ESTIMATION OF DIRECTIONS OF ARRIVAL OF SPATIALLY DISPERSED SIGNALS IN HIGH RESOLUTION ARRAY PROCESSING

By

YAN MENG, B. Sc. Tsinghua University

M. Sc. Academia Sinica

A Thesis

Submitted to the School of Graduate Studies

in Partial Fulfilment of the Requirements

for the Degree

Doctor of Philosophy

McMaster University

July 1995

DOCTOR OF PHILOSOPHY (1995) MCMASTER UNIVERSITY

(Doctor of Philosophy) Hamilton, Ontario

TITLE: **Estimation of Directions of Arrival of Spatially Dispersed Signals in High Resolution Array Processing**

AUTHOR: Yan Meng

B. Sc. Tsinghua University

M. Sc. Academia Sinica

SUPERVISOR: Prof. Kon Max Wong

NUMBER OF PAGES: xiv, 152

ESTIMATION OF DIRECTIONS OF ARRIVAL OF SPATIALLY DISPERSED SIGNALS IN HIGH RESOLUTION ARRAY PROCESSING

To My Beloved

Abstract

Array processing deals with methods for processing the output data of an array of sensors located at different positions in space in a wavefield. The purpose of the array processing is to obtain insight into the structure of the waves carrying information and traversing the array. Array processing methods have been extensively studied, and results can be found in the literature on radio astronomy, electrical engineering, acoustics, geophysics and statistics.

Recently a group of methods known as the high resolution array processing methods have been developed. All these methods make the assumption that the signals arriving at the array are generated by point sources in space. However, in many practical situations, this assumption is invalid leading to the break down of the high r lution methods.

This thesis is devoted to the study on high resolution array processing methods of the location parameters of physically dispersed sources that generate the signals. The formulation and the properties of spatially dispersed sources are presented. The geometric structure of the data space is examined which enables the partitioning of signal and noise subspaces from which several methods for source localization are proposed. The performance of these methods are analyzed. Engineering applications of this research topic are discussed.

Acknowledgement

I owe thanks to many people without whose help this thesis would not come into being. For his introducing me to statistical signal and array processing, his inspiring guidance and enthusiasm through the years of study, I would like to thank my supervisor Professor Max Wong. I am indebted to my supervisory committee members, Professors Pat Yip, Jim Reilly and Tom Luo for their valuable advice and wisdom. Gratitude is also due to Dr. Qiang Wu for his many enlightening suggestions to my research work. Over and above, I would like to give my heartfelt thanks to my family for their endless love and unflagging encouragement through all my life. Finally, I am obliged to DREO (Defence Research Establishment, Ottawa) and TRIO (Telecommunications Research Institute of Ontario) for their financial support.

Notation and Glossary

Vectors and matrices in a real space $\mathcal{R}$ or a complex space C are usually denoted by bold, lower cases and by bold, upper cases respectively. For a complex matrix of dimension $M \times N$, we have

A^T = the transpose of A.

A^* = the complex conjugate of A.

$A^\dagger$ = the conjugate transpose of A.

A^{-1} = the inverse of $A \in C^{M \times M}$.

$A^\#$ = generalized inverse of A.

$\| A \|$ = Norm of A.

$\mathrm{Re}[A]$ = the real part of A.

$\mathrm{Im}[A]$ = the imaginary part of A.

$\mathrm{tr}\{A\}$ = the trace of $A \in C^{M \times M}$.

$\det(A)$ = the determinant of $A \in C^{M \times M}$.

$A[m; n]$ = the mnth element of A.

$\mathrm{vec}(A)$ = $[a_{11} \cdots a_{M1}\, a_{12} \cdots a_{M2} \cdots a_{1N} \cdots a_{MN}]^T = a$,

$\mathrm{unvec}(a)$ = A, given matrix of dimension $M \times N$,

$$A \otimes B = \text{the Kronecker product of } A \text{ and } B$$

$$= \begin{bmatrix} a_{11}B & \cdots & a_{1N}B \\ \cdot & \cdots & \cdot \\ a_{M1}B & \cdots & a_{MN}B \end{bmatrix}.$$

Contents

Abstract iv

Acknowledgement v

Notation and Glossary vi

List of Tables xii

List of Figures xiv

1 INTRODUCTION 1

1.1 Sensor Array Processing 1

1.2 Estimation of the DOA of Point Sources in Spatially White Noise with Properly Calibrated Sensors 3

1.3 DOA Estimation in Non-Ideal Environments 9

1.4 Scope of Thesis 12

2 Signal Model for Spatially Distributed Sources 16

2.1 Phenomena of Spatially Distributed Sources 16

2.2 Mathematical Model 18

2.2.1 Continuous Model 18

2.2.2 Discrete Model 22
2.2.3 Some Examples of Signal Density Functions 24

3 Properties of the Signal Correlation Matrix **27**

3.1 Eigen-Analysis 27
3.2 Eigenvalues and Eigenvectors of a Uniformly Distributed Signal Correlation Matrix 29
3.3 Eigenvalues and Eigenvectors of a Gaussian Distributed Signal Correlation Matrix 30
3.4 Eigenvalues and Eigenvectors of the Correlation Matrix of a Signal with "first-order Butterworth" Distribution 32
3.5 Dimensionality of Signal Subspace of Distributed Sources 32

4 DOA Estimation Based on Second-order Statistics **37**

4.1 Data Model in a White Noise Environment 37
4.2 DISPARE – DOA Estimation for Distributed Signals 38
4.2.1 DISPARE I 41
4.2.2 DISPARE II 42
4.3 Application of the DISPARE Algorithms 42

5 Asymptotic Performance Analysis of DISPARE **55**

5.1 Perturbation Theory 55
5.2 Performance of DISPARE Algorithms 59
5.2.1 Performance of DISPARE I 60
5.2.2 Performance of DISPARE II 62

6 DOA Estimation Based on Higher-order Statistics **67**

6.1 Higher-order Statistics 68
6.1.1 Definitions and Properties 68
6.2 Cumulant Statistics for DOA Estimation in Correlated Gaussian Noise . . . 70
6.2.1 The Fourth-order Cumulant Matrix 71
6.2.2 Rank of the Cumulant Matrix 73
6.3 DOA Estimation with Cumulant Matrix 74
6.4 Asymptotic Performance Analysis 76
6.5 Numerical Experiments 79
6.5.1 Effects of Distribution of signals 79
6.5.2 Further Experiments 82

7 VEC-MUSIC Estimation **90**
7.1 The Estimation Method 91
7.2 Performance Analysis 97
7.2.1 Asymptotic Distribution of the Perturbation in Covariance Matrix . 97
7.2.2 Perturbation in Eigen-projectors 99
7.2.3 Asymptotic Performance of VEC-MUSIC 103
7.3 Choice of Noise Subspaces 105
7.4 Numerical Experiments 107

8 Conclusions **121**
8.1 Summary 121
8.2 Comparison of the Various Methods 123
8.3 Future Work 124

A Prolate Spheroidal Wave Functions **126**
A.1 Prolate Spheroidal Wave Functions of Zero Order 126

A.2 Discrete Prolate Spheroidal Sequences 127
A.3 Relationship between PSWF and DPSS 128

B Proofs of Theorems **129**
B.1 Proof of Theorem 5.1 129
B.2 Proof of Theorem 5.2 133
B.3 Proof of Theorem 7.1 136

C Covariance of the sample Cumulants **140**

List of Tables

3.1 Eigenvalues of a uniformly distributed signal 30
3.2 Eigenvalues of a Gaussian distributed signal 31
3.3 Eigenvalues of a signal with "first-order Butterworth" distribution 33

List of Figures

1.1 An incident point signal and a linear uniform sensor array 14

1.2 Signals transmitted by scattering propagation 15

1.3 Low-angle tracking environments 15

2.1 The coordinate system 18

2.2 The spherical coordinate system (r, θ, ψ) 20

3.1 Eigenvalues of a Uniformly distributed signal 34

3.2 Eigenvalues of a Gaussian distributed signal 35

3.3 Eigenvalues of a signal with "first-order Butterworth" distribution 36

4.1 The estimation errors of mean DOA of a uniformly distributed signal 49

4.2 The estimation errors of mean DOA of a Gaussian distributed signal 50

4.3 The estimation errors of mean DOA of a uniformly distributed signal under the assumption of Gaussian distribution 51

4.4 The estimation errors of mean DOA of a Gaussian distributed signal under the assumption of uniform distribution 52

4.5 The eigenvalues of a uniformly distributed signal given $M = 8$ 53

4.6 The estimation errors of mean DOA assuming both uniformly distributed . 54

5.1 Theoretical evaluation and simulation of RMSE of mean DOA of a uniformly distributed signal with $(\bar{\phi}, \sigma) = (-.5454, .3491)$ arriving at a uniform linear array of 8 sensors; 200 snapshots are used for comparison. 65

5.2 DISPARE spectra 66

6.1 Eigenvalues of a cumulant matrix of a distributed signal given $M = 4$ . . . 85

6.2 The estimation errors of mean DOA of a Gaussian distributed signal 86

6.3 The estimation errors of mean DOA of a uniformly distributed signal 86

6.4 The estimation errors of mean DOA with three methods 88

6.5 Theoretical evaluation and simulation of RMSE of mean DOA of a uniformly distributed signal of 4-QAM with $(\bar{\phi}, \sigma) = (-.2739, .3491)$ arriving at a uniform linear array of 4 sensors: $\rho = 0.5$; 500 snapshots are used for comparison. 89

7.1 The estimation errors of mean DOA of a Gaussian distributed signal 115

7.2 The estimation errors of mean DOA of a uniformly distributed signal 116

7.3 MUSIC and VEC-MUSIC spectra for 2 well-separated signals 117

7.4 MUSIC and VEC-MUSIC spectra for 2 merged signals 118

7.5 VEC-MUSIC spectra for two different noise subspaces 119

7.6 Theoretical evaluation and simulation of RMSE of mean DOA of a uniformly distributed signal with $(\bar{\phi}, \sigma) = (-.5454, .3491)$ arriving at a uniform linear array of 4 sensors; 200 snapshots are used for comparison. 120

Chapter 1

INTRODUCTION

1.1 Sensor Array Processing

Array processing, or more accurately, *sensor array signal processing*, is the processing of the output signals of an array of sensors located at different points in space in a wavefield. The purpose of array processing is to extract useful information from the received signals such as the number and the location of the signal sources, the propagation velocity of the waves, as well as the spectral properties of the signals. Array processing techniques have been employed in various areas in which very different wave phenomena occur. These include seismic exploration, passive sonar, radar, radio astronomy, as well as radio communications.

Common to all these applications, there are, in general, three essential purposes in array processing:

(i) To determine the number of signal sources (decision).

(ii) To estimate the locations of these sources (estimation).

(iii) To extract a desired signal from interference.

The traditional way of tackling the problems in array processing is based on procedures

employing the fast Fourier transform (FFT) [1-6]. This approach is computationally efficient and produces reasonable results for a large class of signal environments. It is also robust in the sense that it does not depend on a signal model. In spite of these advantages, there are several inherent performance limitations of the FFT methods. The most important limitation is that of spatial resolution, i.e., the ability to distinguish the spatial response of two or more signal sources. The spatial resolution of an FFT array processor is roughly inversely proportional to the size of the array. A second limitation is due to the implicit weighting of the sensor outputs (array shading) that occurs when processing with the FFT. Array shading manifests itself as "leakage" in the spatial domain, i.e., energy in the main lobe of a spatial response "spills" into the sidelobes, obscuring and distorting other spatial responses that are present. In fact, weak signal spatial responses can be masked by higher sidelobes from stronger spatial responses. Careful application of tapered sensor apertures can reduce the sidelobe leakage, but always at the expense of reduced resolution.

In an attempt to overcome the inherent limitations of the FFT approach, many alternative array processing procedures, generally known as *high-resolution methods*, have been proposed in the past several years. Both problems of decision and estimation have been addressed and have usually been treated separately [7-14] [15-38]. These methods generally start with a signal and noise model, and, based on the covariance matrix of the received data, the data space is partitioned into the *signal subspace* and the *noise subspace*. Various methods are then applied to estimate the parameters concerning the number and the directions of arrival (DOA) of the signal sources in these subspaces.

In establishing these high resolution array processing methods, there are several assumptions made on the signal and noise model. Different methods may have additional assumptions, but the ones most commonly made are:-

1.1. The signal and noise processes are stationary over the observation interval and are uncorrelated with each other.

1.2. The number of sensors in the array is greater than the number of signal sources.

1.3. The sensors are properly calibrated at the time of the measurement.

1.4. The background noise is isotropic (spatially white), or, in the case of spatially non-white noise, the covariance matrix of the noise is known.

1.5. Signals are propagated from distinct point sources.

1.2 Estimation of the DOA of Point Sources in Spatially White Noise with Properly Calibrated Sensors

We first review the problem of estimating the DOA of far-field point source signals in a spatially white sensor noise environment. The signal model we consider here consists of an M-dimensional complex data vector $\boldsymbol{x}(n) \in \mathcal{C}^M$ which represents the data received by an array of M properly calibrated sensors at the nth snapshot. The data vector is composed of plane-wave incident narrowband signals each of angular frequency ω_0 from K distinct point sources embedded in Gaussian noise. The received data vector at the nth snapshot can be written as

$$\boldsymbol{x}(n) = \boldsymbol{D}(\boldsymbol{\theta})\boldsymbol{s}(n) + \boldsymbol{\nu}(n), \qquad n = 1, \cdots, N \tag{1.1}$$

where N is the total number of snapshots observed. $\boldsymbol{s}(n) \in \mathcal{C}^K$ is the unknown K-dimensional complex vector whose elements represent the phasors of the signals at the nth snapshot. $\boldsymbol{\theta} \in \mathcal{R}^K$ is the K-dimensional real vector the elements of which represent the DOA of the K plane-wave signals such that

$$\boldsymbol{\theta} = [\theta_1 \ \cdots \ \theta_K]^T \tag{1.2}$$

where ^{T} denotes the transpose of a vector or matrix. The vector $\boldsymbol{\theta}$ in Eq.(1.2) is the vector to be estimated. The matrix $\boldsymbol{D}(\boldsymbol{\theta}) \in C^{M\times K}$ has the following structure

$$\boldsymbol{D}(\boldsymbol{\theta}) = [\boldsymbol{d}(\theta_1) \ \cdots \ \boldsymbol{d}(\theta_K)] \triangleq [\boldsymbol{d}_1 \ \cdots \ \boldsymbol{d}_K] \tag{1.3}$$

where $\boldsymbol{d}_k \triangleq \boldsymbol{d}(\theta_k) \in C^M$, $k = 1, \cdots, K$, is an M-dimensional complex vector usually called the *directional vector* or sometimes also the *steering vector* of the signals. The exact form of the steering vector depends on the array configuration. Most of the discussions in this thesis do not require a specific form of $\boldsymbol{d}(\theta_k)$, i.e., they are valid for any general array geometry. However, the uniform linear array, apart from being most commonly used, may also offer advantageous implementation efficiency for some algorithm. A uniform linear array is shown in Fig. 1.1 where the array elements are uniformly spaced. For a velocity of propagation c, the distance between two sensors in a uniform linear array must be $d_0 \leq \pi c/\omega_0$ and the corresponding steering vector is given by

$$\boldsymbol{d}(\theta_k) = [1 \quad e^{j\omega_0 d_0 \sin\theta_k/c} \quad \cdots \quad e^{j(M-1)\omega_0 d_0 \sin\theta_k/c}]^T \tag{1.4}$$

where $\theta_k \in (-\frac{\pi}{2}, \frac{\pi}{2})$ for all k.

The vectors $\boldsymbol{d}(\theta_k), k = 1, \cdots, K$, corresponding to K different values of θ_k are assumed to be linearly independent. This implies that $M > K$, and

$$\text{rank}(\boldsymbol{D}) = K \tag{1.5}$$

The noise vector sequence $\{\boldsymbol{\nu}(n)\}$ is assumed to be zero mean circular Gaussian such that

$$\text{E}\{\boldsymbol{\nu}(n_1)\boldsymbol{\nu}^\dagger(n_2)\} = p_\nu^2 \boldsymbol{I}\delta_{n_1 n_2} \tag{1.6a}$$

and

$$\text{E}\{\boldsymbol{\nu}(n_1)\boldsymbol{\nu}^T(n_2)\} = \boldsymbol{O} \tag{1.6b}$$

where † denotes the conjugate transpose operation,

$$\delta_{n_1 n_2} = \begin{cases} 0 & n_1 \neq n_2 \\ 1 & n_1 = n_2 \end{cases}$$

and p_ν^2 is the noise power. Furthermore, $\nu(n_1)$ and $s(n_2)$ are assumed to be uncorrelated for all n_1 and n_2. Thus, if we form the correlation matrix of the received data $\boldsymbol{x}(n)$, using the above assumptions on the noise vector, we have

$$\boldsymbol{R}_x \triangleq \mathrm{E}\{\boldsymbol{x}\boldsymbol{x}^\dagger\} = \boldsymbol{D}\boldsymbol{R}_s\boldsymbol{D}^\dagger + p_\nu^2\boldsymbol{I} \tag{1.7}$$

where

$$\boldsymbol{R}_s \triangleq \mathrm{E}\{\boldsymbol{s}\boldsymbol{s}^\dagger\} \tag{1.8}$$

and is positive definite.

In practice, the true value of $\boldsymbol{R}_x$ is seldom known. Therefore, an estimate $\hat{\boldsymbol{R}}_x(N)$ is obtained by averaging the outer products of the data over the total number of snapshots N such that

$$\hat{\boldsymbol{R}}_x(N) = \frac{1}{N}\sum_{n=1}^{N} \boldsymbol{x}(n)\boldsymbol{x}^\dagger(n) \tag{1.9}$$

where $\hat{}$ denotes the estimated value of a quantity. In the ensuing materials, the symbol $\hat{}$ appearing over a variable shall indicate the estimate of the variable based on $\hat{\boldsymbol{R}}_x(N)$. It can be shown [11, 14] that under certain conditions the estimate $\hat{\boldsymbol{R}}_x(N)$ converges to $\boldsymbol{R}_x$ with probability 1 and at a rate proportional to $(\log\log N \,/\, N)^{1/2}$, i.e.,

$$\lim_{N\to\infty} \hat{\boldsymbol{R}}_x(N) \underset{a.s.}{=} \boldsymbol{R}_x \tag{1.10}$$

where " $\underset{a.s.}{=}$ " indicates that the equality holds with probability 1, or almost surely. In particular, if the signal s is Gaussian, then under some regular conditions, the convergence of $\hat{\boldsymbol{R}}_x$ to $\boldsymbol{R}_x$ is at the rate of $O(N^{-1/2})$ in probability [47, 70]. In the following text, the explicit dependence of $\hat{\boldsymbol{R}}_x$ on N will be omitted to maintain conciseness of the expression.

In the past decade or two, various high resolution methods of estimating the DOA of the signal sources based upon the above model have been developed. Most of these methods are based on the partition of the space of the observed data into the signal and the noise subspaces assuming that the number of signals K is known. Different techniques are then employed to locate the DOA of the signals using these estimated subspaces. A method that has withstood the test of time and has remained popular among researchers in high resolution array processing is the MUSIC method.

The MUSIC Algorithm

The *Multiple Signal Classification* (MUSIC) algorithm [24] of DOA estimation in high resolution array processing assumes, in addition to Assumptions 1.1 $\sim$ 1.5, that the K signal sources are incoherent with each other; and that the number K has been correctly determined.

We begin by examining the eigenvalues λ_m and the corresponding normalized eigenvectors v_m, $m = 1, \cdots, M$, of $\boldsymbol{R}_x$. Since $\boldsymbol{R}_x$ is Hermitian and positive definite [69], its eigenvalues are real and positive, and the eigenvectors are orthonormal. Furthermore, $\boldsymbol{R}_x$ can be decomposed such that [67-69]

$$\boldsymbol{R}_x = \sum_{m=1}^{M} \lambda_m v_m v_m^{\dagger} \tag{1.11}$$

where $\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_K > \lambda_{K+1} = \cdots = \lambda_M = p_\nu^2$, and

$$v_m^{\dagger} v_n = 0 \quad \text{for } m \neq n \tag{1.12}$$

From the definition of eigenvalues and eigenvectors, we can write

$$(\boldsymbol{R}_x - p_\nu^2 I) v_m = \boldsymbol{O} \qquad m = K+1, \cdots, M \tag{1.13}$$

Using Eq.(1.7), Eq.(1.13) can be rewritten as

$$\boldsymbol{D}\boldsymbol{R}_s\boldsymbol{D}^{\dagger} v_m = \boldsymbol{O} \qquad m = K+1, \cdots, M \tag{1.14}$$

Since the signals are incoherent, $\boldsymbol{R}_s$ is of full rank. Also, since the signal sources are from K distinct directions, $\boldsymbol{D}$ is of full column rank. Thus, we can conclude that

$$\boldsymbol{D}^\dagger \boldsymbol{v}_m = \boldsymbol{O} \quad m = K+1, \cdots, M \tag{1.15}$$

Eq.(1.15) tells us that each of the K directional vectors $\boldsymbol{d}_k$ is orthogonal to the eigenvectors $\{\boldsymbol{v}_m\}$, $m = K+1, \cdots, M$, of the matrix $\boldsymbol{R}_x$. However, from Eq.(1.12) we see that the eigenvectors $\{\boldsymbol{v}_k\}$, $k = 1, \cdots, K$, span the orthogonal complement of the space spanned by $\{\boldsymbol{v}_m,\ m = K+1, \cdots, M\}$. Hence we can deduce that

$$\text{span}\{\boldsymbol{d}_1, \cdots, \boldsymbol{d}_K\} = \text{span}\{\boldsymbol{v}_1, \cdots, \boldsymbol{v}_K\} \tag{1.16}$$

Therefore, we can partition the eigenvectors of $\boldsymbol{R}_x$ into two groups such that

$$\boldsymbol{V}_s \triangleq [\boldsymbol{v}_1\ \boldsymbol{v}_2\ \cdots\ \boldsymbol{v}_K] \tag{1.17a}$$

and

$$\boldsymbol{V}_\nu \triangleq [\boldsymbol{v}_{K+1}\ \boldsymbol{v}_{K+2}\ \cdots\ \boldsymbol{v}_M] \tag{1.17b}$$

From Eqs.(1.15) and (1.16), we see that the eigenvectors of $\boldsymbol{R}_x$ span two disjoint subspaces. The one spanned by the eigenvectors corresponding to the K largest eigenvalues of $\boldsymbol{R}_x$ is called the *signal subspace* because it is also spanned by the directional vectors of the signals. The one spanned by the eigenvectors corresponding to the $(M-K)$ smallest eigenvalues of $\boldsymbol{R}_x$ is called the *noise subspace* because of its association with the noise power p_ν^2. The principle of the MUSIC algorithm is to determine the DOA of the signals by searching for the directional vectors $\boldsymbol{d}_k$, $k = 1, \cdots, K$, that are orthogonal to the noise subspace.

As mentioned before, in practice we can only estimate $\boldsymbol{R}_x$ by $\hat{\boldsymbol{R}}_x$ given by Eq.(1.9). Since $\hat{\boldsymbol{R}}_x$ is still Hermitian and positive definite, its eigenvalues, denoted by $\hat{\lambda}_m$, $m = 1, \cdots, M$, are still positive, and the associated eigenvectors, denoted by $\{\hat{\boldsymbol{v}}_m\}$, $m = 1, \cdots, M$, are still mutually orthonormal. However, the smallest $(M-K)$ eigenvalues $\hat{\lambda}_{K+1}, \cdots, \hat{\lambda}_M$ will not

be equal any more, and the estimated noise subspace spanned by $\{\hat{v}_m\}$, $m = K+1, \cdots, M$, will no longer be orthogonal to the true signal subspace spanned by $\{d_k\}$, $k = 1, \cdots, K$. The best that we can do is to search for the directional vectors which are closest to being orthogonal to the noise subspace. Accordingly, we form an *orthogonal projector* [66-68] such that

$$\hat{P}_\nu = \hat{V}_\nu (\hat{V}_\nu^\dagger \hat{V}_\nu)^{-1} \hat{V}_\nu^\dagger = \hat{V}_\nu \hat{V}_\nu^\dagger \tag{1.18}$$

where $\hat{V}_\nu = [\hat{v}_{K+1}, \cdots, \hat{v}_M]$. The second equality in Eq.(1.18) is due to the orthonormality of the eigenvectors. We note that

$$\hat{P}_\nu^\dagger = \hat{P}_\nu \tag{1.19a}$$

and

$$\hat{P}_\nu^\dagger \hat{P}_\nu = \hat{P}_\nu \tag{1.19b}$$

We can now use $\hat{P}_\nu$ to project the general directional vector $d(\theta)$ onto the estimated noise subspace spanned by the column vectors of $\hat{V}_\nu$. The values of $\hat{\theta}_k$, $k = 1, \cdots, K$, which give the K highest peaks of the quantity $1/\| \hat{P}_\nu d(\theta) \|^2$ are the estimated DOA of the K signal sources. In other words, the estimates of DOA for the MUSIC algorithm can be written as

$$\hat{\theta}_k = \arg\min_{\vartheta_k} \{\| \hat{P}_\nu d(\theta_k) \|^2\} = \arg\min_{\vartheta_k} \{d^\dagger(\theta_k) \hat{P}_\nu d(\theta_k)\} \tag{1.20}$$

where $\arg\min_{\theta_k}(\cdot)$ denotes the minimization of the bracketed quantity with respect to the argument θ_k.

Suggestions have been made to weight the projection $\hat{P}_\nu d(\theta_k)$ in different ways resulting in various criteria of *weighted MUSIC* algorithms [20, 27]. The MUSIC method remains one of the most popular methods of DOA estimation in high resolution array processing because of its relative simplicity in computation as well as its relative robustness and efficiency [39-44, 46-48].

1.3 DOA Estimation in Non-Ideal Environments

Analyses and simulation tests on MUSIC and other high resolution methods have been carried out [39-48] and it has been shown that under appropriate assumptions, these methods have superior performance compared to the FFT approach in the cases involving multiple and closely spaced signal sources. However, when the environment departs from its ideal assumptions, the performance of these methods generally deteriorates.

The five assumptions generally made in high resolution array processing methods have different practical implications. The first assumption can usually be satisfied, at least to a close approximation, in practice. The second assumption does not represent a harsh practical restriction either. However Assumptions 1.3, 1.4 and 1.5 are not easily satisfied in practice.

Assumption 1.3 is not easy to satisfy because, in practice, calibrating an antenna array can be time consuming and expensive. In addition to the problem of initial calibration, there is the problem of maintaining proper calibration of the sensors. The factors contributing to the gradual changes in the response of the sensors and their associated circuits are many and varied: Changes in the environment around the array may affect the response (e.g. coupling) of the sensors, changes in the location of the sensors may introduce perturbation of the sensor positions, aging of components, moisture, and temperature variation all contribute to changes in the response of the sensors. It is impossible in many practical situations to maintain array calibration to the accuracy required for the proper operation of these high-resolution methods. Analyses and simulation tests [49-51] show that improperly calibrated sensor arrays may lead to unacceptable estimation results. Self-calibration algorithms [52-56] for high resolution array processing have been proposed to alleviate the problems introduced by imprecise array calibration. These methods generally use the received signals to fine-tune the array calibration and are usually based on the simultaneous

estimation of the DOA of the signal sources and of the unknown array parameters. However, estimation of both signal and array parameters is a) computationally complex, b) hampered by potential identifiable problems.

Assumption 1.4, on the other hand, is often invalid in practice. Background noise, or *ambient noise*, refers to the noise that remains after all easily identifiable signal sources are eliminated. The sources of ambient noise can be both natural and man-made, with different sources exhibiting different directional and spectral characteristics. In a passive sonar system for instance [56, 57], the natural sources of noise include seismic disturbances, agitation of the sea surfaces by wind, thermal activity of the water molecules, as well as biological sources such as snapping shrimps, dolphins, and other fishes and ocean mammals. The principal man-made component of ambient noise, on the other hand, is the sound generated by distant shipping. Hence, the ambient noise characteristics for such cases are highly dependent on geographic location, the transmission medium, season of the year, and weather, and are, in general, unknown. Thus, Assumption 1.4 is often an over-simplified model which may not represent the true spatial characteristics of the noise.

The problem of DOA estimation in unknown correlated noise has been studied by various authors [58-65]. The most complete and effective methods, however, have been proposed by Wong, Wu, and Stoica [71-73] in which two arrays of sensors are used and the *canonical correlation decomposition* is applied to the two sets of data obtained. This leads to very accurate estimates of the signal and noise subspace from which various high resolution estimation methods can be established.

Assumption 1.5 is, at best, an approximation of the practical environment. The "size" of a source depends on its distance from the array of sensors. In addition, signal sources, in practice, may often be transmitted by reflection and the reflective medium may often be dispersed and thus violating the "point source" assumption. In the case of communication

in the Arctic environment [80, 81] for instance, the transmission of radio waves is often performed by ionospheric scattering. Fig.1.2 is a simplified diagram illustrating such a situation. The signal having undergone reflection through a scattering volume reaches the receiver and would appear to be from a distributed source. The receiver, in this application, has to determine the DOA of this distributed signal. A second example arises in the case of a radar system performing low angle tracking [101]. This case is illustrated in Fig.1.3 in which the aircraft, if it is sufficiently far away, can be approximately modelled as a point source. However, the diffused components caused by the reflection from a rough sea surface would appear as a distributed source. Here, the objective may be to estimate the DOA of a point source in the presence of distributed sources. Since the existing methods of high resolution array processing rely on the assumption that the signals are from point sources, the application of these methods to such problems will show grave deterioration in performance.

Motivated by the applications of high resolution array processing in radio communications, radar processing, ocean engineering etc., we study the problem of high resolution DOA estimation for spatially dispersed signal sources. It is on this particular problem that our attention is focused in this thesis.

We approach the problem by developing a signal model for distributed signal sources relative to the array of sensors. From this model, the spatial correlation matrix of the data received by the sensors is established. The important properties of this correlation matrix are analyzed and examined. Based upon these properties, several methods of estimating the DOA of distributed and point sources are developed. Computer simulations are then carried out to test and compare the effectiveness of these methods. Asymptotic performance of each proposed estimation method is analyzed.

1.4 Scope of Thesis

The thesis is organized as follows.

Chapter 1 introduces what is sensor array processing, especially, DOA estimation under ideal and non-ideal environments; and from the latter, initiates the problem of DOA estimation for spatially distributed signals.

In Chapter 2, we formulate the problem by developing a signal model for distributed sources relative to the array of sensors. We treat a spatially distributed signal source as a combination of a large number of closely spaced point sources. Thus, by making some reasonable assumptions, the spatial correlation matrix of the signals received by the sensors is established.

The important properties of the correlation matrix are studied in Chapter 3 revealing that even though the matrix is of full rank (being equal to the number of sensors) which renders the conventional high resolution array processing methods inapplicable in this case, the dimensionality of the signal subspace can be approximated by a number usually much smaller than the number of sensors.

From the study in Chaper 3 and utilizing the orthogonality of the signal and noise subspaces, two algorithms of DISPARE to estimate the DOAs of distributed sources have been developed in Chapter 4. Numerical experiments are carried out examining the effectiveness of the DOA estimation methods and revealing many insightful observations.

In Chapter 5, we introduce a new approach with algebraic simplicity for performance analysis. The asymptotic distribution of the eigen-projectors from the perturbed covariance matrix is examined and utilized to determine the effect of this perturbation on the DOA estimates of DISPARE.

A DOA estimation method based on higher-order statistics (order > 2) of the data received by a sensor array has been proposed in Chapter 6. This method exploits the

information hidden in higher-order moments of the data when signals are non-Gaussian. Furthermore, this method has the advantage of suppressing Gaussian noise by using higher-order cumulants.

In Chapter 7, we also explore a different scheme of using second-order statistics and generate the VEC-MUSIC algorithm for DOA estimation by applying the principles of DISPARE and MUSIC algorithms. VEC-MUSIC has the flexibility to choose various noise subspaces and provides the ability to distinguish distributed signals and point signals.

Although conceptually VEC-MUSIC utilizes only the first-order and second-order statistics of the data, its performance analysis is rather involved. However, based on its performance analysis, we can obtain the asymptotic performance of the estimation method based on the fourth-order cumulants of the received data.

In the final chapter, some observations are drawn from the extensive studies of this topic in previous chapters. Future work and potential engineering applications on the subject are discussed.

Some technical details including proofs of theorems and calculations are given in the appendices.

