_i;\Theta)].$$

Therefore,

$$\frac{\partial}{\partial\Theta}L(\Theta) = \frac{\partial}{\partial\Theta}\log\left\{f[X(t_1),...,X(t_N);\Theta]\right\}$$

$$= \sum_{i=1}^{N}\frac{\partial}{\partial\Theta}\log\left\{f[X(t_i);\Theta]\right\}$$

$$= \sum_{i=1}^{N} (1/f\,[X(t_i);\Theta]\,)\,\frac{\partial}{\partial\Theta} f\,[X(t_i);\Theta]$$

$$= \sum_{i=1}^{N} \left\{ \frac{\partial}{\partial\Theta} A_2(t_i;\Theta) \right\}. \qquad (5.5.3)$$

Then the Fisher Information Matrix can be computed as

$$I(\Theta) = \int [\frac{\partial}{\partial\Theta} L(\Theta)][\frac{\partial}{\partial\Theta} L(\Theta)]^T \; dF \;, \qquad (5.5.4)$$

and the covariance matrix of the parameter vector Θ is given by

$$Cov(\Theta) = [I(\Theta)]^{-1} \;. \qquad (5.5.5)$$

5.5.2. Cramer Rao Lower Bound (CRLB) for Contaminated Gaussian Noise Case

Since the distribution Q in equation (5.3.1) is a Gaussian with zero mean and η times the variance of the dominant Gaussian,

$$f\,[X(t_i);\Theta] = (1-\varepsilon)A_1 exp\,[\,A_2(t_i;\Theta)] + \varepsilon\frac{A_1}{\eta^{M/2}} exp\,[\,A_2(t_i;\Theta)/\eta\,].$$

$$= A_1 \exp\,[A_2(t_i;\Theta)]\,A_3(t_i;\Theta), \qquad (5.5.6)$$

where A_1, $A_2(t_i;\Theta)$ are defined in equation (5.5.1), and

$$A_3(t_i;\Theta)=1-\varepsilon+\varepsilon\eta^{-\frac{M}{2}} exp\,[(\frac{1}{\eta}-1)A_2(t_i;\Theta)]. \qquad (5.5.7)$$

Hence

$$\frac{\partial}{\partial\Theta} L(\Theta) = \sum_{i=1}^{N} \frac{\partial}{\partial\Theta} \log\left(f[X(t_i);\Theta]\right)$$

$$= \sum_{i=1}^{N} \left[\left\{ \frac{\partial}{\partial\Theta} A_2(t_i;\Theta) \right\} + \left\{ \frac{\partial}{\partial\Theta} A_3(t_i;\Theta) \right\} / A_3(t_i;\Theta) \right] \qquad (5.5.8)$$

The Fisher Information Matrix and the covariance matrix of the parameter vector thus can be computed by (5.5.4) and (5.5.5), respectively.

Figure 5.1 shows the plot of CRLB's *vs.* SNR for a pure Gaussian and a contaminated Gaussian noise case. The simulation involves two uncorrelated emitters impinging from 30° and 33°, with one hundred snapshots taken from eight equi-spaced linear array of sensors. Note the approximately forty percent increase in the CRLB caused by replacing five percent of the Gaussian data with outliers that has ten times the variance of the parent Gaussian distribution.

5.5.3. Variance of the Robust Estimates

The expression for the variance of robust estimates can be obtained using Influence Function (IF), where the details of the derivation can be found in [52]. In our case, the covariance matrix of the robust DOA estimates are given by

$$Var(\hat{\Theta}) = \frac{ave\left[\psi^2(r(i,j;\hat{\Theta}))\right]}{\left\{ave\left[\psi'(r(i,j;\hat{\Theta}))\, r(i,j;\hat{\Theta})\right]\right\}^2} \left[\sum_{i=1}^{N}\sum_{j=1}^{M} V(i,j;\hat{\Theta})\, V^T(i,j;\hat{\Theta})\right]^{-1} \qquad (5.5.9)$$

where

$$r(i,j;\hat{\Theta}) = y_j(t_i;\Theta)/\hat{\sigma},$$

$$V(i,j;\hat{\Theta}) = -\partial/\partial\Theta r(i,j;\hat{\Theta}),$$

and

$$\psi'(x) = \partial/\partial x\ \psi(x).$$

$y_j(t_i)$ and $\psi(x)$ are defined in (5.2.5) and (5.3.5), respectively. Here $\hat{\Theta}$ and $\hat{\sigma}$ indicates the estimated value of Θ and σ.

Figure 5.2 shows the theoretical RMSE (sum of bias square and the variance derived above) *vs.* SNR plot of the Gaussian ML estimates and the robust estimates when the underlying noise distribution is purely Gaussian, whereas Figure 5.3 shows the theoretical RMSE *vs.* SNR plot of the Gaussian ML estimates and the robust estimates after replacing five percent of the Gaussian data with outliers that has ten times the variance of the original Gaussian distribution. Figure 5.4 and Figure 5.5 shows the results of similar experiments shown in Figure 5.2 and Figure 5.3 but with coherent sources.

Figure 5.6 shows the theoretical RMSE *vs.* Snapshots plot of the Gaussian ML estimates and the robust estimates when the underlying noise distribution is purely Gaussian, and Figure 5.7 shows the plot after replacing five percent of the Gaussian data with outliers that has ten times the variance of the original Gaussian distribution. These plots were made for the fixed signal to noise ratio (SNR) of 12dB.

Table 5.1 and Table 5.2 shows the twenty-run averages of $\hat{\theta}_1$ and $\hat{\theta}_2$, and the resulting RMSE's computed for different values of SNR for the Gaussian noise case and for the contaminated Gaussian noise case, respectively. On the other hand, Table 5.3 and Table 5.4 shows the results for different numbers of snapshots taken.

Note that the performance of the robust estimator is almost as good as that of the Gaussian ML estimator in the Gaussian noise, but much better in the presence of outliers regardless of whether the sources are incoherent or coherent.

5.5.4. Testing the Presence of Two Closely Spaced Sources

In the context of direction-of-arrival (DOA) estimation, resolution usually refers to the ability of an algorithm in resolving two closely spaced sources. Thus, in estimating the DOA's of closely spaced sources, an important problem is determining whether there is just one source or there are two dominant sources with very close DOA's.

Suppose the estimates of the two unknown DOA's, $\theta^0{}_1$ and $\theta^0{}_2$, are obtained from q independent experiments, i.e., $\{\hat{\theta}_1^{(1)},\hat{\theta}_2^{(1)}\}$,..., $\{\hat{\theta}_1^{(q)},\hat{\theta}_2^{(q)}\}$. Let the difference between the two estimates from an arbitrary experiment, $\hat{\theta}_1-\hat{\theta}_2$, obeys a Gaussian distribution with unknown mean μ and variance σ^2, i.e.,

$$\hat{\theta}_1-\hat{\theta}_2 \sim N(\mu,\sigma^2)\,. \tag{5.5.10}$$

It is also important to be reminded that each of the robust DOA estimates, $\hat{\theta}_1$ and $\hat{\theta}_2$, obeys a Gaussian distribution with unknown mean and variance even when the noise process at the array sensors is an outlier contaminated Gaussian. Also define

$$R_0{}^2 = \sum_{k=1}^{q}\left\{(\hat{\theta}_1^{(k)}-\hat{\theta}_2^{(k)})-\frac{1}{q}\sum_{k=1}^{q}(\hat{\theta}_1^{(k)}-\hat{\theta}_2^{(k)})\right\}^2 . \tag{5.5.11}$$

Then

$$R_0{}^2 \sim \sigma^2\chi^2(q-1)\,, \tag{5.5.12}$$

and it can be easily shown that $\hat{\theta}_1-\hat{\theta}_2$, the difference between the two estimates from any of the q independent experiments, and $R_0{}^2$ are independent. Therefore, if $s^2 = R_0{}^2/(q-1)$, then

$$t = \left[\frac{(\hat{\theta}_1 - \hat{\theta}_2) - \mu}{\sigma}\right] / \left[\frac{s}{\sigma}\right] = \frac{(\hat{\theta}_1 - \hat{\theta}_2) - \mu}{s} \sim S(q-1) \qquad (5.5.13)$$

which is Student's distribution on $(q-1)$ degrees of freedom.

If t_α is the α probability point of $|t|$, that is, $P(|t| > t_\alpha) = \alpha$, then

$$P\left[\frac{|(\hat{\theta}_1 - \hat{\theta}_2) - \mu|}{s} \leq t_\alpha\right] = 1 - \alpha, \qquad (5.5.14)$$

that is

$$P\left[(\hat{\theta}_1 - \hat{\theta}_2) - s\,t_\alpha \leq \mu \leq (\hat{\theta}_1 - \hat{\theta}_2) + s\,t_\alpha\right] = 1 - \alpha. \qquad (5.5.15)$$

Let the null hypothesis be $\mu = 0$, i.e., $E[\hat{\theta}_1] = E[\hat{\theta}_2]$. If there is no *a priori* information as to whether the true value of θ_1 is greater than that of θ_2 or otherwise, a large value of t in either direction would be evidence against the null hypothesis. If the null hypothesis is true, then using (5.5.5), we have

$$P\left[\frac{|\hat{\theta}_1 - \hat{\theta}_2|}{s} \leq t_\alpha\right] = 1 - \alpha. \qquad (5.5.16)$$

The null hypothesis, which assumes the presence of only one source, is thus rejected at α level of significance if, for observed $\hat{\theta}_1$ and $\hat{\theta}_2$, we have

$$\frac{|\hat{\theta}_1 - \hat{\theta}_2|}{s} > t_\alpha. \qquad (5.5.17)$$

When there are actually two closely spaced sources, rejecting the null hypothesis is equivalent to resolving the two closely spaced dominant sources at α level of significance.

The following simulations have been carried out to compare such resolving capability of the proposed robust DOA estimation algorithm with that of the Gaussian ML DOA estimation scheme. The resolution probability, i.e., the probability that the

two sources with close DOA's are identified as two separate sources, have been estimated at $1-\alpha$ level of significance, where α is chosen to be 0.05. The resolution probability thus can be estimated by taking the number of experiments with successful resolution divided by the total number of experiments. Forty one independent experiments, i.e., q=41, have been performed, each with twenty five snapshots taken from eight equispaced sensors.

Figure 5.8 shows the estimated resolution probability *vs.* SNR plot of the Gaussian ML estimates and the robust estimates when the underlying noise distribution is purely Gaussian, whereas Figure 5.9 shows a similar plot after replacing five percent of the Gaussian data with outliers which has ten times the variance of the parent Gaussian distribution. The true DOA's are chosen as 30.0° and 30.2°.

Figure 5.10 shows the estimated resolution probability *vs.* angular separation between the two sources when the underlying noise distribution is purely Gaussian, whereas Figure 5.11 shows the corresponding plot after replacing five percent of the Gaussian data with outliers. Here, the SNR is fixed at 18dB, and the smaller value of the two true DOA's is chosen as 30.0°.

Again, the resolving capability of the robust DOA estimation algorithm can be seen almost as good as that of the Gaussian ML estimation scheme in the pure Gaussian noise, and much better in the presence of outliers.

Table 5.1. *Gaussian Noise Case*: Twenty-run averages of θ_1, θ_2, and the RMSE's (shown in parentheses) for different values of SNR. The Gaussian ML DOA estimates and the robust DOA estimates are shown.

Average of the Estimates (& RMSE)				
	Gaussian ML		robust	
SNR(dB)	θ_1	θ_2	θ_1	θ_2
21.0	30.0059 (0.0165)	33.0038 (0.0132)	30.0057 (0.0162)	33.0037 (0.0130)
18.0	30.0034 (0.0165)	32.9985 (0.0291)	30.0035 (0.0169)	32.9997 (0.0289)
15.0	29.9980 (0.0305)	33.0053 (0.0356)	29.9976 (0.0302)	33.0051 (0.0362)
12.0	30.0026 (0.0485)	32.9894 (0.0529)	30.0015 (0.0480)	32.9901 (0.0536)
9.0	30.0082 (0.0719)	33.0152 (0.0477)	30.0069 (0.0719)	33.0142 (0.0483)
6.0	29.9928 (0.0978)	33.0053 (0.0726)	29.9912 (0.0949)	33.0039 (0.0709)
3.0	29.9876 (0.1195)	32.9881 (0.1356)	29.9851 (0.1165)	32.9899 (0.1392)
0.0	30.0228 (0.1485)	33.0402 (0.1812)	30.0232 (0.1499)	33.0457 (0.1798)
-3.0	29.8938 (0.2455)	32.8686 (0.2760)	29.8938 (0.2429)	32.8659 (0.2822)
-6.0	29.8448 (0.3432)	32.9578 (0.2846)	29.8554 (0.3343)	32.9761 (0.2886)
-9.0	29.8168 (0.4684)	33.1310 (0.5434)	29.8188 (0.4711)	33.1321 (0.5394)

Table 5.2. *Contaminated Gaussian Noise Case*: Twenty-run averages of θ_1, θ_2, and the RMSE's (shown in parentheses) for different values of SNR. The Gaussian ML DOA estimates and the robust DOA estimates are shown. The contamination is caused by replacing five percent of the Gaussian data with outliers which has ten times the variance of the parent Gaussian distribution.