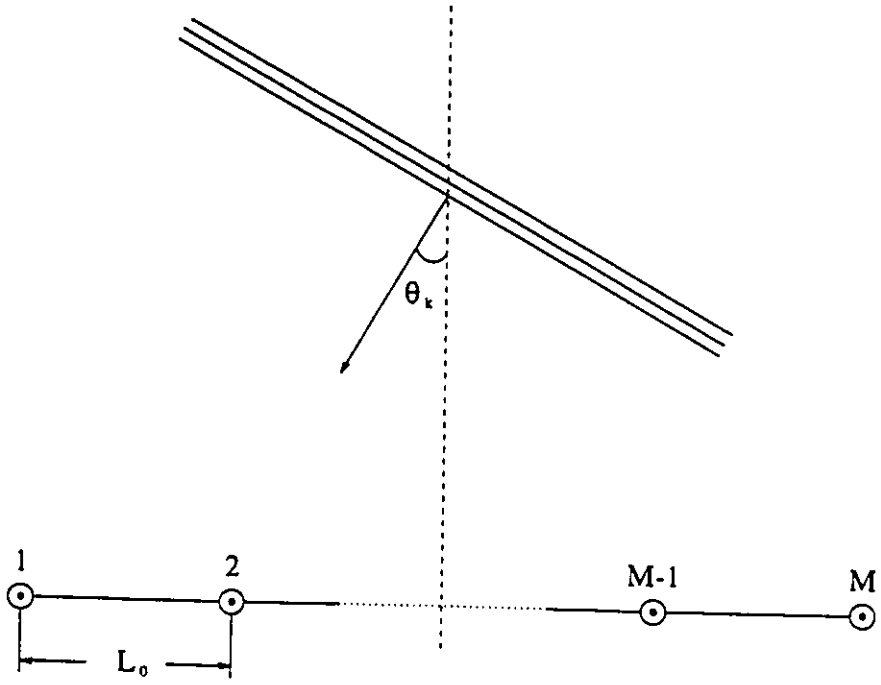


Figure 1.1: An incident point signal and a linear uniform sensor array

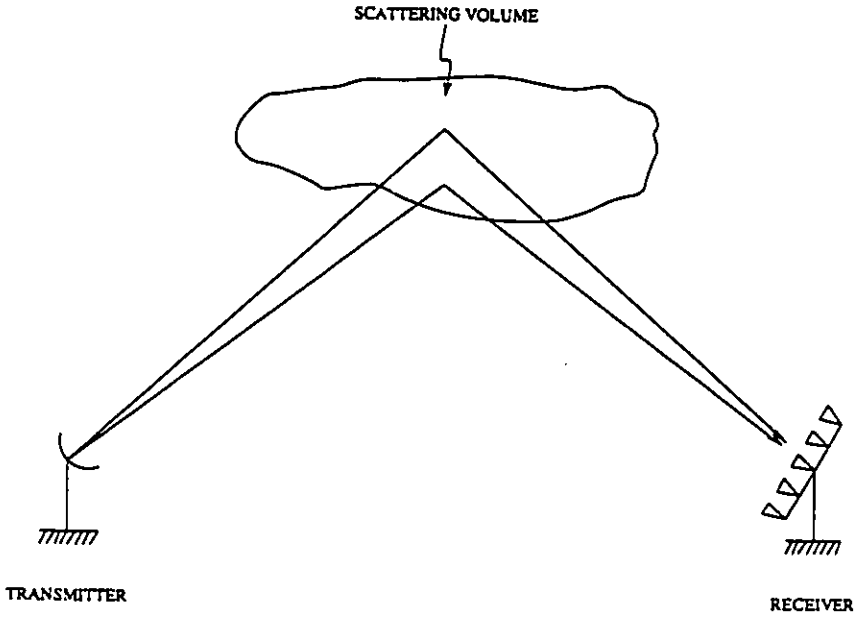


Figure 1.2: Signals transmitted by scattering propagation

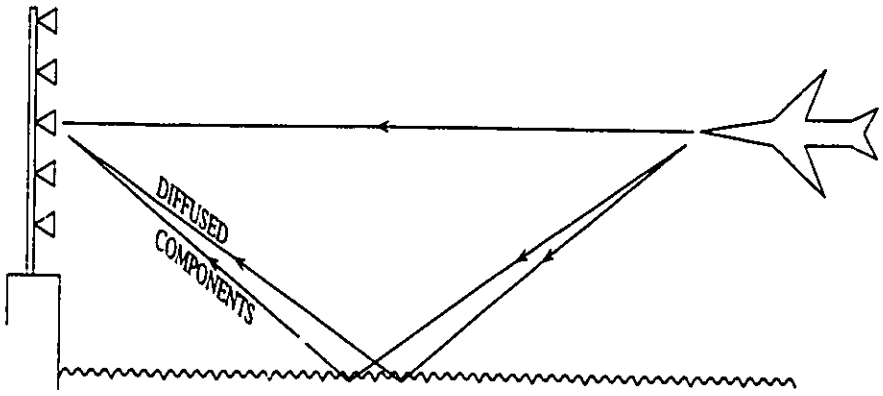


Figure 1.3: Low-angle tracking environments

Chapter 2

Signal Model for Spatially Distributed Sources

In high resolution array processing it is frequently assumed that the signals of interest are generated by point sources. A point source is a mathematical abstraction, which is a valid representation in many cases of practical applications such as in sonar and radar array processing, in radio astronomy, or in seismic exploration. However, in many other applications, as they will be discussed in the succeeding section, a point source model is no longer valid.

In this chapter, we will first introduce a number of engineering applications in which the dispersed sources occur. Then we will focus on establishing a mathematical model for spatially distributed signal sources.

2.1 Phenomena of Spatially Distributed Sources

In ocean exploration such as seafloor mapping or seabed classification, multibeam echo sounders as an active acoustic system are used to transmit acoustic pulse and receive the

echoes. However, the pulse always scatters at the seabed causing spatial extension of the source, and the pulse penetrates into the seabed and scatters from the lower layers increasing the extension [75].

In mobile communications, especially in an urban area, the signal received by the base station from a mobile terminal can be considered as resulting mainly from reflections of the radio waves. Usually, none of these different signal paths will dominate, unless there is free sight between the mobile and the base station. Therefore, the received signal from one emitter (mobile) at the base station is approximated by a sum of a number of reflections with random phases and amplitudes from the vicinity of the emitter [76], i.e., the signal impinging on the antenna array appears to be from a "cluster" as seen from the base station. The larger the distance is between the emitter and the base station, the smaller the size of the cluster will be, and the better is a "point-source" approximation of the scenario.

In radar processing [77] in which short signal pulses are used, the received signal is a superposition of the reflections of the pulse from different parts of the target. If the target is spread in range, it appears to be a distributed source, which is usually called range-spread or dispersive target. Furthermore, the reflected signals from different ranges can be assumed statistically independent, as long as the surface of the target is rough [1] compared to the carrier wavelength.

A source distribution in space can also be observed in the transmission of radio waves through ionospheric and tropospheric scatter links, in low-angle tracking at a rough sea environment, and in the propagation of audio signals in a reverberant room.

[1] According to the Rayleigh criterion [79], a surface may be considered rough or smooth depending on $h \sin\theta > \lambda/8$ or $h \sin\theta < \lambda/8$ respectively where h is the standard deviation of the height variation of the surface, θ is the incident angle measured from the normal of the surface, and λ is the wavelength of the reflected signal.

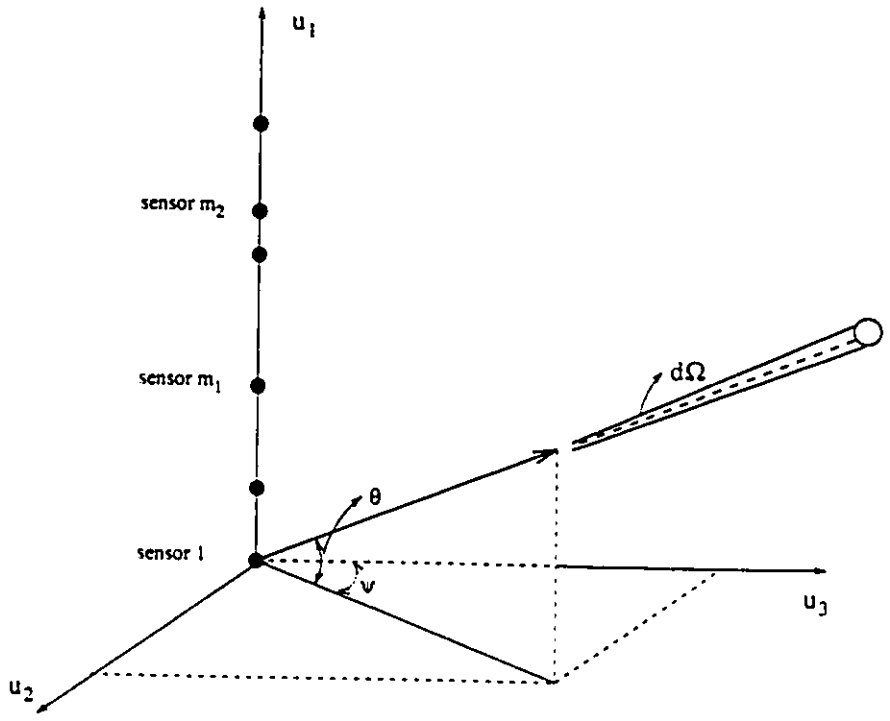


Figure 2.1: The coordinate system

2.2 Mathematical Model

2.2.1 Continuous Model

Consider a linear array of M sensors with arbitrary geometry monitoring a spatially distributed source originating from a far distance, therefore the signal waves incident upon the array can be regarded as plane waves.

We use a coordinate system as shown in Fig.2.1 to depict the scenario in which the first sensor is at the origin. The signal from one elemental source of the scattering volume, which subtends a solid angle $d\Omega$, is coming from the direction defined by elevation angle θ and azimuth angle ψ.

Given a narrow band signal

$$s(t) = ae^{j\omega_0 t} \tag{2.1}$$

where a is the random complex envelope of the signal and ω_0 is the carrier frequency, the signal density from the elemental source at the direction (θ, ψ) received by sensor m_1 with reference to sensor 1 is given by

$$s_{m_1}(\theta, \psi, t) = a(\theta, \psi)e^{j\omega_0[t-\tau_{m_1}(\theta,\psi)]} \tag{2.2}$$

where $\tau_{m_1}(\theta, \psi)$ is the time delay between the wave front propagation from sensor 1 to sensor m_1.

The total signal received by sensor m_1 can be obtained by integrating the signal density with respect to the entire scattering volume, i.e.,

$$s_{m_1}(t) = \int_{4\pi} a(\theta, \psi)e^{j\omega_0[t-\tau_{m_1}(\theta,\psi)]}d\Omega \tag{2.3}$$

Similarly, we can obtain the signal received by sensor m_2.

The spatial correlation between signals received by sensors m_1 and m_2 is given by

$$\begin{aligned} & R_{s_{m_1}s_{m_2}}(\tau, m_1, m_2) \\ = \; & \mathrm{E}\{s_{m_1}(t)s^*_{m_2}(t)\} \\ = \; & \int_{4\pi}\int_{4\pi} \mathrm{E}\{a(\theta_i, \psi_i)a^*(\theta_j, \psi_j)\}e^{j\omega_0[\tau_{m_2}(\theta_j,\psi_j)-\tau_{m_1}(\theta_i,\psi_i)]}d\Omega_i d\Omega_j \end{aligned} \tag{2.4}$$

We assume that the signal from different elemental sources are uncorrelated, i.e.,

$$\mathrm{E}\{a(\theta_i, \psi_i)a^*(\theta_j, \psi_j)\} = \mathrm{E}\{|a(\theta_i, \psi_i)|^2\}\delta_{ij} \tag{2.5}$$

where δ_{ij} is the Kronecker delta. This assumption usually holds in the transmission of radio waves by ionospheric scattering as well as in the reception of reflected signals from rough surfaces [78, 79] and is also supported by recent experimental results in the cases of tropospheric and ionospheric scattering [80, 81]. Under this assumption, the spatial correlation of Eq.(2.4) can, by using Eq.(2.5), be written as

$$R_{s_{m_1}s_{m_2}}(m_1, m_2) = \int_{4\pi} \mathrm{E}\{|a(\theta_i, \psi_i)|^2\}e^{j\omega_0[\tau_{m_2}(\theta_i,\psi_i)-\tau_{m_1}(\theta_i,\psi_i)]}d\Omega_i \tag{2.6}$$

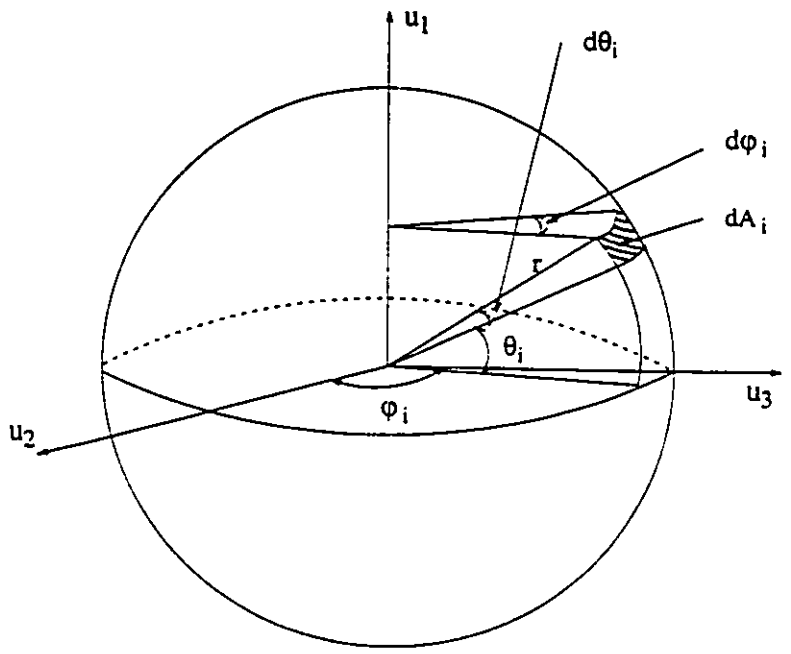


Figure 2.2: The spherical coordinate system (r, θ, φ)

To further develop the model we take the special case of a uniform linear array because uniform linear arrays are most commonly used and the signal model [74] for a uniform linear array can be simplified by confining our attention on a plane in which the array and the distributed signal lie. Furthermore, most of the time, DOA estimation algorithms have many computational advantages for uniform linear arrays.

For a uniform linear array which lies on u_1-axis with the first sensor at the origin, we can see that, the propagation time delay between sensors only depends on the angle θ, different angle ψ will not introduce time delay among the sensors. Thus, the delay is given by

$$\tau_{m_1}(\theta) = d_0(m_1 - 1)\sin\theta/c \tag{2.7}$$

where c is the velocity of the wave front and d_0 is the distance between two adjacent sensors usually set at half of the wavelength of the signal. Hence, Eq.(2.6) can be simplified to

$$R_{s_{m_1} s_{m_2}}(m_1, m_2) = \int_{4\pi} \mathrm{E}\{|a(\theta_i, \psi_i)|^2\} e^{j\omega_0(m_2 - m_1)d_0 \sin\theta_i/c} d\Omega_i \tag{2.8}$$

The elemental solid angle $d\Omega_i$ is defined by the area of subtending elemental surface divided by radius square. By using the spherical coordinate system (r, θ, ψ), as seen in Fig.2.2, we have

$$d\Omega_i = \frac{dA_i}{r^2} = \frac{r\cos\theta_i d\psi_i r d\theta_i}{r^2} = \cos\theta_i d\theta_i d\psi_i \tag{2.9}$$

Therefore, by substituting Eq.(2.9) into Eq.(2.8), and using the fact that d_0 is set at half the wavelength of the transmitted signal the spatial correlation function can be written as

$$\begin{aligned} & R_{s_{m_1} s_{m_2}}(m_1, m_2) \\ = & \int_{-\pi/2}^{\pi/2}\int_{-\pi}^{\pi} \mathrm{E}\{|a(\theta_i, \psi_i)|^2\} d\psi_i \; e^{j(m_2-m_1)\omega_0 d_0 \sin\theta_i/c} d(\sin\theta_i) \\ = & p_s^2 \int_{-\pi}^{\pi} \gamma(\phi) e^{j(m_2-m_1)\phi} d\phi \end{aligned} \tag{2.10}$$

where p_s^2 denotes the power of the distributed signal and we define (using $d_0 = \lambda/2$)

$$\phi \triangleq \omega_0 d_0 \sin\theta_i / c = \pi \sin\theta_i \tag{2.11}$$

and

$$\gamma(\phi) \triangleq \frac{1}{p_s^2} \int_{-\pi}^{\pi} \mathrm{E}\{|\, a(\theta_i, \psi_i)\,|^2\} d\psi_i \tag{2.12}$$

The normalized signal density $\gamma(\phi)$ defined in Eq.(2.12) describes the distribution of the signal source when the source and the array are coplanar. If the distribution $\gamma(\phi)$ can be defined by two parameters $\bar{\phi}$ and σ which are, respectively, the mean angle of arrival and the "spread" of the signal, then the correlation matrix can be reduced to one which depends only on these two parameters.

For a linear array, we can confine our attention to a one-dimensional DOA estimation. For instance, the linear array which lies on u_1-axis is only used to estimate the elevation angle. For an array with arbitrary geometry, except when the distributed signal is flat and coplanar with the array, DOA estimation usually involves both angles of elevation and azimuth. We concentrate on the estimation of a single angle although the study can be extended to estimation of both angles.

2.2.2 Discrete Model

We can formulate the signal received by a linear array from a distributed source in another way by considering the signal distribution to be comprised of a collection of discrete sources (point sources) closely clustered together and developing a discrete model which may facilitate computation. Furthermore, we can easily extend the formulation to multiple distributed signal sources. In the discrete model, we again confine our consideration to distributed sources which are coplanar with the array of sensors so that the angle of arrival of interest is measured on the plane of the array.

In the presence of K distributed sources, the signal received by the array at nth snapshots is given by

$$s(n) = \sum_{k=1}^{K} \sum_{l=1}^{L_k} d(\phi_{kl}) s_{kl}(n) \tag{2.13}$$

where $s(n)$ is an M-dimensional vector containing the received signals of the M sensors in the array. L_k is the number of discrete "points" in the kth distributed source, whose lth discrete "point" component emits energy at some angle parameter ϕ_{kl} with respect to the normal of the array, and whose intensity (standard deviation) is s_{kl}, $l = 1, \cdots, L_k$; $k = 1, \cdots, K$. $d(\phi_{kl})$ is the direction manifold being an M-dimensional complex vector characterized by the parameter ϕ_{kl} associated with the lth component of the kth source. $d(\phi_{kl})$ depends on the geometry of the array. For a uniform linear array with M sensors separated by d_0, for instance, if the angle of arrival of the lth discrete source component of the kth distributed signal is θ_{kl} with respect to the normal of the array, then $d(\phi_{kl}) = [1 \; e^{j\phi_{kl}} \; \cdots \; e^{j(M-1)\phi_{kl}}]^T$ where $\phi_{kl} = \omega_0 d_0 \sin\theta_{kl}/c$.

If the K distributed sources are statistically independent of each other, then the correlation matrix R_s of the signal received by the array of sensors is

$$R_s = \mathrm{E}\{s(n)s^\dagger(n)\} = \sum_{k=1}^{K} R_k \tag{2.14}$$

where R_k is the covariance matrix of the received signal due to the kth distributed signal source and is given by

$$R_k = \sum_{l=1}^{L_k} d(\phi_{kl}) d^\dagger(\phi_{kl}) \gamma_{kl} \tag{2.15}$$

with

$$\gamma_{kl} = \mathrm{E}\{s_{kl} s_{kl}^*\} \tag{2.16}$$

which is the average power of the lth discrete source component of the kth distributed signal. If these discrete source components of the same diffused signal are closely clustered together, we replace each of the discrete source by the source average power p_{sk}^2 and an incremental contribution from the component $\gamma_k(\phi_l)\Delta\phi$, where $\gamma_k(\phi)$ is a normalized density function spreading over the extent of the distributed source Φ_k. Then as the incremental angle $\Delta\phi \to 0$, Eq.(2.15) becomes

$$\begin{aligned} R_k &= p_{sk}^2 \sum_{l=1}^{L_k} d(\phi_{kl}) d^\dagger(\phi_{kl}) \gamma(\phi_{kl}) \Delta\phi \\ &= p_{sk}^2 \int_{\phi \in \Phi_k} d(\phi) d^\dagger(\phi) \gamma_k(\phi) d\phi \end{aligned} \tag{2.17}$$

For the specific case of a uniform linear array, Eq.(2.17) can be shown to be the same as the continuous model in Eq.(2.10) when the appropriate array manifold $d(\phi)$ is employed. Here, $\gamma_k(\phi)$ is a normalized positive density function characterizing the distribution of the kth signal such that

$$\int_{\phi \in \Phi_k} \gamma_k(\phi) d\phi = 1 \tag{2.18}$$

with Φ_k being the region within which the distributed signal spreads.

A point source can be regarded as a distributed source with $L_k = 1$ in the discrete model, or a distributed source with an impulse power density function in the continuous model.

In this section and the previous section (Section 2.2.1), we developed the expressions for the elements of the covariance matrix of the data received by the array. The expressions are

derived based on different simplifying assumptions. Eq.(2.6) is the most general expression for a three-dimensional signal and a three-dimensional array. For the particular case of a uniform linear array being used, this expression can be simplified to Eq.(2.10). On the other hand, the research interest of this thesis is on planar arrays and distributed signals which are co-planar with the array. The angle of interest is the angle of arrival of the signal measured on the plane of the array. In this case, Eq.(2.17) is valid. As has been mentioned, if the appropriate array manifold of the uniform linear array is applied to Eq.(2.17), then we obtain Eq.(2.10), showing that both derivations are valid.

2.2.3 Some Examples of Signal Density Functions

A particular distribution function $\gamma_k(\phi)$ gives rise to a particular covariance matrix of the signals at the sensor outputs. In general, a distribution function has two important parameters $\bar{\phi}_k$ and σ_k where $\bar{\phi}_k$ denotes the mean DOA of the kth signal with respect to the array and σ_k is a measure of the "spread" of the signal.

The following are three examples of signal correlation matrix of different distributed sources.

Example 1: Uniform Power Density

Suppose that we have a linear array having M sensors uniformly spaced at half the wavelength of the narrow-band signal and that the kth signal power density is uniformly distributed within $[\bar{\phi}_k - \sigma_k, \bar{\phi}_k + \sigma_k]$, i.e.,

$$\gamma_k(\phi) = \begin{cases} \frac{1}{2\sigma_k} & |\phi - \bar{\phi}_k| \leq \sigma_k \\ 0 & \text{elsewhere} \end{cases} \tag{2.19}$$

then using Eq.(2.17) for which $\boldsymbol{d}(\phi) = [1 \;\; e^{j\phi} \;\; \cdots \;\; e^{j(M-1)\phi}]^T$, the mnth element of the correlation matrix $\boldsymbol{R}_k$ is given by

$$r_{kmn} = \int_{-\pi}^{\pi} e^{j(m-n)\phi} \gamma_k(\phi) d\phi = \frac{\sin(m-n)\sigma_k}{(m-n)\sigma_k} e^{j(m-n)\bar{\phi}_k} \tag{2.20}$$

Hence, the correlation matrix of the signal received by the array is

$$R_k = \begin{bmatrix} 1 & \frac{\sin\sigma_k}{\sigma_k}e^{-j\bar{\phi}_k} & \cdots & \frac{\sin(M-1)\sigma_k}{(M-1)\sigma_k}e^{-j(M-1)\bar{\phi}_k} \\ \frac{\sin\sigma_k}{\sigma_k}e^{j\bar{\phi}_k} & 1 & \cdots & \frac{\sin(M-2)\sigma_k}{(M-2)\sigma_k}e^{-j(M-2)\bar{\phi}-k} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\sin(M-1)\sigma_k}{(M-1)\sigma_k}e^{j(M-1)\bar{\phi}_k} & \cdot & \cdots & 1 \end{bmatrix} \tag{2.21}$$

Example 2: Gaussian Power Density

If the kth signal power density is Gaussian distributed such that

$$\gamma_k(\phi) = \frac{1}{\sqrt{2\pi\sigma_k^2}}e^{-\frac{(\phi-\bar{\phi}_k)^2}{2\sigma_k^2}} \tag{2.22}$$

where we assume that the spread width $\sigma_k \ll \pi$. Then again using Eq.(2.17), the mnth entry of the correlation matrix R_k for the same uniform linear array as in Example 1 is given by

$$r_{k\,mn} = \int_{-\pi}^{\pi} e^{j(m-n)\phi}\gamma_k(\phi)d\phi = e^{-\frac{(m-n)^2\sigma_k^2}{2}}e^{j(m-n)\bar{\phi}_k} \tag{2.23}$$

Hence, the signal spatial correlation matrix is given as

$$R_k = \begin{bmatrix} 1 & e^{-\frac{\sigma_k^2}{2}}e^{-j\bar{\phi}_k} & \cdots & e^{-\frac{(M-1)^2\sigma_k^2}{2}}e^{-j(M-1)\bar{\phi}_k} \\ e^{-\frac{\sigma_k^2}{2}}e^{j\bar{\phi}_k} & 1 & \cdots & e^{-\frac{(M-2)^2\sigma_k^2}{2}}e^{-j(M-2)\bar{\phi}_k} \\ \vdots & \vdots & \ddots & \vdots \\ e^{-\frac{(M-1)^2\sigma_k^2}{2}}e^{j(M-1)\bar{\phi}_k} & e^{-\frac{(M-2)^2\sigma_k^2}{2}}e^{j(M-2)\bar{\phi}_k} & \cdots & 1 \end{bmatrix} \tag{2.24}$$

Example 3: "First-order Butterworth" Distribution

If the signal power density is given by

$$\gamma_k(\phi) = \frac{\sigma_k}{\sigma_k^2 + (\phi - \bar{\phi}_k)^2} \tag{2.25}$$

where σ_k is the $-3dB$ "width" of the spectrum of $\gamma_k(\phi)$ and we assume $\sigma_k \ll \pi$. Again by using Eq.(2.17), the mnth entry of R_k for the same uniform linear array is given as

$$r_{k\,mn} = \int_{-\pi}^{\pi} e^{j(m-n)\phi}\gamma_k(\phi)d\phi = e^{-|m-n|\sigma_k}e^{j(m-n)\bar{\phi}_k} \tag{2.26}$$

Hence, the correlation matrix of the signal

$$R_k = \begin{bmatrix} 1 & z^* & z^{*2} & \cdots & z^{*(M-1)} \\ z & 1 & z^* & \cdots & z^{*(M-2)} \\ z^2 & z & 1 & \cdots & z^{*(M-3)} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ z^{M-1} & z^{M-2} & z^{M-3} & \cdots & 1 \end{bmatrix} \tag{2.27}$$

where $z = e^{-\sigma_k} e^{j\bar{\phi}_k}$

From the above three examples, we can see that for a distributed signal with symmetric power density about the mean DOA $\bar{\phi}$, each entry of the signal correlation matrix has both a modulus part which is dependent on the spread σ_ϕ and an argument part which is dependent on the mean DOA. For the case of an unsymmetric density, we can similarly separate each element of the correlation matrix into two parts, one that depends on the spread only, and the other that depends on the mean DOA only. However, the spread-dependent part will not be real in this case.

In the above three examples, if the width of the spread in each power density function is reduced to zero, then the power density function tends to a Dirac δ-function. Accordingly, the correlation matrix R_k in each case becomes

$$R_k = \begin{bmatrix} 1 & e^{-j\bar{\phi}_k} & \cdots & e^{-j(M-1)\bar{\phi}_k} \\ e^{j\bar{\phi}_k} & 1 & \cdots & e^{-j(M-2)\bar{\phi}_k} \\ \vdots & \vdots & \ddots & \vdots \\ e^{j(M-1)\bar{\phi}_k} & e^{j(M-2)\bar{\phi}_k} & \cdots & 1 \end{bmatrix} \tag{2.28}$$

which is the same as the correlation matrix for a uniform linear array with M sensors in a conventional point source data model.

Chapter 3

Properties of the Signal Correlation Matrix

In this chapter, we will examine the properties of the signal model we established in Chapter 2, i.e., the spatial correlation matrix for a distributed source. We study the correlation matrix $\boldsymbol{R}_k$ given by Eq.(2.15) in the specific case of a uniform linear array which is one of the most commonly used arrays.

3.1 Eigen-Analysis

As we have mentioned in Chapter 2, if a distributed signal power density is symmetric to the mean DOA, then the correlation matrix $\boldsymbol{R}_k$ for a uniform linear array with M sensors for the kth distributed signal is generally of the following form

$$\boldsymbol{R}_k = \begin{bmatrix} r_{11} & r_{12}e^{-j\tilde{\phi}_k} & \cdots & r_{1M}e^{-j(M-1)\tilde{\phi}_k} \\ r_{21}e^{j\tilde{\phi}_k} & r_{22} & \cdots & r_{2M}e^{-j(M-2)\tilde{\phi}_k} \\ \vdots & \vdots & \ddots & \vdots \\ r_{M1}e^{j(M-1)\tilde{\phi}_k} & r_{M2}e^{j(M-2)\tilde{\phi}_k} & \cdots & r_{MM} \end{bmatrix} = \boldsymbol{L}_k \boldsymbol{R}_r \boldsymbol{L}_k^{\dagger} \qquad (3.1)$$

where

$$\boldsymbol{R}_r = \begin{bmatrix} r_{11} & r_{12} & \cdots & r_{1M} \\ r_{21} & r_{22} & \cdots & r_{2M} \\ \vdots & \vdots & \ddots & \vdots \\ r_{M1} & r_{M2} & \cdots & r_{MM} \end{bmatrix} \tag{3.2}$$

and

$$\boldsymbol{L}_k = \begin{bmatrix} 1 & & & \boldsymbol{O} \\ & e^{j\bar{\phi}_k} & & \\ & & \ddots & \\ \boldsymbol{O} & & & e^{j(M-1)\bar{\phi}_k} \end{bmatrix} \tag{3.3}$$

with r_{mn} being real quantities such that $r_{mn} = r_{nm}$ and being only dependent on the spread, and $\bar{\phi}_k$ being the mean DOA. If the signal power density is not symmetric to the mean DOA, then r_{mn} are generally complex quantities and $r_{mn} = r_{nm}^*$. We note that, in general, $\boldsymbol{R}_r$ is of full rank. Also, since $\boldsymbol{R}_r$ is Hermitian, its eigenvalues are real and non-negative, and its eigenvectors are mutually orthogonal.

Theorem 3.1 *If matrix $\boldsymbol{R}_r$ has eigenvalues $\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_M$ with corresponding eigenvectors $\boldsymbol{v}_1, \boldsymbol{v}_2, \cdots, \boldsymbol{v}_M$, then $\boldsymbol{R}_k$ has the same eigenvalues $\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_M$ with corresponding eigenvectors*

$$\boldsymbol{L}_k \boldsymbol{v}_m = \begin{bmatrix} 1 & & & \boldsymbol{O} \\ & e^{j\bar{\phi}_k} & & \\ & & \ddots & \\ \boldsymbol{O} & & & e^{j(M-1)\bar{\phi}_k} \end{bmatrix} \boldsymbol{v}_m, \quad m = 1, 2, \cdots, M \tag{3.4}$$

Proof:

By eigen-decomposition, matrix $\boldsymbol{R}_r$ can be written as

$$\boldsymbol{R}_r = \boldsymbol{V} \boldsymbol{\Lambda} \boldsymbol{V}^\dagger$$

where $\Lambda = \text{diag}(\lambda_1, \lambda_2, \cdots, \lambda_M)$ is the eigenvalue matrix of R_r and $V = [v_1\ v_2\ \cdots\ v_M]$ is the eigenvector matrix of R_r.

From Eq.(3.1), we know

$$\begin{aligned} R_k &= L_k R_r L_k^\dagger \\ &= L_k V \Lambda V^\dagger L_k^\dagger \\ &= (L_k V)\Lambda(L_k V)^\dagger \end{aligned}$$

where L_k is given by Eq.(3.3).

It is easy to verify that $L_k V$ is a unitary matrix. Therefore, Λ is the eigenvalue matrix and $L_k V$ is the eigenvector matrix of matrix R_k, i.e., R_k has the eigenvalues $\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_M$ with corresponding eigenvectors $L_k v_1$, $L_k v_2$, $\cdots$, $L_k v_M$. □

This theorem says that we only need to examine the real matrix R_r if we want to obtain the eigenvalues and eigenvectors of the signal correlation matrix R_k.