Average of the Estimates (& RMSE)				
	Gaussian ML		robust	
SNR(dB)	θ_1	θ_2	θ_1	θ_2
21.0	29.9999 (0.0361)	32.9815 (0.0369)	29.9989 (0.0206)	32.9876 (0.0191)
18.0	29.9948 (0.0588)	33.0104 (0.0501)	29.9911 (0.0292)	33.0017 (0.0243)
15.0	29.9630 (0.0983)	32.9521 (0.0634)	29.9917 (0.0465)	32.9925 (0.0327)
12.0	30.0321 (0.1124)	32.9800 (0.1186)	30.0056 (0.0478)	32.9923 (0.0508)
9.0	29.9090 (0.1624)	33.0654 (0.1718)	29.9665 (0.0864)	33.0143 (0.0640)
6.0	30.0149 (0.2135)	32.9667 (0.2125)	30.0035 (0.1093)	32.9767 (0.0951)
3.0	30.0328 (0.3311)	33.0701 (0.4286)	30.0086 (0.1641)	33.0092 (0.1797)
0.0	29.9624 (0.4844)	33.0048 (0.3854)	30.0084 (0.1961)	33.0521 (0.1982)
-3.0	30.3364 (0.5868)	32.9346 (0.5440)	30.1757 (0.2838)	32.9839 (0.2833)
-6.0	29.8573 (0.8512)	32.6346 (0.9258)	29.9833 (0.4219)	32.8909 (0.4585)
-9.0	29.7591 (1.1717)	32.8647 (1.3988)	30.0571 (0.5535)	31.9184 (0.7358)

Table 5.3. *Gaussian Noise Case*: Twenty-run averages of θ_1, θ_2, and the RMSE's (shown in parentheses) for different Number of Snapshots taken. The Gaussian ML DOA estimates and the robust DOA estimates are shown.

Average of the Estimates (& RMSE)				
	Gaussian ML		robust	
Snapshots	θ_1	θ_2	θ_1	θ_2
1	30.0999 (0.4354)	33.0667 (0.5099)	30.0996 (0.4351)	33.0680 (0.5095)
5	29.8950 (0.4212)	33.0322 (0.2535)	29.8911 (0.2424)	33.0297 (0.2525)
10	29.9847 (0.1547)	32.9855 (0.1400)	29.9806 (0.1567)	32.9866 (0.1393)
15	29.9842 (0.0903)	33.0207 (0.0874)	29.9831 (0.0924)	33.0195 (0.0846)
20	29.9966 (0.0644)	33.0191 (0.0989)	29.9958 (0.0671)	33.0181 (0.0996)
25	29.9906 (0.0865)	32.9689 (0.1016)	29.9984 (0.0883)	32.9680 (0.1006)

Table 5.4. *Contaminated Gaussian Noise Case*: Twenty-run averages of θ_1, θ_2, and the RMSE's (shown in parentheses) for different values of SNR. The Gaussian ML DOA estimates and the robust DOA estimates are shown. The contamination is caused by replacing five percent of the Gaussian data with outliers which has ten times the variance of the parent Gaussian distribution.

Average of the Estimates (& RMSE)				
	Gaussian ML		robust	
Snapshots	θ_1	θ_2	θ_1	θ_2
1	29.5093 (2.8417)	32.4328 (1.6070)	30.0240 (0.5264)	32.6880 (1.1279)
5	29.9498 (0.4242)	32.7784 (0.4779)	29.9931 (0.2278)	32.9358 (0.2524)
10	30.0399 (0.3361)	33.1074 (0.3920)	30.0200 (0.1607)	33.0312 (0.1576)
15	30.0741 (0.3130)	33.0654 (0.2424)	29.9872 (0.1500)	33.0070 (0.1396)
20	29.9875 (0.2798)	32.9387 (0.2257)	29.9922 (0.1343)	32.9627 (0.1107)
25	30.0410 (0.2003)	33.0242 (0.1598)	30.0199 (0.0916)	33.0434 (0.1051)

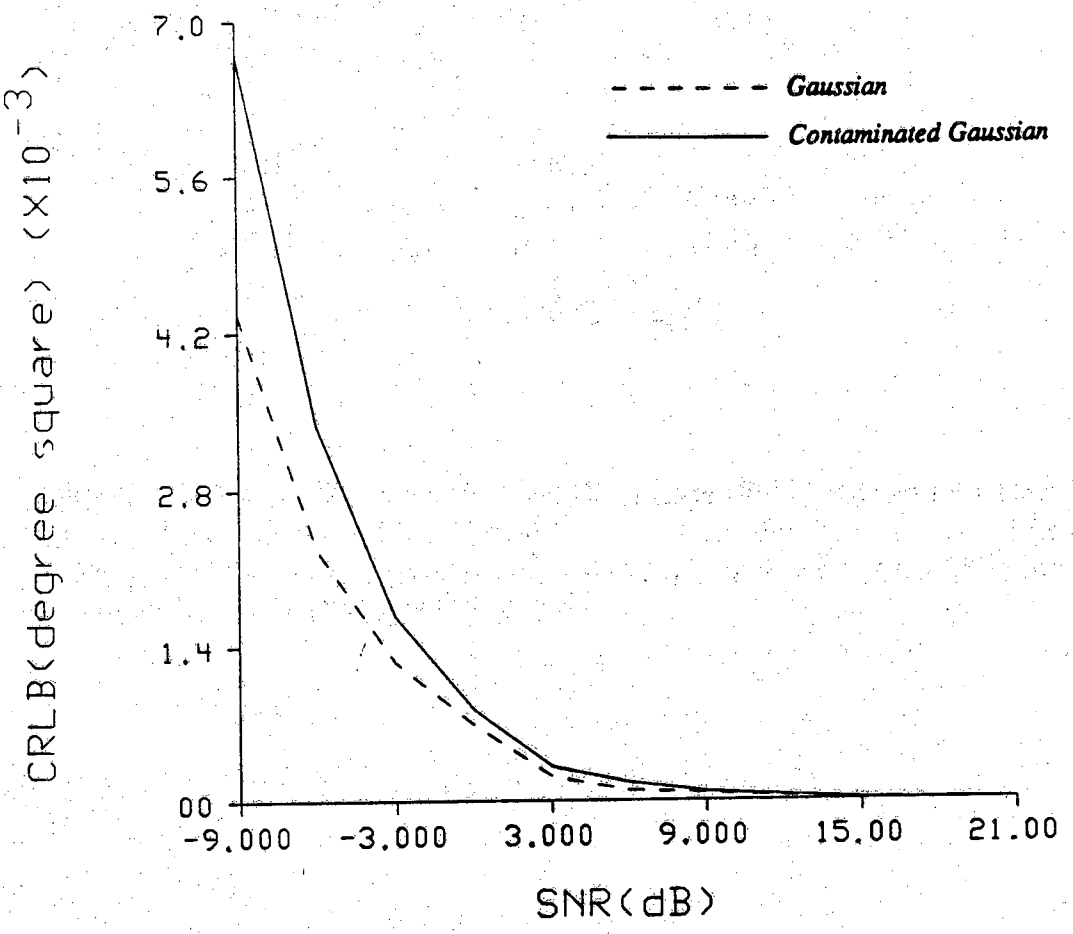


Figure 5.1. *CRLB vs. SNR* for a Gaussian Noise Case (solid line) and for a Contaminated Gaussian Noise Case (dotted line). The contamination is caused by replacing five percent of the Gaussian data with outliers which has ten times the variance of the parent Gaussian distribution.

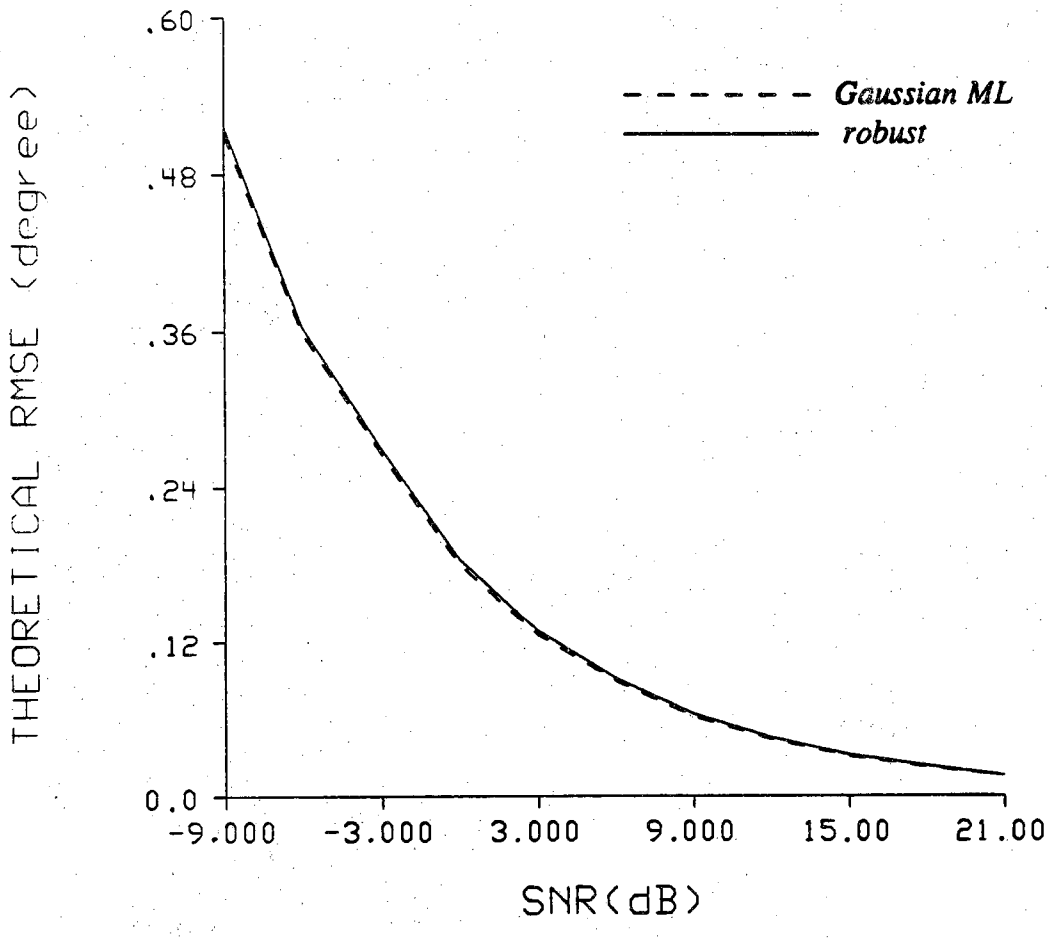


Figure 5.2. *Incoherent Sources & Gaussian Noise*: Theoretical RMSE's vs. SNR for the ML DOA estimates (dotted line) and for the robust DOA estimates (solid line).

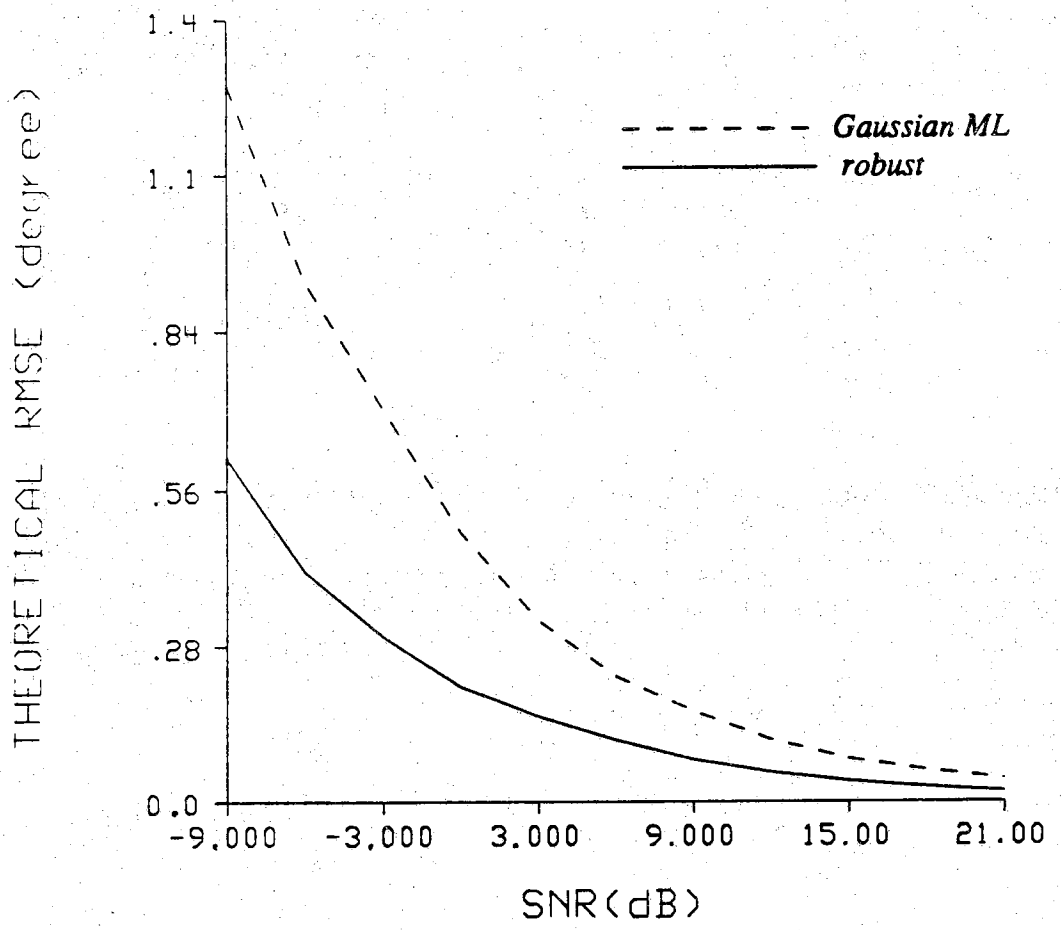


Figure 5.3. *Incoherent Sources & Contaminated Gaussian Noise*: Theoretical RMSE's vs. SNR for the ML DOA estimates (dotted line) and for the robust DOA estimates (solid line). The contamination is caused by replacing five percent of the Gaussian data with outliers which has ten times the variance of the parent Gaussian distribution.

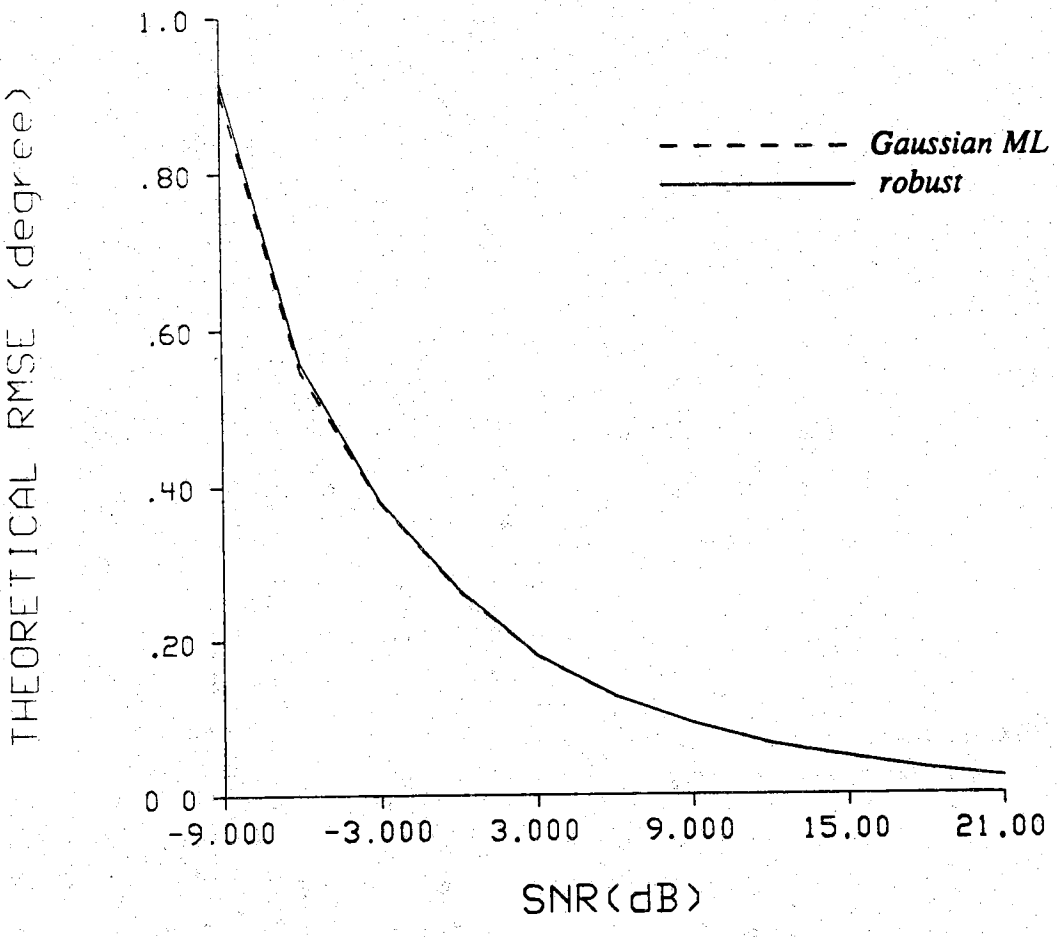


Figure 5.4. *Coherent Sources & Gaussian Noise*: Theoretical RMSE's vs. SNR for the ML DOA estimates (dotted line) and for the robust DOA estimates (solid line).

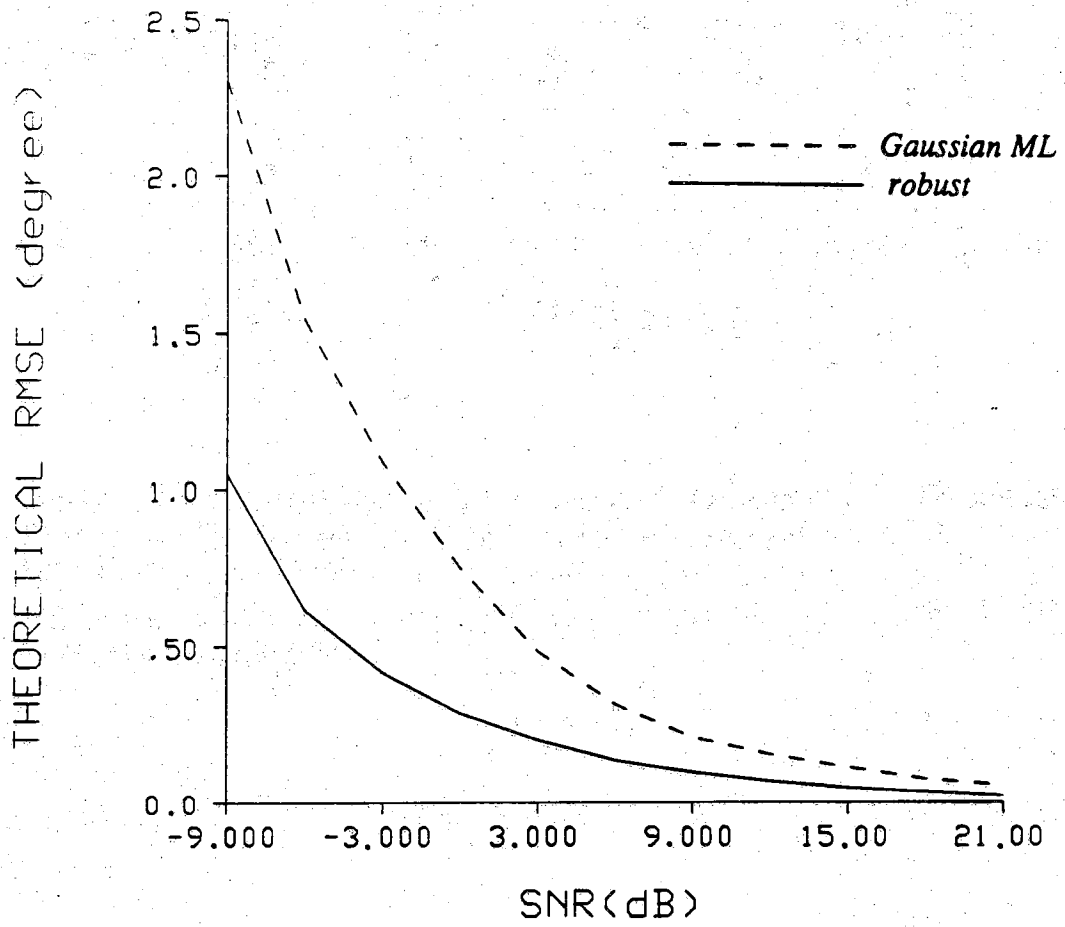


Figure 5.5. *Coherent Sources & Contaminated Gaussian Noise:* Theoretical RMSE's vs. SNR for the ML DOA estimates (dotted line) and for the robust DOA estimates (solid line). The contamination is caused by replacing five percent of the Gaussian data with outliers which has ten times the variance of the parent Gaussian distribution.

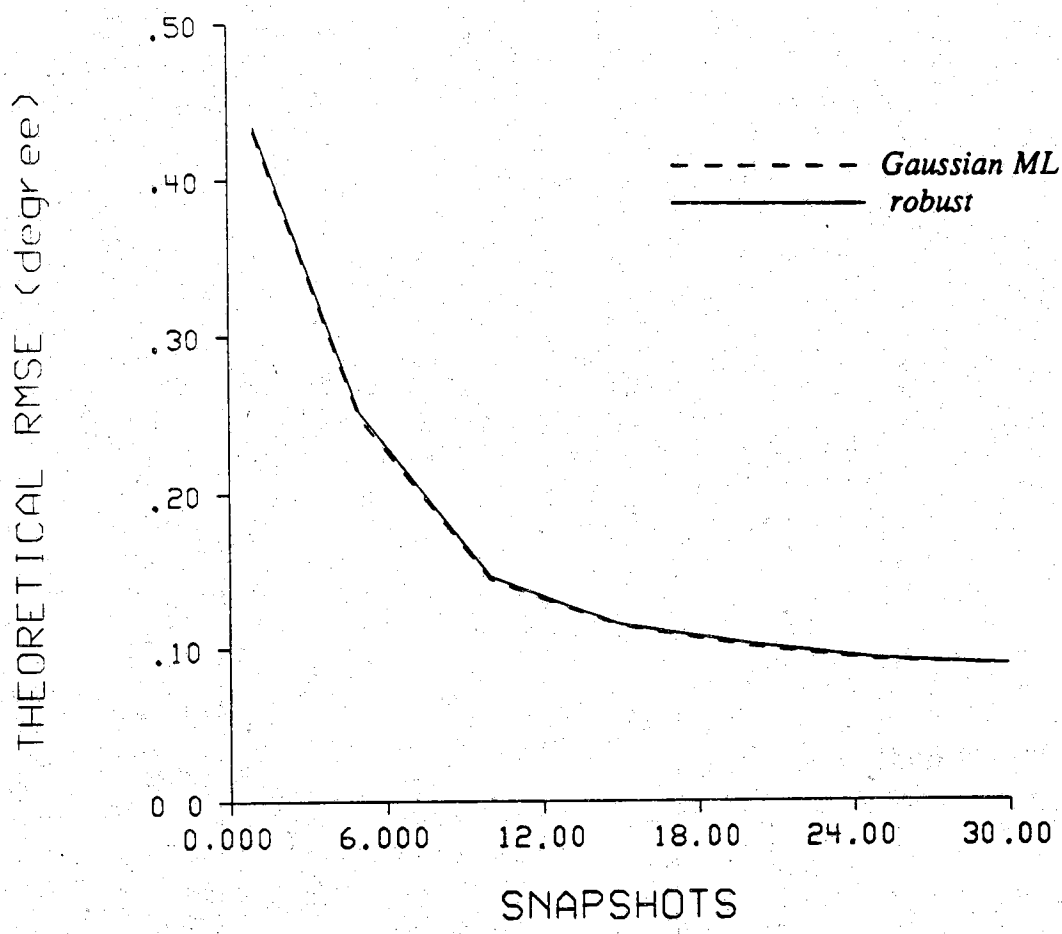


Figure 5.6. *Incoherent Sources & Gaussian Noise*: Theoretical RMSE's vs. Number of Snapshots for the ML DOA estimates (dotted line) and for the robust DOA estimates (solid line).