3.2 Eigenvalues and Eigenvectors of a Uniformly Distributed Signal Correlation Matrix

From Example 1 in Chapter 2, we can see that for the case of uniformly distributed signal source, the *mnth* element of the real matrix R_r is given by

$$r_{mn} = \frac{\sin(m-n)\sigma_k}{(m-n)\sigma_k} \quad m, n = 1, 2, \cdots, M \tag{3.5}$$

where $2\sigma_k$ is the total width of the distribution. Then the mth eigenvectors of the matrix R_r is formed by the elements of the *discrete prolate spheroidal sequences* $\zeta_m^{(i)}(\sigma_k, M)$ (see Appendix A), and the corresponding eigenvalue $\lambda_m(\sigma_k)$ is given by

$$\lambda_m(\sigma_k) = \frac{\pi \sum_{i=1}^{M} |\, \zeta_m^{(i)}(\sigma_k, M)\,|^2}{\sigma_k \sum_{i=-\infty}^{\infty} |\, \zeta_m^{(i)}(\sigma_k, M)\,|^2} \quad m = 1, \cdots, M \tag{3.6}$$

Table 3.1: Eigenvalues of a uniformly distributed signal

σ_k	.1745	.3491	.5236
λ_{k1}	7.5935e-00	6.5799e-00	5.3928e-00
λ_{k2}	4.0301e-01	1.3654e-00	2.3460e-00
λ_{k3}	3.4498e-03	5.4000e-02	2.5322e-01
λ_{k4}	1.0676e-05	6.8853e-04	7.8551e-03
λ_{k5}	1.5762e-08	2.3514e-06	5.4649e-05
λ_{k6}	1.1449e-11	1.2518e-08	7.4453e-07
λ_{k7}	5.0777e-13	1.6884e-11	2.5461e-09
λ_{k8}	5.0319e-13	2.0202e-12	1.1466e-12

The eigenvalues are plotted in Fig.3.1a to Fig.3.1c, where the eigenvalues are normalized to the largest eigenvalue for $M = 20$. Table 3.1 shows the numerical value of $\lambda_m(\sigma_k)$ for various width of σ_k when $M = 8$. We notice that after $m > m_0 = \lceil \frac{\sigma_k M}{2} \rceil$, where σ_k is in radian, $\lambda_m(\sigma_k)$ is negligible being less than 5% of the total energy. Here $\lceil x \rceil$ denotes the smallest integer containing x. This means that most of the signal energy are concentrated in the first m_0 eigenvalues, therefore, the first m_0 eigenvectors corresponding to these eigenvalues span a subspace which is close to the real signal subspace with an approximation error of $\sum_{m=m_0+1}^{M} \lambda_m(\sigma_k)$ if measured by energy loss.

3.3 Eigenvalues and Eigenvectors of a Gaussian Distributed Signal Correlation Matrix

For the case of Gaussian distributed signal source, the elements of the real matrix R_r are given by

$$r_{mn} = e^{-(m-n)^2\sigma_k^2/2} \quad m, n = 1, 2, \cdots, M \tag{3.7}$$

where σ_k is the width of the distribution.

The eigenvalues $\lambda_m(\sigma_k)$ and the corresponding eigenvectors $\xi_m(\sigma_k)$ can be obtained by

Table 3.2: Eigenvalues of a Gaussian distributed signal

σ_k	.0583	.1163	.1745
λ_{k1}	7.6969e-00	7.4822e-00	6.9524e-00
λ_{k2}	2.9846e-01	5.0369e-01	9.8349e-01
λ_{k3}	4.5987e-03	1.3898e-02	6.1805e-02
λ_{k4}	4.1575e-05	2.2466e-04	2.2810e-03
λ_{k5}	2.4320e-07	2.3514e-06	5.4649e-05
λ_{k6}	9.2058e-10	1.5953e-08	8.5213e-07
λ_{k7}	1.7032e-12	6.4096e-11	7.9857e-09
λ_{k8}	3.7873e-13	5.5602e-13	3.3071e-11

solving the following equation by standard numerical subroutines,

$$\lambda_m(\sigma_k)\xi_m(\sigma_k) = \boldsymbol{R}_r(\sigma_k, M)\xi_m(\sigma_k) \tag{3.8}$$

where $\boldsymbol{R}_r(\sigma_k, M)$ has elements given by Eq.(3.7). The eigenvalues are plotted in Fig.3.2a to Fig.3.2c for different σ_k, where again, the eigenvalues are normalized to the largest eigenvalue for $M = 20$. We can see that $\lambda_1(\sigma_k) \geq \lambda_2(\sigma_k) \geq \cdots \geq \lambda_M(\sigma_k)$ and more and more eigenvalues $\lambda_m(\sigma_k)$ have significant value when M increases. We can also examine the numerical values of eigenvalues $\lambda_m(\sigma_k)$ from Table 3.2 for the case of $M = 8$, and we note that for a fixed value of σ_k, $\lambda_m(\sigma_k)$ fall off to zero rapidly with increasing m once m has exceeds $m_0 = \lceil \frac{\pi\sigma_k M}{2} \rceil$, where again, $\lceil \frac{\pi\sigma_k M}{2} \rceil$ is the first integer greater than or equal to $\frac{\pi\sigma_k M}{2}$. Therefore, we can choose the first m_0 eigenvalues as the signal eigenvalues and the corresponding eigenvectors as the basis spanning the approximated signal subspace.

3.4 Eigenvalues and Eigenvectors of the Correlation Matrix of a Signal with "first-order Butterworth" Distribution

For the case of a signal source with "first-order" Butterworth distribution, the elements of the real matrix $\boldsymbol{R}_r$ are given by

$$r_{mn} = e^{-|m-n|\sigma_k} \quad m, n = 1, 2, \cdots, M \tag{3.9}$$

Table 3.3 shows the numerical values of the eigenvalues λ_{km} of the matrix $\boldsymbol{R}_r$ for various "spread" of σ_k when $M = 8$. Again, the eigenvalues $\lambda_m(\sigma_k)$ and the corresponding eigenvectors $\xi_m(\sigma_k)$ can be obtained by solving eigen equation by standard numerical subroutines.

The eigenvalues are plotted in Fig.3.3a to Fig.3.3c for different σ_k, where the eigenvalues are normalized to the largest eigenvalue for $M = 20$. We can also examine the numerical values of eigenvalues $\lambda_m(\sigma_k)$ from Table 3.3. As we can see both in the plots and the table that, for a fixed value of σ_k, $\lambda_m(\sigma_k)$ fall off to relatively small magnitudes abruptly once m has exceeds $m_0 = \lceil 10\sigma_k M \rceil$, where again, $\lceil 10\sigma_k M \rceil$ is the first integer greater than or equal to $10\sigma_k M$. Therefore, we can choose the first m_0 eigenvalues as the signal eigenvalues and the corresponding eigenvectors as the basis spanning the approximate signal subspace.

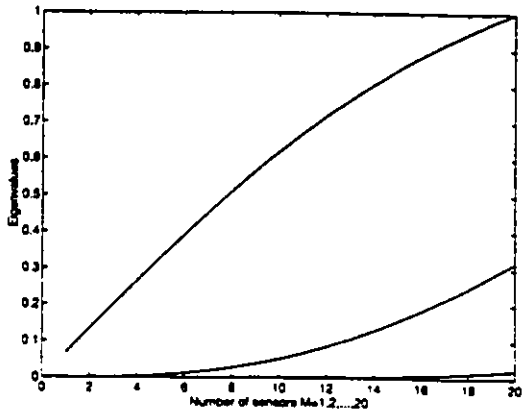
3.5 Dimensionality of Signal Subspace of Distributed Sources

From the examples in the preceding sections, we observe that, in general, the number of eigenvalues in which a dominant proportion of signal energy concentrates depends on the number of sensors M, the extent to which the signal spreads (σ_k), and to a lesser degree, the distribution of the signal. The larger is M, and/or the larger is the signal spread, the more eigenvalues have to be accounted for the concentration of signal energy. Figs.3.1 $\sim$ 3.3 show the plots of the eigenvalues against the number of the sensors ($M = 1, \cdots, 20$) for the cases of

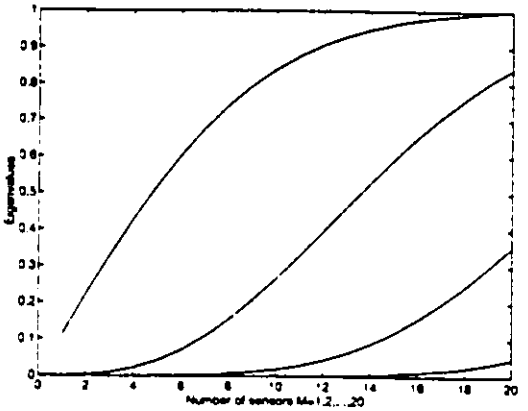
Table 3.3: Eigenvalues of a signal with "first-order Butterworth" distribution

σ_k	.0009	.0175	.0262
λ_{k1}	7.8201e-00	7.6466e-00	7.4794e-00
λ_{k2}	1.1152e-01	2.1710e-01	3.1692e-01
λ_{k3}	2.9598e-02	5.8823e-02	8.7633e-02
λ_{k4}	1.4099e-02	2.8133e-02	4.2083e-02
λ_{k5}	8.7152e-03	1.7415e-02	2.6089e-02
λ_{k6}	6.3069e-03	1.2611e-02	1.8904e-02
λ_{k7}	5.1099e-03	1.0221e-02	1.5326e-02
λ_{k8}	4.5348e-03	9.0716e-03	1.3605e-02

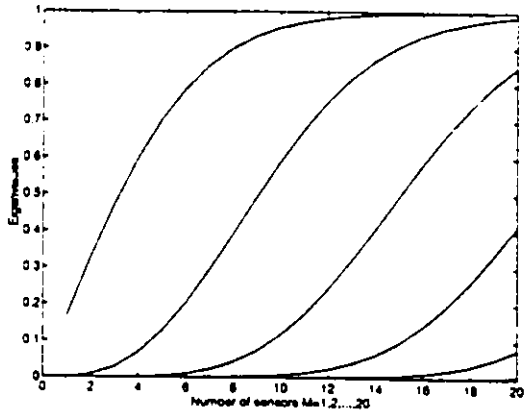
uniformly, Gaussian and "first-order" Butterworth distributed signals with different spreads. It is further observed that with the increase of M and/or σ_k, the energy concentration spreads more quickly to other eigenvalues in the case of a uniformly distributed signal than signals with other distributions. From the above observations, we can infer that even though the matrix $\boldsymbol{R}_k$ is of full-rank, because of the energy concentration in only m_{0k} ($< M$) of the dominant eigenvalues, the dimension of the signal subspace induced by a distributed signal can be approximated by m_{0k}. This *quasi signal subspace* is spanned by the eigenvectors corresponding to the largest m_{0k} eigenvalues of $\boldsymbol{R}_k$. The evaluation of the approximate dimension of the signal subspace by examining the concentration of signal energy in the dominant eigenvalues is a common technique and has been used by several researchers [83, 84]. Since the uniform distribution represents the widest spread of signal energy among the dominant eigenvalues, i.e., its signal subspace has the largest approximate dimension, we can treat this as "worst case" so that for a given number of sensors and an estimated spread, we can examine from graphs such as Figs.3.1 the number of dominant eigenvalues and determine the approximate signal subspace dimension accordingly. We further note that if there are K uncorrelated distributed signals, then the number of sensors must satisfy $M > \sum_{k=1}^{K} m_{0k}$.



(a) $\sigma_k = .1745 rad.$

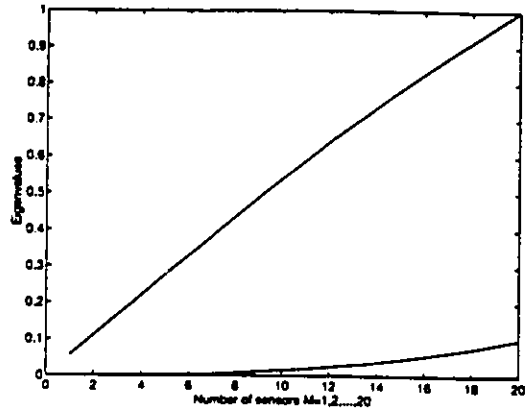


(b) $\sigma_k = .3491 rad.$

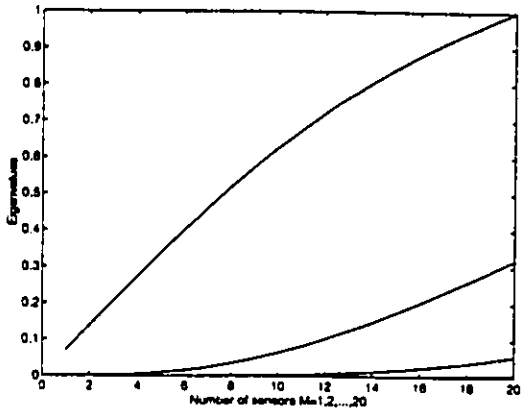


(c) $\sigma_k = .5236 rad.$

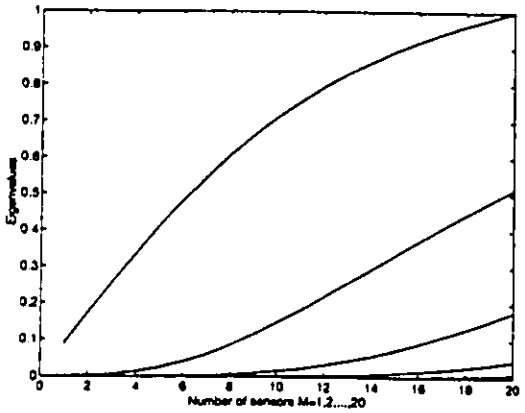
Figure 3.1: Eigenvalues of a Uniformly distributed signal



(a) $\sigma_k = .0583rad.$

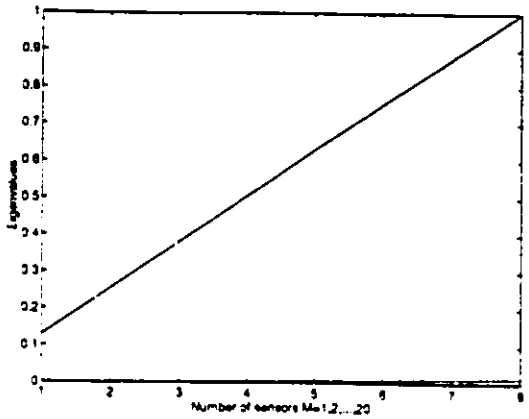


(b) $\sigma_k = .1163rad.$

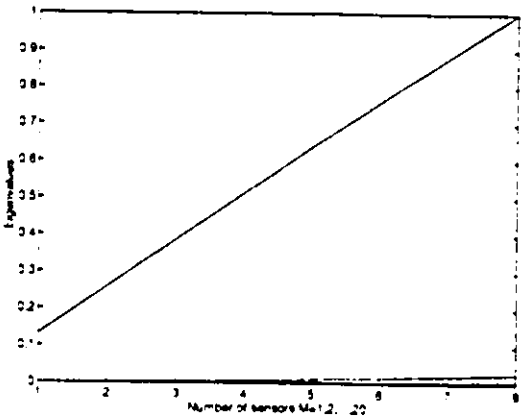


(c) $\sigma_k = .1745rad.$

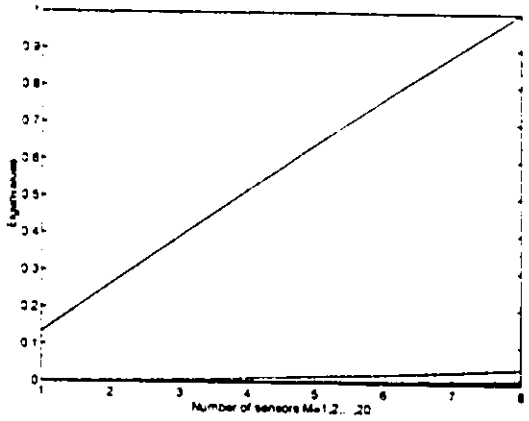
Figure 3.2: Eigenvalues of a Gaussian distributed signal



(a) $\sigma_k = .0009rad.$



(b) $\sigma_k = .0175rad.$



(c) $\sigma_k = .0262rad.$

Figure 3.3: Eigenvalues of a signal with "first-order Butterworth" distribution

Chapter 4

DOA Estimation Based on Second-order Statistics

In this chapter, we develop the DOA estimation methods [85] for distributed signal sources in a white sensor noise environment which is assumed to be statistically independent of signals. The development of the algorithms is based on the signal model established in Chapter 2 and the properties examined in Chapter 3 play an important part. We assume that the number of distributed signal sources, K, has been correctly determined.

4.1 Data Model in a White Noise Environment

In the presence of K distributed sources, and when the sensor noise is spatially white, the array output at the nth snapshots can be modelled as

$$\boldsymbol{x}(n) = \boldsymbol{s}(n) + \boldsymbol{\nu}(n) \tag{4.1}$$

where $\boldsymbol{x}(n)$ is an M-dimensional vector containing the output data of the M sensors in the array, $\boldsymbol{s}(n)$ is the received signal components which is modelled as shown in Eq.(2.13), and

$\nu(n)$ is an $M \times 1$ complex random noise vector assumed to be ergodic, independent of the signals, with zero mean and covariance $p_\nu^2 I$ where p_ν is an unknown constant and I is the identity matrix.

Therefore, the observation data correlation matrix R_x is given by

$$R_x = R_s + p_\nu^2 I \tag{4.2}$$

where R_s is the correlation matrix due to the K distributed signals. If we further assume that the number of distributed signals K is known and the signals are uncorrelated with each other, by using Eq.(2.14), then the data correlation matrix can be rewritten as

$$R_x = \sum_{k=1}^{K} R_k + p_\nu^2 I \tag{4.3}$$

Now, we are ready to derive a parametric DOA estimation method for distributed signals.

4.2 DISPARE – DOA Estimation for Distributed Signals

From the signal model described in Chapter 2, we see that there are two essential parameters $\bar{\theta}_k$ and σ_k which characterize the mean location and the spread of the kth distributed signal, and have to be estimated using the data covariance matrix R_x given by Eq.(4.3).

We can first express the covariance matrix due to the K distributed signal as

$$R_s \triangleq \sum_{k=1}^{K} R_k = \sum_{k=1}^{K} \{ \sum_{m=1}^{M} \lambda_{km} v_{km} v_{km}^\dagger \} \tag{4.4a}$$

However, given that each of the distributed signals is of a particular distribution and has m_{0k} dominant eigenvalues, we now define the correlation matrix formed by the eigenvectors associated with only the dominant eigenvalues such that

$$S \triangleq \sum_{k=1}^{K} \{ \sum_{m=1}^{m_{0k}} \lambda_{km} v_{km} v_{km}^\dagger \} = \sum_{k=1}^{K} V_{sk} \Lambda_k V_{sk}^\dagger \tag{4.4b}$$

where $\Lambda_k = \text{diag}(\lambda_{k1}, \cdots, \lambda_{km_{0k}})$.

Let

$$M_0 = \sum_{k=1}^{K} m_{0k} < M \tag{4.5}$$

since λ_{km} are negligible for $M_0 < m < M$, we can say that

$$\boldsymbol{R}_s \simeq \boldsymbol{S} \tag{4.6}$$

Thus, we can say that the rank of $\boldsymbol{R}_s$ is approximately M_0. If we let $\{\boldsymbol{w}_1, \cdots, \boldsymbol{w}_{M_0}\}$ be the eigenvectors corresponding to the largest M_0 eigenvalues of $\boldsymbol{R}_s$, then each $\boldsymbol{w}_m$, $m = 1, \cdots, M_0$, is a linear combination of $\{\boldsymbol{v}_{km};\ m = 1, \cdots, m_{0k};\ k = 1, \cdots, K\}$. In developing the above argument, we have tacitly assumed that none of the λ_{km} that we have neglected in the approximation from Eqs.(4.4a) to (4.4b) is larger than any of the λ_{km} we have retained. This set of eigenvectors $\{\boldsymbol{w}_1, \cdots, \boldsymbol{w}_{M_0}\}$ spans the *quasi signal subspace* $\mathcal{S}$.

Now, the covariance matrix $\boldsymbol{R}_x$ of the received data is the sum of $\boldsymbol{R}_s$ and the covariance matrix of the spatially white noise as indicated by Eq.(4.2). Therefore, corresponding to the M eigenvalues $\mu_1, \cdots, \mu_M$, the eigenvectors of $\boldsymbol{R}_x$ are given by $\{\boldsymbol{w}_1, \cdots, \boldsymbol{w}_M\}$, where the first M_0 eigenvectors $\{\boldsymbol{w}_m\}$, $m = 1, \cdots, M_0$ are identical to those of $\boldsymbol{R}_s$, whereas the last $(M - M_0)$ eigenvectors corresponding to the least significant eigenvalues of $\boldsymbol{R}_x$ span the *quasi noise subspace* $\mathcal{N}$.

If we form the quasi signal eigenvector matrix such that

$$\boldsymbol{W}_s = [\boldsymbol{w}_1 \ \cdots \ \boldsymbol{w}_{M_0}] \tag{4.7a}$$

and the quasi noise eigenvector matrix such that

$$\boldsymbol{W}_\nu = [\boldsymbol{w}_{M_0+1} \ \cdots \ \boldsymbol{w}_M] \tag{4.7b}$$

we have the orthogonality condition that

$$\boldsymbol{W}_s^\dagger \boldsymbol{W}_\nu = \boldsymbol{O} \tag{4.8}$$

Since the columns of W_s, are linear combinations of $\{v_{km}\}$, we also have the orthogonality condition that

$$V_{sk}^{\dagger} W_{\nu} = O \tag{4.9}$$

where

$$V_{sk} = [v_{k1} \cdots v_{km_0}], \quad k = 1, \cdots, K \tag{4.10}$$

with $v_{k1}, \cdots, v_{km_{0k}}$ being the eigenvectors corresponding to the m_{0k} most significant eigenvalues of R_k due to the kth distributed signal.

In practice, we do not know R_x exactly; we can only obtain an estimate of it by averaging over N sample observations such that

$$\hat{R}_x = \frac{1}{N} \sum_{n=1}^{N} x(n) x^{\dagger}(n) \tag{4.11}$$

It has been proved [11, 14] that

$$\lim_{N \to \infty} \hat{R}_x \underset{a.s.}{=} R_x, \text{ at a rate of } N^{-\frac{1}{2}} \tag{4.12}$$

where "$\underset{a.s.}{=}$" indicates that the equality holds with probability 1.

From $\hat{R}_x$ we can obtain the eigenvectors $\{\hat{w}_m\}$ corresponding to the eigenvalues $\{\hat{\mu}_m\}$, $m = 1, \cdots, M$. Since $\hat{R}_x$ converges asymptotically to R_x, we can write

$$\lim_{N \to \infty} W_s^{\dagger} \hat{W}_{\nu} \underset{a.s.}{=} O \tag{4.13a}$$

and

$$\lim_{N \to \infty} V_{sk}^{\dagger} \hat{W}_{\nu} \underset{a.s.}{=} O, \quad k = 1, \cdots, K \tag{4.13b}$$

where

$$\hat{W}_{\nu} = [\hat{w}_{M_0+1} \cdots \hat{w}_M] \tag{4.14}$$

These asymptotic orthogonality properties are the bases for the development of the algorithms for the estimation of the mean DOA of the distributed signals in this chapter. Two

different algorithms of estimating the parameters $\bar{\phi}_k$ and σ_k of the kth distributed signal are derived. These are called the Distributed Signal Parametric Estimation (DISPARE) methods.

4.2.1 DISPARE I

Let us examine the weighted inner products of eigenvectors $\boldsymbol{R}_k$ and the estimated quasi noise eigenvectors,

$$\gamma_{kim} = \lambda_{km}^{1/2} \boldsymbol{v}_{km}^{\dagger} \hat{\boldsymbol{w}}_i, \;\; i = M_0 + 1, \cdots, M; \;\; k = 1, \cdots, K; \;\; m = 1, \cdots, M \tag{4.15}$$

For a reasonable large number of snapshots N, we note that γ_{kim} should be a very small quantity. This is because when $m = 1, \cdots, M_0$, $\boldsymbol{v}_{km}$ and $\hat{\boldsymbol{w}}_i$ are nearly orthogonal by virtue of Eq.(4.13b) and thus have a very small inner product. On the other hand, for $m = M_0 + 1, \cdots, M$, the eigenvalues λ_{km} are negligibly small leading to a negligible value of γ_{kim}. The DISPARE I algorithm seeks to minimize the sum of the squares of these inner products to locate the mean DOA $\bar{\phi}_k$ and the spread σ_k of the kth signal, i.e.,

$$\begin{aligned} \{\hat{\bar{\phi}}_k, \hat{\sigma}_k\} &= \arg \min_{\bar{\phi}_k, \sigma_k} \sum_{i,m} \mid \gamma_{kim} \mid^2 \\ &= \arg \min_{\bar{\phi}_k, \sigma_k} \mathrm{tr}\{\hat{\boldsymbol{W}}_{\nu}^{\dagger} \boldsymbol{R}_k \hat{\boldsymbol{W}}_{\nu}\} \end{aligned} \tag{4.16}$$

where $\arg\min_{\bar{\phi}_k, \sigma_k}(.)$ denotes the minimization of the quantity with respect to the parameters $\bar{\phi}_k$ and σ_k, and tr(.) denotes the trace of a square matrix. Eq.(4.16) expresses the criterion of DISPARE I which, when minimized will yield the values of the estimated parameters $\bar{\phi}_k$ and σ_k of the kth distributed signal. We note that if $\sigma_k = 0$, $\boldsymbol{R}_k = \boldsymbol{d}(\phi_k)\boldsymbol{d}^{\dagger}(\phi_k)$, and Eq.(4.16) is reduced to the MUSIC algorithm as described in Chapter 1.

4.2.2 DISPARE II

We may further exploit the orthogonal relationship of Eq.(4.13b). Due to the asymptotic orthogonality between $\boldsymbol{V}_{sk}$ and $\hat{\boldsymbol{W}}_\nu$, we can write

$$\lim_{N\to\infty} \boldsymbol{V}_{sk}\boldsymbol{\Lambda}_k\boldsymbol{V}_{sk}^\dagger\hat{\boldsymbol{W}}_\nu \overset{a.s.}{=} \boldsymbol{O} \tag{4.17}$$

But from the approximation of the rank of the signal covariance matrices as indicated by Eq.(4.4b), Eq.(4.17) can be rewritten as

$$\lim_{N\to\infty} \boldsymbol{R}_k\hat{\boldsymbol{W}}_\nu \simeq \boldsymbol{O} \tag{4.18}$$

where $\boldsymbol{R}_k$ is the covariance matrix due to the kth distributed signal. Thus, we can conceive an algorithm in which we adjust the parameters $\bar{o}_k$ and σ_k in $\boldsymbol{R}_k$ until Eq.(4.18) is satisfied. Since the square of the Frobenius norm of a matrix is the measure of the sum of the square of each element of the matrix, we can arrive at the estimates such that

$$\{\hat{\bar{o}}_k, \hat{\sigma}_k\} = \arg\min_{\bar{o}_k,\sigma_k} \|\boldsymbol{R}_k\hat{\boldsymbol{W}}_\nu\|_F^2 \tag{4.19}$$

where $\|.\|_F$ denotes the Frobenius norm. Eq.(4.19) yields the DISPARE II estimates of the parameters $\bar{o}_k$ and σ_k of the kth signal. It can easily be shown by using the relation that $\|\boldsymbol{R}_k\hat{\boldsymbol{W}}_\nu\|_F^2 = \mathrm{tr}\{\hat{\boldsymbol{W}}_\nu^\dagger\boldsymbol{R}_k\boldsymbol{R}_k^\dagger\hat{\boldsymbol{W}}_\nu\}$ that the DISPARE II criterion amounts to the minimization of $\sum_{i,m}\lambda_{km} \mid \gamma_{kim} \mid^2$ where γ_{kim} is given by Eq.(4.15), i.e., DISPARE II employs a weighting on the inner products of the eigenvectors of $\boldsymbol{R}_k$ and the estimated quasi noise eigenvectors different from that used in DISPARE I.

4.3 Application of the DISPARE Algorithms

Let us first test the effectiveness of the two DISPARE algorithms by carrying out some numerical experiments. In each of these examples, an 8-sensor linear array with uniform

spacing of half-wavelength of the signal source is used. All sensors are omnidirectional and have unit gain. Isotropic Gaussian noise with zero mean is added to the received signal. The number of snapshots is taken to be 200. In each case an exhaustive search for the minimum of the respective criterion is located in the $\bar{\phi}_k \sim \sigma_k$ plane to obtain the estimates $\hat{\bar{\phi}}_k$ and $\hat{\sigma}_k$. The root mean square (RMS) of the estimation error (average over 500 trials) are plotted for different signal-to-noise ratios (SNR).

Experiment 4.1

In this experiment we have the situation in which there is only one uniformly distributed signal arriving at a mean DOA of $\bar{\phi}_k = -.5454$ *rad.* corresponding to a physical angle of $-10°$ from the normal of the array. We study two cases in which the signal has spreads $\sigma_k = .1745$ *rad.* and $\sigma_k = .3491$ *rad.* corresponding to total physical width of 6.4° and 13.4° respectively. Examining the eigenvalues of the covariance matrix $\boldsymbol{R}_k$ from plots in Figs.3.1a and 3.1b, we determine that the approximate dimensions of the signal subspace in each case are 1 and 2 respectively. We now apply the two criteria given by Eqs.(4.15) and (4.18) to obtain the estimates $\hat{\bar{\phi}}_k$ and $\hat{\sigma}_k$. Since we are, in general, only interested in the DOA, we only plot the RMS values of the estimation error for the two cases. These are shown in Figs.4.1a and 4.1b respectively for the two different signal spreads at various SNR. As a comparison, we also employ MUSIC to estimate the DOA (given the knowledge that $K = 1$). It can be observed that for small signal spread all the three methods perform almost identically (Fig.4.1a, the performance of the three methods coincide). However, for a large signal spread (Fig.4.1b), while there is little difference between the performance of the two DISPARE algorithms, MUSIC completely breaks down, yielding unacceptably large errors over the whole range of SNR.

Experiment 4.2

In this experiment, there is again only one signal arriving at a mean DOA of $\bar{\phi}_k =$

$-.5454$ *rad.*. However, here the signal is Gaussian distributed. The spread is chosen to be .0583 *rad.* and .1163 *rad.* corresponding to a physical spread of 1.1° and 2.2° respectively. We note that 99% of the energy of the signal energy is collected in the range within $(\bar{\phi}_k \pm 3\sigma_k)$ making the energy spread of the signals about the same as the corresponding ones in Experiment 4.1. From the plots of the eigenvalues (Figs.3.2a and 3.2b), the approximate dimensions of the signal subspace are 1 and 2 respectively. The two criteria of the DISPARE algorithms are applied and the RMS of the estimation error of $\hat{\bar{\phi}}_k$ for the two cases are shown in Figs.4.2a and 4.2b. Also shown that the RMS of the estimation error using the MUSIC method. Results similar to those in Experiment 4.1 are observed here. For small signal spread, all the three methods perform almost identically (Fig.4.2a). For a large signal spread, however, the MUSIC method again yields unacceptably large errors.

The results in Experiment 4.1 and Experiment 4.2 are not surprising. When the spread of a signal is small, it behaves like a point source (with approximate signal dimension equal to one). Thus, the methods developed here will have similar performance as MUSIC which is based on the assumption of a point source. When the spread of a signal is large, i.e., $m_{0k} > 1$, there will be multiple peaks in the MUSIC spectrum leading to erroneous estimation of the mean DOA and MUSIC will not be effective at all.

In the development of the DISPARE algorithms and in the two experiments above, we have assumed that the distribution of the kth dispersed signal is known. This enables us to establish the functional form of the signal covariance matrix $\boldsymbol{R}_k$ so that the optimization criteria of Eqs.(4.15) and (4.18) can be correspondingly set up. Furthermore, we also assumed that the spread of the signal is at least approximately known so that the approximate dimension of the signal subspace can be determined. In practice, neither the distribution nor the spread of the signal is usually known for us. In the following experiment, we will test how the DISPARE algorithms perform when these assumptions are not exact.

Experiment 4.3

In this experiment we study the effect of making a wrong assumption on the distribution of the spread signal. The scenario is identical to that in Experiment 4.1 which a uniformly distributed signal of spread σ_k arrives at a mean DOA of $\bar{\phi}_k = -.5454$ *rad.* to the normal of a uniform linear array. However, we assume that the signal is Gaussian distributed and accordingly we apply the signal covariance matrix $\boldsymbol{R}_k$ as given by Eq.(2.24) and employ the DISPARE criteria of Eqs.(4.15) and (4.18) to estimate the parameters $\bar{\phi}_k$, σ_k. We assume that we know roughly the signal spread σ_k. We tested two cases in which the true σ_k are .1745 *rad.* and .3491 *rad.* respectively. Now, to encompass 99% of the signal energy within this range of spread under the assumption of a Gaussian distribution, we assume that the Gaussian distribution in the two cases to have σ_k equal to .0583 *rad.* and .1163 *rad.* respectively and therefore the respective approximate dimension of the signal subspaces are 1 and 2. The RMS errors of the estimates $\hat{\bar{\phi}}_k$ are plotted in Fig.4.3. Compared to the results shown in Fig.4.1, we observe that while the performance of both DISPARE algorithms have much deteriorated, the failure of the performance of DISPARE I is especially severe. The wider is the signal spread, the larger are the estimation errors.