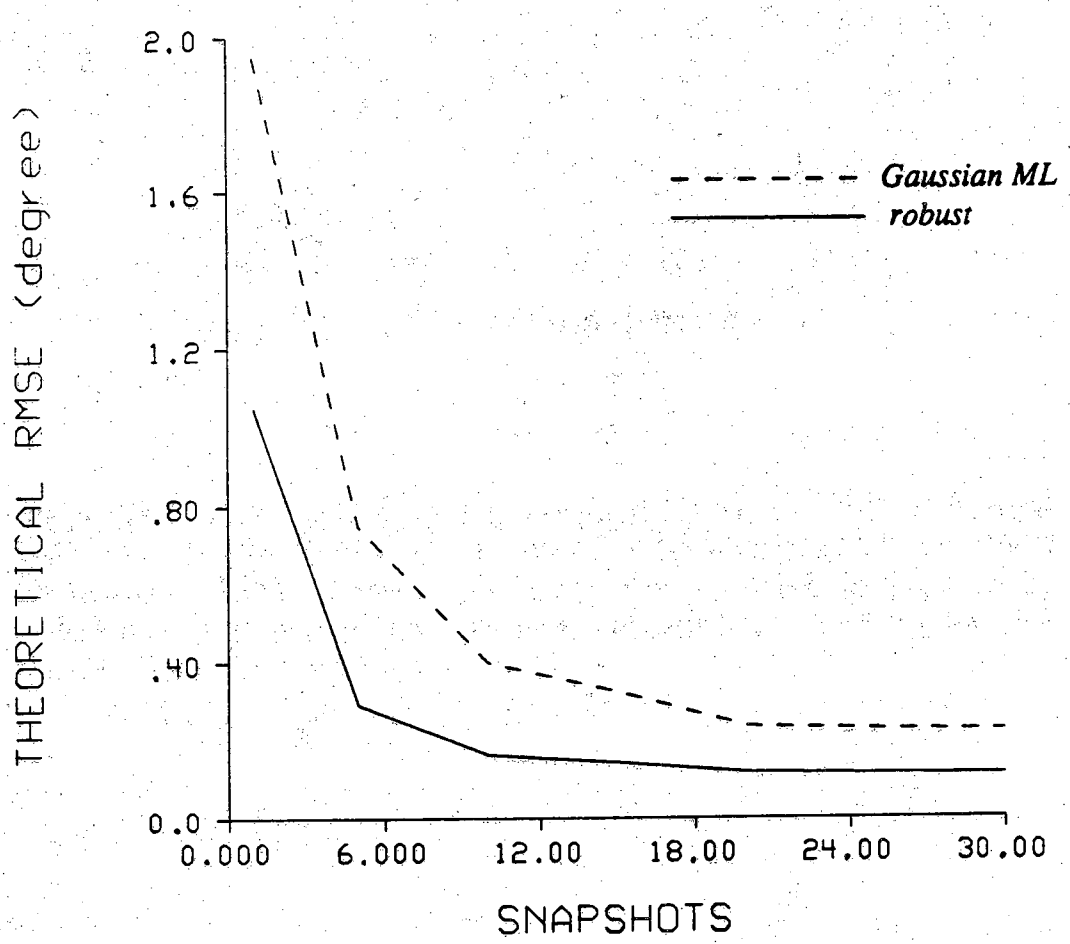


Figure 5.7. ***Incoherent Sources & Contaminated Gaussian Noise*****:** Theoretical RMSE's vs. Number of Snapshots for the ML DOA estimates (dotted line) and for the robust DOA estimates (solid line). The contamination is caused by replacing five percent of the Gaussian data with outliers which has ten times the variance of the parent Gaussian distribution.

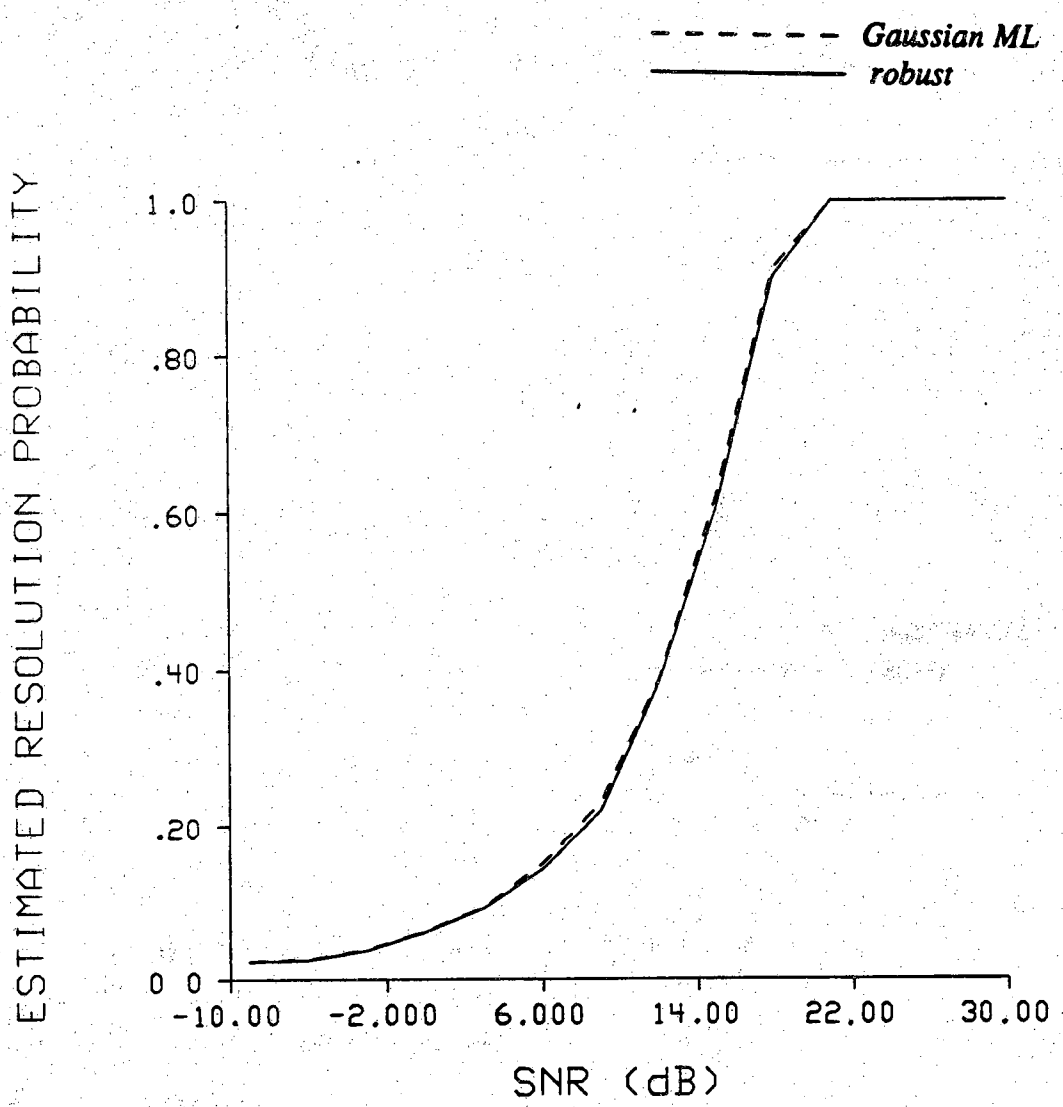


Figure 5.8. ***Resolution vs. SNR in Gaussian Noise***: Estimated Resolution Probability (no. of successful resolution / no. of trials) vs. SNR for the ML DOA estimates (dotted line) and for the robust DOA estimates (solid line). The true DOA's are 30.0° and 30.2°.

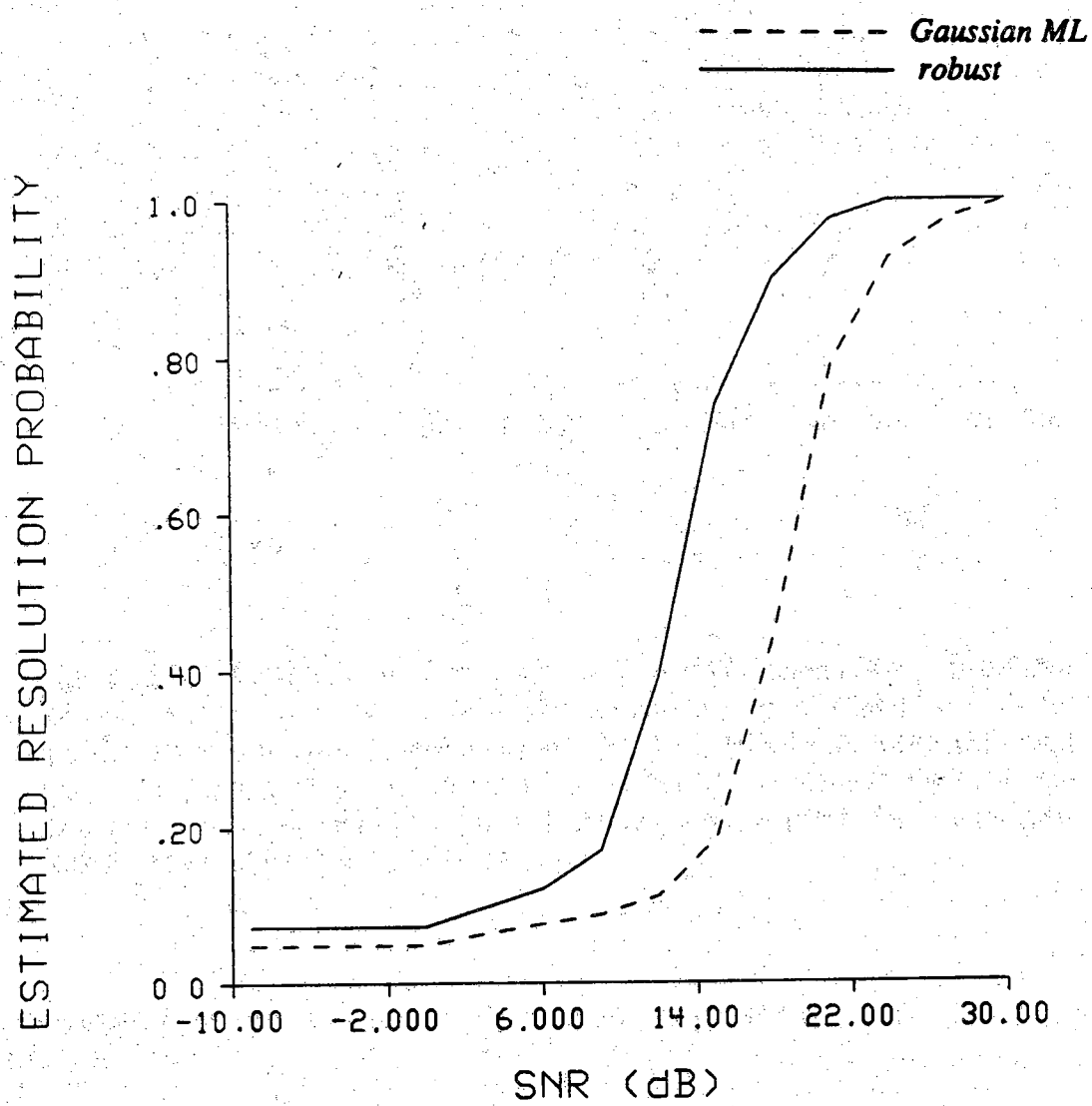


Figure 5.9. *Resolution vs. SNR in Contaminated Gaussian Noise*: Estimated Resolution Probability (no. of successful resolution / no. of trials) vs. SNR for the ML DOA estimates (dotted line) and for the robust DOA estimates (solid line). The contamination is caused by replacing five percent of the Gaussian data with outliers which has ten times the variance of the parent Gaussian distribution. The true DOA's are 30.0° and 30.2°.