Experiment 4.4

We continue to examine the effect of making a wrong assumption on the distribution of the spread signal in this experiment. This time, the signal is Gaussian distributed. The scenario is identical to that in Experiment 4.2 in which the signal with spread σ_k arrives at a mean DOA of $-.5454$ *rad.* to the normal of the array. However, we assume that the signal is uniformly distributed and limited to a spread of $\bar{\phi}_k \pm 3\sigma_k$. The two cases we study are $\sigma_k = .0583$ *rad.* and $\sigma_k = .1163$ *rad.*, with corresponding approximate dimension of the signal subspace being 1 and 2. The RMS of the errors in these two cases are plotted in Fig.4.4. Compared to the results in Fig.4.2, it is observed that the performance of

the two algorithms show hardly any deterioration. This is true regardless of the spread of the distribution. Many symmetric distributions other than Gaussian have been tested with assumption that the signal is uniformly distributed. Similar observations of negligible deterioration in performance persist.

The results in Experiment 4.3 and 4.4 are of great significance. It tells us that no matter what shape the signal distribution may be, if it is assumed to be uniformly distributed, the estimation error will be about the same as in the case where the distribution is known. Other assumptions may lead to deterioration of performance if the assumption is incorrect.

We now examine the problem of assuming the approximate dimension of the signal subspace. For a given number of sensors, the number of significant eigenvalues of the correlation matrix due to a dispersed signal depends on the spread of the signal. In many applications such as ionospheric scattering, we can often roughly estimate how much the spread σ_k may be [81]. Furthermore, the number of significant eigenvalues chosen usually is not too sensitive to the estimation error of the spread of the signal. This can be illustrated by Fig.4.5 in which we show the magnitudes of the eigenvalues of R_k for various signal spread values. The signal is uniformly distributed and the number of sensors in the uniform linear array is eight. For a practically wide range of signal spread, the signal energy is dominated by the first two eigenvalues. Thus, even if there exists some error in the rough estimate of the signal spread, the number of dominant eigenvalues in R_x will not be affected.

Equipped with the above observations, the procedure of applying the DISPARE algorithms to estimate the mean DOA of dispersed signals is clear and is summarized as follows:-

1. We assume that the K dispersed signals are all uniformly distributed and are independent of each other.

2. The spread of each signal is roughly estimated so that, for M sensors in the array and

a particular array geometry, the number of dominant eigenvalues m_{0k} can be obtained from charts similar to Fig.4.5.

3. The dimension of the quasi noise subspace can now be determined. An eigen-decomposition of the estimated covariance matrix of the data is performed, and the orthogonal projector onto the estimated quasi noise subspace can be formulated using the noise eigenvector matrix $\hat{\boldsymbol{W}}_\nu$.

4. Either one of the DISPARE algorithms can now be applied to estimate the mean DOA of the dispersed signals. Since all the signals are assumed to be uniformly distributed, the functional forms of the covariance matrix $\boldsymbol{R}_k$ due to each signal will be identical, and is completely pre-determined using Eq.(2.21) if a uniform linear array is used. The criteria for DISPARE I and II can be rewritten as

$$\{\hat{\bar{\phi}}_k, \hat{\sigma}_k\} = \arg\min_{\bar{\phi}_k,\sigma_k} f_I(\bar{\phi},\sigma) \triangleq \mathrm{tr}\{\hat{\boldsymbol{W}}_\nu^\dagger \boldsymbol{R}_u(\bar{\phi},\sigma)\hat{\boldsymbol{W}}_\nu\}] \tag{4.20}$$

and

$$\begin{aligned}\{\hat{\bar{\phi}}_k, \hat{\sigma}_k\} &= \arg\min_{\bar{\phi}_k,\sigma_k} f_{II}(\bar{\phi},\sigma) \\ &\triangleq \arg\min_{\bar{\phi}_k,\sigma_k} \|\boldsymbol{R}_u(\bar{\phi},\sigma)\hat{\boldsymbol{W}}_\nu\|_F^2 \\ &= \arg\min_{\bar{\phi}_k,\sigma_k} \mathrm{tr}\{\hat{\boldsymbol{W}}_\nu^\dagger \boldsymbol{R}_u(\bar{\phi},\sigma)\boldsymbol{R}_u^\dagger(\bar{\phi},\sigma)\hat{\boldsymbol{W}}_\nu\}]\end{aligned} \tag{4.21}$$

where $\boldsymbol{R}_u(\bar{\phi},\sigma)$ is the correlation matrix due to a uniformly distributed signal.

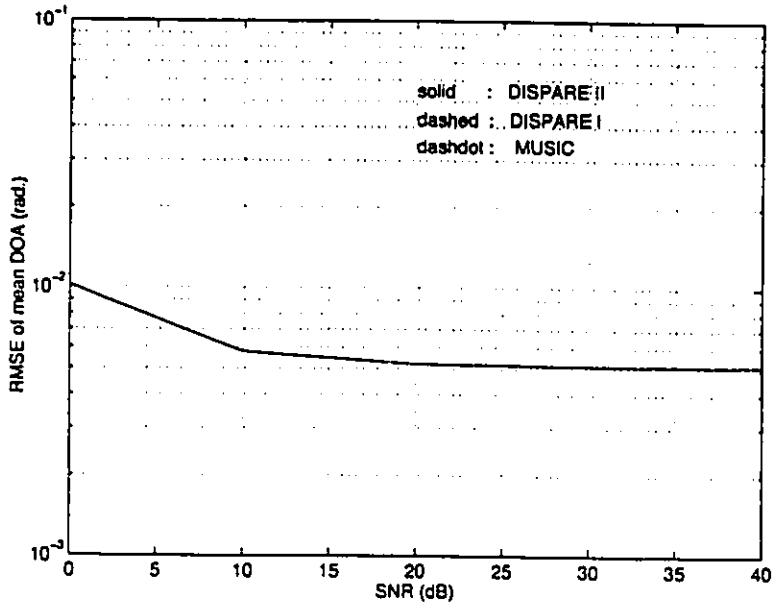
The mean DOA $\bar{\phi}$ can then be estimated by locating the K minima of Eqs.(4.20) and (4.21).

The following experiment will illustrate this procedure.

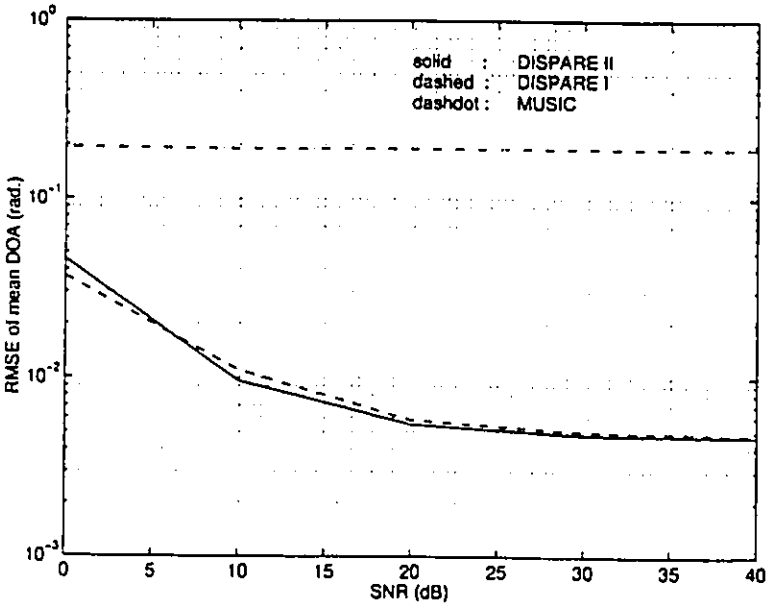
Experiment 4.5

In this experiment, we apply the procedure as outlined above to estimate the mean DOA of two independent narrow-band dispersed signals, one uniformly distributed ($\bar{\phi}_1$ =

.5454 *rad.*, σ_1 = .3491 *rad.*), the other Gaussian distributed ($\bar{\phi}_2$ = 2.4063 *rad.*, σ_2 = .0583 *rad.*). Figs.4.6a and 4.6b show the respective RMS errors of DOA estimation for the two signals. From these two figures, we can see that both algorithms have similar performance for the two signals. Compared to the Figs.4.1b and 4.2a, the estimation accuracy for two signals in this case show very slight deterioration. This is because there are more parameters to be estimated in this experiment than in the other two. As a result of having more parameters to estimate, more uncertainty is introduced [12, 99].

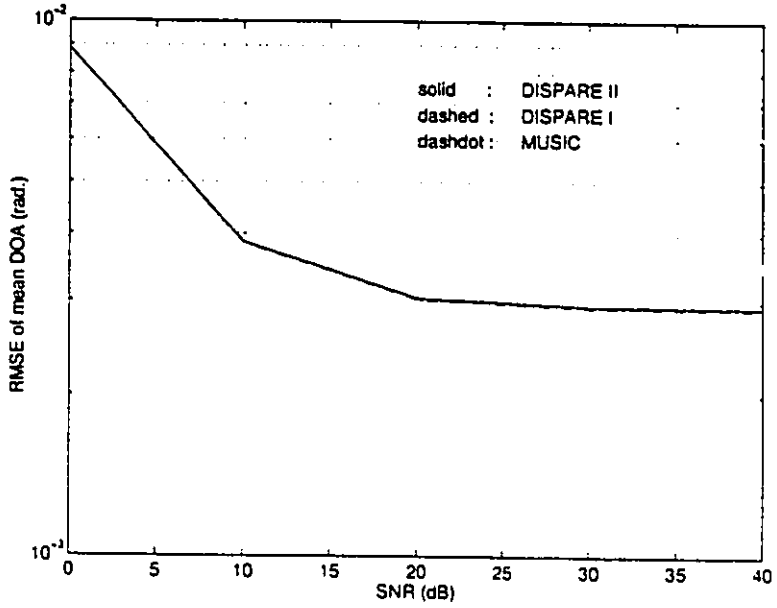


(a) $\sigma_k = .1745\ rad.$

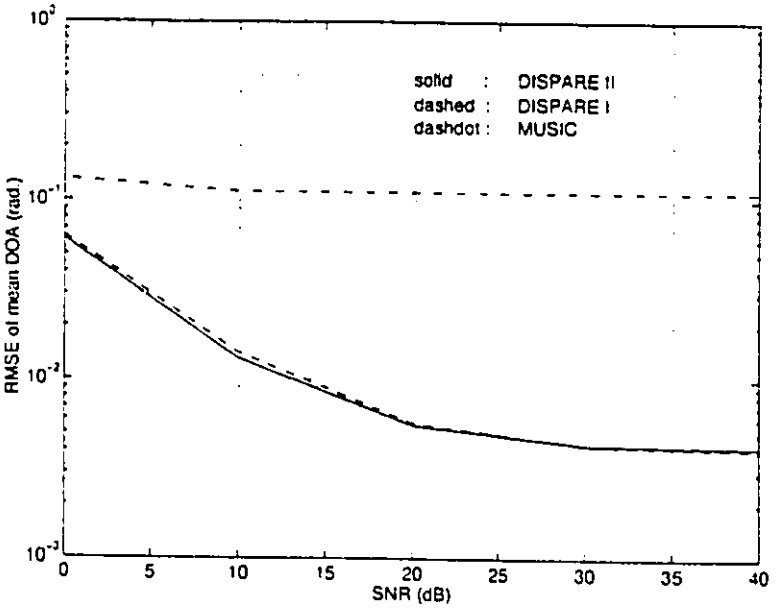


(b) $\sigma_k = .3491\ rad.$

Figure 4.1: The estimation errors of mean DOA of a uniformly distributed signal

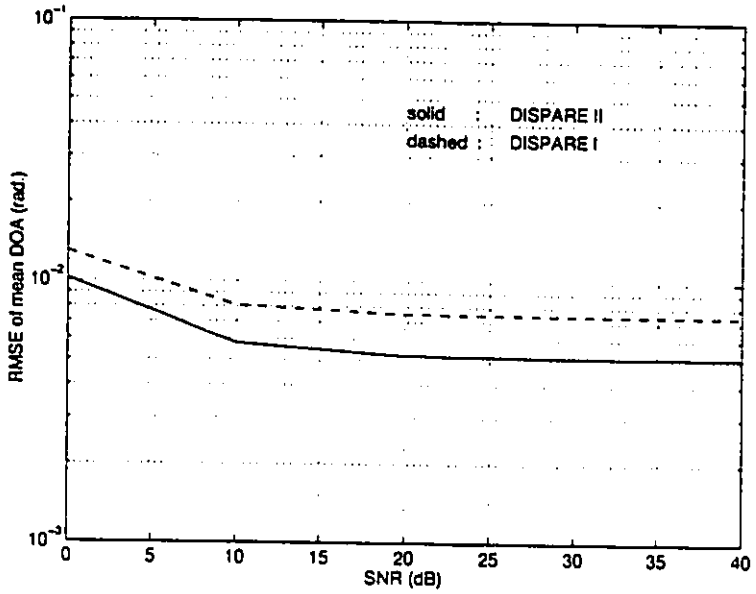


(a) $\sigma_k = .0583\ rad.$

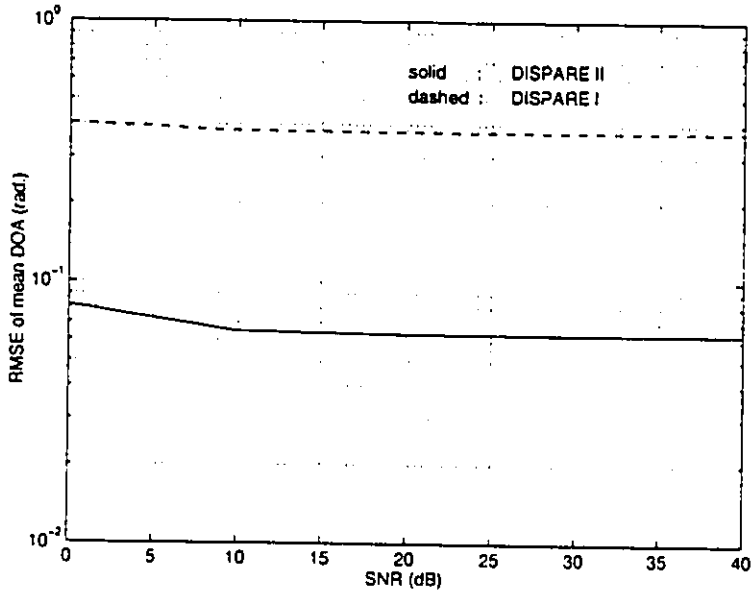


(b) $\sigma_k = .1163\ rad.$

Figure 4.2: The estimation errors of mean DOA of a Gaussian distributed signal

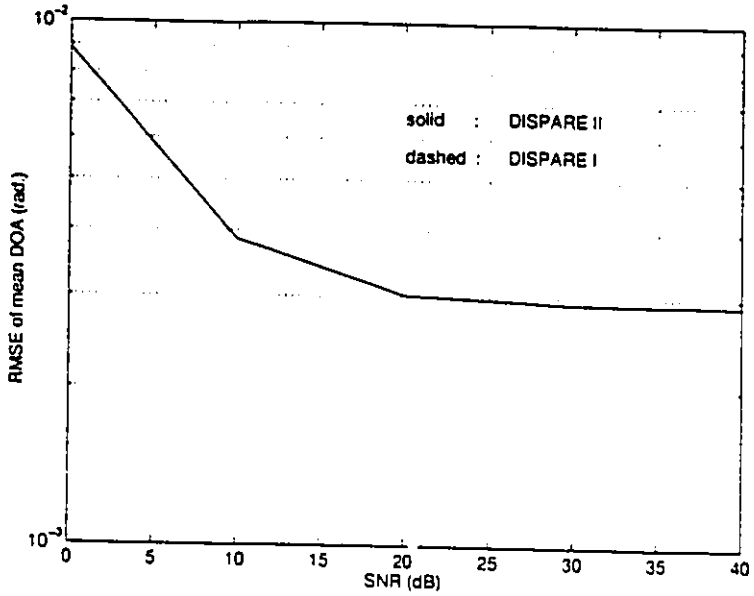


(a) $\sigma_k = .1745\ rad.$

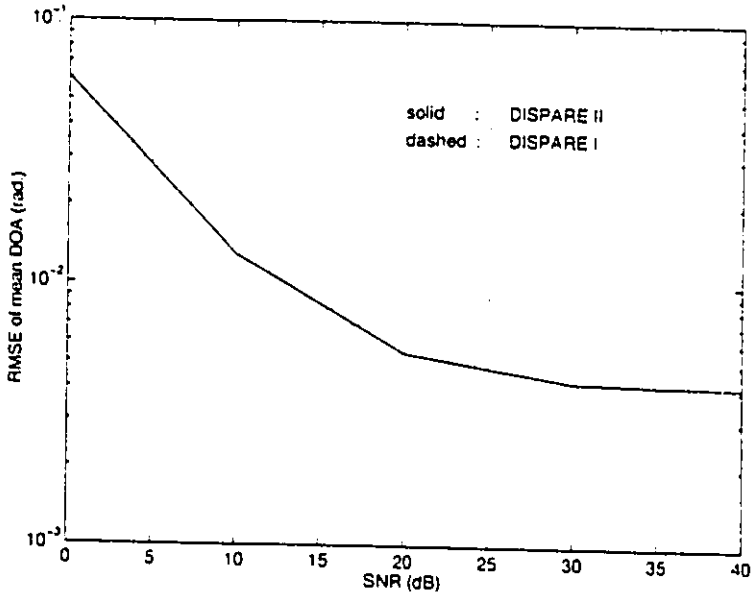


(b) $\sigma_k = .3491\ rad.$

Figure 4.3: The estimation errors of mean DOA of a uniformly distributed signal under the assumption of Gaussian distribution



(a) $\sigma_k = .0583\ rad.$



(b) $\sigma_k = .1163\ rad.$

Figure 4.4: The estimation errors of mean DOA of a Gaussian distributed signal under the assumption of uniform distribution

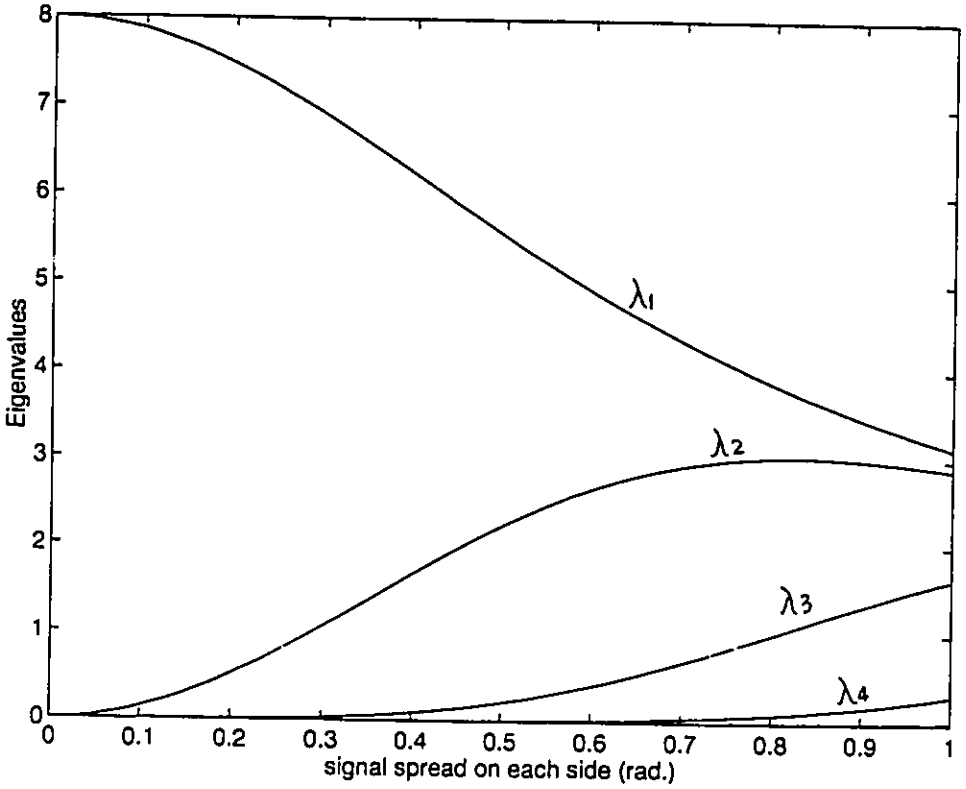
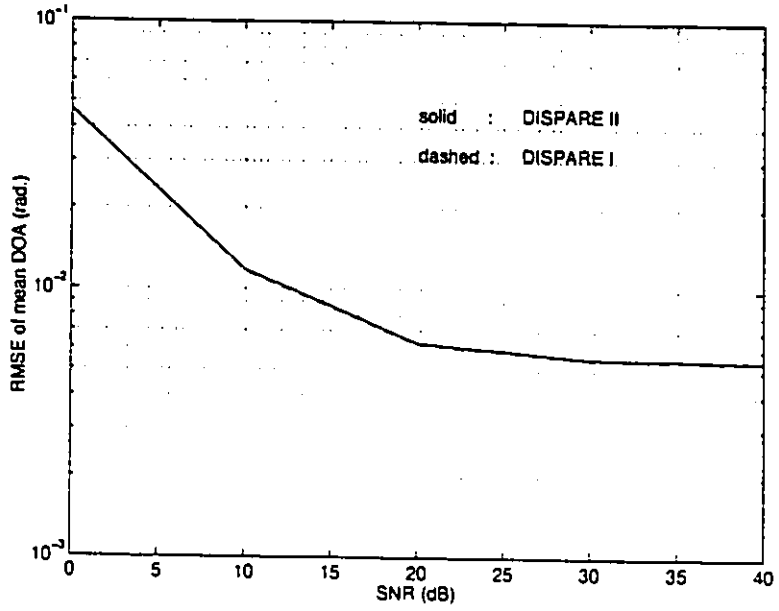
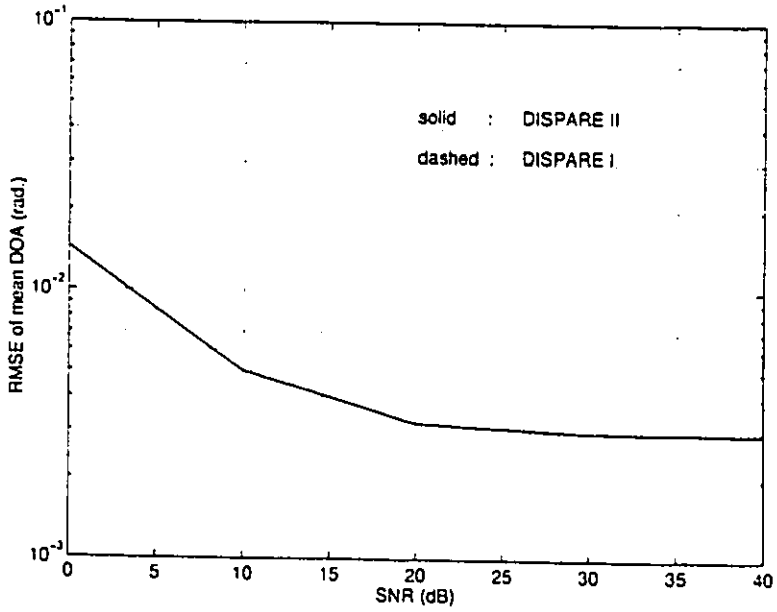


Figure 4.5: The eigenvalues of a uniformly distributed signal given $M = 8$



(a) Uniformly distributed signal with $(\bar{\phi}_1, \sigma_1) = (-.5454, .3491)$



(b) Gaussian distributed signal with $(\bar{\phi}_2, \sigma_2) = (-2.4063, .0583)$

Figure 4.6: The estimation errors of mean DOA assuming both uniformly distributed

Chapter 5

Asymptotic Performance Analysis of DISPARE

In this Chapter, we establish the asymptotic Gaussianity and derive close-form expressions for the mean DOA estimates based on second-order statistics proposed in Chapter 4.

The statistical properties of the perturbation in the estimated covariance matrix $\hat{R}_x$ is analyzed and is utilized to determine the effects of this perturbation on the DOA estimates. A series expansion is performed on the estimated projectors. This technique has the advantage of algebraic simplicity in the perturbation analysis of the projectors.

5.1 Perturbation Theory

This section is devoted to introducing the theorems related to a unifying framework for asymptotic performance analysis.

The data correlation matrix can be rewritten in terms of its eigenvalues and eigenvectors, such that

$$R_x = R_s + p_\nu^2 I$$

$$\simeq \sum_{i=1}^{M_0} \mu_i \boldsymbol{w}_i \boldsymbol{w}_i^\dagger + p_\nu^2 \sum_{i=M_0+1}^{M} \boldsymbol{w}_i \boldsymbol{w}_i^\dagger \tag{5.1a}$$

where M_0 is the dimension of quasi signal subspace. Thus,

$$\boldsymbol{S} = \sum_{i=1}^{M_0} (\mu_i - p_\nu^2) \boldsymbol{w}_i \boldsymbol{w}_i^\dagger \tag{5.1b}$$

Let

$$\boldsymbol{P}_s = \boldsymbol{W}_s \boldsymbol{W}_s^\dagger = \sum_{i=1}^{M_0} \boldsymbol{w}_i \boldsymbol{w}_i^\dagger = \text{projector on the quasi signal subspace} \tag{5.2a}$$

$$\boldsymbol{P}_\nu = \boldsymbol{W}_\nu \boldsymbol{W}_\nu^\dagger = \sum_{i=M_0+1}^{M} \boldsymbol{w}_i \boldsymbol{w}_i^\dagger = \text{projector on the quasi noise subspace} \tag{5.2b}$$

$$\hat{\boldsymbol{P}}_s = \hat{\boldsymbol{W}}_s \hat{\boldsymbol{W}}_s^\dagger = \sum_{i=1}^{M_0} \hat{\boldsymbol{w}}_i \hat{\boldsymbol{w}}_i^\dagger = \text{projector on the estimated quasi signal subspace} \tag{5.2c}$$

$$\hat{\boldsymbol{P}}_\nu = \hat{\boldsymbol{W}}_\nu \hat{\boldsymbol{W}}_\nu^\dagger = \sum_{i=M_0+1}^{M} \hat{\boldsymbol{w}}_i \hat{\boldsymbol{w}}_i^\dagger = \text{projector on the estimated quasi noise subspace} \tag{5.2d}$$

The corresponding errors of estimation are denoted by

$$\Delta \boldsymbol{R}_x = \hat{\boldsymbol{R}}_x - \boldsymbol{R}_x \tag{5.3a}$$

$$\Delta \boldsymbol{P}_s = \hat{\boldsymbol{P}}_s - \boldsymbol{P}_s \tag{5.3b}$$

$$\Delta \boldsymbol{P}_\nu = \hat{\boldsymbol{P}}_\nu - \boldsymbol{P}_\nu \tag{5.3c}$$

We note that

$$\Delta \boldsymbol{P}_s = -\Delta \boldsymbol{P}_\nu \tag{5.3d}$$

Statistical Properties of $\Delta \boldsymbol{R}_x$

Given that the observation consists of N independent complex, zero mean, circular Gaussian random vectors, we have the following theorem which characterizes the perturbation in the estimation of the correlation matrix.

Theorem 5.1 *The elements of $\triangle R_x$ are asymptotically Gaussian distributed with the first- and second-order moments given by*

$$E\{vec(\triangle R_x)\} = \boldsymbol{O} \tag{5.4a}$$

$$E\{vec(\triangle R_x)vec^{\dagger}(\triangle R_x)\} = \frac{1}{N} R_x^T \otimes R_x \triangleq \Sigma_x \tag{5.4b}$$

$$E\{vec(\triangle R_x)vec^{T}(\triangle R_x)\} = \frac{1}{N}(R_x^T \otimes R_x)\Upsilon = \Sigma_x \Upsilon \tag{5.4c}$$

where $\otimes$ denotes the Kronecker product [96] and Υ is a permutation matrix defined as

$$\Upsilon = \sum_{l=1}^{M}\sum_{m=1}^{M} E_{lm} \otimes E_{ml} \tag{5.5}$$

with E_{lm} being an $M \times M$ matrix having unity as its lmth element and zero elsewhere.

Proof: see Appendix B.1. □

Series Expansion of the Projectors

The series expansions of $\hat{P}_s$ and $\hat{P}_\nu$, with respect to the elements of $\triangle R_x$, is given by the following theorem.

Theorem 5.2 *The expansion of $\hat{P}_s$ and $\hat{P}_\nu$ with respect to $\triangle R$ are given by*

$$\hat{P}_s = P_s + \delta P + \cdots + \delta^n P + \cdots \tag{5.6a}$$

$$\hat{P}_\nu = P_\nu - \delta P - \cdots - \delta^n P - \cdots \tag{5.6b}$$

where the terms of the series are related by the following recurrence

$$\begin{aligned}
\delta P &= P_\nu \triangle R_x S^{\#} + S^{\#} \triangle R_x P_\nu \\
&\vdots \\
\delta^n P &= -P_\nu(\delta^{n-1}P)\triangle R_x S^{\#} + P_\nu \triangle R_x(\delta^{n-1}P)S^{\#} \\
&\quad -S^{\#}\triangle R_x(\delta^{n-1}P)P_\nu + S^{\#}(\delta^{n-1}P)\triangle R_x P_\nu \\
&\quad -\sum_{i=1}^{n-1} P_s(\delta^i P)P_\nu(\delta^{n-i}P)P_s + \sum_{i=1}^{n-1} P_\nu(\delta^i P)P_\nu(\delta^{n-i}P)P_\nu \quad (n>1)
\end{aligned} \tag{5.7}$$

and $S^{\#}$ denotes the pseudo-inverse of S.

Proof:

This theorem first appeared in [86]. An alternative proof is given in Appendix B.2. □

Statistical Properties of the Eigen-projection

Given a set of vectors belonging to the subspace spanned by the column vectors of $\boldsymbol{R}_s(\bar{\phi},\sigma)$, the following theorem describes the statistical properties of the projection of these vectors onto the quasi-noise subspace.