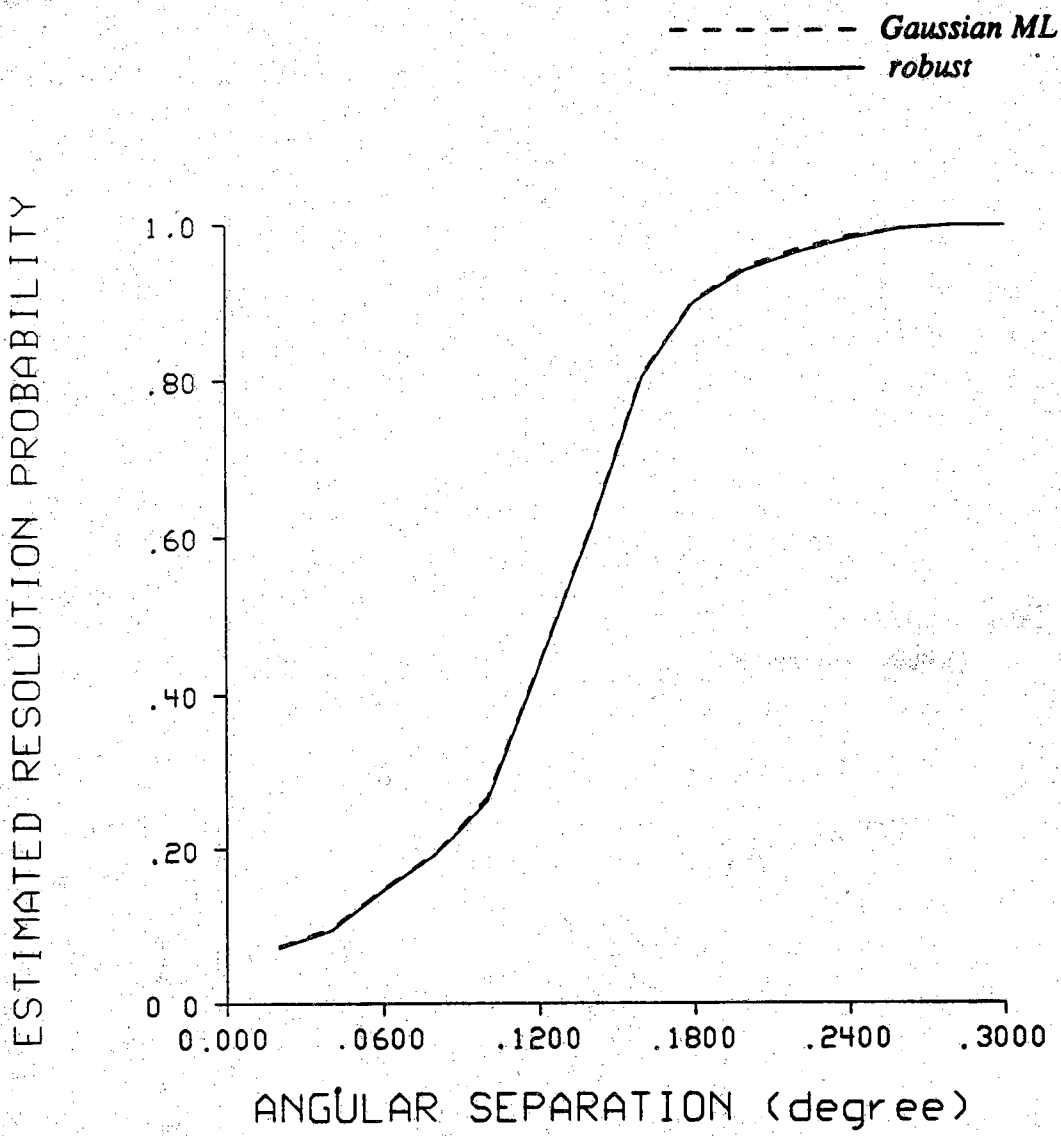


Figure 5.10. *Resolution vs. Separation in Gaussian Noise*: Estimated Resolution Probability (no. of successful resolution / no. of trials) vs. Angular Separation for the ML DOA estimates (dotted line) and for the robust DOA estimates (solid line). The SNR is fixed at 18dB, and the smaller value of the two true DOA's is chosen as 30.0°.

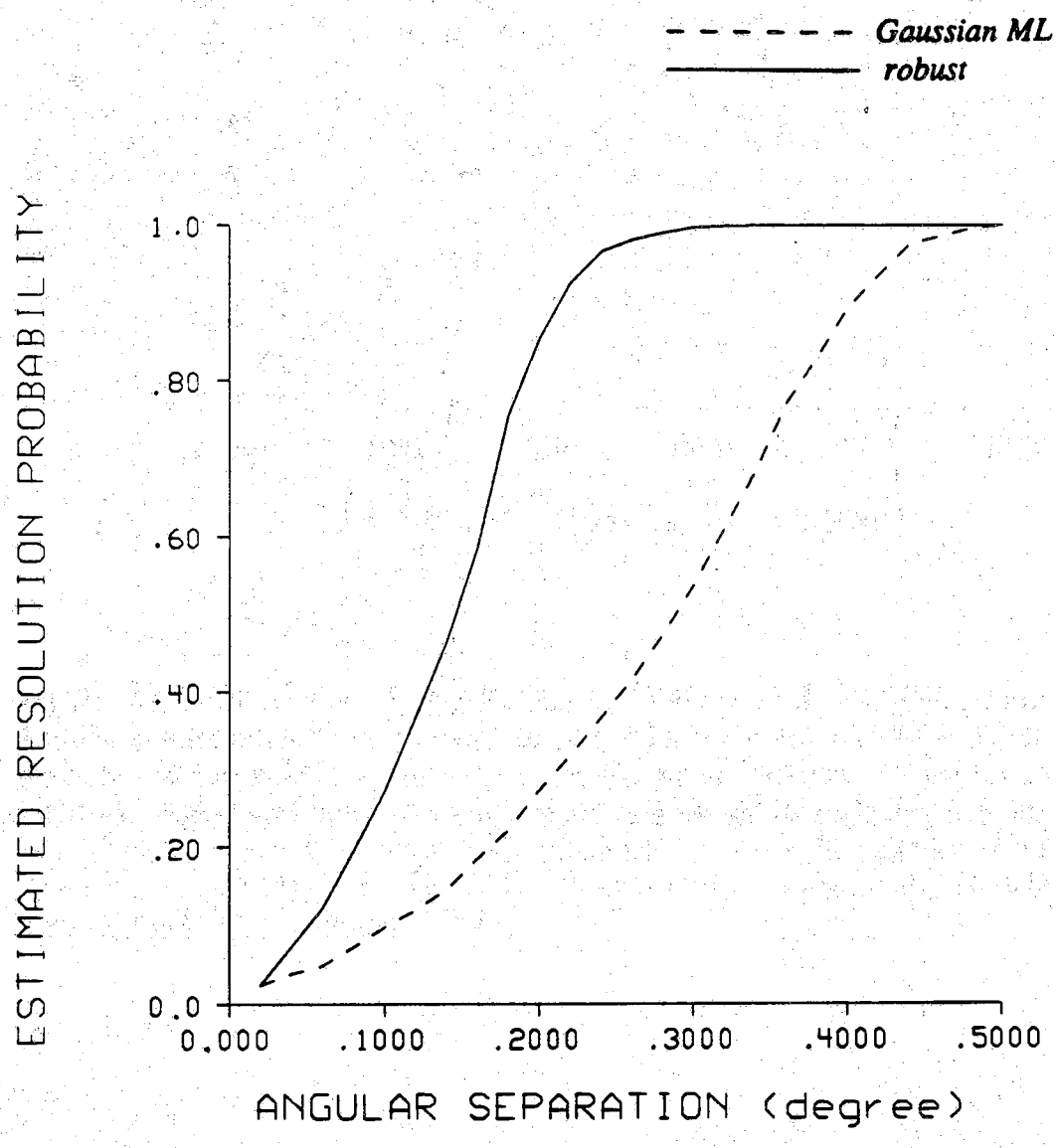


Figure 5.11. ***Resolution vs. Separation in Contaminated Gaussian Noise***: Estimated Resolution Probability (no. of successful resolution / no. of trials) vs. Angular Separation for the ML DOA estimates (dotted line) and for the robust DOA estimates (solid line). The contamination is caused by replacing five percent of the Gaussian data with outliers which has ten times the variance of the parent Gaussian distribution. The SNR is fixed at 18dB, and the smaller value of the two true DOA's is chosen as 30.0°.

5.6. Conclusions

We have presented a robust direction-of-arrival (DOA) estimation algorithm which performs almost as well as a maximum likelihood (ML) DOA estimation scheme in the pure Gaussian noise, and much better in the presence of outliers. The DOA's are estimated by a robust technique based on the so called M-estimators, a generalization of classical ML estimator by Huber [27]. The technique is equally applicable to single snapshot cases and coherent signals. Performances of the estimator in both the Gaussian and outlier contaminated Gaussian noise have been evaluated using the Cramer Rao Lower Bound (CRLB) and the variance derived from the Influence Function (IF), followed by resolution analysis regarding the ability of the algorithm in resolving two closely spaced sources with equal power.

Computational experience has also shown that the cost function may have many local minima, especially when the SNR is low. Thus for SNR's less than about 0dB, in spite of using sophisticated optimization techniques, the minimization algorithm do not always converge to the global minimum. One can always obtain the global minimum by repeating the algorithm with several different starting points and choosing the one with the least value.

CHAPTER 6

GENERALIZATION OF EIGENSPACE METHODS FOR BEARING ESTIMATION USING MAXIMUM LIKELIHOOD

6.1. Introduction

Under the condition that the observation period is long and signal to noise ratio (SNR) is not too low, the eigenspace approach, also called eigenvector methods, has previously been shown to have substantially higher resolution in estimating DOA's than the conventional beamformer, Capon's MLM [11], and autoregressive (AR) spectral estimators [15]. It is also known, however, that the performance of eigenspace methods is is usually inferior to that of the maximum likelihood (ML) technique. As was also pointed out by Ziskind et al. [87], the inferiority is especially conspicuous in the threshold region, namely, when the signal to noise ratio (SNR) is small, or alternatively, when the number of snapshots is small. Moreover, these techniques cannot handle the case of coherent signals, while the ML techniques can. The coherent case appears, for an example, in specular multipath propagation problems, which is of great practical importance.

In this chapter, a maximum likelihood (ML) direction-of-arrival (DOA) estimation problem is considered where the source signals are treated as sample functions of

Gaussian random processes, instead of the unknown deterministic sequences as assumed in most of the previous ML approaches. As with the Multiple Signal Classification algorithm (MUSIC), one of the eigenvector methods, the ML DOA estimation problem considered here only requires previous knowledge of the number of sources, e.g., the amplitudes and frequencies of the source envelopes need not be estimated for estimating DOA's. As a matter of fact, this is one of the reasons for the recent popularity of eigenvector algorithms.

The study reveals a relationship between this ML DOA estimation scheme and eigenvector methods for estimating DOA's. In particular, the focus is on interconnecting the notions of DOA estimation using eigenvector methods to a more quantitative Gaussian ML approach, i.e., choosing the DOA estimates to be in the directions of the eigenvectors which corresponds to the largest eigenvalues in the signal subspace.

When the number of sources is one, it is shown that maximizing the likelihood function with respect to the DOA angle is identical to choosing the steering vector to be in the direction of the eigenvector which corresponds to the largest eigenvalue in the signal subspace. In the case of multiple sources, however, the equality does not hold exactly. The similarities and differences between this ML method and eigenvector methods for two source cases are also investigated.

6.2. The Single Source Case

Consider a linear array composed of M identical equi-spaced sensors. Assume that a source, centered around a known frequency with wavelength λ, impinge on the array from the direction θ. Since the narrow-bandness in the sensor array means that the propagation delays of the signal along the array are much smaller than the

reciprocal of the bandwidth of the signal, the envelope of the source signal received by the array can be expressed as follows. Let $x_m(t_i)$ and $u_m(t_i)$ denote the ith snapshots of the mth sensor output and noise, respectively. Then $x_m(t_i)$ can be expressed as

$$x_m(t_i) = \alpha \exp[j\omega t_i] \exp[j\,\pi\,(m-1)(2d/\lambda)\sin\theta] + u_m(t_i)$$

$$i=1,...,N \;;\; m=1,...,M, \qquad (6.2.1)$$

where α and ω are the amplitude and frequency of the source envelope, respectively. d is the spacing between sensors, θ is the unknown DOA, and N is the total number of snapshots. The vector of the received signals $X(t_i)$ can be expressed more compactly as

$$X(t_i) = \alpha\, exp\,[j\omega t_i]\, f(\theta) + U(t_i) \qquad i=1,...,N \qquad (6.2.2)$$

where

$$X(t_i) = col.[\, x_1(t_i), x_2(t_i), ..., x_M(t_i)\,],$$

$$U(t_i) = col.[\, u_1(t_i), u_2(t_i), ..., u_M(t_i)\,],$$

and

$$f(\theta) = col.[\, 1,\ \exp\{j\,\pi\,(2d/\lambda)\sin\theta\}.....,$$

$$.....,\ \exp\{j\,\pi\,(M-1)(2d/\lambda)\sin\theta\,\}\,]. \qquad (6.2.3)$$

It is also assumed that α is a zero mean Complex Gaussian random variable with

$$E[\alpha\alpha^*] = \sigma^2,$$

while $U(t_i)$ is a zero mean Complex Gaussian random vector with

$$E[U(t_i)U^H(t_i)] = \rho I,$$

where * and H denotes the conjugate and conjugate transpose, respectively.