Theorem 5.3 *If* $\boldsymbol{Y} \subseteq span(\boldsymbol{R}_s(\bar{\phi},\sigma))$, *then* $vec(\hat{\boldsymbol{P}}_\nu \boldsymbol{Y})$ *is asymptotically Gaussian distributed with mean and second-order moments given by*

$$E\{vec(\hat{\boldsymbol{P}}_\nu \boldsymbol{Y})\} \doteq \boldsymbol{O} \tag{5.8a}$$

$$E\{vec(\hat{\boldsymbol{P}}_\nu \boldsymbol{Y})vec^\dagger(\hat{\boldsymbol{P}}_\nu \boldsymbol{Y})\} = (\boldsymbol{Y}^T \otimes \boldsymbol{I})\boldsymbol{C}_\omega \boldsymbol{\Sigma}_x \boldsymbol{C}_\omega^\dagger (\boldsymbol{Y}^* \otimes \boldsymbol{I}) \tag{5.8b}$$

$$E\{vec(\hat{\boldsymbol{P}}_\nu \boldsymbol{Y})vec^T(\hat{\boldsymbol{P}}_\nu \boldsymbol{Y})\} = (\boldsymbol{Y}^T \otimes \boldsymbol{I})\boldsymbol{C}_\omega \boldsymbol{\Sigma}_x \boldsymbol{\Upsilon} \boldsymbol{C}_\omega^T (\boldsymbol{Y} \otimes \boldsymbol{I}) \tag{5.8c}$$

where $\boldsymbol{C}_\omega$ *is defined as*

$$\boldsymbol{C}_\omega = \boldsymbol{S}^{\#T} \otimes \boldsymbol{P}_\nu + \boldsymbol{P}_\nu^T \otimes \boldsymbol{S}^{\#} \tag{5.9}$$

Proof:

From Theorem 5.1, we note that $\Delta\boldsymbol{R}$ converges to $\boldsymbol{O}$ at the rate of $N^{-\frac{1}{2}}$, therefore Eq.(5.6b) can be rewritten as

$$\hat{\boldsymbol{P}}_\nu = \boldsymbol{P}_\nu - \boldsymbol{P}_\nu(\Delta\boldsymbol{R})\boldsymbol{S}^{\#} - \boldsymbol{S}^{\#}(\Delta\boldsymbol{R})\boldsymbol{P}_\nu + O(N^{-1})$$

Let $\boldsymbol{C}_\omega = \boldsymbol{S}^{\#T} \otimes \boldsymbol{P}_\nu + \boldsymbol{P}_\nu^T \otimes \boldsymbol{S}^{\#}$, then

$$\begin{aligned}
vec(\hat{\boldsymbol{P}}_\nu \boldsymbol{Y}) &= vec(\boldsymbol{P}_\nu \boldsymbol{Y} - \boldsymbol{P}_\nu(\Delta\boldsymbol{R})\boldsymbol{S}^{\#}\boldsymbol{Y} - \boldsymbol{S}^{\#}(\Delta\boldsymbol{R})\boldsymbol{P}_\nu \boldsymbol{Y}) + O(N^{-1}) \\
&= vec(\boldsymbol{P}_\nu \boldsymbol{Y}) - (\boldsymbol{Y}^T \otimes \boldsymbol{I})(\boldsymbol{S}^{\#T} \otimes \boldsymbol{P}_\nu)vec(\Delta\boldsymbol{R}) - (\boldsymbol{Y}^T \otimes \boldsymbol{I})(\boldsymbol{P}_\nu^T \otimes \boldsymbol{S}^{\#})vec(\Delta\boldsymbol{R}) \\
&\quad + O(N^{-1}) \\
&= vec(\boldsymbol{P}_\nu \boldsymbol{Y}) - (\boldsymbol{Y}^T \otimes \boldsymbol{I})\boldsymbol{C}_\omega vec(\Delta\boldsymbol{R}) + O(N^{-1})
\end{aligned}$$

By using Eq.(5.4a) and $\forall \boldsymbol{Y} \subseteq \text{span}(\boldsymbol{R}_s(\bar{\phi}, \sigma))$, we have

$$\mathrm{E}\{\mathrm{vec}(\hat{\boldsymbol{P}}_\nu \boldsymbol{Y})\} = \mathrm{vec}(\hat{\boldsymbol{P}}_\nu \boldsymbol{Y}) - (\boldsymbol{Y}^T \otimes \boldsymbol{I})\boldsymbol{C}_\omega \mathrm{E}\{\mathrm{vec}(\Delta \boldsymbol{R})\} \doteq \boldsymbol{O}$$

Similarly, by using Eqs.(5.4b) and (5.4c), we have

$$\begin{aligned}
\mathrm{cov}\{\mathrm{vec}(\hat{\boldsymbol{P}}_\nu \boldsymbol{Y}), \mathrm{vec}^\dagger(\hat{\boldsymbol{P}}_\nu \boldsymbol{Y})\} &= (\boldsymbol{Y}^T \otimes \boldsymbol{I})\boldsymbol{C}_\omega \mathrm{E}\{\mathrm{vec}(\Delta \boldsymbol{R})\mathrm{vec}^\dagger(\Delta \boldsymbol{R})\}\boldsymbol{C}_\omega^\dagger(\boldsymbol{Y}^* \otimes \boldsymbol{I}) \\
&= (\boldsymbol{Y}^T \otimes \boldsymbol{I})\boldsymbol{C}_\omega \boldsymbol{\Sigma}_x \boldsymbol{C}_\omega^\dagger(\boldsymbol{Y}^* \otimes \boldsymbol{I}) \\
\mathrm{cov}\{\mathrm{vec}(\hat{\boldsymbol{P}}_\nu \boldsymbol{Y}), \mathrm{vec}^T(\hat{\boldsymbol{P}}_\nu \boldsymbol{Y})\} &= (\boldsymbol{Y}^T \otimes \boldsymbol{I})\boldsymbol{C}_\omega \mathrm{E}\{\mathrm{vec}(\Delta \boldsymbol{R})\mathrm{vec}^T(\Delta \boldsymbol{R})\}\boldsymbol{C}_\omega^T(\boldsymbol{Y} \otimes \boldsymbol{I}) \\
&= (\boldsymbol{Y}^T \otimes \boldsymbol{I})\boldsymbol{C}_\omega \boldsymbol{\Sigma}_x \boldsymbol{\Upsilon} \boldsymbol{C}_\omega^T(\boldsymbol{Y} \otimes \boldsymbol{I})
\end{aligned}$$

□

Theorem 5.3 will facilitate the analysis of the estimation algorithms of DISPARE in the remainder of this chapter.

5.2 Performance of DISPARE Algorithms

To evaluate the performance of the DISPARE algorithms, we introduce the notation that $(.)_k$, $\dot{(.)}_k$ and $(.)'_k$ denote respectively a quantity, and its partial derivative with respect to $\bar{\phi}$ and σ evaluated at $\bar{\phi} = \bar{\phi}_k$ and $\sigma = \sigma_k$ where $\bar{\phi}_k$ and σ_k are the true values of the kth signal parameters. Extending this notation to second order derivatives, we have

$$\ddot{A}_k = \left.\frac{\partial^2 A}{\partial \bar{\phi} \partial \bar{\phi}}\right|_{\substack{\bar{\phi}=\bar{\phi}_k \\ \sigma=\sigma_k}} \tag{5.10a}$$

$$A''_k = \left.\frac{\partial^2 A}{\partial \sigma \partial \sigma}\right|_{\substack{\bar{\phi}=\bar{\phi}_k \\ \sigma=\sigma_k}} \tag{5.10b}$$

$$\dot{A}'_k = \left.\frac{\partial^2 A}{\partial \bar{\phi} \partial \sigma}\right|_{\substack{\bar{\phi}=\bar{\phi}_k \\ \sigma=\sigma_k}} \tag{5.10c}$$

Since the performance criteria for for both DISPARE algorithms are quite similar, the procedure of their analysis is very similar as well. We proceed by obtaining a generalized

form of the two results first. Let $f(\vec{\phi},\sigma)$ be a generalized DISPARE criterion which may represent $f_I(\vec{\phi},\sigma)$ or $f_{II}(\vec{\phi},\sigma)$ in Eqs.(4.19) and (4.20) as the case may be. We want to examine the errors of estimation denoted by

$$\Delta\vec{\phi}_k = \hat{\vec{\phi}}_k - \vec{\phi}_k \tag{5.11a}$$

$$\Delta\sigma_k = \hat{\sigma}_k - \sigma_k \tag{5.11b}$$

where $\hat{\vec{\phi}}_k$ and $\hat{\sigma}_k$ are the coordinates of the kth minimum on the $\vec{\phi}\sim\sigma$ plane and $\vec{\phi}_k$ and σ_k are the true parameters of the kth signal. Using the notation of Eqs.(5.10), we can expand the partial derivatives of f at $(\vec{\phi}_k,\sigma_k)$ in a Taylor expansion and equate them to zero, i.e.,

$$\begin{bmatrix} \dot{f}(\hat{\vec{\phi}}_k,\hat{\sigma}_k) \\ f'(\hat{\vec{\phi}}_k,\hat{\sigma}_k) \end{bmatrix} = \begin{bmatrix} \dot{f}_k \\ f'_k \end{bmatrix} + \begin{bmatrix} \ddot{f}_k & \dot{f}'_k \\ \dot{f}'_k & \ddot{f}_k \end{bmatrix} \begin{bmatrix} \Delta\vec{\phi}_k \\ \Delta\sigma_k \end{bmatrix} + \cdots = \boldsymbol{O} \tag{5.12}$$

Assuming $\Delta\vec{\phi}_k$ and $\Delta\sigma_k$ are small, we can ignore the terms involving the second and higher orders of $\Delta\vec{\phi}_k$ and $\Delta\sigma_k$ in Eq.(5.12) so that we can write

$$\mathrm{E}\begin{bmatrix} \Delta\vec{\phi}_k \\ \Delta\sigma_k \end{bmatrix} = -\boldsymbol{H}_k \mathrm{E}\begin{bmatrix} \dot{f}_k \\ f'_k \end{bmatrix} \tag{5.13}$$

where $\boldsymbol{H}_k = \begin{bmatrix} \ddot{f}_k & \dot{f}'_k \\ \dot{f}'_k & \ddot{f}_k \end{bmatrix}^{-1}$ and therefore the covariance matrix of $\Delta\vec{\phi}_k$ and $\Delta\sigma_k$ can be expressed as

$$\mathrm{cov}\{\begin{bmatrix} \Delta\vec{\phi}_k \\ \Delta\sigma_k \end{bmatrix}\} = \boldsymbol{H}_k \begin{bmatrix} \mathrm{cov}\{\dot{f}_k,\dot{f}_k\} & \mathrm{cov}\{\dot{f}_k,f'_k\} \\ \mathrm{cov}\{f'_k,\dot{f}_k\} & \mathrm{cov}\{f'_k,f'_k\} \end{bmatrix} \boldsymbol{H}_k^{\dagger} \tag{5.14}$$

Using the general formulas of Eqs.(5.13) and (5.14), we can now evaluate the performance of DISPARE I and DISPARE II algorithms.

5.2.1 Performance of DISPARE I

For DISPARE I, from Eq.(4.19), since $\mathrm{tr}\{\boldsymbol{AB}\} = \mathrm{tr}\{\boldsymbol{BA}\}$ and $\dot{\boldsymbol{P}}_\nu = \dot{\boldsymbol{W}}_\nu \dot{\boldsymbol{W}}_\nu^{\dagger}$

$$\dot{f}_{Ik} = \mathrm{tr}\{\dot{\boldsymbol{P}}_\nu \dot{\boldsymbol{R}}_{uk}^{\dagger}\} = 2\mathrm{Re}[\mathrm{tr}\{\dot{\boldsymbol{P}}_\nu \boldsymbol{G}_{uk} \dot{\boldsymbol{G}}_{uk}^{\dagger}\}] \tag{5.15a}$$

$$f'_{Ik} = \mathrm{tr}\{\hat{P}_\nu R'^{\dagger}_{uk}\} = 2\mathrm{Re}[\mathrm{tr}\{\hat{P}_\nu G_{uk} G'^{\dagger}_{uk}\}] \tag{5.15b}$$

where $R_{uk} = G_{uk}G_{uk}^{\dagger}$ with G_{uk} being the Cholesky factor of R_{uk}. Since the columns of G_{uk} are in the same subspace spanned by the columns of R_{uk}, therefore we can use Eq.(5.8a) of Theorem 5.2 and Eq.(5.13) to arrive at

$$\mathrm{E}\begin{bmatrix} \Delta\bar{\phi}_k \\ \Delta\sigma_k \end{bmatrix} = -2H_{Ik}\begin{bmatrix} \mathrm{Retr}\{\mathrm{E}[\hat{P}_\nu G_{uk}]\dot{G}^{\dagger}_{uk}\} \\ \mathrm{Retr}\{\mathrm{E}[\hat{P}_\nu G_{uk}]G'^{\dagger}_{uk}\} \end{bmatrix} \simeq O \tag{5.16}$$

where H_{Ik} consists of elements similar to those of H_k with the obvious substitution of f by f_I. The second derivatives of $f_I(\bar{\phi}, \sigma)$ can also be written as

$$\begin{aligned} \ddot{f}_{Ik} &= \mathrm{tr}\{\hat{P}_\nu \ddot{R}_{uk}\} \\ f''_{Ik} &= \mathrm{tr}\{\hat{P}_\nu R''_{uk}\} \\ \dot{f}'_{Ik} &= \mathrm{tr}\{\hat{P}_\nu \dot{R}'_{uk}\} \end{aligned} \tag{5.17}$$

which constitute the elements of the matrix H_{Ik}. Thus, the only quantities to be evaluated are the elements of the central matrix on RHS of Eq.(5.14). Now, from Eq.(5.15a),

$$\begin{aligned} \mathrm{cov}\{\dot{f}_{Ik}, \dot{f}_{Ik}\} &= 4\mathrm{cov}\{\mathrm{Re}[\mathrm{tr}\{\hat{P}_\nu G_{uk}\dot{G}^{\dagger}_{uk}\}], \mathrm{Re}[\mathrm{tr}\{\hat{P}_\nu G_{uk}\dot{G}^{\dagger}_{uk}\}]\} \\ &= 2\mathrm{Re}[\mathrm{cov}\{\mathrm{tr}\{\hat{P}_\nu G_{uk}\dot{G}^{\dagger}_{uk}\}, \mathrm{tr}\{\hat{P}_\nu G_{uk}\dot{G}^{\dagger}_{uk}\}\} \\ &\quad +\mathrm{cov}\{\mathrm{tr}\{\hat{P}_\nu G_{uk}\dot{G}^{\dagger}_{uk}\}, (\mathrm{tr}\{\hat{P}_\nu G_{uk}\dot{G}^{\dagger}_{uk}\})^*\}] \end{aligned} \tag{5.18}$$

where in the last step, we have used the identity

$$\mathrm{E}\{\mathrm{Re}[u - \mathrm{E}(u)]\mathrm{Re}[u - \mathrm{E}(u)]\} = \frac{1}{2}\mathrm{E}\{\mathrm{Re}[u - \mathrm{E}(u)]^2 + \mathrm{Re}[u - \mathrm{E}(u)][u - \mathrm{E}(u)]^*\} \tag{5.19}$$

Now, the first term on RHS of Eq.(5.18) can be written as (using the fact that $\mathrm{tr}\{AB\} = \mathrm{vec}^{\dagger}(B^{\dagger})\mathrm{vec}(A)$),

$$\begin{aligned} &\mathrm{cov}\{\mathrm{tr}\{\hat{P}_\nu G_{uk}\dot{G}^{\dagger}_{uk}\}, \mathrm{tr}\{\hat{P}_\nu G_{uk}\dot{G}^{\dagger}_{uk}\}\} \\ = \ &\mathrm{cov}\{\mathrm{vec}^{\dagger}(\dot{G}^{\dagger}_{uk})\mathrm{vec}(\hat{P}_\nu G_{uk}), \mathrm{vec}^{\dagger}(\dot{G}^{\dagger}_{uk})\mathrm{vec}(\hat{P}_\nu G_{uk})\} \\ = \ &\mathrm{vec}^{\dagger}(\dot{G}^{\dagger}_{uk})\mathrm{cov}\{\mathrm{vec}(\hat{P}_\nu G_{uk}), \mathrm{vec}(\hat{P}_\nu G_{uk})\}\mathrm{vec}(\dot{G}^{\dagger}_{uk}) \end{aligned}$$

$$= \quad \mathrm{vec}^\dagger(\dot{G}_{uk}G_{uk}^\dagger)C_\omega \Sigma_x C_\omega^\dagger \mathrm{vec}(\dot{G}_{uk}G_{uk}^\dagger) \tag{5.20a}$$

where, in the last step, Eq.(5.8b) has been used. Similarly, using Eq.(5.8c), we can write

$$\mathrm{cov}\{\mathrm{tr}\{\dot{P}_\nu G_{uk}\dot{G}_{uk}^\dagger\}, (\mathrm{tr}\{\dot{P}_\nu G_{uk}\}\dot{G}_{uk}^\dagger\})^*\} = \mathrm{vec}^\dagger(\dot{G}_{uk}G_{uk}^\dagger)C_\omega \Sigma_x \Upsilon C_\omega^T \mathrm{vec}^*(\dot{G}_{uk}G_{uk}^\dagger) \tag{5.20b}$$

Hence, Eq.(5.19) can be rewritten as

$$\begin{aligned}\mathrm{cov}\{\dot{f}_{Ik}, \dot{f}_{Ik}\} = 2\mathrm{Re}[&\mathrm{vec}^\dagger(\dot{G}_{uk}G_{uk}^\dagger)C_\omega \Sigma_x C_\omega^\dagger \mathrm{vec}(\dot{G}_{uk}G_{uk}^\dagger) \\ &+ \mathrm{vec}^\dagger(\dot{G}_{uk}G_{uk}^\dagger)C_\omega \Sigma_x \Upsilon C_\omega^T \mathrm{vec}^*(\dot{G}_{uk}G_{uk}^\dagger)]\end{aligned} \tag{5.21a}$$

Similarly,

$$\begin{aligned}\mathrm{cov}\{\dot{f}_{Ik}, f'_{Ik}\} = 2\mathrm{Re}[&\mathrm{vec}^\dagger(\dot{G}_{uk}G_{uk}^\dagger)C_\omega \Sigma_x C_\omega^\dagger \mathrm{vec}(G'_{uk}G_{uk}^\dagger) \\ &+ \mathrm{vec}^\dagger(\dot{G}_{uk}G_{uk}^\dagger)C_\omega \Sigma_x \Upsilon C_\omega^T \mathrm{vec}^*(G'_{uk}G_{uk}^\dagger)]\end{aligned} \tag{5.21b}$$

$$\begin{aligned}\mathrm{cov}\{f'_{Ik}, f'_{Ik}\} = 2\mathrm{Re}[&\mathrm{vec}^\dagger(G'_{uk}G_{uk}^\dagger)C_\omega \Sigma_x C_\omega^\dagger \mathrm{vec}(G'_{uk}G_{uk}^\dagger) \\ &+ \mathrm{vec}^\dagger(G'_{uk}G_{uk}^\dagger)C_\omega \Sigma_x \Upsilon C_\omega^T \mathrm{vec}^*(G'_{uk}G_{uk}^\dagger)]\end{aligned} \tag{5.21c}$$

Substituting Eqs(5.17), (5.21) into Eq.(5.14), the covariance matrix of $[\Delta\bar{\phi}, \Delta\sigma]^T$ for DISPARE I estimates can be evaluated.

5.2.2 Performance of DISPARE II

For DISPARE II, the performance analysis is similar to that of DISPARE I. We have

$$\dot{f}_{IIk} = 2\mathrm{Re}[\mathrm{tr}\{\dot{P}_\nu R_{uk}\dot{R}_{uk}^\dagger\}] \tag{5.22a}$$

$$f'_{IIk} = 2\mathrm{Re}[\mathrm{tr}\{\dot{P}_\nu R_{uk}R'^\dagger_{uk}\}] \tag{5.22b}$$

and similar to Eq.(5.16), we have

$$\mathrm{E}\begin{bmatrix}\Delta\bar{\phi}_k \\ \Delta\sigma_k\end{bmatrix} = -2H_{IIk}\begin{bmatrix}\mathrm{Retr}\{\mathrm{E}[\dot{P}_\nu R_{uk}]\dot{R}_{uk}^\dagger\} \\ \mathrm{Retr}\{\mathrm{E}[\dot{P}_\nu R_{uk}]R'^\dagger_{uk}\}\end{bmatrix} \simeq 0 \tag{5.23}$$

The second derivatives of $f_{IIk}(\bar{\phi},\sigma)$ are given by

$$\ddot{f}_{IIk} = 2\mathrm{Re}[\mathrm{tr}\{\hat{P}_\nu(R_{uk}\ddot{R}_{uk}^\dagger + \dot{R}_{uk}\dot{R}_{uk}^\dagger)\}] \tag{5.24a}$$

$$\dot{f}'_{IIk} = 2\mathrm{Re}[\mathrm{tr}\{\hat{P}_\nu(R_{uk}\dot{R}'^\dagger_{uk} + \dot{R}_{uk}R'^\dagger_{uk})\}] \tag{5.24b}$$

$$f''_{IIk} = 2\mathrm{Re}[\mathrm{tr}\{\hat{P}_\nu(R_{uk}R''^\dagger_{uk} + R'_{uk}R'^\dagger_{uk})\}] \tag{5.24c}$$

Following exactly the same procedure as in DISPARE I above, we arrive at

$$\begin{aligned}\mathrm{cov}\{\dot{f}_{IIk},\dot{f}_{IIk}\} = 2\mathrm{Re}[&\mathrm{vec}^\dagger(\dot{R}_{uk}R_{uk}^\dagger)C_\omega\Sigma_x C_\omega^\dagger\mathrm{vec}(\dot{R}_{uk}R_{uk}^\dagger)\\ &+ \mathrm{vec}^\dagger(\dot{R}_{uk}R_{uk}^\dagger)C_\omega\Sigma_x\Upsilon C_\omega^T\mathrm{vec}^*(\dot{R}_{uk}R_{uk}^\dagger)]\end{aligned} \tag{5.25a}$$

$$\begin{aligned}\mathrm{cov}\{\dot{f}_{IIk},f'_{IIk}\} = 2\mathrm{Re}[&\mathrm{vec}^\dagger(\dot{R}_{uk}R_{uk}^\dagger)C_\omega\Sigma_x C_\omega^\dagger\mathrm{vec}(R'_{uk}R_{uk}^\dagger)\\ &+ \mathrm{vec}^\dagger(\dot{R}_{uk}R_{uk}^\dagger)C_\omega\Sigma_x\Upsilon C_\omega^T\mathrm{vec}^*(R'_{uk}R_{uk}^\dagger)]\end{aligned} \tag{5.25b}$$

$$\begin{aligned}\mathrm{cov}\{f'_{IIk},f'_{IIk}\} = 2\mathrm{Re}[&\mathrm{vec}^\dagger(R'_{uk}R_{uk}^\dagger)C_\omega\Sigma_x C_\omega^\dagger\mathrm{vec}(R'_{uk}R_{uk}^\dagger)\\ &+ \mathrm{vec}^\dagger(R'_{uk}R_{uk}^\dagger)C_\omega\Sigma_x\Upsilon C_\omega^T\mathrm{vec}^*(R'_{uk}R_{uk}^\dagger)]\end{aligned} \tag{5.25c}$$

Substituting Eqs(5.24) and (5.25) into Eq.(5.14), we can evaluate the covariance matrix of $[\Delta\bar{\phi},\Delta\sigma]^T$ for DISPARE II.

Fig.5.1 shows the comparison between the theoretical evaluation of the performance and the simulation results of the two DISPARE methods. The comparison is based on the same scenario as that in Experiment 4.1 in Chapter 4.

The above performance analysis on the two DISPARE algorithms are based on analysis in the neighbourhood of the local minima. As such it may not be able to provide a global view of the optimization problem. Figs.5.2a and 5.2b show a typical "profiles" of the respective criteria of DISPARE I and II in which, we plot $-\log f(\bar{\phi},\sigma)$ of the respective algorithms for various values of $\bar{\phi}$ and σ. The profiles are plotted for a dispersed signal of uniform distribution with $\bar{\phi}_k = -0.5454$ and $\sigma_k = 0.3491$. The SNR for both cases is

$30dB$. For DISPARE I as shown in Fig.5.2a, around the point where true $\bar{\phi}_k$ and σ_k lie, there is a local maximum of the profile. However, this local maximum is overshadowed by the existence of two dominant peaks at some other points. Thus, if DISPARE I is applied to estimate $\bar{\phi}_k$ and σ_k without prior knowledge of the approximate values of $\bar{\phi}_k$ and σ_k, an erroneous peak may be taken as the estimate causing unacceptable performance errors. On the other hand, for DISPARE II, as shown in Fig.5.2b, around the point of $(\bar{\phi}_k, \sigma_k)$, there is a global maximum, which ensures that if a search procedure is enforced in optimization, the estimates will be close to the true values of the parameters. Many other profiles under various SNR and signal distribution have been plotted. Similar observations persist. Thus, it can be concluded that DISPARE II is generally more robust.

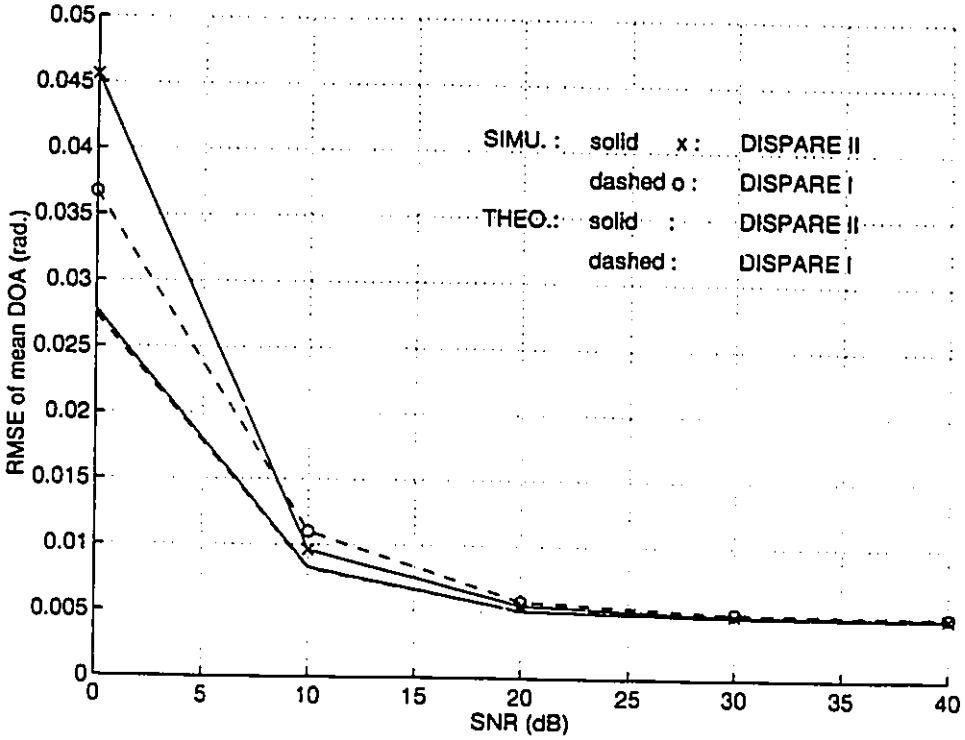
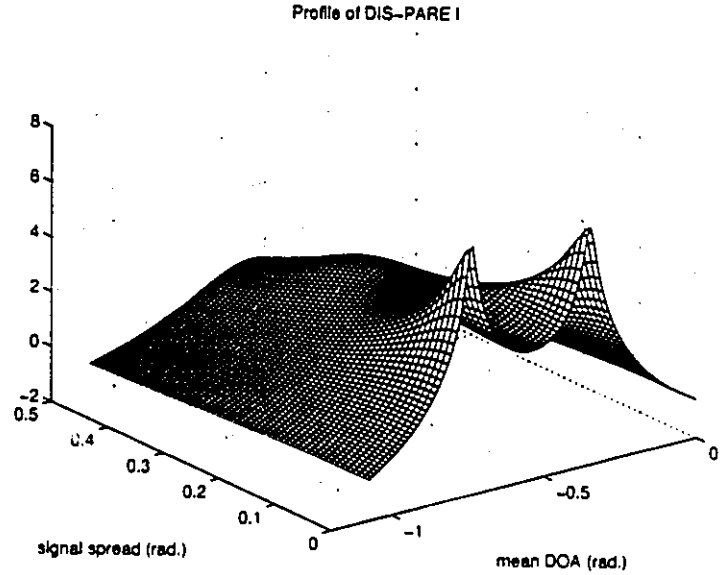
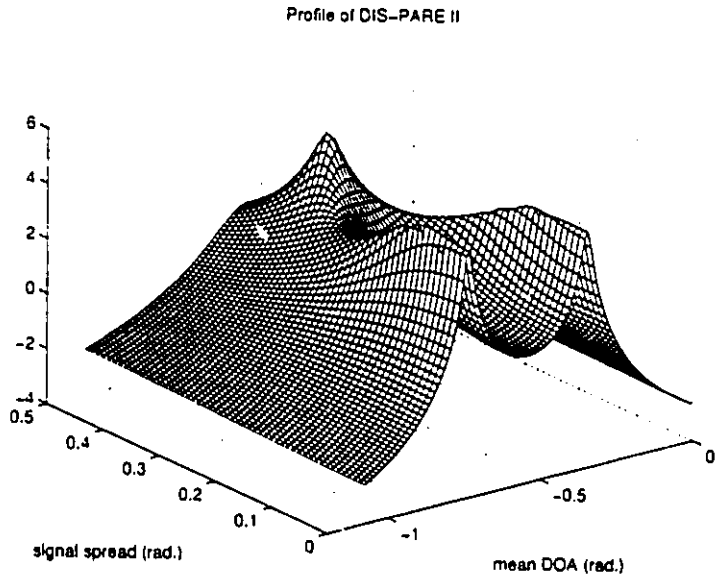


Figure 5.1: Theoretical evaluation and simulation of RMSE of mean DOA of a uniformly distributed signal with $(\bar{\phi}, \sigma) = (-.5454, .3491)$ arriving at a uniform linear array of 8 sensors; 200 snapshots are used for comparison.



(a) DISPARE I (30dB)



(b) DISPARE II (30dB)

Figure 5.2: DISPARE spectra

Chapter 6

DOA Estimation Based on Higher-order Statistics

Currently, high resolution array processing techniques are based on second-order statistics of the received signals (covariance, periodograms, etc.). The techniques of DISPARE proposed in Chapter 4 are of this type. Methods based on second-order statistics assume, either explicitly or implicitly, that the signals of interest are Gaussian. However, there are fairly common applications where signals are non-Gaussian, particularly in digital communications.

Non-Gaussian processes contain valuable statistical information in their higher-order statistics (HOS) (orders greater than two). This information is not used by "conventional" array processing algorithms, and therefore optimal performance may not be achieved. In fact, it has been shown that estimation methods which exploit the non-Gaussianity of the signals have some inherent advantages over methods based on second-order statistics [87, 88]. These include:

(i) the ability to reconstruct the phase of non-minimum phase systems [87].

(ii) the ability to separate non-Gaussian signals from Gaussian noise (sensor noise or interference) [88].

(iii) the ability to separate statistically independent sources, provided they have different statistical properties [87].

(iv) provide improved estimation of signal parameters (such as direction of arrival) [88].

Owing to these potential advantages, it is of interest to develop array processing techniques which are designed to make use of this higher-order information. In this chapter, we consider the application of high resolution array processing in ionospheric scattering in which the signals involved are of quadrature amplitude modulation (QAM) and are therefore non-Gaussian. In these cases the use of higher-order statistics appears to be of advantage.

6.1 Higher-order Statistics

6.1.1 Definitions and Properties

In this section, we introduce the definitions, properties and computation of higher-order statistics, i.e., moments and cumulants, and their corresponding spectra.

If $\{x(n)\}$, $n = 0, \pm 1, \pm 2, \cdots$ is a real stationary discrete-time signal and its moments up to order i exist, then its ith-order moment is defined as [89]

$$m_i^x(\tau_1, \tau_2, \cdots, \tau_{i-1}) = \mathrm{E}\{x(n)x(n+\tau_1)\cdots x(n+\tau_{i-1})\} \tag{6.1}$$

The ith-order cumulant of a non-Gaussian stationary random signal sequence $\{x(n)\}$ is defined as (for $i = 3, 4$ only)

$$c_i^x(\tau_1, \tau_2, \cdots, \tau_{i-1}) = m_i^x(\tau_1, \tau_2, \cdots, \tau_{i-1}) - m_i^G(\tau_1, \tau_2, \cdots, \tau_{i-1}) \tag{6.2}$$

where $m_i^G(\tau_1, \tau_2, \cdots, \tau_{i-1})$ is the ith-order moment of a Gaussian signal having the same

mean value and autocorrelation function as $\{x(n)\}$. Clearly, if $\{x(n)\}$ is Gaussian, then $c_i^x(\tau_1, \tau_2, \cdots, \tau_{i-1}) = 0$.

Higher-order spectra are defined in terms of their cumulants (e.g., cumulant spectra) or moments (e.g., moment spectra) [89]. Simply stated, higher-order spectra are multi-dimensional Fourier transforms of higher-order statistics. Thus, the ith-order moment spectrum is defined as

$$\begin{aligned} & P_{i-1}^m(\omega_1, \omega_2, \cdots, \omega_{i-1}) \\ \triangleq & \sum_{\tau_1=-\infty}^{\infty} \sum_{\tau_2=-\infty}^{\infty} \cdots \sum_{\tau_{i-1}=-\infty}^{\infty} m_{i-1}^x(\tau_1, \tau_2, \cdots, \tau_{i-1}) e^{-j(\omega_1\tau_1+\omega_2\tau_2+\cdots+\omega_{i-1}\tau_{i-1})}, \\ & |\omega_1| \le \pi, \ |\omega_2| \le \pi, \ \cdots, \ |\omega_{i-1}| \le \pi, \ |\omega_1+\omega_2+\cdots+\omega_{i-1}| \le \pi \end{aligned} \tag{6.3}$$

where $m_{i-1}^x(\tau_1, \tau_2, \cdots, \tau_{i-1})$ is the ith-order moment sequence of $\{x(n)\}$. And the ith-order cumulant spectrum is defined as

$$\begin{aligned} & P_{i-1}^c(\omega_1, \omega_2, \cdots, \omega_{i-1}) \\ \triangleq & \sum_{\tau_1=-\infty}^{\infty} \sum_{\tau_2=-\infty}^{\infty} \cdots \sum_{\tau_{i-1}=-\infty}^{\infty} c_{i-1}^x(\tau_1, \tau_2, \cdots, \tau_{i-1}) e^{-j(\omega_1\tau_1+\omega_2\tau_2+\cdots+\omega_{i-1}\tau_{i-1})}, \\ & |\omega_1| \le \pi, \ |\omega_2| \le \pi, \ \cdots, \ |\omega_{i-1}| \le \pi, \ |\omega_1+\omega_2+\cdots+\omega_{i-1}| \le \pi \end{aligned} \tag{6.4}$$

where $c_{i-1}^x(\tau_1, \tau_2, \cdots, \tau_{i-1})$ is the ith-order cumulant sequence of $\{x(n)\}$.

Cumulants and cumulant spectra are more useful in the processing of random signals than moments and moment spectra [89], because:

(i) cumulants of order $i > 2$ are zero if the signal is Gaussian and thus non-zero cumulants provide a measure of the extent of non-Gaussianity;

(ii) the higher-order cumulant of white noise is a multi-dimensional impulse function and its polyspectrum is multi-dimensionally flat;

(iii) the cumulant of two statistically independent random process equals the sum of the cumulants of the individual random process, whereas the same is not true for higher-order moments [89].

This last property allows us to work with cumulant more easily as an operator.

A QAM signal can be represented (assuming the sampling is no faster than the symbol rate) as an i.i.d. sequence of complex numbers, belonging to a discrete set φ. The probabilities of all symbol values are assumed to be equal. A typical symbol set is $\varphi = \{a + jb; a, b = \pm 1, \pm 3, \cdots\}$. Due to the inherent symmetry, the odd-order statistical moments of QAM signals are zero [90]. Therefore, we will be concerned with even-order moments only, in particular, with the fourth-order moments. Assuming the sensor noise is Gaussian and statistically independent of the signals, we use the fourth-order cumulants as the statistics to perform the DOA estimation [91].

6.2 Cumulant Statistics for DOA Estimation in Correlated Gaussian Noise

From Chapter 2.2.2 we know that, in the presence of distributed sources, the array output at the nth snapshot is given as

$$\boldsymbol{x}(n) = \sum_{k=1}^{K} \sum_{l=1}^{L_k} \boldsymbol{d}(\phi_{kl}) \sqrt{\gamma_{kl}} e_{kl}(n) + \boldsymbol{\nu}(n) \tag{6.5}$$

where again $\boldsymbol{x}(n)$ is an M-dimensional vector of the outputs from all M sensors; K is the number of independent distributed sources; L_k is the number of discrete components in the

kth distributed source; ϕ_{kl} is the DOA of the lth discrete component in the kth distributed source and $\boldsymbol{d}(\phi_{kl})$ is the direction manifold of the array, $\sqrt{\gamma_{kl}}$ is the intensity (standard deviation) of this discrete source; $e_{kl}(n)$, $l = 1, \cdots, L_k$; $k = 1, \cdots, K$ are random variables with unity variance; $\boldsymbol{\nu}(n)$ is an $M \times 1$ complex noise vector assumed to be zero-mean, Gaussian, ergodic and independent of signals. Here, we do not confine our consideration to spatially white noise only and allow $\boldsymbol{\nu}(n)$ to have an unknown covariance matrix.

6.2.1 The Fourth-order Cumulant Matrix

Given a vector process there exists a variety of definition of higher-order cumulants of the vector process. In this chapter, we define the fourth-order cumulant matrix of the data received by the array as

$$\boldsymbol{Q} = \mathrm{cum}_4\{\boldsymbol{x}\} \triangleq \mathrm{E}\{(\boldsymbol{x}^* \otimes \boldsymbol{x})(\boldsymbol{x}^* \otimes \boldsymbol{x})^\dagger\} - \mathrm{E}\{\boldsymbol{x}^* \otimes \boldsymbol{x}\}(\mathrm{E}\{\boldsymbol{x}^* \otimes \boldsymbol{x}\})^\dagger - (\mathrm{E}\{\boldsymbol{x}\boldsymbol{x}^\dagger\})^* \otimes \mathrm{E}\{\boldsymbol{x}\boldsymbol{x}^\dagger\} \quad (6.6)$$

Here the obvious dependence of $\boldsymbol{x}$ on n is omitted. Denoting the ith element of $\boldsymbol{x}$ by x_i, the elements of matrix $\boldsymbol{Q}$ are given by

$$\begin{aligned} & \boldsymbol{Q}[(i-1)M + j;\ (u-1)M + v] \\ = \ & \mathrm{cum}_4\{x_i^*, x_j, x_u, x_v^*\} \\ = \ & \mathrm{E}\{x_i^* x_j x_u x_v^*\} - \mathrm{E}\{x_i^* x_j\}\mathrm{E}\{x_u x_v^*\} - \mathrm{E}\{x_i^* x_u\}\mathrm{E}\{x_j x_v^*\} \\ \triangleq \ & \boldsymbol{Q}[i, j;\ u, v], \quad 1 \le i, j, u, v \le M \end{aligned} \quad (6.7)$$

Here, we use the notation $\boldsymbol{Q}(i, j;\ u, v)$ to denote an element of matrix whose row and column counts depend on (i, j) and (u, v) respectively. The advantages of the definition in Eq.(6.6) are twofold: first, we can use matrix to represent cumulants, resulting in much simpler notations than a four-dimensional array; second, we are able to use linear algebra of vectors and matrices for analysis and computation.

As we have noted in the preceding section, the cumulant of sum of independent processes is the sum of the cumulants of the individual processes and the fourth-order cumulants of Gaussian random process are zero. Thus, the cumulant matrix of the kth distributed signal composed of L_k discrete sources each arriving from a direction ϕ_{kl}, assuming $e_{kl}, l = 1, 2, \cdots, L_k; k = 1, \cdots, K$ are i.i.d. random variables with fourth-order cumulants denoted by κ, is given by

$$\begin{aligned}
Q_k &= \sum_{l=1}^{L_k} \mathrm{cum}_4\{\boldsymbol{d}(\phi_{kl})\sqrt{\gamma_{kl}}e_{kl}\} \\
&= \sum_{l=1}^{L_k} \mathrm{E}\{[(\boldsymbol{d}(\phi_{kl})\sqrt{\gamma_{kl}}e_{kl})^* \otimes (\boldsymbol{d}(\phi_{kl})\sqrt{\gamma_{kl}}e_{kl})][(\boldsymbol{d}(\phi_{kl})\sqrt{\gamma_{kl}}e_{kl})^* \otimes (\boldsymbol{d}(\phi_{kl})\sqrt{\gamma_{kl}}e_{kl})]^\dagger\} \\
&\quad -\mathrm{E}\{(\boldsymbol{d}(\phi_{kl})\sqrt{\gamma_{kl}}e_{kl})^* \otimes (\boldsymbol{d}(\phi_{kl})\sqrt{\gamma_{kl}}e_{kl})\}(\mathrm{E}\{(\boldsymbol{d}(\phi_{kl})\sqrt{\gamma_{kl}}e_{kl})^* \otimes (\boldsymbol{d}(\phi_{kl})\sqrt{\gamma_{kl}}e_{kl})\})^\dagger \\
&\quad -[\mathrm{E}\{(\boldsymbol{d}(\phi_{kl})\sqrt{\gamma_{kl}}e_{kl})(\boldsymbol{d}(\phi_{kl})\sqrt{\gamma_{kl}}e_{kl})^\dagger\}]^* \otimes \mathrm{E}\{(\boldsymbol{d}(\phi_{kl})\sqrt{\gamma_{kl}}e_{kl})(\boldsymbol{d}(\phi_{kl})\sqrt{\gamma_{kl}}e_{kl})^\dagger\} \\
&= \sum_{l=1}^{L_k} (\boldsymbol{d}^*(\phi_{kl}) \otimes \boldsymbol{d}(\phi_{kl})\gamma_{kl})(\mathrm{E}\{e_{kl}^* e_{kl} e_{kl} e_{kl}^*\} - 2[\mathrm{E}\{e_{kl}^* e_{kl}\}]^2)(\boldsymbol{d}^*(\phi_{kl}) \otimes \boldsymbol{d}(\phi_{kl})\gamma_{kl})^\dagger \\
&= \sum_{l=1}^{L_k} (\boldsymbol{d}^*(\phi_{kl}) \otimes \boldsymbol{d}(\phi_{kl})\gamma_{kl})\mathrm{cum}_4\{e_{kl}\}(\boldsymbol{d}^*(\phi_{kl}) \otimes \boldsymbol{d}(\phi_{kl})\gamma_{kl})^\dagger \\
&= \kappa \sum_{l=1}^{L_k} (\boldsymbol{d}(\phi_{kl})\boldsymbol{d}^\dagger(\phi_{kl})\gamma_{kl})^* \otimes (\boldsymbol{d}(\phi_{kl})\boldsymbol{d}^\dagger(\phi_{kl})\gamma_{kl}) \\
&= \kappa \sum_{l=1} (\boldsymbol{d}(\phi_{kl})\boldsymbol{d}^\dagger(\phi_{kl}))^* \otimes (\boldsymbol{d}(\phi_{kl})\boldsymbol{d}^\dagger(\phi_{kl}))\gamma_{kl}^2
\end{aligned} \tag{6.8}$$

If these discrete source components of the same distributed signal are closely clustered together, we replace each γ_{kl}^2 of the discrete source by the square of a density function spreading over the extent of the distributed source Φ_k. Then as the incremental angle $\Delta\phi \to 0$, Eq.(6.8) becomes

$$\begin{aligned}
Q_k &= \kappa \sum_{l=1}^{L_k} (\boldsymbol{d}(\phi_{kl})\boldsymbol{d}^\dagger(\phi_{kl}))^* \otimes (\boldsymbol{d}(\phi_{kl})\boldsymbol{d}^\dagger(\phi_{kl}))\gamma_k^2(\phi_l)\Delta\phi \\
&= p_{sk}^4 \kappa \int_{\phi \in \Phi_k} (\boldsymbol{d}(\phi)\boldsymbol{d}^\dagger(\phi))^* \otimes (\boldsymbol{d}(\phi)\boldsymbol{d}^\dagger(\phi))\gamma_k^2(\phi)d\phi
\end{aligned} \tag{6.9}$$

Here, p_{sk}^2 is the power of the kth distributed signal and $\gamma_k(\phi)$ is a normalized positive density

function characterizing the distribution of the kth signal, as is defined in Eq.(2.18).

6.2.2 Rank of the Cumulant Matrix

We now examine the rank of matrix $\boldsymbol{Q}_k$. We know that

$$\begin{aligned}
&\sum_{l=1}^{L_k}(\boldsymbol{d}(\phi_{kl})\boldsymbol{d}^\dagger(\phi_{kl})\gamma_{kl})^* \otimes \sum_{l=1}^{L_k}(\boldsymbol{d}(\phi_{kl})\boldsymbol{d}^\dagger(\phi_{kl})\gamma_{kl}) \\
= &\sum_{l=1}^{L_k}(\boldsymbol{d}(\phi_{kl})\boldsymbol{d}^\dagger(\phi_{kl})\gamma_{kl})^* \otimes (\boldsymbol{d}(\phi_{kl})\boldsymbol{d}^\dagger(\phi_{kl})\gamma_{kl}) \\
&+\sum_{\substack{l=1 \\ l\neq p}}^{L_k}\sum_{p=1}^{L_k}(\boldsymbol{d}(\phi_{kl})\boldsymbol{d}^\dagger(\phi_{kl})\gamma_{kl})^* \otimes (\boldsymbol{d}(\phi_{kp})\boldsymbol{d}^\dagger(\phi_{kp})\gamma_{kp}) \qquad (6.10)
\end{aligned}$$

where the first term on the RHS of Eq.(6.10) is identified as $\boldsymbol{Q}_k$ by virtue of Eq.(6.9). We note that the two terms on the RHS of Eq.(6.11) are positive semi-definite. Now, for two matrices $\boldsymbol{A}$ and $\boldsymbol{B}$ which are positive semi-definite, we have

$$\text{rank}\{\boldsymbol{A}+\boldsymbol{B}\} \geq \max(\text{rank}\{\boldsymbol{A}\},\ \text{rank}\{\boldsymbol{B}\}) \qquad (6.11)$$

The property of Eq.(6.11) can be easily seen from the fact that if $\mathcal{O}_A$ and $\mathcal{O}_B$ are respectively the null spaces for the positive semi-definite matrices A and B, then the null space for $A+B$ is given by $\mathcal{O}_A \cap \mathcal{O}_B$. Thus, applying this property to Eq.(6.10), we have

$$\begin{aligned}
&\text{rank}\{\sum_{l=1}^{L_k}(\boldsymbol{d}(\phi_{kl})\boldsymbol{d}^\dagger(\phi_{kl})\gamma_{kl})^* \otimes (\boldsymbol{d}(\phi_{kl})\boldsymbol{d}^\dagger(\phi_{kl})\gamma_{kl})\} \\
\leq\ &\text{rank}\{\sum_{l=1}^{L_k}(\boldsymbol{d}(\phi_{kl})\boldsymbol{d}^\dagger(\phi_{kl})\gamma_{kl})^* \otimes \sum_{l=1}^{L_k}(\boldsymbol{d}(\phi_{kl})\boldsymbol{d}^\dagger(\phi_{kl})\gamma_{kl})\} \qquad (6.12)
\end{aligned}$$

From Eq.(2.15), we know

$$\boldsymbol{R}_k = \sum_{l=1}^{L_k}\boldsymbol{d}(\phi_{kl})\boldsymbol{d}^\dagger(\phi_{kl})\gamma_{kl}$$

In Chapter 3, we have found that, depending on the spread of the source, the eigenvalues of $\boldsymbol{R}_k$ are usually negligibly small after the first few. Therefore, if the approximated rank of $\boldsymbol{R}_k$ is m_{0k}, then

$$\text{rank}\{\boldsymbol{Q}_k\} \leq \text{rank}\{\boldsymbol{R}_k^* \otimes \boldsymbol{R}_k\} = m_{0k}^2 \qquad (6.13)$$

where we have used the property [96] that rank$\{\boldsymbol{A} \otimes \boldsymbol{B}\}$ = rank$\{\boldsymbol{A}\}$ × rank$\{\boldsymbol{B}\}$. Thus, if we know that the approximate rank m_{0k} of $\boldsymbol{R}_k$, then we can use Eq.(6.13) to estimate the approximate rank of $\boldsymbol{Q}_k$. (However, this may sometimes introduce errors in the estimation of the dimensionality of the quasi signal subspace as we will see in Section 6.5.)

For the received data of Eq.(6.5), the cumulant matrix of the data is given by

$$\boldsymbol{Q} = \sum_{k=1}^{K} \boldsymbol{Q}_k \tag{6.14}$$

and if we use Eq.(6.13) to estimate the dimensionality, then $\boldsymbol{Q}$ will have a quasi signal subspace of dimension $M_0 \simeq \sum_{k=1}^{K} m_{0k}^2$ ($< M^2$). The remaining ($M^2 - M_0$) dimensional subspace will be designated the quasi noise subspace. By carrying out the eigen-decomposition on the cumulant matrix $\boldsymbol{Q}$,

$$\boldsymbol{Q} = [\boldsymbol{U}_s, \boldsymbol{U}_\nu] \begin{bmatrix} \boldsymbol{\Lambda}_Q & \boldsymbol{O} \\ \boldsymbol{O} & \boldsymbol{O}^+ \end{bmatrix} \begin{bmatrix} \boldsymbol{U}_s^\dagger \\ \boldsymbol{U}_\nu^\dagger \end{bmatrix} \tag{6.15}$$

where $\boldsymbol{\Lambda}_Q$ is an M_0-dimensional diagonal matrix of eigenvalues and $\boldsymbol{O}^+$ can be approximated as a zero-matrix by neglecting the last ($M^2 - M_0$) eigenvalues, we can clearly see that the cumulant matrix $\boldsymbol{Q}$ is in the quasi signal subspace.

6.3 DOA Estimation with Cumulant Matrix

As discussed above, the fourth-order cumulant matrix is rank defective, and the quasi signal subspace corresponds to the range space of this matrix. In contrast with the methods based on second-order statistics, this property holds true regardless of the noise spatial covariance; estimates of the DOA can be obtained without any *a priori* information about the noise spatial structure as long as noise is Gaussian and is independent of the signals. This property appears to be the strongest motivation for using higher-order statistics in the DOA estimation problem [90, 91], [93-95].

From Eq.(6.15), we know that the cumulant matrix $\boldsymbol{Q}$ is in the quasi signal subspace which is orthogonal to the quasi noise eigenvectors, i.e.,

$$\boldsymbol{Q}\boldsymbol{U}_\nu = \boldsymbol{U}_s \boldsymbol{\Lambda}_Q \boldsymbol{U}_s^\dagger \boldsymbol{U}_\nu \simeq \boldsymbol{O} \tag{6.16}$$

Since from Eq.(6.14) the row of $\boldsymbol{Q}$ is a linear combination of the rows of $\boldsymbol{Q}_k$, $k = 1, \cdots, K$, then the rows of $\boldsymbol{Q}$ span the same subspace as that spanned by the rows of $\boldsymbol{Q}_k$, $\forall k$. Therefore, we have the orthogonal condition that

$$\boldsymbol{Q}_k \boldsymbol{U}_\nu \simeq \boldsymbol{O} \tag{6.17}$$

where $\boldsymbol{Q}_k$ is the cumulant matrix induced by the kth distributed signal source.

In practice, we measure the sample cumulant matrix $\hat{\boldsymbol{Q}}$ by

$$\begin{aligned}\hat{\boldsymbol{Q}} &= \frac{1}{N}\sum_{n=1}^{N}[\boldsymbol{x}^*(n)\otimes\boldsymbol{x}(n)][\boldsymbol{x}^*(n)\otimes\boldsymbol{x}(n)]^\dagger - \frac{1}{N^2}\sum_{n=1}^{N}[\boldsymbol{x}^*(n)\otimes\boldsymbol{x}(n)]\sum_{n=1}^{N}[\boldsymbol{x}^*(n)\otimes\boldsymbol{x}(n)]^\dagger \\ &\quad - \frac{1}{N^2}\sum_{n=1}^{N}[\boldsymbol{x}(n)\boldsymbol{x}^\dagger(n)]^*\otimes\sum_{n=1}^{N}\boldsymbol{x}(n)\boldsymbol{x}^\dagger(n)\end{aligned} \tag{6.18}$$

where N is the number of snapshots. And we assign the last $(M^2 - M_0)$ eigenvectors of $\hat{\boldsymbol{Q}}$ as the estimated quasi noise eigenvectors $\hat{\boldsymbol{U}}_\nu$.

Now, $\hat{\boldsymbol{Q}}$ converges asymptotically to $\boldsymbol{Q}$ [97], i.e.,

$$\lim_{N\to\infty}\hat{\boldsymbol{Q}} \underset{a.s.}{=} \boldsymbol{Q} \tag{6.19}$$

where "$\underset{a.s.}{=}$" indicates that the equality holds with probability 1. Thus, we have

$$\lim_{N\to\infty}\hat{\boldsymbol{U}}_\nu \underset{a.s.}{=} \boldsymbol{U}_\nu \tag{6.20}$$

Using (6.17) we obtain

$$\lim_{N\to\infty}\boldsymbol{Q}_k\hat{\boldsymbol{U}}_\nu \underset{a.s.}{\simeq} \boldsymbol{O} \tag{6.21}$$

Thus, we can conceive an algorithm in which we adjust the parameters $\bar{\phi}_k$ and σ_k in $\boldsymbol{Q}_k$ until Eq.(6.21) is satisfied. Since the square of the Frobenius norm of a matrix is the measure of

the sum of the square of each element of the matrix, we arrive at an algorithm to obtain the estimates such that

$$\{\hat{\bar{\phi}}_k, \hat{\sigma}_k\} = \arg \min_{\bar{\phi}_k, \sigma_k} \|\boldsymbol{Q}_k \hat{\boldsymbol{U}}_\nu\|_F^2 \tag{6.22}$$

where $\|\cdot\|_F$ denotes the Frobenius norm. Therefore, we locate the K minima of Eq.(6.22) for the estimation of mean DOA $\bar{\phi}_k$ and the spread σ_k for K distributed signals.

6.4 Asymptotic Performance Analysis

The asymptotic performance of DOA estimation based on fourth-order cumulant matrix can be established along similar lines as in analyzing the performance of DISPARE II because in both cases, the DOA estimates are obtained from the sample estimates $\hat{\boldsymbol{R}}_x$ and $\hat{\boldsymbol{Q}}$ according to similar procedures. That is, we minimize the Frobenius norm of the projection of $\boldsymbol{R}_k$ and $\boldsymbol{Q}_k$ respectively onto estimated noise subspace in each case. Therefore, our first step consists of evaluating the asymptotic distribution of the sample statistics of interest, i.e., the fourth-order cumulant matrix $\hat{\boldsymbol{Q}}$.

Theorem 6.1 *[97] The elements of the sample cumulant matrix $\hat{\boldsymbol{Q}}$ are asymptotically Gaussian distributed with the mean $\boldsymbol{Q}$ and covariance $\boldsymbol{\Xi}_Q$, i.e.,*

$$\sqrt{N}\, vec(\hat{\boldsymbol{Q}} - \boldsymbol{Q}) \rightarrow \mathcal{N}(\boldsymbol{O}, \boldsymbol{\Xi}_Q) \tag{6.23}$$

Here, the covariance matrix $\boldsymbol{\Xi}_Q$ is a quantity with eight indices, whose entries are defined as

$$\begin{aligned} & \boldsymbol{\Xi}_Q[(i-1)M^3 + (j-1)M^2 + (k-1)M + l;\ (u-1)M^3 + (v-1)M^2 + (s-1)M + t] \\ \triangleq\ & \boldsymbol{\Xi}_Q[i,j,k,l;\ u,v,s,t] \\ =\ & \lim_{N\to\infty} N\mathrm{E}\{(\hat{\boldsymbol{Q}}[i,j;\ k,l] - \boldsymbol{Q}[i,j;\ k,l])(\hat{\boldsymbol{Q}}[u,v;\ s,t] - \boldsymbol{Q}[u,v;\ s,t])^*\} \\ & \qquad\qquad 1 \le i,j,k,l,u,v,s,t \le M \end{aligned} \tag{6.24}$$

The expansion of Ξ_Q in terms of the elements of the cumulant matrix Q is given in Appendix C.

In the following, we will examine the asymptotic distribution of eigen-projections. We define

$$\Pi_{Q_s} = U_s U_s^\dagger = \text{projector on the quasi signal subspace} \tag{6.25a}$$

$$\Pi_{Q_\nu} = U_\nu U_\nu^\dagger = \text{projector on the quasi noise subspace} \tag{6.25b}$$

The corresponding errors of estimation are denoted by

$$\Delta Q = \hat{Q} - Q \tag{6.26a}$$

$$\Delta\Pi_{Q_s} = \hat{\Pi}_{Q_s} - \Pi_{Q_s} \tag{6.26b}$$

$$\Delta\Pi_{Q_\nu} = \hat{\Pi}_{Q_\nu} - \Pi_{Q_\nu} \tag{6.26c}$$

We note that

$$\Delta\Pi_{Q_s} = -\Delta\Pi_{Q_\nu} \triangleq \Delta\Pi_Q \tag{6.26d}$$

Theorem 6.2 *The matrix $\Delta\Pi_Q$ can be approximated in the first order by*

$$\Delta\Pi_Q = \Delta_Q + O(\Delta Q^2) \tag{6.27}$$

where

$$\Delta_Q = \Pi_{Q_\nu}(\Delta Q)Q^\# + Q^\#(\Delta Q)\Pi_{Q_\nu} \tag{6.28}$$

Proof:

The proof of Theorem 6.2 can be carried out similarly to that given in Appendix B.2 with the obvious substitution of the covariance matrix by the cumulant matrix. □

Covariance of the DOA Estimates

From Eq.(6.22), we know that the criterion for DOA estimation based on fourth-order cumulants is given by

$$f_Q(\bar{\phi}, \sigma) = \|Q_k(\bar{\phi}, \sigma)\hat{U}_\nu\|_F^2 = 2\text{Re}[\text{tr}\{\hat{\Pi}_{Q_\nu} Q_k Q_k^\dagger\}] \tag{6.29}$$

We have the first derivatives

$$\dot{f}_{Qk} = 2\mathrm{Re}[\mathrm{tr}\{\dot{\hat{\Pi}}_{Q\nu}\boldsymbol{Q}_k\dot{\boldsymbol{Q}}_k^\dagger\}] \tag{6.30a}$$

$$f'_{Qk} = 2\mathrm{Re}[\mathrm{tr}\{\dot{\hat{\Pi}}_{Q\nu}\boldsymbol{Q}_k\boldsymbol{Q}'^\dagger_k\}] \tag{6.30b}$$

and similar to Eq.(5.23), we have

$$\mathrm{E}\begin{bmatrix}\Delta\bar{\phi}_k\\ \Delta\sigma_k\end{bmatrix} = -2\boldsymbol{H}_{Qk}\begin{bmatrix}\mathrm{Retr}\{\mathrm{E}[\dot{\hat{\Pi}}_{Q\nu}\boldsymbol{Q}_k]\dot{\boldsymbol{Q}}_k^\dagger\}\\ \mathrm{Retr}\{\mathrm{E}[\dot{\hat{\Pi}}_{Q\nu}\boldsymbol{Q}_k]\boldsymbol{Q}'^\dagger_k\}\end{bmatrix} = \boldsymbol{O} \tag{6.31}$$

where $\boldsymbol{H}_{Qk}$ is defined as

$$\boldsymbol{H}_{Qk} \triangleq \begin{bmatrix}\ddot{f}_{Qk} & \dot{f}'_{Qk}\\ \dot{f}'_{Qk} & \ddot{f}_{Qk}\end{bmatrix}^{-1} \tag{6.32}$$

and the second derivatives of $f_Q(\bar{\phi},\sigma)$ are given as

$$\ddot{f}_{Qk} = 2\mathrm{Re}[\mathrm{tr}\{\dot{\hat{\Pi}}_{Q\nu}(\boldsymbol{Q}_k\ddot{\boldsymbol{Q}}_k^\dagger + \dot{\boldsymbol{Q}}_k\dot{\boldsymbol{Q}}_k^\dagger)\}] \tag{6.33a}$$

$$\dot{f}'_{Qk} = 2\mathrm{Re}[\mathrm{tr}\{\dot{\hat{\Pi}}_{Q\nu}(\boldsymbol{Q}_k\dot{\boldsymbol{Q}}'^\dagger_k + \dot{\boldsymbol{Q}}_k\boldsymbol{Q}'^\dagger_k)\}] \tag{6.33b}$$

$$f''_{Qk} = 2\mathrm{Re}[\mathrm{tr}\{\dot{\hat{\Pi}}_{Q\nu}(\boldsymbol{Q}_k\boldsymbol{Q}''^\dagger_k + \boldsymbol{Q}'_k\boldsymbol{Q}'^\dagger_k)\}] \tag{6.33c}$$

Furthermore, we have

$$\mathrm{cov}\{\begin{bmatrix}\Delta\bar{\phi}_k\\ \Delta\sigma_k\end{bmatrix}\} = \boldsymbol{H}_{Qk}\begin{bmatrix}\mathrm{cov}\{\dot{f}_{Qk},\dot{f}_{Qk}\} & \mathrm{cov}\{\dot{f}_{Qk},f'_{Qk}\}\\ \mathrm{cov}\{f'_{Qk},\dot{f}_{Qk}\} & \mathrm{cov}\{f'_{Qk},f'_{Qk}\}\end{bmatrix}\boldsymbol{H}_{Qk}^\dagger \tag{6.34}$$

Recalling from Eqs.(5.24) and (5.25) which give the covariance of $[\dot{f}_{IIk}, f'_{IIk}]^T$ for DISPARE II with obvious substitution of $\boldsymbol{P}_\nu$ and $\boldsymbol{R}_{uk}$ by $\boldsymbol{\Pi}_{Q\nu}$ and $\boldsymbol{Q}_k$, and, using the first order approximation of $\Delta\boldsymbol{\Pi}_Q$ given in Eq.(6.28) and the asymptotic distribution of $\Delta\boldsymbol{Q}$ given by Theorem 6.1, we can evaluate the following covariance as, by straightforward mathematical manipulations,

$$\mathrm{cov}\{\dot{f}_{Qk},\dot{f}_{Qk}\} = 2\mathrm{Re}[\mathrm{vec}^\dagger(\dot{\boldsymbol{Q}}_k\boldsymbol{Q}_k^\dagger)\boldsymbol{C}_Q\boldsymbol{\Xi}_Q\boldsymbol{C}_Q^\dagger\mathrm{vec}(\dot{\boldsymbol{Q}}_k\boldsymbol{Q}_k^\dagger)] \tag{6.35a}$$

$$\mathrm{cov}\{\dot{f}_{Qk}, f'_{Qk}\} = 2\mathrm{Re}[\mathrm{vec}^{\dagger}(\dot{Q}_k Q_k^{\dagger}) C_Q \Xi_Q C_Q^{\dagger} \mathrm{vec}(Q'_k Q_k^{\dagger})] \tag{6.35b}$$

$$\mathrm{cov}\{f'_{Qk}, f'_{Qk}\} = 2\mathrm{Re}[\mathrm{vec}^{\dagger}(Q'_k Q_k^{\dagger}) C_Q \Xi_Q C_Q^{\dagger} \mathrm{vec}(Q'_k Q_k^{\dagger})] \tag{6.35c}$$

where

$$C_Q \triangleq \Pi_{Q\nu}^T \otimes Q^{\#} + Q^{\#T} \otimes \Pi_{Q\nu} \tag{6.36}$$

Substituting Eqs.(6.35) and (6.36) into Eq.(6.34), we can evaluate the covariance matrix of the estimation errors.

6.5 Numerical Experiments

6.5.1 Effects of Distribution of signals

So far we have not made any assumption on the distribution of signal sources. In the first numerical experiment, we will verify the rank evaluation of the matrix Q_k given by Eq.(6.13) for different distributions. In the second and the third experiments, we will examine the effect of distribution of signals on the performance of the DOA estimation via computer simulation.

Experiment 6.1

In this experiment, we numerically evaluate the rank of the matrix Q_k in two cases where the signal source is of uniform distribution and Gaussian distribution respectively.

For a uniformly distributed source with power density function as

$$\gamma_k(\phi) = \begin{cases} \frac{1}{2\sigma_k} & |\phi - \bar{\phi}_k| \leq \sigma_k \\ 0 & \text{elsewhere} \end{cases}$$

and a uniform linear array with M sensors, by using Eq.(6.9) the mnth entries of the cumulant matrix Q_k are given as,

$$Q_k[m;n] = \frac{1}{2\sigma_k} \frac{\sin(n_1 - n_2 - (m_1 - m_2))\sigma_k}{(n_1 - n_2 - (m_1 - m_2))\sigma_k} e^{j(n_1 - n_2 - (m_1 - m_2))\bar{\phi}_k} \tag{6.37}$$

where $m = (m_1-1)M+n_1$, $n = (m_2-1)M+n_2$, and $m_i, n_i = 1, 2, \cdots, M$, $i = 1, 2$. Fig.6.1a shows the dominant eigenvalues of $\boldsymbol{Q}_k$ given by Eq.(6.37) for a uniformly distributed signal with various spread width σ_k, given $M = 4$. The eigenvalues are normalized to the sum of total eigenvalues for each σ_k. We can use graphs similar to Fig.6.1a to determine the approximate rank of $\boldsymbol{Q}_k$ rather than using Eq.(6.13) which may only be a rough approximation. This can be demonstrated by observing from Fig.3.1c that for an array of 4 sensors and a signal spread of $\sigma_k = .5236 rad.$, $\boldsymbol{R}_k$ is of approximate rank of 2. If we use Eq.(6.13) to estimate the approximate rank of $\boldsymbol{Q}_k$, then $m_{0k}^2 = 4$. However from Fig.6.1a for an array of 4 sensors and a signal spread of $.5236 rad.$, the approximate rank of $\boldsymbol{Q}_k$ is only 3. Thus we can conclude that if we know the distribution of the signal, a more reliable way is to plot the eigenvalues of the matrix $\boldsymbol{Q}_k$ directly. From this plot, the approximate rank of $\boldsymbol{Q}_k$ can be obtained more accurately than the use of Eq.(6.13).

For a Gaussian distributed source with power density function given by

$$\gamma_k(\phi) = \frac{1}{\sqrt{2\pi\sigma_k^2}} e^{-\frac{(\phi-\bar{\phi}_k)^2}{2\sigma_k^2}}$$

and again a linear array of M sensors with equal spacing, by using Eq.(6.9), the mnth entries of the cumulant matrix $\boldsymbol{Q}_k$ are given as

$$\boldsymbol{Q}_k[m;n] = \frac{1}{2\sqrt{\pi}\sigma_k} e^{\frac{\sigma_k^2}{4}(n_1-n_2-(m_1-m_2))^2} e^{j(n_1-n_2-(m_1-m_2))\bar{\phi}_k} \tag{6.38}$$

where $m = (m_1 - 1)M + n_1$, $n = (m_2 - 1)M + n_2$, and $m_i, n_i = 1, 2, \cdots, M$, $i = 1, 2$. Fig.6.1b plots the dominant eigenvalues of $\boldsymbol{Q}_k$ given by Eq.(6.38) for a Gaussian distributed signal with various σ_k, given $M = 4$. The eigenvalues are also normalized to the sum of total eigenvalues for each σ_k. Once again, we can use this graph to estimate the approximate rank of the matrix $\boldsymbol{Q}_k$ of a Gaussian distributed signal more accurately than using Eq.(6.13).