Thus

$$E[X(t_i)] = col.[0.,....,0.]$$

and

$$E[X(t_i)X^H(t_i)]$$

$$= E[\{\alpha \exp[j\omega t_i] f(\theta) + U(t_i)\}\{\alpha^* \exp[-j\omega t_i] f^H(\theta) + U^H(t_i)\}]$$

$$= \rho I + \sigma^2 f(\theta) f^H(\theta) \qquad (6.2.4)$$

Let us denote the signal to noise ratio (SNR) as K, i.e., $K = \sigma^2/\rho$, then the log-likelihood terms, is given by

$$L(\theta) = log\,[p\{X(t_1), X(t_2), ..., X(t_N)\,;\, \theta, \rho, K\}]$$

$$= -\frac{1}{2}\sum_{i=1}^{N} X^H(t_i)[\rho I + K\rho f(\theta) f^H(\theta)]^{-1} X(t_i)$$

$$- \frac{N}{2} log\,[\det|\rho I + K\rho f(\theta) f^H(\theta)\,|] + constant. \qquad (6.2.5)$$

After some algebraic simplifications, the loglikelihood is given by

$$L(\theta) = -\frac{1}{2\rho}\sum_{i=1}^{N} X^H(t_i)[\,I + Kf(\theta) f^H(\theta)\,]^{-1} X(t_i)$$

$$- \frac{NM}{2} \log\rho - \frac{N}{2} log\,(1+KM) + constant\,. \qquad (6.2.6)$$

To compute the Maximum Likelihood (ML) estimate of θ, we need to maximize (6.2.6) with respect to the unknown θ, ρ, and K. Maximizing with respect to ρ after fixing θ and K, we obtain $\hat{\rho}$, the estimate of ρ in terms of θ and K.

$$\hat{\rho} = \frac{1}{NM} \sum_{i=1}^{N} X^H(t_i)[\, I + K\, f\,(\theta) f^H(\theta)\,]^{-1} X\,(t_i) \tag{6.2.7}$$

Substituting this result back into equation (6.2.6), again ignoring constant terms, the estimates can be obtained by solving the following minimization problem.

$$\min_{\theta, K} \; NM \log \left\{ \frac{1}{NM} \sum_{i=1}^{N} X^H(t_i)[\, I + K\, f\,(\theta) f^H(\theta)\,]^{-1} X\,(t_i) \right\}$$

$$+ N \log (1+KM) \tag{6.2.8}$$

Since the logarithm is a monotonic function, it is equivalent to

$$\min_{\theta, K} \left[(1+KM)^{\frac{1}{M}} \sum_{i=1}^{N} X^H(t_i) \left\{ I + K\, f\,(\theta) f^H(\theta) \right\}^{-1} X\,(t_i) \right]. \tag{6.2.9}$$

Here, one can observe a special relationship between this ML estimation approach and the eigenvector methods for estimating DOA's. In particular, the focus is on interconnecting the notions of DOA estimation using eigenvector methods to more quantitative Gaussian ML approach, i.e., choosing the DOA estimates to be in the directions of the eigenvectors which correspond to the largest eigenvalues in the signal subspace.

For any arbitrarily fixed value of K in (6.2.9), the problem is equivalent to minimizing

$$J(\theta) = \sum_{i=1}^{N} X^H(t_i)[\, I + K\, f\,(\theta) f^H(\theta)\,]^{-1} X\,(t_i) \tag{6.2.10}$$

with respect to θ, and $J\,(\theta)$ can be rewritten as the following.

$$J\,(\theta) = trace \left\{ [\, I + K\, f\,(\theta) f^H(\theta)\,]^{-1} \sum_{i=1}^{N} X\,(t_i)\, X^H(t_i) \right\}$$

$$= trace\left\{ [I - \frac{K}{1+KM} f(\theta) f^H(\theta)] \sum_{m=1}^{M} \lambda_m \beta_m \beta^H{}_m \right\}, \qquad (6.2.11)$$

where λ_m and β_m, $m = 1,...,M$ denote the M eigenvalues and the corresponding normalized eigenvectors of $\sum_{i=1}^{N} X(t_i) X^H(t_i)$. $J(\theta)$ is simplified further to

$$J(\theta) = \sum_{m=1}^{M} \lambda_m - \frac{K}{1+KM} \sum_{m=1}^{M} \lambda_m |f^H(\theta)\beta_m|^2 . \qquad (6.2.12)$$

Therefore, minimizing $J(\theta)$ with respect to θ is indeed equivalent to choosing $f(\theta)$ to be in the direction of the eigenvector which corresponds to the largest eigenvalue in the signal subspace.

6.3. The Case of Two Sources

Now, $x_m(t_i)$ can be expressed as

$$x_m(t_i) = \alpha_1 \exp[j\omega_1 t_i] \exp[j\,\pi\,(m-1)(2d/\lambda)\sin\theta_1]$$

$$+\ \alpha_2 \exp[j\omega_2 t_i] \exp[j\,\pi\,(m-1)(2d/\lambda)\sin\theta_2] + u_m(t_i) . \qquad (6.3.1)$$

The vector of the received signals $X(t_i)$ is then expressed as

$$X(t_i) = F\,S(t_i) + U(t_i) \qquad i=1,...,N \qquad (6.3.2)$$

where

$$F = [f(\theta_1), f(\theta_2)],$$

and

$$S(t_i) = col.[\alpha_1 \exp\{j\omega_1 t_i\}, \alpha_2 \exp\{j\omega_2 t_i\}] \qquad (6.3.3)$$

where $X(t_i)$, $U(t_i)$ and $f(\theta)$ are the same as defined before. In addition, α_1 and α_2 are Complex Gaussian random variables with zero mean with

$$E[\alpha_1 \alpha_1^*] = \sigma_1^2$$

and

$$E[\alpha_2 \alpha_2^*] = \sigma_2^2. \tag{6.3.4}$$

Let us denote the individual signal to ratio (SNR) as K_1 and K_2, respectively, i.e., $K_1 = \sigma_1^2/\rho$ and $K_2 = \sigma_2^2/\rho$. Then the log-likelihood is given by

$$\log[p\{X(t_1), X(t_2), \ldots, X(t_N); \theta_1, \theta_2, K_1, K_2, \rho\}]$$

$$= -\frac{1}{2}\sum_{i=1}^{N} X^H(t_i)[\rho I + K_1 \rho f(\theta_1) f^H(\theta_1) + K_2 \rho f(\theta_2) f^H(\theta_2)]^{-1} X(t_i)$$

$$-\frac{N}{2}\log[\det|\rho I + K_1 \rho f(\theta_1) f^H(\theta_1) + K_2 \rho f(\theta_2) f^H(\theta_2)|]$$

$$+ \textit{constant}. \tag{6.3.5}$$

After eliminating ρ as before, the problem is now to find θ_1, θ_2, K_1, and K_2 by minimizing

$$J(\theta_1, \theta_2, K_1, K_2) = [\det|I + K_1 f(\theta_1) f^H(\theta_1) + K_2 f(\theta_2) f^H(\theta_2)|]^{\frac{1}{M}} \times$$

$$\sum_{i=1}^{N} X^H(t_i)[I + K_1 f(\theta_1) f^H(\theta_1) + K_2 f(\theta_2) f^H(\theta_2)]^{-1} X(t_i) \tag{6.3.6}$$

with respect to θ_1, θ_2, K_1, and K_2.

Note that

$$[I + K_1 f(\theta_1) f^H(\theta_1)]^{-1} = 1 - \frac{K_1}{1 + K_1 M} f^H(\theta_1) f(\theta_1)$$

and

$$[\, I + K_1 f(\theta_1) f^H(\theta_1) + K_2 f(\theta_2) f^H(\theta_2) \,]^{-1}$$

$$= [\, I + K_1 f(\theta_1) f^H(\theta_1) \,]^{-1}$$

$$- \frac{K_2 [\, I + K_1 f(\theta_1) f^H(\theta_1) \,]^{-1} f(\theta_2) f^H(\theta_2) [\, I + K_1 f(\theta_1) f^H(\theta_1) \,]^{-1}}{1 + K_2 f^H(\theta_2) [\, I + K_1 f(\theta_1) f^H(\theta_1) \,]^{-1} f(\theta_2)} \tag{6.3.7}$$

After some algebraic simplifications, it can be shown that,

$$[\, I + K_1 f(\theta_1) f^H(\theta_1) + K_2 f(\theta_2) f^H(\theta_2) \,]^{-1}$$

$$= I - B_1 f(\theta_1) f^H(\theta_1) - B_2 f(\theta_2) f^H(\theta_2)$$

$$+ B_3 f(\theta_1) f^H(\theta_2) + B_4 f(\theta_2) f^H(\theta_1) \tag{6.3.8}$$

where

$$B_1 = \frac{K_1 (1 + K_2 M)}{(1 + K_1 M)(1 + K_2 M) - K_1 K_2 |f^H(\theta_1) f(\theta_2)|^2}$$

$$B_2 = \frac{K_2 (1 + K_1 M)}{(1 + K_1 M)(1 + K_2 M) - K_1 K_2 |f^H(\theta_1) f(\theta_2)|^2}$$

$$B_3 = \frac{K_1 K_2 f^H(\theta_1) f(\theta_2)}{(1 + K_1 M)(1 + K_2 M) - K_1 K_2 |f^H(\theta_1) f(\theta_2)|^2}$$

and

$$B_4 = \frac{K_1 K_2 f^H(\theta_2) f(\theta_1)}{(1 + K_1 M)(1 + K_2 M) - K_1 K_2 |f^H(\theta_1) f(\theta_2)|^2} \tag{6.3.9}$$

Let us also denote

$$B_0 = [\det |I + K_1 f(\theta_1) f^H(\theta_1) + K_2 f(\theta_2) f^H(\theta_2)| \,]^{\frac{1}{M}} \tag{6.3.10}$$

Then equation (6.3.6) can be rewritten as

$$J(\theta_1,\theta_2,K_1,K_2)$$

$$= B_0 \; trace\{ \sum_{i=1}^{N} X^H(t_i)[\, I + K_1 f(\theta_1) f^H(\theta_1) + K_2 f(\theta_2) f^H(\theta_2)\,]^{-1} X(t_i) \;\}$$

$$= B_0 \; trace\{ \;[\, I + K_1 f(\theta_1) f^H(\theta_1) + K_2 f(\theta_2) f^H(\theta_2)\,]^{-1} \sum_{i=1}^{N} X(t_i) X^H(t_i) \;\}$$

$$= B_0 \; trace\{ \;[\, I - B_1 f(\theta_1) f^H(\theta_1) - B_2 f(\theta_2) f^H(\theta_2)$$

$$+ B_3 f(\theta_1) f^H(\theta_2) + B_4 f(\theta_2) f^H(\theta_1)\,] \sum_{m=1}^{M} \lambda_m \beta_m \beta^H{}_m \;\}$$

$$= B_0 \sum_{m=1}^{M} \lambda_m - B_0 B_1 \sum_{m=1}^{M} \lambda_m \left| f^H(\theta_1)\beta_m \right|^2$$

$$- B_0 B_2 \sum_{m=1}^{M} \lambda_m \left| f^H(\theta_2)\beta_m \right|^2$$

$$+ B_0 B_3 \sum_{m=1}^{M} \lambda_m \{ f^H(\theta_2)\beta_m \, \beta^H{}_m f(\theta_1) \}$$

$$+ B_0 B_4 \sum_{m=1}^{M} \lambda_m \{ f^H(\theta_1)\beta_m \, \beta^H{}_m f(\theta_2) \} \qquad (6.3.11)$$

where λ_m and β_m, $m = 1,...,M$ again denote the M eigenvalues and the corresponding normalized eigenvectors of $\sum_{i=1}^{N} X(t_i) X^H(t_i)$.

Note that the coefficients B_0, B_1, B_2, B_3, and B_4 as defined in equations (6.3.9) and (6.3.10) are functions of $\sin\theta_1 - \sin\theta_2$ only, if the values of K_1 and K_2 are fixed. For arbitrarily chosen values of K_1 and K_2, it can be clearly seen from equations (6.3.9) through (6.3.11) that minimizing $J(\theta_1,\theta_2)$ with respect to θ_1 and θ_2 is not exactly the same as choosing $f(\theta_1)$ and $f(\theta_2)$ to be in the directions of the

eigenvectors which correspond to the two largest eigenvalues in the signal subspace.

6.4. Conclusions

A maximum likelihood (ML) direction-of-arrival (DOA) estimation problem is considered where the source signals are treated as sample functions of random processes instead of unknown deterministic sequences as assumed in most of the previous approaches. The study revealed a relationship between this ML DOA estimation scheme and eigenvector methods for estimating DOA's.

When the number of sources is one, it has been clearly shown that maximizing the likelihood function with respect to the DOA angle is exactly equal to choosing the steering vector to be in the direction of the eigenvector which corresponds to the largest eigenvalue in the signal subspace. In cases of multiple sources this equality does not hold exactly. The similarities and differences between this ML method and eigenvector methods have been shown for the case of two sources.