Experiment 6.2

In this experiment, a 4-sensor linear array with uniform spacing of half-wavelength of

the signal source is used. All sensors are omnidirectional and have unity gain. A spatially Gaussian distributed signal of 4-QAM with mean DOA $\bar{\phi} = -.2739rad.$ and spread $\sigma = .1745rad.$ (corresponding to physical angle of arrival $[-5°, 3.3°]$) arrives at the array. Gaussian noise with spatial correlation between sensor m and sensor n equal to $\rho^{|m-n|}$ is added to the received signal given $\rho = 0.5$. The number of snapshots is taken to be 500. Examining the eigenvalues of the cumulant matrix $\boldsymbol{Q}_k$ from the plots in Fig.6.1b, we determine the approximate dimension of signal subspace to be 3. In the estimation of $\bar{\phi}_k$, we employ Eq.(6.22) to estimate the DOA. We assume that the source is uniformly distributed. Under this assumption, we perform the DOA estimation using Eq.(6.22) with $\boldsymbol{Q}_k$ corresponding to that given by a uniformly distributed signal. Fig.6.2 plots the RMS errors of the mean DOA estimate for the assumption of uniform distribution (solid line). The RMS error for the correctly assumed Gaussian distribution is also plotted by the dashed line for comparison. We observe that under the assumption of uniform distribution the estimation algorithm gives rise to a performance very close to that under the correct assumption of Gaussian distribution.

Experiment 6.3

In this experiment, the scenario is as same as that in Experiment 6.2 except that the signal of 4-QAM is spatially in uniform distribution and arrives at the array with mean DOA $\bar{\phi} = -.2739rad.$ and spread $\sigma = .5236rad.$ (corresponding to physical angle of arrival $[-5°, 9.9°]$) on each side of mean DOA. Gaussian noise with spatial correlation between sensor m and sensor n equal to $\rho^{|m-n|}$ is added to the received signal given $\rho = 0.5$. The number of snapshots is 500. Examining the eigenvalues of the cumulant matrix $\boldsymbol{Q}_k$ from the plots in Fig.6.1a, we determine the approximate dimension of signal subspace to be 3. In the estimation of $\bar{\phi}_k$, we employ Eq.(6.22) to estimate the DOA. We now assume that the source is Gaussian distributed, and compare the performance to that when a

correct assumption is made. Again Eq.(6.22) is used to obtain $\bar{\phi}_k$. Fig.6.3 plots the RMS errors of mean DOA estimates under the assumptions of Gaussian distribution (plotted by dashed line). The performance under the correct assumption of uniform distribution is also plotted (solid line). We observe that under the assumption of Gaussian distribution the estimation algorithm yields large error compared to the results under the correct assumption of distribution

Experiments 6.2 and 6.3 indicate that assuming uniform distribution the DOA estimation algorithm gives rise to a good estimation even when the signal is not in uniform distribution. Other symmetric distributions are tested and the same observation of robustness of uniform distribution persists. Therefore, the procedure of employing Eq.(6.22) for the estimation of DOA of a distributed signal is to assume the signal to be uniformly distributed and then apply Eq.(6.22) using the corresponding expression of Q_k to locate the minima.

6.5.2 Further Experiments

We carry out more numerical experiments to further illustrate the performance of the algorithm proposed in Section 6.3. In each of these simulations, again, a 4-sensor linear array with uniform spacing of half-wavelength of the signal source is used. All sensors are omnidirectional and have unity gain. We consider signals with uniform distribution. Gaussian noise with spatial correlation between sensor m and sensor n equal to $\rho^{|m-n|}$ is added to the received signal, where the cases of $\rho = 0$, 0.5 and 0.8 are examined. The number of snapshots is 500 and 100 Monte Carlo trials are run to evaluate the root mean square error (RMSE) of the estimated mean DOA under various SNR.

Experiment 6.4

In this experiment, one spatially uniformly distributed signal with spread $.3491 rad.$ on

each side of the mean DOA of $-.2739rad.$ arrives at the array (corresponding to physical angle of arrival $[-10°, 6.7°]$), the signal type is 4-QAM. Examining the eigenvalues of the cumulant matrix Q_k from the plots in Fig.6.1a, we determine the approximate dimension of signal subspace to be 2. We apply the criterion given by Eq.(6.22) to obtain the estimates $\tilde{\phi}_k$ and σ_k. Since we are, in general, only interested in the mean DOA, we only plot the RMSE of the estimation of mean DOA (in radian) for SNR from $0 \sim 20dB$. As a comparison, we also employ DISPARE II proposed in Chapter 4.2.2 and cumulant-MUSIC [90] to estimate the DOA. (Cumulant-MUSIC is a DOA estimation algorithm for point source which uses the fourth-order cumulants of the received data to form the so-called null-spectrum

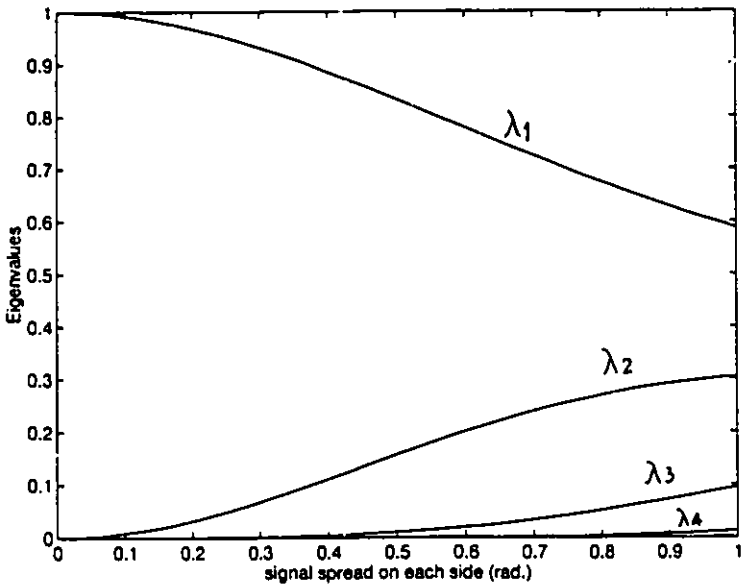
$$f_M(\phi_0) = \|(d(\phi_0) \otimes d^*(\phi_0))^\dagger \hat{U}_\nu\|^2 \tag{6.39}$$

and estimate the DOA by searching for the minima of $f_M(\phi_0)$.) Figs.6.4a, 6.4b and 6.4c plot the RMS errors of mean DOA of the above three methods in three different noise environments with $\rho = 0$, 0.5 and 0.8, respectively, where the RMSE of the new algorithm, DISPARE II and cum-MUSIC are plotted by solid line, dashed line and dashdot line respectively. It can be observed that the new algorithm yields similar performance in different noise environments (for $\rho = 0, 0.5$ and 0.8), and outperforms DISPARE II when SNR $< 15dB$; DISPARE II is sensitive to the spatially correlated noise, however, when SNR is above $15dB$ DISPARE II gives rise to better performance than the new algorithm. On the other hand, although cumulant-MUSIC is not sensitive to the noise correlation, it yields much larger errors for the DOA estimation of the distributed signal.

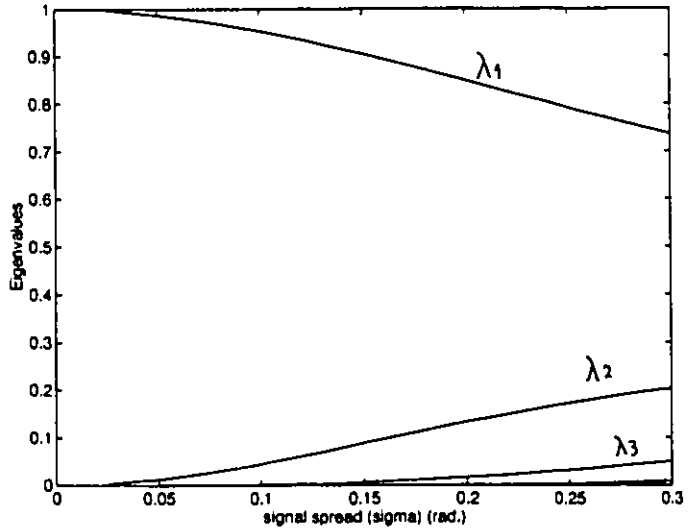
Experiment 6.5

The final experiment of this chapter is for comparison between the theoretical evaluation and the computer simulation of the estimation accuracy. The comparison is based on the same scenario as that in Experiment 6.4, where the signal type is 4-QAM, arriving at the array at mean DOA $-.2739rad.$ with spread $.3491rad.$ on each side of the mean DOA.

Gaussian noise with spatial correlation between sensor m and sensor n equals to $\rho^{|m-n|}$ is added to the received signal, where $\rho = 0.5$. The number of snapshots is 500. We use Eqs.(6.31) and (6.34) to evaluate the asymptotic performance of the DOA estimates. Fig.6.5 plots the RMS errors of the mean DOA estimation under various SNR (solid line), and as a comparison, the simulation result is plotted again (crossed line). We observe that, the performance analysis indicates that the algorithm based on fourth-order cumulants can achieve higher accuracy in estimating mean DOA of the signal when SNR increases; however, computer simulation shows the leveling off in the accuracy after SNR being greater than $15dB$. This discrepancy may be due to the convergent speed of sample cumulants to true ones and that sampled estimates of cumulants may not be optimal for non-Gaussian processes. This inaccuracy in the sample cumulants may not show up under lower SNR due to the dominance of noise. However, at high SNR, this inaccuracy becomes the dominant part and represents an irreducible error in the DOA estimation. Hence, other estimates of cumulants rather than sampled estimates may provide better results.



(a) A signal with uniform distribution



(b) A signal with Gaussian distribution

Figure 6.1: Eigenvalues of a cumulant matrix of a distributed signal given $M = 4$

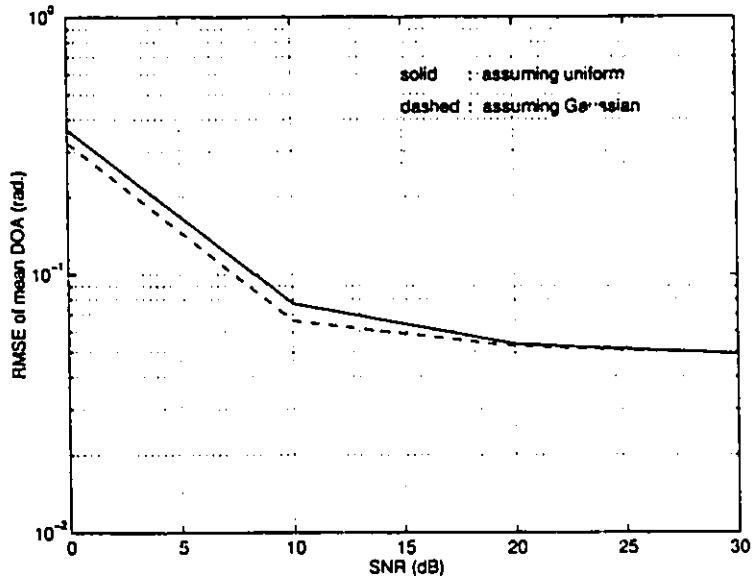


Figure 6.2: The estimation errors of mean DOA of a Gaussian distributed signal

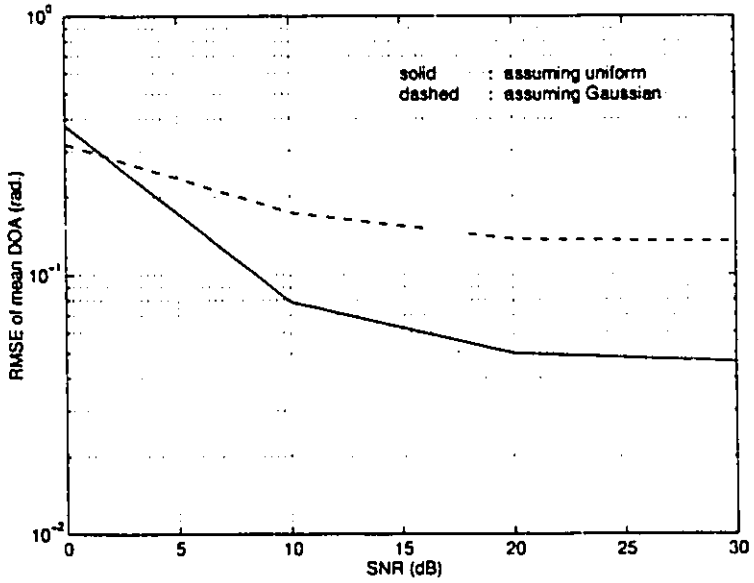
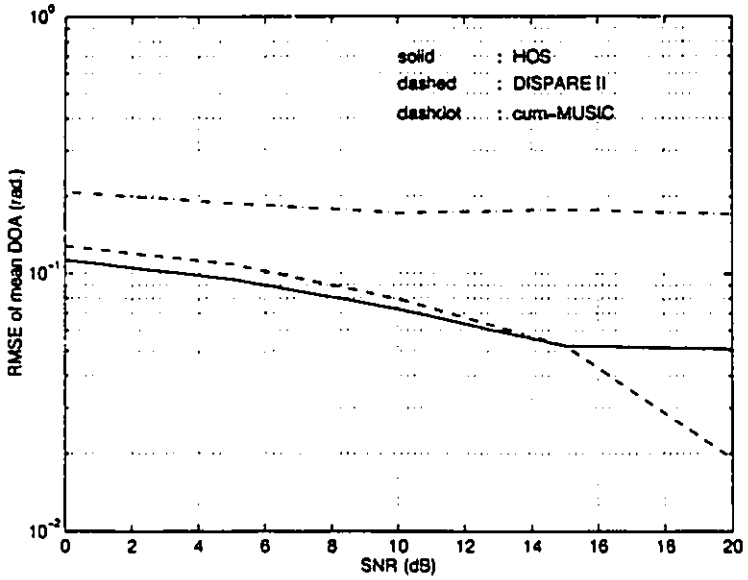
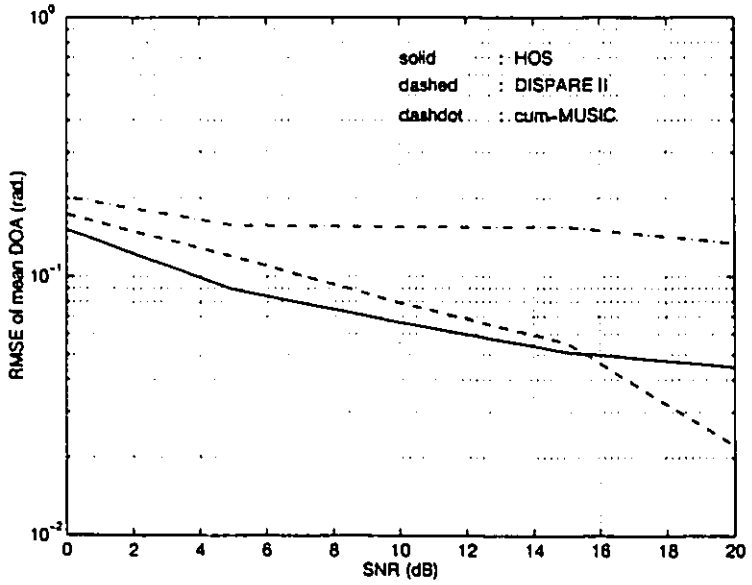


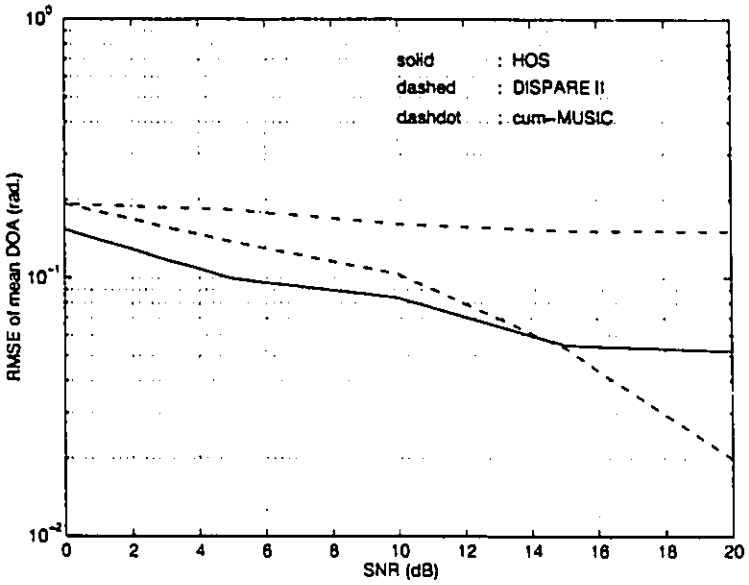
Figure 6.3: The estimation errors of mean DOA of a uniformly distributed signal



(a) $\rho = 0$



(b) $\rho = 0.5$



(c) $\rho = 0.8$

Figure 6.4: The estimation errors of mean DOA with three methods

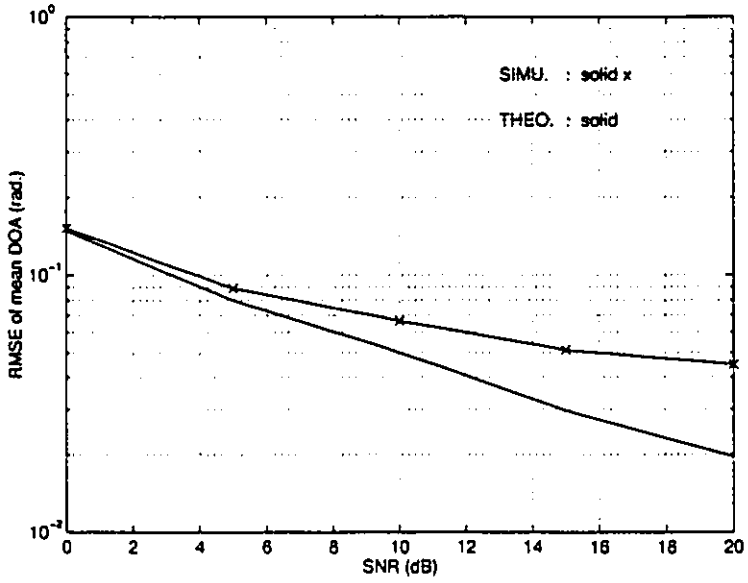


Figure 6.5: Theoretical evaluation and simulation of RMSE of mean DOA of a uniformly distributed signal of 4-QAM with $(\bar{\phi}, \sigma) = (-.2739, .3491)$ arriving at a uniform linear array of 4 sensors; $\rho = 0.5$; 500 snapshots are used for comparison.

Chapter 7

VEC-MUSIC Estimation

As we discussed in Chapter 2, we can consider point sources as a special case of distributed signal sources, or in other words, the model of a distributed sources is a more general form of the signals (which include point source signals) we encounter in practical applications. It is based on this perception that, we develop the DOA estimation methods by using the second-order statistics and the higher-order statistics in Chapters 4 and 6. These methods can be applied to the general distributed signal model, estimate the mean DOA and the spread of the sources. The obvious distinction between point and distributed sources is the estimated values of spreads.

Stimulated by the intention of developing a method which can distinguish a point source from a distributed source even when the point source is merged with the distributed one, in this chapter, we will propose another novel method which can jointly estimate the DOA of point sources and distributed sources [98].

7.1 The Estimation Method

We start from our discrete signal model and examine the case when one distributed source and one point source are received without background noise. We confine ourselves to Gaussian signals again. The received data by an array are given by

$$\begin{aligned} \boldsymbol{x}(n) &= \sum_{l=1}^{L} \boldsymbol{d}(\phi_l)\sqrt{\gamma_l}e_l(n) + \boldsymbol{d}(\phi_0)s_0(n) \\ &= \boldsymbol{D}\boldsymbol{\Gamma}\begin{bmatrix} e_1(n) \\ \vdots \\ e_L(n) \end{bmatrix} + \boldsymbol{d}(\phi_0)s_0(n) \end{aligned} \tag{7.1}$$

The first term as has been observed in Eq.(6.5), is the received signal component corresponding to a distributed source in which there are L closely clustered point sources; however, $e_l(n)$, $l = 1, \cdots, L$ here are independently identically distributed (i.i.d.) zero-mean Gaussian random processes with unity variance; and $\boldsymbol{D} = [\boldsymbol{d}(\phi_1) \cdots \boldsymbol{d}(\phi_L)]$, $\boldsymbol{\Gamma} = \text{diag}(\sqrt{\gamma_1}, \cdots, \sqrt{\gamma_L})$ with $\sqrt{\gamma_l}$ being the strength of the signal from the lth point source. The second term is due to a point signal in which $\boldsymbol{d}(\phi_0)$ is the array steering vector at the incident angle ϕ_0 of the point source and $s_0(n)$ is the signal waveform.

We see that the outer product of the received data $\boldsymbol{x}(n)$ is given by

$$\begin{aligned} \boldsymbol{x}(n)\boldsymbol{x}^{\dagger}(n) &= \boldsymbol{D}\boldsymbol{\Gamma}\begin{bmatrix} e_1(n) \\ \vdots \\ e_L(n) \end{bmatrix}[e_1^*(n) \cdots e_L^*(n)]\boldsymbol{\Gamma}\boldsymbol{D}^{\dagger} + \boldsymbol{D}\boldsymbol{\Gamma}\begin{bmatrix} e_1(n) \\ \vdots \\ e_L(n) \end{bmatrix} s_0^*(n)\boldsymbol{d}^{\dagger}(\phi_0) \\ &\quad + \boldsymbol{d}(\phi_0)s_0(n)[e_1^*(n) \cdots e_L^*(n)]\boldsymbol{\Gamma}\boldsymbol{D}^{\dagger} + \boldsymbol{d}(\phi_0)s_0(n)s_0^*(n)\boldsymbol{d}^{\dagger}(\phi_0) \end{aligned} \tag{7.2}$$

We take the vec operation (see Notation and Glossary) on $\boldsymbol{x}(n)\boldsymbol{x}^{\dagger}(n)$, by using the identity [96]

$$\text{vec}(\boldsymbol{ABC}) = (\boldsymbol{C}^T \otimes \boldsymbol{A})\text{vec}(\boldsymbol{B}) \tag{7.3}$$

where again $\otimes$ is the Kronecker product. Then we have

$$\begin{aligned}
\text{vec}[\boldsymbol{x}(n)\boldsymbol{x}^{\dagger}(n)] &= (\boldsymbol{D}^{*}\boldsymbol{\Gamma}) \otimes (\boldsymbol{D}\boldsymbol{\Gamma})\text{vec}\left\{\begin{bmatrix} e_1(n) \\ \vdots \\ e_L(n) \end{bmatrix} [e_1^{*}(n) \cdots e_L^{*}(n)]\right\} \\
&\quad + \boldsymbol{d}^{*}(\phi_0) \otimes (\boldsymbol{D}\boldsymbol{\Gamma})\text{vec}\left\{\begin{bmatrix} e_1(n) \\ \vdots \\ e_L(n) \end{bmatrix} s_0^{*}(n)\right\} \\
&\quad + (\boldsymbol{D}^{*}\boldsymbol{\Gamma}) \otimes \boldsymbol{d}(\phi_0)\text{vec}\left\{s_0(n)[e_1^{*}(n) \cdots e_L^{*}(n)]\right\} \\
&\quad + \boldsymbol{d}^{*}(\phi_0) \otimes \boldsymbol{d}(\phi_0)\text{vec}(s_0(n)s_0^{*}(n)) \qquad (7.4)
\end{aligned}$$

We now evaluate the covariance matrix of $\text{vec}[\boldsymbol{x}(n)\boldsymbol{x}^{\dagger}(n)]$, we have

$$\begin{aligned}
\bar{\boldsymbol{R}}_x &\triangleq \text{cov}\left\{\text{vec}[\boldsymbol{x}(n)\boldsymbol{x}^{\dagger}(n)]\right\} \\
&= (\boldsymbol{D}^{*}\boldsymbol{\Gamma}) \otimes (\boldsymbol{D}\boldsymbol{\Gamma})\text{cov}\left\{\text{vec}\left(\begin{bmatrix} e_1(n) \\ \vdots \\ e_L(n) \end{bmatrix} [e_1^{*}(n) \cdots e_L^{*}(n)]\right)\right\}(\boldsymbol{\Gamma}\boldsymbol{D}^{T}) \otimes (\boldsymbol{\Gamma}\boldsymbol{D}^{\dagger}) \\
&\quad + \boldsymbol{d}^{*}(\phi_0) \otimes (\boldsymbol{D}\boldsymbol{\Gamma})\text{cov}\left\{\text{vec}\left(\begin{bmatrix} e_1(n) \\ \vdots \\ e_L(n) \end{bmatrix} s_0^{*}(n)\right)\right\}\boldsymbol{d}^{T}(\phi_0) \otimes (\boldsymbol{\Gamma}\boldsymbol{D}^{\dagger}) \\
&\quad + (\boldsymbol{D}^{*}\boldsymbol{\Gamma}) \otimes \boldsymbol{d}(\phi_0)\text{cov}\{\text{vec}(s_0(n)[e_1^{*}(n) \cdots e_L^{*}(n)])\}(\boldsymbol{\Gamma}\boldsymbol{D}^{T}) \otimes \boldsymbol{d}^{\dagger}(\phi_0) \\
&\quad + (\boldsymbol{d}^{*}(\phi_0) \otimes \boldsymbol{d}(\phi_0))\text{cov}\{s_0(n)s_0^{*}(n)\}(\boldsymbol{d}^{T}(\phi_0) \otimes \boldsymbol{d}^{\dagger}(\phi_0)) \qquad (7.5)
\end{aligned}$$

Since $e_1(n), \cdots, e_L(n)$ are i.i.d. Gaussian processes with zero mean and unity variance and $s_0(n)$ is an i.i.d. Gaussian zero mean process with a variance of p_0^2 and $s_0(n)$ is uncorrelated

with $e_1(n), \cdots, e_L(n)$, from Eq.(B.15) we know that

$$\text{cov}\left\{\text{vec}\left(\begin{bmatrix} e_1(n) \\ \vdots \\ e_L(n) \end{bmatrix}[e_1^*(n), \cdots, e_L^*(n)]\right)\right\} = I_{L^2} \otimes I_{L^2} \tag{7.6a}$$

$$\text{cov}\left\{\text{vec}\left(\begin{bmatrix} e_1(n) \\ \vdots \\ e_L(n) \end{bmatrix} s_0^*(n)\right)\right\} = p_0^2 I_L \tag{7.6b}$$

where I_L and I_{L^2} are identity matrices with dimension $L \times L$ and $L^2 \times L^2$, respectively.

Therefore, by using Eqs.(7.6), we can further simplify Eq.(7.5) to

$$\begin{aligned} \bar{R}_x &= (D^* \Gamma^2 D^T) \otimes (D \Gamma^2 D^\dagger) + p_0^2 (d^* d^T) \otimes (D \Gamma^2 D^\dagger) \\ &\quad + p_0^2 (D^* \Gamma^2 D^T) \otimes (d d^\dagger) + p_0^4 (d^* d^T) \otimes (d d^\dagger) \end{aligned} \tag{7.7}$$

where we omit the explicit dependence of d on the parameter ϕ_0 and the identity

$$(A \otimes B)(C \otimes D) = (AC) \otimes (BD) \tag{7.8}$$

has been employed, given matrices A, B, C and D have compatible dimensions.

To interpret the structure of the subspaces of $\bar{R}_x$, we examine the rank of each term in Eq.(7.7). We notice that

$$D \Gamma^2 D^\dagger = \sum_{l=1}^{L} d(\phi_l) d^\dagger(\phi_l) \gamma_l \tag{7.9}$$

which is the correlation matrix of the distributed signal. As we have examined in Chapter 3, the approximate dimensionality of signal subspace of a distributed signal is $m_0 (< M)$, i.e., the rank of matrix $D \Gamma^2 D^\dagger$ is approximately equal to m_0. Since we also know that [96]

$$\text{rank}\{A \otimes B\} = \text{rank}\{A\} \times \text{rank}\{B\} \tag{7.10}$$

and

$$\text{rank}\{d d^\dagger\} = 1 \tag{7.11}$$

Therefore, we have

$$\left.\begin{aligned} &\text{rank}\{(D^*\Gamma^2D^T)\otimes(D\Gamma^2D^\dagger)\}\simeq m_0^2\\ &\text{rank}\{(D^*\Gamma^2D^T)\otimes(dd^\dagger)\}\simeq m_0\\ &\text{rank}\{(d^*d^T)\otimes(D\Gamma^2D^\dagger)\}\simeq m_0\\ &\text{rank}\{(d^*d^T)\otimes dd^\dagger)\}=1 \end{aligned}\right\} \tag{7.12}$$

We can use the rank of the matrices shown in Eq.(7.12) to establish the dimensionality of our signal and noise subspaces. Before the vec operation, for M sensors, the dimensionality of the data space, the quasi signal subspace, and the quasi noise subspace are M, (m_0+1), and $M-(m_0+1)$ respectively. Here we assume $1\leq m_0\leq M-1$. Now, after the vec operation, the dimension of $\bar{R}_x$ in Eq.(7.7) is M^2. Thus the space in which the columns of $\bar{R}_x$ exists is of dimension M^2 and is designated the *extended data space*. However, the rank of $\bar{R}_x$ as indicated by Eq.(7.12) is approximately $m_0^2+2m_0+1=(m_0+1)^2$. Hence we can regard $\bar{R}_x$ as a rank deficient matrix. The rank of $\bar{R}_x$ is induced by the distributed and the point signals together with their cross-products. Hence the approximate dimensionality of the *extended signal subspace* is $(m_0+1)^2$. The remaining dimensionality is $M^2-(m_0+1)^2$ and is for a subspace which contains a mixture of signals and noise. This subspace is referred to as the *extended pseudo noise subspace* which includes an *extended quasi noise subspace* of dimension $(M-(m_0+1))^2$. For example, in the case when $M=4$, $K_1=K_2=1$ and $m_0=2$, then the dimensionality of the *extended data space* is $M^2=16$, the dimensionality of the *extended signal subspace* is $(m_0+1)^2=9$, the dimensionality of the *extended quasi noise subspace* is $(M-m_0-1)^2=1$, and the dimensionality of the *extended pseudo noise subspace* is $16-9=7$.

In practice, we can only have the estimated matrix $\hat{\bar{R}}_x$, hence, we can use the last few eigenvectors of $\hat{\bar{R}}_x$, which are $\{\hat{\bar{v}}_{(m_0+1)^2+1},\cdots,\hat{\bar{v}}_{M^2}\}$, corresponding to the eigenvectors spanning the extended pseudo noise subspace. We use $\hat{\bar{V}}_\nu$ to denote the estimated extended

pseudo noise eigenvector matrix

$$\hat{\bar{V}}_\nu = [\hat{\bar{v}}_{(m_0+1)^2+1} \cdots \hat{\bar{v}}_{M^2}] \tag{7.13}$$

We let

$$\bar{R} = (D^* \Gamma^2 D^T) \otimes (D \Gamma^2 D^\dagger) \tag{7.14}$$

and

$$\bar{d} = d^* \otimes d \tag{7.15}$$

We note that $\bar{d}$ and $\bar{R}$ only involve the point signal and distributed signal, respectively. Employing an idea similar to that of DISPARE II to find the DOA of the distributed signal, we establish the VEC-MUSIC algorithm for the DOA estimation of a distributed source such that

$$\{\hat{\bar{\phi}}, \hat{\sigma}\} = \arg \min_{\bar{\phi}, \sigma} \| \bar{R} \hat{\bar{V}}_\nu \|_F^2 \tag{7.16}$$

where $\| \cdot \|_F$ is the Frobenius norm and $\bar{R}$ is an explicit function of the parameters to be estimated. Eq.(7.16) describes a criterion for the estimation of the mean DOA and the spread of a distributed signal.

On the other hand, to estimate a point source, we can utilize the component $\bar{d}$ which as we have noted, involves only the point source. Thus, we can employ the orthogonality between $\bar{d}$ and $\hat{\bar{V}}_\nu$, and formulate a projection onto the estimated extended pseudo noise subspace such that

$$\hat{\bar{P}}_\nu = \hat{\bar{V}}_\nu \hat{\bar{V}}_\nu^\dagger \tag{7.17}$$

Applying this projector on $\bar{d}$ and changing the parameter ϕ_0 in the steering vector, we can locate the DOA of a point source by

$$\hat{\phi}_0 = \arg \min_{\phi_0} \bar{d}^\dagger \hat{\bar{P}}_\nu \bar{d} \tag{7.18}$$

Eq.(7.18) describes the criterion for estimating the DOA of a point source signal. The pair of Eqs. (7.16) and (7.18) are called the VEC-MUSIC algorithm.

If there are K_1 and K_2 distributed and point sources respectively, then the total dimensionality of the signal subspace is $M_0 + K_2$ ($< M$) where $M_0 = \sum_{k=1}^{K_1} m_{0k}$ with m_{0k} being the approximate signal subspace dimensionality induced by the kth distributed source before the vector operation is taken. The dimensionality of the noise subspace is therefore $M - M_0 - K_2$. After the vector operation has been taken, the total dimensionality of the *extended data space* is M^2. The total dimensionality of the *extended signal subspace* is $(M_0 + K_2)^2$. The total remaining dimensionality is $M^2 - (M_0 + K_2)^2$ for the *extended pseudo noise subspace* which contains a mixture of noise and signals and which includes an *extended quasi noise subspace* of dimension $(M - M_0 - K_2)^2$.