CHAPTER 7

CONCLUSIONS AND SUGGESTIONS FOR FUTURE RESEARCH

7.1. Conclusions

Several robust methods of estimating directions-of-arrival (DOA) using arrays of sensors were proposed. The received source signals at the arrays may be narrow band or wide band, and also incoherent or coherent. The noise at each sensor of the arrays may be uncorrelated from those of nearby sensors, or correlated with unknown correlation structures. The main emphasis was on the property of robustness, which refers to insensitivity against a small deviation in the underlying Gaussian noise assumption.

The contribution of the research can be summarized as follows. First of all, a robust narrow band DOA estimation technique has been developed by reconstructing the correlation matrix utilizing a multivariate time series modeling of the array data. Many eigenvector DOA estimation algorithms can be robustified by replacing the usual sample correlation matrix estimate with the reconstructed correlation matrix mentioned before.

Secondly, a robust decentralized DOA estimation scheme have been considered. A notable feature is the robust combining procedure for estimates of the number of sources and the corresponding DOA's. Estimating parameters at each subarray site is

a totally independent process from that of other subarray sites, and the fusion center can recognize data from malfunctioning subarray sites or at least minimize the harmful effects of the estimates from such subarray sites.

A robust method of DOA estimation which can handle both the narrow band and the wide band sources have been developed. The method requires very little information about the types of sources, frequencies of signals, and the noise correlation. The proposed scheme utilizes a 2-D spectrum estimation technique utilizing 1-D autoregressive (AR) models in the Radon space.

Lastly, a robust technique is considered for maximum likelihood (ML) narrow band DOA estimation against outliers and distributional uncertainties. The algorithm employs a robustified Gaussian ML estimator based on the so called M-estimators, a generalization of classical ML estimator. It is equally capable of handling coherent sources as well as the single snapshot cases.

In the last chapter, a maximum likelihood (ML) direction-of-arrival (DOA) estimation problem is considered where the source signals are treated as sample functions of Gaussian random processes, instead of the unknown deterministic sequences as assumed in most of the previous ML approaches. In particular, the focus was on interconnecting the notions of DOA estimation using eigenvector methods to more quantitative Gaussian ML approach, i.e., choosing the DOA estimates to be in the directions of the eigenvectors which corresponds to the largest eigenvalues in the signal subspace.

Table 7.1 summarizes various conventional DOA estimation techniques which are based on Gaussian noise assumption and the corresponding robust techniques developed in the report.

Table 7.1. Summary of the conventional DOA estimation techniques based on Gaussian noise assumption *vs.* the corresponding robust techniques.

Source	Noise	Methods Devised	References
narrow band & incoherent	Gaussian	MUSIC & Related Methods ESPIRIT	Schmidt [65], CH6 Paulraj et al. [54]
	mixture	*Robust* Correlation Matrix Reconstruction & Decentralized Processing	Lee et al. [43] CH2 CH3
narrow band & coherent	Gaussian	ML Estimation Coherent Signal-Subspace	Haykin [24], CH6 Cadzow [9]
	mixture	*Robust* ML Estimation	Lee et al. [42], CH5
wide band & incoherent or coherent	Gaussian	2-D PSD Estimation Focusing Operation Divide into Narrow Bands	Halpney et al. [19] Wang et al. [82] Wax et al. [86]
	mixture	*Robust* Method using Radon Transform	Srinivasa et al. [72] CH4

7.2. Suggestions for Future Research

Related to the robust direction-of-arrival (DOA) estimation techniques developed in this report, several suggestions can be made.

7.2.1. Robust Direction-of-Arrival Estimation with Non-Uniform Linear Array Spacing

In many practical applications of sonar array signal processing, an array of sensors attached to a line is towed by a moving ship. The spacings between the presumably linear array are not supposed to be the same. Suppose that the exact locations of

sensor arrays are known or can be estimated, the problem of estimating direction-of-arrivals (DOA) can be handled by the DOA estimation technique using Radon Transform which was presented in Chapter 4. What remains to be solved is the question of selecting the weights in the equation (4.3.1), or the interpolation of the 2-D data array in the equation (4.2.1).

The choice of the weights can have many alternatives. For an example, the weights can be made proportional to the length of the intersection as shown in Figure 4.1 multiplied by the spacing between the corresponding sensors. On the other hand, the weights can be chosen to be proportional to the area of the intersection also shown in Figure 4.1 multiplied by the spacing between them. Instead of choosing weights according to the spacing between the corresponding sensors, one may fix the weights and use the interpolated array data.

The DOA estimation technique using Radon Transform can be extended to higher dimensional spaces also, i.e., simultaneous estimation of azimuth and elevation angles. In this case one can again utilize the central slice theorem for stationary random fields (SRF), for 3-dimensional space. The Radon Transform has to be applied twice: once for 3-D to 2-D transformation, and then from 2-D to 1-D transformation.

7.2.2. Robust Direction-of-Arrival Estimation using Least Median of Squares Criterion

Classical least squares regression consists of minimizing the sum of the squared residuals. But in spite of its mathematical beauty and computational simplicity, this estimator is being criticized more and more for its dramatic lack of robustness. In this connection, Hampel [21] introduced the notion of the *breakdownpoint*, which is the smallest percentage of contaminated data that can cause the estimator to take on

arbitrarily large aberrant values. In the least squares, the *breakdownpoint* is zero. The generalized M-estimators [27] have a breakdown point of at most $1/(p+1)$ where p is the dimension of the data.

All of this raises the question whether robust regression with a high breakdown point is at all possible. The least median of squares (LMS) technique has been proposed by Rousseeuw [61]. It replace the sum by by median, which is very robust, and this yields the LMS estimator given by

$$\min_{\hat{\theta}} \underset{i}{med}\, r_i^2 \tag{7.2.1}$$

where r_i is the residual. The proposal is essentially based on an idea of Hampel [21]. It is known that the LMS estimator has the breakdown point of 0.5, but has a very low efficiency. Utilization of the LMS criterion can provide a breakthrough in estimating the directions-of-arrival (DOA's) when the percentage of outliers is very large, i.e., fifty percent, so that the conventional robust techniques based on the M-estimators are no longer appropriate.

7.2.3. Robust Direction-of-Arrival Estimation using Neural Networks

In direction finding, one tries to estimate the directions-of-arrival (DOA) from plane waves impinging on an array of sensors. The output signal at each sensor is completely determined by the frequency of the signal, the propagation of the signal, the geometry of the sensors and the DOA. Several robust algorithms have been developed in this report for the estimation of DOA's. A drawback of such DOA algorithms, whether traditional or robust, is that they depend on computationally burdensome algebraic techniques thus do not deliver a real time performance.

With the current advances made in VLSI technology, a number of parallel architectures have been proposed to alleviate the computational burden of traditional DOA estimation techniques, and make real time application possible. Recently much interest has been focused on so called neural networks [41,46]. A neural network is an array of highly interconnected simple analogue, non-linear, processing units. The strength of the neural network lies in the collective computational ability it possesses. Hopefield et al. [26] have shown that a neural network can be used to rapidly find a good solution to a difficult optimization problem, and Rastogi et al. [59] have shown that the neural network algorithm could have significant benefits over classical approaches for the bearing estimation problem. Furthermore, Jha et al. [30] extends the work of Rastogi et al. [59] by adapting the neural network algorithm to increase its convergence to the global minima, by such techniques as iterated descent and gain annealing.

The robust DOA estimation require the system to converge to the global minimum. Computational experience with the robust DOA estimation algorithms, however, have shown that the cost function may have many local minima, especially when the SNR is low. Thus for SNR's less than about 0dB, in spite of using sophisticated optimization techniques, the minimization algorithm do not always converge to the global minimum. Application of neural networks to the robust DOA estimation problem not only provide real time performance but may also benefit the convergence as well.

LIST OF REFERENCES

LIST OF REFERENCES

1. A. J. Barabell, J. Capon, D. F. Delong, J. R. Johnson, and K. Senne, "Performance Comparison of Superresolution Array Processing Algorithms", *Technical Report TST-72, Lincoln Laboratory, M.I.T., 1984.*

2. V. Barnett and T. Lewis, *Outliers in Statistical Data, 2nd ed.*, John Wiley & Sons, 1984.

3. U. K. Bhargava and R. L. Kashyap, "Robust Parametric Approach for Impulse Response Estimation", *IEEE Trans. Acoust., Speech, Signal Processing*, vol. ASSP-36, pp. 1592-1601, Oct. 1988.

4. G. Bienvenu and L. Kopp, "Source Power Estimation Method Associated with High Resolution Bearing Estimation", in *Signal Processing*, New York: Academic Press, pp. 577-590, 1973.

5. R. B. Blackman and J. W. Tukey, *The Measurement of Power Spectra*, New York: Dover, 1958.

6. J. Boheme, "Estimating the Source Parameters by Maximum Likelihood and Nonlinear Regression," *Proc. ICASSP'84 Conf.*, pp. 7.3.1-7.3.4, 1984.

7. Y. Bresler and A. Macovski, "Exact Maximum Likelihood Parameter Estimation of Superimposed Exponential Signals in Noise", *IEEE Trans. Acoust., Speech, Signal Processing*, vol. ASSP-34, pp. 1081-1089, Oct. 1986.

8. H. P. Bucker, "High-Resolution Cross-Sensor Beamforming for a Uniform Line Array," *J. Acoust. Soc. Amer.*, vol. 63, pp. 420-424, 1978.

9. J. A. Cadzow, "A High Resolution Direction-of-Arrival Algorithm for Narrow-Band Coherent and Incoherent Sources", *IEEE Trans. Acoust., Speech, Signal Processing*, vol. ASSP-36, pp. 965-979, July 1988.

10. J. Capon, "High Resolution Frequency Wavenumber Spectral Analysis," *Proc. IEEE.*, vol. 57, pp. 1408-1418, 1969.

11. J. Capon, "Maximum-Likelihood Spectral Estimation", in *Nonlinear Methods of Spectral Analysis*, S. Haykin, Ed. New York: Springer, pp. 155-179, 1979.

12. C. Chatterjee, R. L. Kashyap, and G. Boray, "Estimation of Close Sinusoids in Colored Noise and Model Discrimination", *IEEE Trans. Acoust., Speech, Signal Processing*, vol. ASSP-35, pp. 328-337, March 1987.

13. R. E. Collins and F. J. Zucker, *Antenna Theory, Part I*, McGraw-Hill, New York, 1969.

14. R. S. Elliott, *Antenna Theory and Design*, Prentice-Hall, Englewood Cliffs, N.J., 1981.

15. J. E. Evans, J. R. Johnson, and D. F. Sun, "Application of Advanced Signal Processing Techniques to Angle of Arrival Estimation in ATC Navigation and Surveillance Systems", *Lincoln Laboratory, MIT, Tech. Rep. 582, June 1982.*

16. W. F. Gabriel, "Spectral Analysis and Adaptive Array Superresolution Techniques", *Proc. IEEE.*, vol 68, pp. 654-666, 1981.

17. D. M. Goodman, "NLS: A System Identification Package for Transient Signals", *Lawrence Livermore Laboratory Report, UCID-19767*, March 1983.

18. Fred Haber and M. Zoltowski, "Spatial Spectrum Estimation in a Coherent Signal Environment Using an Array in Motion", *IEEE Trans. Antennas and Propagation*, vol. AP-34, pp. 301-310, March 1986.

19. O. S. Halpney and D. G. Childers, "Composite Wavefront Decomposition via Multidimensional Digital Filtering of Array Data," *IEEE Trans. Circuits and Systems*, vol. CAS-22, pp. 552-562, June 1975.

20. F. R. Hampel, "The Influence Curve and Its Role in Robust Estimation", *J. Amer. Statist. Assoc.*, vol. 69, pp. 383-393, 1974.

21. F. R. Hampel, "Beyond Location Parameters: Robust Concepts and Methods", *Bulletin of the International Statistical Institute, 46*, pp. 375-382, 1975.

22. F. R. Hampel, E. M. Ronchetti, P. J. Rousseeuw, and W. A. Stahel, *Robust Statistics: The Approach Based on Influence Functions*, John Wiley & Sons, 1986.

23. R. R. Hansen, Jr. and R. Chellappa, "Two-Dimensional Robust Spectrum Estimation", *IEEE Trans. Acoust., Speech, Signal Processing*, vol. ASSP-36, No.7, pp. 1051-1066, July 1988.

24. S. Haykin, *Array Signal Processing*, Prentice Hall, Englewood Cliffs, 1985.

25. J. J. Hopfield, "Neural Networks and Physical Systems with Emergent Collective Computational Abilities", *Proc. Natl. Acad. Sci.* April 1982.

26. J. J. Hopfield and D. W. Tank, "Neural Computation of Decisions in Optimization Problems", Biological Cybern., vol. 52, 1985.

27. P. J. Huber, *Robust Statistics*, John Wiley and Sons, pp. 153-191, 1981.

28. L. B. Jackson and H. C. Chien, "Frequency and Bearing Estimation by Two Dimensional Linear Prediction", *Proc. ICASSP'79 Conf.*, Washington DC, pp. 665-668, April 1979.

29. A. K. Jain and S. Ansari, "Radon Transform Theory for Random Fields and Optimum Image Reconstruction from Noisy Projections", *Proc. ICASSP'84 Conf.*, California, March 1984.

30. S. Jha, R. Chapman, and T. S. Durrani, "Bearing Estimation using Neural Networks", *Proc. ICASSP'88 Conf.*, New York City, April 1988.

31. D. H. Johnson, "The Application of Spectral Estimation Methods to Bearing Estimation Problems," *Proc. IEEE*, vol. 70, No. 9, pp. 1018-1028, September 1982.

32. D. H. Johnson and S. DeGraff, "Improving the Resolution of Bearing in Passive Sonar Arrays by Eigenvalue Analysis," *IEEE Trans. Acoust., Speech, Signal Processing*, vol. ASSP-30, pp. 638-647, Aug. 1982.

33. T. Kailath, *Linear Systems*, Prentice Hall, pp. 662-663, 1980.

34. R. L. Kashyap, "Optimal Choice of AR and MA Parts in Autoregressive Moving Average Models", *IEEE Trans. Pattern Analysis and Machine Intelligence*, vol. PAMI-4, No.2, March 1982.

35. R. L. Kashyap and A. R. Rao, *Dynamic Stochastic Models from Empirical Data*, New York: Academic, 1976.

36. R. L. Kashyap and David D. Lee, "Robust Decentralized Direction of Arrival Estimation in the Presence of Outlier Contaminated Noise", *Proc. IEEE Fourth ASSP Workshop on Spectrum Estimation and Modeling*, Minneapolis, MN, pp. 117-122, Aug. 1988.

37. R. L. Kashyap and R. E. Nasburg, "Parameter Estimation in Multivariate Stochastic Difference Equations", *IEEE Trans. Automatic Control*, vol. AC-19, No.6, Dec. 1974.

38. R. L. Kashyap, S. G. Oh, and R. N. Madan, "Robust Estimation of Sinusoidal Signal and Colored Noise using Decentralized Processing", *Proc. IEEE Conf. on Decision and Control*, Dec. 1987.

39. M. Kaveh and A. J. Barabell, "The Statistical Performance of MUSIC and the Minimum-Norm Algorithms in Resolving Plane Waves in Noise", *IEEE Trans. Acoust., Speech, Signal Processing*, vol. ASSP-34, pp. 331-341, April 1986.

40. S. W. Lang and J. H. McClellan, "Frequency Estimation with Maximum Entropy Spectral Estimators", *IEEE Trans. Acoust., Speech, Signal Processing*, vol. ASSP-28, pp. 716-724, 1980.

41. A. Lapedes and R. Farber, "Nonlinear Signal Processing using Neural Networks", *Proc. IEEE Conf. on Neural Information Processing Systems - Natural and Synthetic*, Denver, November 1987.

42. David D. Lee and R. L. Kashyap, "Robust Direction of Arrival Estimations", *Proc. Twenty-Seventh Annual Allerton Conference on Communication, Control, and Computing*, University of Illinois at Urbana-Champaign, September 1989.

43. David D. Lee, R. L. Kashyap, and Rabinder N. Madan, "Robust Decentralized Direction-of-Arrival Estimation in Contaminated Noise", *(to be published in the IEEE Trans. Acoust., Speech, Signal Processing.)*

44. B. C. Levy and M. B. Adams, "Global Optimization with Stochastic Neural Networks", *IEEE First Intl. Conference on Neural Networks*, San diego, California, June 21-24, 1987.

45. W. S. Liggett, "Passive Sonar: Fitting Models to Multiple Time Series", *in Signal Processing*, J. W. R. Griffith et al., Eds., New York: Academic, 1973.

46. R. P. Lippmann, "An Introduction to Computing with Neural Nets", *IEEE ASSP Magazine, 4:* pp. 4-22, April, 1987.

47. R. N. McDonough, "Application of the Maximum-Likelihood Method and the Maximum Entropy Method in Array Processing", *in Nonlinear Methods of Spectral Analysis*, S. Haykin, Ed. New York: Springer, pp. 181-243, 1979.

48. J. Makhoul, "Linear Prediction: A Tutorial Review", *Proc. IEEE*, vol. 63, pp. 561-580, 1975.

49. L. Marple, "New Autoregressive Spectrum Analysis Algorithm", *IEEE Trans. Acoustics, Speech, and Signal processing*, vol. ASSP-28, No. 4, pp. 441-454, August 1980.

50. J. N. McDonough, "Maximum-Likelihood Method in Array Processing", *in Nonlinear Methods of Spectral Analysis*, S. Haykin, Ed. New York: Springer, pp. 281-293. 1979.

51. A. Nehorai, G. Su, and M. Morf, "Estimation of Time Differences of Arrival by Pole Decomposition", *IEEE Trans. Acoust., Speech, Signal Processing*, vol. ASSP-31, pp. 1478-1491, Dec.1983.

52. Sang G. Oh and R. L. Kashyap, "Robust Approach for High Resolution Frequency Estimation", *(Submitted to the IEEE Trans. Acoustics, Speech, and Signal Processing, for Publication)*.

53. Sang Geun Oh and R. L. Kashyap, "Robust Frequency Estimation", *Proc. Int. Conf. Acoustics, Speech, and Signal Processing*, vol. ICASSP-E, May 1989.

54. A. Paulraj, R. Roy, and T. Kailath, "A Subspace Rotation Approach to Signal Parameter Estimation", *Proc. IEEE*, pp. 1044-1045, July 1986.

55. P. Perreto, "Collective Properties of Neural Networks : A Statistical Physics Approach", *Biol. Cybern. 50*, pp. 51-62, 1984.

56. B. Porat and B. Friedlander, "Estimation of Spatial and Spectral Parameters of Multiple Sources", *IEEE Trans. Information Theory*, vol. IT-29, pp. 412-425, May 1983.

57. S. J. Press, *Applied Multivariate Analysis*, 2nd ed., New York: R. E. Krieger, 1982.

58. C. R. Rao, *Linear Statistical Inference and Its Applications, 2nd edition*, John Wiley & Sons, pp. 181-183, 1973.

59. R. Rastogi, P. K. Gupta and R. Kumaresan, "Array Signal Processing with Interconnected Neuron-Like Elements", *Proc. Int. Conf. Acoust., Speech, Signal Processing*, pp. 2328-2331, 1987.

60. E. A. Robinson, "Spectral Approach to Geophysical Inversion by Lorentz, Fourier, and Radon Transform", *Proc. IEEE*, vol. 70, pp. 1039-1054, Sept. 1982.

61. Peter J. Rousseeuw, "Least Median of Squares Regression", *Journal of the American Statistical Association*, vol. 79, No. 388 Dec 1984.

62. D. E. Rumelhart, J. L. McClelland, and the PDP Research Group, *Parallel Distributed Processing (PDP): Exploration in the Microstructure of Cognition (Vol. 1)*, MIT Press, Cambridge, Massachusetts, 1986.

63. D. J. Scheibner, and T. W. Parks, "Slowness Aliasing in the Discrete Radon Transform: A Multirate System Approach to Beamforming", *IEEE Trans. Acoust. Speech, Signal Processing*, vol. ASSP-32, pp. 1160-1165, Dec. 1984.

64. R. O. Schmidt, "Multiple Emitter Location and Signal Parameter Estimation", *Proc. RADC Spectrum Estimation Workshop*, Oct. 1979.

65. R. O. Schmidt, "Multiple Emitter Location and Signal Parameter Estimation", *IEEE Trans. Antennas, and Propagation*, vol. ASSP-34, pp. 276-280, March 1986.

66. T. J. Shan, M. Wax and T. Kailath, "On Spatial Smoothing for Direction-of-Arrival Estimation of Coherent Sources", *IEEE Trans. Acoust., Speech, Signal Processing*, vol. ASSP-33, pp. 806-811, 1985.

67. G. Sharma and R. Chellappa, "Two-Dimensional Spectrum Estimation Using Noncausal Autoregressive Models", *IEEE Trans. on Information Theory*, vol. IT-32, No. 2, pp. 268-275, March 1986.

68. K. Sharman and T. S. Durrani, "A Comparative Study of Modern Eigenstructure Methods for Bearing Estimation - a New High Performance Approach", *Proc. 25th IEEE Conf. Dec. Contr.*, Athens, Greece, pp. 1737-1742, Dec. 1986.

69. T. Soderstrom, "Convergence Properties of the Generalized Least Squares Identification Method", *Automatica*, vol. 10, pp. 617-626, 1974.

70. T. Soderstrom, "On the Uniqueness of Maximum Likelihood Identification", *Automatica*, vol. 11, pp. 193-197, 1975.

71. N. Srinivasa, "Application of Linear Prediction Modeling and Filtering in the Radon Space", *Ph.D Dissertation,* Indian Institute of Science, Bangalore, 1988.

72. N. Srinivasa, David D. Lee, and R. L. Kashyap, "Direction of Arrival Estimation using Radon Transform", *(submitted to the IEEE Trans. Acoust. Speech, Signal Processing for publication.)*

73. N. Srinivasa, David D. Lee, and R. L. Kashyap, "Robust 2-D Spectrum Estimation using Radon Transform", *Proc. IEEE ASSP Sixth Workshop on Multidimensional Signal Processing,* Pacific Grove, CA, September, 1989.

74. N. Srinivasa, David D. Lee, and R. L. Kashyap, "Direction of Arrival Estimation for Wide Band Signals", *(submitted to the IEEE 1990 International Conference on IEEE Trans. on ASSP, Albuquerque, New Mexico.)*

75. N. Srinivasa, K. R. Ramakrishnan, and K. Rajgopal, "Two-Dimensional Spectral Estimation: A Radon Transform Approach", *IEEE Journal of Oceanic Engineering,* vol. OE-12, No.1, pp. 90-96, January 1987.

76. R. G. Staudte, Robust Estimation, *Queen's Papers in Pure Appl. Math.,* No.53, 1980.

77. G. Su, and M. Morf, "Modal Decomposition Signal Subspace Algorithms", *IEEE Trans. Acoust. Speech, Signal Processing,* vol. ASSP-34, pp. 585-602, June 1986.

78. R. R. Tenny and N. R. Sandell, "Detection with Distributed sensors", *IEEE Trans. Aerosp. Electron. Syst.,* vol. AES-17, pp. 501-510, July 1981.

79. D. W. Tufts, and R. Kumaresan, "Estimation of Frequencies of Multiple Sinusoids: Making Linear Prediction Perform like Maximum Likelihood", *Proc. IEEE,* vol. 70, September 1982.

80. J. W. Tukey, *A Survey of Sampling from Contaminated Distributions, in : Contributions to Probability and Statistics,* I.Olkin, Ed., Stanford University Press, 1960.

81. B. D. Van Veen and K. M. Buckley, "Beamforming : A Versatile Approach to Spatial Filtering", *IEEE ASSP Mag.,* pp. 4-24, April 1988.

82. H. Wang, and M. Kaveh, "Coherent Signal-Subspace Processing for the Detection and Estimation of Angles of Arrival of Multiple Wide-Band Sources", *IEEE Trans. Acoust., Speech, Signal Processing*, vol. ASSP-33, pp. 823-831, August 1985.

83. M. Wax and T. Kailath, "Decentralized Processing in Sensor Arrays", *IEEE Trans. Acoust., Speech, Signal Processing*, vol. ASSP-33, pp. 1123-1129, Oct. 1985.

84. M. Wax and T. Kailath, "Detection of Signals by Information Theoretic Criteria", *IEEE Trans. Acoust., Speech, Signal Processing*, vol. ASSP-33, pp. 387-392, April 1985.

85. M. Wax and T. Kailath, "Optimum Localization of Multiple Sources by Passive Arrays", *IEEE Trans. Acoust., Speech, Signal Processing*, vol. ASSP-31, pp. 1210-1218, Oct. 1983.

86. M. Wax, T. J. Shan, and T. Kailath, "Source Location and Spectral Density Estimation of Multiple Sources", *Proc. 16th Asilomar Conf. Cir., System., Comp.*, 1982.

87. I. Ziskind and M. Wax, "Maximum Likelihood Localization of Multiple Sources by Alternating Projection", *IEEE Trans. Acoust., Speech, Signal Processing*, vol. ASSP-36, pp. 1553-1560, Oct. 1988.

88. M. Zoltowski and Fred Haber, "A vector Space Approach to Direction Finding in a Coherent Multipath Environment", *IEEE Trans. Antennas and Propagation*, vol. AP-34, pp. 1069-1079, Sept. 1986.