For the multiple distributed and point signal source case, there are K_1 and K_2 distributed and point sources respectively. Then by carrying out an eigen-decomposition on the estimated matrix $\hat{\bar{R}}_x$ we can make use of the last $M^2 - (M_0 + K_2)^2$ eigenvectors of $\hat{\bar{R}}_x$ which are $\{\hat{\bar{w}}_{(M_0+K2)^2+1}, \cdots, \hat{\bar{w}}_{M^2}\}$ spanning the extended pseudo noise subspace. We use $\hat{\bar{W}}_\nu$ to denote the estimated extended pseudo noise eigenvector matrix

$$\hat{\bar{W}}_\nu = [\hat{\bar{w}}_{(M_0+K_2)^2+1} \cdots \hat{\bar{w}}_{M^2}] \tag{7.19}$$

and thus we can form the projection onto the estimated extended pseudo noise subspace as

$$\hat{\bar{\Pi}}_\nu = \hat{\bar{W}}_\nu \hat{\bar{W}}_\nu^\dagger \tag{7.20}$$

We let

$$\bar{R}_k = (D_k^* \Gamma_k^2 D_k^T) \otimes (D_k \Gamma_k^2 D_k^\dagger) \tag{7.21}$$

and

$$\bar{d}_k = d_k^* \otimes d_k \tag{7.22}$$

where $\bar{R}_k$ and $\bar{d}_k$ only involve the kth distributed signals and kth point signals respectively. Hence, the VEC-MUSIC algorithm for the DOA estimation of K_1 distributed signal sources is given as

$$\{\hat{\bar{\phi}}_k, \hat{\sigma}_k\} = \arg\min_{\bar{\phi}_k, \sigma_k} \| \bar{R}_k \hat{\bar{W}}_\nu \|_F^2 \tag{7.23}$$

where $\| \cdot \|_F$ is the Frobenius norm and $\bar{R}_k$ is an explicit function of the parameters to be estimated for kth distributed signal source. The mean DOA and the spreads of the sources can be obtained by locating the K_1 minima of Eq.(7.23). The VEC-MUSIC algorithm for the DOA estimation of K_2 point signals is given as

$$\hat{\phi}_{0k} = \arg\min_{\phi_{0k}} \bar{d}_k^\dagger \hat{\bar{\Pi}}_\nu \bar{d}_k \tag{7.24}$$

The DOA for the point signals can be obtained by locating the K_2 minima of Eq.(7.24). The pair of Eqs. (7.23) and (7.24) describe the VEC-MUSIC algorithm for DOA estimation of multiple distributed and point sources.

7.2 Performance Analysis

In the analysis presented in this section, the statistical properties of the perturbation in the estimated covariance matrix $\bar{R}_x$ is examined and is utilized to determine the estimation accuracy of VEC-MUSIC algorithm.

7.2.1 Asymptotic Distribution of the Perturbation in Covariance Matrix

Unlike the statistics used in the DOA estimation methods in Chapter 4 and Chapter 6, from Eq.(7.5) the statistics used in VEC-MUSIC are

$$\begin{aligned}
\bar{R}_x &= \text{cov}\{\text{vec}(x(n)x^\dagger(n))\} \\
&= \text{cov}\{x^*(n) \otimes x(n)\} \\
&= \text{E}\{(x^*(n) \otimes x(n))(x^*(n) \otimes x(n))^\dagger\} - \text{E}\{x^*(n) \otimes x(n)\}\text{E}^\dagger\{x^*(n) \otimes x(n)\}
\end{aligned}$$

Recalling from Eq.(6.6) which defines the fourth-order cumulant matrix, i.e.,

$$\boldsymbol{Q} = \mathrm{cum}_4\{\boldsymbol{x}\} \triangleq \mathrm{E}\{(\boldsymbol{x}^* \otimes \boldsymbol{x})(\boldsymbol{x}^* \otimes \boldsymbol{x})^\dagger\} - \mathrm{E}\{\boldsymbol{x}^* \otimes \boldsymbol{x}\}[\mathrm{E}\{\boldsymbol{x}^* \otimes \boldsymbol{x}\}]^\dagger - [\mathrm{E}\{\boldsymbol{x}\boldsymbol{x}^\dagger\}]^* \otimes \mathrm{E}\{\boldsymbol{x}\boldsymbol{x}^\dagger\}$$

and the fact that the fourth-order cumulants of a Gaussian signal are zero, we have the following identity:

$$\begin{aligned} \bar{\boldsymbol{R}}_x &= [\mathrm{E}\{\boldsymbol{x}(n)\boldsymbol{x}^\dagger(n)\}]^* \otimes \mathrm{E}\{\boldsymbol{x}(n)\boldsymbol{x}^\dagger(n)\} \\ &= \boldsymbol{R}_x^* \otimes \boldsymbol{R}_x \end{aligned} \tag{7.25}$$

Given a finite number of snapshots, we estimate $\bar{\boldsymbol{R}}_x$ by

$$\hat{\bar{\boldsymbol{R}}}_x = \frac{1}{N}\sum_{n=1}^{N}[\mathrm{vec}(\boldsymbol{x}(n)\boldsymbol{x}^\dagger(n)) - \mathrm{vec}\hat{\boldsymbol{R}}_x][\mathrm{vec}(\boldsymbol{x}(n)\boldsymbol{x}^\dagger(n)) - \mathrm{vec}\hat{\boldsymbol{R}}_x]^\dagger \tag{7.26}$$

The following theorem gives the asymptotic distribution of $\hat{\bar{\boldsymbol{R}}}_x$.

Theorem 7.1 *The elements of* $\triangle\bar{\boldsymbol{R}}_x = \hat{\bar{\boldsymbol{R}}}_x - \bar{\boldsymbol{R}}_x$ *are asymptotically Gaussian distributed with first- and second-order moments given by*

$$E\{vec(\triangle\bar{\boldsymbol{R}}_x)\} = 0 \tag{7.27a}$$

$$E\{vec(\triangle\bar{\boldsymbol{R}}_x)vec^\dagger(\triangle\bar{\boldsymbol{R}}_x) = \frac{1}{N}\boldsymbol{\Xi}_{\bar{R}} \tag{7.27b}$$

$$E\{vec(\triangle\bar{\boldsymbol{R}}_x)vec^T(\triangle\bar{\boldsymbol{R}}_x) = \frac{1}{N}\boldsymbol{\Xi}_{\bar{R}}\bar{\boldsymbol{\Upsilon}} \tag{7.27c}$$

where

$$\boldsymbol{\Xi}_{\bar{R}} = unvec(\boldsymbol{c}_x) - vec(\bar{\boldsymbol{R}}_x)vec^\dagger(\bar{\boldsymbol{R}}_x) \tag{7.28}$$

with

$$\begin{aligned} \boldsymbol{c}_x &= E\{(\boldsymbol{x}_n^* \otimes \boldsymbol{x}_n - vec\boldsymbol{R}_x) \otimes (\boldsymbol{x}_n \otimes \boldsymbol{x}_n^* - vec\boldsymbol{R}_x^*) \\ &\quad \otimes(\boldsymbol{x}_n \otimes \boldsymbol{x}_n^* - vec\boldsymbol{R}_x^*) \otimes (\boldsymbol{x}_n^* \otimes \boldsymbol{x}_n - vec\boldsymbol{R}_x)\} \end{aligned} \tag{7.29}$$

and $\bar{\Upsilon}$ is a permutation matrix defined as

$$\bar{\Upsilon} = \sum_{l=1}^{M^2} \sum_{m=1}^{M^2} \bar{E}_{lm} \otimes \bar{E}_{ml} \tag{7.30}$$

with $\bar{E}_{lm}$ being an $M^2 \times M^2$ matrix having unity as its lmth element and zero elsewhere.

Proof: see Appendix B.3. □

7.2.2 Perturbation in Eigen-projectors

Now, we examine the perturbation in eigen-projectors. Let

$$\Pi_s = \sum_{i=1}^{(M_0+K_2)^2} \bar{w}_i \bar{w}_i^\dagger = \text{projector on the extended signal subspace} \tag{7.31a}$$

$$\Pi_\nu = \sum_{i=(M_0+K_2)^2+1}^{M^2} \bar{w}_i \bar{w}_i^\dagger = \text{projector on the extended pseudo noise subspace} \tag{7.31b}$$

$$\hat{\Pi}_s = \sum_{i=1}^{(M_0+K_2)^2} \hat{\bar{w}}_i \hat{\bar{w}}_i^\dagger = \text{projector on estimated extended signal subspace} \tag{7.31c}$$

$$\hat{\Pi}_\nu = \sum_{i=(M_0+K_2)^2+1}^{M^2} \hat{\bar{w}}_i \hat{\bar{w}}_i^\dagger = \text{projector on estimated extended pseudo noise subspace} \tag{7.31d}$$

We denote the corresponding errors of estimation by

$$\Delta \bar{R}_x = \hat{\bar{R}}_x - \bar{R}_x \tag{7.32a}$$

$$\Delta \Pi = \hat{\Pi}_s - \Pi_s = -(\hat{\Pi}_\nu - \Pi_\nu) \tag{7.32b}$$

Theorem 7.2 *Under the condition when there is no noise, the matrix $\Delta\Pi$ can be approximated in the first order by*

$$\Delta \Pi = \delta \Pi + O(\Delta \bar{R}_x^2) \tag{7.33a}$$

and

$$\delta \Pi = \Pi_\nu (\Delta \bar{R}_x) \bar{S}^\# + \bar{S}^\# (\Delta \bar{R}_x) \Pi_\nu \tag{7.33b}$$

where $\bar{S}^{\#}$ denotes the pseudo-inverse of $\bar{S}$ which is defined as

$$\bar{S} \triangleq \sum_{i=1}^{(M_0+K_2)^2} \bar{\mu}_i \bar{w}_i \bar{w}_i^{\dagger}, \quad \bar{\mu}_i \text{ are the eigenvalues of } \bar{R}_x. \tag{7.34}$$

Proof:

If we denote the nth-order variational of $\hat{\Pi}_s$ and $\hat{\Pi}_\nu$ by $\delta^n \Pi$, then we may express $\hat{\Pi}_s$ and $\hat{\Pi}_\nu$ as

$$\hat{\Pi}_s = \Pi_s + \delta\Pi + \cdots + \delta^n \Pi + \cdots = \Pi_s + \Delta\Pi \tag{7.35a}$$

$$\hat{\Pi}_\nu = \Pi_\nu - \delta\Pi - \cdots - \delta^n \Pi - \cdots = \Pi_\nu - \Delta\Pi \tag{7.35b}$$

where $\Delta\Pi$ is given by Eq.(7.32b).

If we define the pseudo-inverse of $\bar{S}$ as

$$\bar{S}^{\#} \triangleq \sum_{i=1}^{(M_0+K_2)^2} \frac{\bar{w}_i \bar{w}_i^{\dagger}}{\bar{\mu}_i} \tag{7.36}$$

then we have

$$\bar{S}^{\#} \bar{S} = \bar{S} \bar{S}^{\#} = \Pi_s \tag{7.37}$$

Since orthonormal projectors are self-adjoint [67, 68], i.e.,

$$\Pi_s^{\dagger} = \Pi_s, \quad \Pi_\nu^{\dagger} = \Pi_\nu \tag{7.38a}$$

and idempotent, i.e.,

$$\Pi_s \Pi_s = \Pi_s, \quad \Pi_\nu \Pi_\nu = \Pi_\nu \tag{7.38b}$$

Replacing Π_s and Π_ν in Eq.(7.38b) by their expressions in Eqs.(7.35), we have

$$(\Pi_s + \delta\Pi + \cdots + \delta^n \Pi + \cdots)(\Pi_s + \delta\Pi + \cdots + \delta^n \Pi + \cdots) = \Pi_s + \delta\Pi + \cdots + \delta^n \Pi + \cdots \tag{7.39a}$$

$$(\Pi_\nu - \delta\Pi - \cdots - \delta^n \Pi - \cdots)(\Pi_\nu - \delta\Pi - \cdots - \delta^n \Pi - \cdots) = \Pi_\nu - \delta\Pi - \cdots - \delta^n \Pi - \cdots \tag{7.39b}$$

We equate the first-order terms on both side of Eqs.(7.39) to obtain the identities:

$$\delta\Pi = (\delta\Pi)\Pi_s + \Pi_s(\delta\Pi) \tag{7.40a}$$

$$\delta\Pi = (\delta\Pi)\Pi_\nu + \Pi_\nu(\delta\Pi) \tag{7.40b}$$

Premultiplying and postmultiplying Eq.(7.40a) by Π_s and likewise for Eq.(7.40b) by Π_ν, we obtain that

$$\Pi_s(\delta\Pi)\Pi_s = O \tag{7.41a}$$

$$\Pi_\nu(\delta\Pi)\Pi_\nu = O \tag{7.41b}$$

Since $\delta\Pi$ can be expressed as a sum of alternating projections, i.e.,

$$\begin{aligned}
\delta\Pi &= I(\delta\Pi)I \\
&= (\Pi_s + \Pi_\nu)(\delta\Pi)(\Pi_s + \Pi_\nu) \\
&= \Pi_s(\delta\Pi)\Pi_s + \Pi_s(\delta\Pi)\Pi_\nu + \Pi_\nu(\delta\Pi)\Pi_s + \Pi_\nu(\delta\Pi)\Pi_\nu \\
&= \Pi_s(\delta\Pi)\Pi_\nu + \Pi_\nu(\delta\Pi)\Pi_s
\end{aligned} \tag{7.42}$$

where the identities in Eqs.(7.41) are used.

We know

$$\hat{\Pi}_s\hat{\bar{R}}_x = \sum_{i=1}^{(M_0+K_2)^2} \hat{\bar{\mu}}_i\hat{\bar{w}}_i\hat{\bar{w}}_i^\dagger = \hat{\bar{R}}_x\hat{\Pi}_s \tag{7.43}$$

thus,

$$(\bar{R}_x + \Delta\bar{R}_x)\hat{\Pi}_s = \hat{\Pi}_s(\bar{R}_x + \Delta\bar{R}_x) \tag{7.44}$$

Under the condition when there is no noise, $\bar{R}_x$ can be approximated as

$$\bar{R}_x \simeq \sum_{i=1}^{(M_0+K_2)^2} \bar{\mu}_i\bar{w}_i\bar{w}_i^\dagger = \bar{S} \tag{7.45}$$

therefore,

$$(\bar{S} + \Delta\bar{R}_x)\hat{\Pi}_s = \hat{\Pi}_s(\bar{S} + \Delta\bar{R}_x) \tag{7.46}$$

with the first-order approximation of $\Delta\Pi$ in $\hat{\Pi}_s$, Eq.(7.46) can be rewritten as

$$(\bar{S} + \Delta\bar{R}_x)(\Pi_s + \delta\Pi) = (\Pi_s + \delta\Pi)(\bar{S} + \Delta\bar{R}_x) \tag{7.47}$$

The first-order terms of Eq.(7.47) yield an equation

$$\bar{S}(\delta\Pi) + \Delta\bar{R}_x\Pi_s = (\delta\Pi)\bar{S} + \Pi_s\Delta\bar{R}_x \tag{7.48}$$

Since $\bar{S}^{\#}\bar{S} = \bar{S}\bar{S}^{\#} = \Pi_s$, we premultiply Eq.(7.48) by $\bar{S}^{\#}$ and postmultiply by Π_ν and obtain

$$\Pi_s(\delta\Pi)\Pi_\nu = \bar{S}^{\#}\Delta\bar{R}_x\Pi_\nu \tag{7.49a}$$

Similarly, by premultiplying Eq.(7.48) with Π_ν and postmultiplying with $\bar{S}^{\#}$, we have

$$\Pi_\nu(\delta\Pi)\Pi_s = \Pi_\nu\Delta\bar{R}_x\bar{S}^{\#} \tag{7.49b}$$

Substituting Eqs.(7.49) into Eq.(7.42), we obtain

$$\delta\Pi = \bar{S}^{\#}\Delta\bar{R}_x\Pi_\nu + \Pi_\nu\Delta\bar{R}_x\bar{S}^{\#}$$

□

This theorem informs us that the expansion of $\Delta\Pi$ in Eqs.(7.33) is valid when there is no noise. However, if a spatially white noise is presented, i.e.,

$$R_x \simeq S + p_\nu^2 I_M$$

where S is the quasi signal subspace with dimensionality $M_0 + K_2$, then $\bar{R}_x$ will include both the covariance matrices of signal and noise. Then Eq.(7.44) becomes

$$\begin{aligned}&(S^* \otimes S + S^* \otimes p_\nu^2 I_M + p_\nu^2 I_M \otimes S + p_\nu^4 I_{M^2} + \Delta\bar{R}_x)\hat{\Pi}_s \\ = \;& \hat{\Pi}_s(S^* \otimes S + S^* \otimes p_\nu^2 I_M + p_\nu^2 I_M \otimes S + p_\nu^4 I_{M^2} + \Delta\bar{R}_x)\end{aligned} \tag{7.50}$$

where I_M and I_{M^2} denote the identity matrices with dimension $M \times M$ and $M^2 \times M^2$ respectively.

Since in general the equality that

$$(S^* \otimes p_\nu^2 I_M + p_\nu^2 I_M \otimes S)\hat{\Pi}_s = \hat{\Pi}_s(S^* \otimes p_\nu^2 I_M + p_\nu^2 I_M \otimes S) \tag{7.51}$$

cannot be established, therefore, the conclusion of Eq.(7.46) will not hold in general in the presence of noise. Under very high SNR, the expansion of Eqs.(7.33) can be regarded as a close approximation.

7.2.3 Asymptotic Performance of VEC-MUSIC

The criteria of VEC-MUSIC algorithm for the DOA estimation of distributed and point signals can be rewritten as

$$f_d(\bar{\phi},\sigma) = ||\bar{R}_k(\bar{\phi},\sigma)\hat{\bar{W}}_\nu||_F^2 = \mathrm{tr}\{\hat{\Pi}_\nu \bar{R}_k \bar{R}_k^\dagger\} \tag{7.52}$$

$$f_p(\phi_0) = \bar{d}_k^\dagger \hat{\Pi}_\nu \bar{d}_k = \mathrm{tr}\{\hat{\Pi}_\nu \bar{d}_k \bar{d}_k^\dagger\} \tag{7.53}$$

where $\bar{R}_k$ and $\bar{d}_k$ are respectively defined in Eqs.(7.21) and (7.22). To evaluate the performance using Eq.(7.52) for the kth distributed signal, we have the first derivatives

$$\dot{f}_{dk} = 2\mathrm{Re}[\mathrm{tr}\{\hat{\Pi}_\nu \bar{R}_k \dot{\bar{R}}_k^\dagger\}] \tag{7.54a}$$

$$f'_{dk} = 2\mathrm{Re}[\mathrm{tr}\{\hat{\Pi}_\nu \bar{R}_k \bar{R}_k'^\dagger\}] \tag{7.54b}$$

and second derivatives

$$\ddot{f}_{dk} = 2\mathrm{Re}[\mathrm{tr}\{\hat{\Pi}_\nu(\bar{R}_k \ddot{\bar{R}}_k^\dagger + \dot{\bar{R}}_k \dot{\bar{R}}_k^\dagger)\}] \tag{7.55a}$$

$$f''_{dk} = 2\mathrm{Re}[\mathrm{tr}\{\hat{\Pi}_\nu(\bar{R}_k \bar{R}_k''^\dagger + \bar{R}'_k \bar{R}_k'^\dagger)\}] \tag{7.55b}$$

$$\dot{f}'_{dk} = 2\mathrm{Re}[\mathrm{tr}\{\hat{\Pi}_\nu(\bar{R}_k \dot{\bar{R}}_k'^\dagger + \dot{\bar{R}}_k \bar{R}_k'^\dagger)\}] \tag{7.55c}$$

Analogous to the procedure described in Chapter 5.2.1, we have

$$\mathrm{E}\begin{bmatrix} \hat{\bar{\phi}}_k - \bar{\phi}_k \\ \hat{\sigma}_k - \sigma_k \end{bmatrix} = -\boldsymbol{H}_{dk} \begin{bmatrix} 2\mathrm{Re}[\mathrm{tr}\{\hat{\Pi}_\nu \bar{R}_k \dot{\bar{R}}_k^\dagger\}] \\ 2\mathrm{Re}[\mathrm{tr}\{\hat{\Pi}_\nu \bar{R}_k \bar{R}_k'^\dagger\}] \end{bmatrix} \tag{7.56a}$$

$$\mathrm{cov}\begin{bmatrix} \hat{\bar{\phi}}_k - \bar{\phi}_k \\ \hat{\sigma}_k - \sigma_k \end{bmatrix} = \boldsymbol{H}_{dk} \begin{bmatrix} \mathrm{cov}\{\dot{f}_{dk}, \dot{f}_{dk}\} & \mathrm{cov}\{\dot{f}_{dk}, f'_{dk}\} \\ \mathrm{cov}\{f'_{dk}, \dot{f}_{dk}\} & \mathrm{cov}\{f'_{dk}, f'_{dk}\} \end{bmatrix} \boldsymbol{H}_{dk}^\dagger \tag{7.56b}$$

where

$$H_{dk} = \begin{bmatrix} \ddot{f}_{dk} & \dot{f}'_{dk} \\ \dot{f}'_{dk} & f''_{dk} \end{bmatrix}^{-1} \tag{7.57}$$

Recalling from Eqs.(5.24) and (5.25) which give the covariance of $[\dot{f}_{IIk}, f'_{IIk}]^T$ with obvious substitution of P_ν and R_{uk} by Π_ν and $\bar{R}_k$, and using Theorems 7.1 and 7.2, we have

$$\begin{aligned}
\mathrm{cov}\{\dot{f}_{dk}, \dot{f}_{dk}\} &= 2\mathrm{Re}[\mathrm{vec}^\dagger(\dot{\bar{R}}_k \bar{R}_k^\dagger) C_{\bar{R}} \bar{\Upsilon} \frac{1}{N} \Xi_{\bar{R}} C_{\bar{R}}^T \mathrm{vec}^*(\dot{\bar{R}}_k \bar{R}_k^\dagger)] \\
&\quad + 2\mathrm{Re}[\mathrm{vec}^\dagger(\dot{\bar{R}}_k \bar{R}_k^\dagger) C_{\bar{R}} \frac{1}{N} \Xi_{\bar{R}} C_{\bar{R}}^\dagger \mathrm{vec}(\dot{\bar{R}}_k \bar{R}_k^\dagger)] \\
\mathrm{cov}\{\dot{f}'_{dk}, \dot{f}'_{dk}\} &= 2\mathrm{Re}[\mathrm{vec}^\dagger(\bar{R}'_k \bar{R}_k^\dagger) C_{\bar{R}} \bar{\Upsilon} \frac{1}{N} \Xi_{\bar{R}} C_{\bar{R}}^T \mathrm{vec}^*(\bar{R}'_k \bar{R}_k^\dagger)] \\
&\quad + 2\mathrm{Re}[\mathrm{vec}^\dagger(\bar{R}'_k \bar{R}_k^\dagger) C_{\bar{R}} \frac{1}{N} \Xi_{\bar{R}} C_{\bar{R}}^\dagger \mathrm{vec}(\bar{R}'_k \bar{R}_k^\dagger)] \\
\mathrm{cov}\{\dot{f}_{dk}, \dot{f}'_{dk}\} &= 2\mathrm{Re}[\mathrm{vec}^\dagger(\dot{\bar{R}}_k \bar{R}_k^\dagger) C_{\bar{R}} \bar{\Upsilon} \frac{1}{N} \Xi_{\bar{R}} C_{\bar{R}}^T \mathrm{vec}^*(\bar{R}'_k \bar{R}_k^\dagger)] \\
&\quad + 2\mathrm{Re}[\mathrm{vec}^\dagger(\dot{\bar{R}}_k \bar{R}_k^\dagger) C_{\bar{R}} \frac{1}{N} \Xi_{\bar{R}} C_{\bar{R}}^\dagger \mathrm{vec}(\bar{R}'_k \bar{R}_k^\dagger)] \\
&= \mathrm{cov}\{\dot{f}'_{dk}, \dot{f}_{dk}\}
\end{aligned} \tag{7.58}$$

and

$$C_{\bar{R}} \triangleq \bar{S}^{\#T} \otimes \Pi_\nu + \Pi_\nu^T \otimes \bar{S}^{\#} \tag{7.59}$$

Similarly, we can analyze the DOA estimation error of a point signal. We expand the first-order derivatives of $f_p(\bar{\phi}_0)$ in a Taylor series up to the first-order terms of the error and equate it to zero since the VEC-MUSIC algorithm seeks the minima of f_p, so that

$$\dot{f}_p(\hat{\phi}_{0k}) = \dot{f}_p(\phi_{0k}) + \ddot{f}_p(\phi_{0k})(\hat{\phi}_{0k} - \phi_{0k}) + O((\hat{\phi}_{0k} - \phi_{0k})^2) \tag{7.60}$$

where

$$\dot{f}_{pk} = 2\mathrm{Re}[\mathrm{tr}\{\hat{\Pi}_\nu \bar{d}_k \dot{\bar{d}}_k^\dagger\}] \tag{7.61}$$

$$\ddot{f}_{pk} = 2\mathrm{Re}[\mathrm{tr}\{\hat{\Pi}_\nu (\dot{\bar{d}}_k \dot{\bar{d}}_k^\dagger + \bar{d}_k \ddot{\bar{d}}_k^\dagger)\}] \tag{7.62}$$

Here, $\dot{\bar{d}}_k$ and $\ddot{\bar{d}}_k$ denote the first and second derivatives of $\bar{d}_k$ with respect to ϕ_0 at the point where $\phi_0 = \phi 0k$, respectively.

Parallel to Eq.(7.56), we can write

$$\hat{\phi}_{0k} - \phi_{0k} = -\dot{f}_{pk}/\ddot{f}_{pk} = -\frac{\mathrm{Re}[\mathrm{tr}\{\hat{\Pi}_\nu \bar{d}_k \dot{\bar{d}}_k^\dagger\}]}{\mathrm{Re}[\mathrm{tr}\{\Pi_\nu(\dot{\bar{d}}_k \dot{\bar{d}}_k^\dagger + \bar{d}_k \ddot{\bar{d}}_k^\dagger)\}]} \tag{7.63}$$

Therefore, using the property that $\mathrm{E}\{\hat{\Pi}_\nu\} = \Pi_\nu$ and that $\bar{d}_k \perp \Pi_\nu$,

$$\mathrm{E}\{\hat{\phi}_{0k} - \phi_{0k}\} = 0 \tag{7.64a}$$

and

$$\mathrm{cov}\{\hat{\phi}_{0k} - \phi_{0k}\} = \frac{1}{N}\frac{\mathrm{vec}^\dagger(\dot{\bar{d}}_k \bar{d}_k)(\bar{S}^{\#T} \otimes \Pi_\nu)\Xi_R(\bar{S}^{\#T} \otimes \Pi_\nu)^\dagger \mathrm{vec}(\dot{\bar{d}}_k \bar{d}_k)}{\mathrm{Re}[\mathrm{tr}\{\Pi_\nu(\dot{\bar{d}}_k \dot{\bar{d}}_k^\dagger + \bar{d}_k \ddot{\bar{d}}_k^\dagger)\}]} \tag{7.64b}$$

Eqs.(7.64a) and (7.64b) provide the expression of the bias and variance of the distributed and point signal estimates using VEC-MUSIC.

7.3 Choice of Noise Subspaces

Now, we scrutinize the eigen-subspaces of $\bar{R}_x$. We again consider K_1 distributed signals each inducing a dimensionality of m_{0k} for the signal subspace mixed with K_2 point signals. For an array of M sensors, let $\bar{W}$ denote the eigenvector matrix and $\bar{\Lambda}$ the eigenvalue matrix of $\bar{R}_x$, i.e.,

$$\bar{W} = [\bar{w}_1 \ \cdots \ \bar{w}_{M^2}] \tag{7.65}$$

$$\bar{\Lambda} = \mathrm{diag}(\bar{\lambda}_1, \cdots, \bar{\lambda}_{M^2}) \tag{7.66}$$

Since

$$\bar{R}_x = R_x^* \otimes R_x$$

then by carrying out an eigen-decomposition on both sides of the above equation, we have

$$\begin{aligned} \bar{W}\bar{\Lambda}\bar{W}^\dagger &= ([W_s|W_\nu]\Lambda[W_s|W_\nu]^\dagger)^* \otimes ([W_s|W_\nu]\Lambda[W_s|W_\nu]^\dagger) \\ &= ([W_s|W_\nu]^* \otimes [W_s|W_\nu])(\Lambda \otimes \Lambda)([W_s|W_\nu]^* \otimes [W_s|W_\nu])^\dagger \end{aligned}$$

$$= [\boldsymbol{W}_s^* \otimes \boldsymbol{W}_s | \boldsymbol{W}_s^* \otimes \boldsymbol{W}_\nu | \boldsymbol{W}_\nu^* \otimes \boldsymbol{W}_s | \boldsymbol{W}_\nu^* \otimes \boldsymbol{W}_\nu](\boldsymbol{\Lambda} \otimes \boldsymbol{\Lambda})$$
$$[\boldsymbol{W}_s^* \otimes \boldsymbol{W}_s | \boldsymbol{W}_s^* \otimes \boldsymbol{W}_\nu | \boldsymbol{W}_\nu^* \otimes \boldsymbol{W}_s | \boldsymbol{W}_\nu^* \otimes \boldsymbol{W}_\nu]^\dagger \tag{7.67}$$

where

$$\boldsymbol{R}_x = [\boldsymbol{W}_s | \boldsymbol{W}_\nu] \boldsymbol{\Lambda} [\boldsymbol{W}_s | \boldsymbol{W}_\nu]^\dagger \tag{7.68}$$

and $\boldsymbol{W}_s$ and $\boldsymbol{W}_\nu$ are signal and quasi-noise eigenvector matrices of $\boldsymbol{R}_x$ as defined in Eq.(4.7).

As described in Section 7.1, the extended pseudo noise subspace includes the extended quasi noise subspace. Now, we can choose either the extended quasi noise subspace or the extended pseudo noise subspace in forming the noise projector $\boldsymbol{\Pi}_\nu$ when implementing VEC-MUSIC algorithm. If we choose extended quasi noise subspace, since $\hat{\bar{\boldsymbol{R}}}_x$ asymptotically converges to $\bar{\boldsymbol{R}}_x$, then the estimated quasi noise eigenvector matrix $\bar{\boldsymbol{W}}_{\nu q}$, which is given by

$$\hat{\bar{\boldsymbol{W}}}_{\nu q} = [\hat{\bar{\boldsymbol{w}}}_{M^2-(M-M_0-K_2)^2+1} \cdots \hat{\bar{\boldsymbol{w}}}_{M^2}] \tag{7.69}$$

where $M_0 = \sum_{k=1}^{K_1} m_{0k}$ converges to the quasi noise eigenvectors of $\bar{\boldsymbol{R}}_x$, i.e.,

$$\lim_{N \to \infty} \hat{\bar{\boldsymbol{W}}}_{\nu q} \underset{a.s.}{=} \bar{\boldsymbol{W}}_{\nu q} = \boldsymbol{W}_\nu^* \otimes \boldsymbol{W}_\nu \tag{7.70}$$

By using Eq.(2.15), we obtain

$$\bar{\boldsymbol{R}}_k = \boldsymbol{R}_k^* \otimes \boldsymbol{R}_k \tag{7.71}$$

where $\bar{\boldsymbol{R}}_k$ is defined in Eq.(7.21). Thus

$$\lim_{N \to \infty} \|\bar{\boldsymbol{R}}_k \hat{\bar{\boldsymbol{W}}}_{\nu q}\|_F^2 \underset{a.s.}{=} \|(\boldsymbol{R}_k^* \otimes \boldsymbol{R}_k)(\boldsymbol{W}_\nu^* \otimes \boldsymbol{W}_\nu)\|_F^2$$
$$= \|\boldsymbol{R}_k \boldsymbol{W}_\nu\|_F^4 \tag{7.72}$$

We realize that VEC-MUSIC sprectrum for distributed signal source asymptotically tends to that of DISPARE II if the extended quasi noise subspace is employed.

On the other hand, from Eq.(7.67), we can see that there are several ways to choose noise subspace. We see that the extended pseudo noise subspace is spanned by columns of

the matrix

$$\bar{W}_{\nu p} = [W_s^* \otimes W_\nu | W_\nu^* \otimes W_s | W_\nu^* \otimes W_\nu] \tag{7.73}$$

of which the extended quasi noise subspace is only a part. Thus we can choose any combination of the eigenvectors in Eq.(7.73) to form the noise projector for VEC-MUSIC. We will choose the entire extended pseudo noise subspace in the following in comparison with the choice of only the extended quasi noise subspace.

We note that since $\bar{R}_x$ is Hermitian, its eigenvectors are orthogonal, i.e.,

$$(W_s^* \otimes W_s)^\dagger (W_s^* \otimes W_\nu) = (W_s^* \otimes W_s)^\dagger (W_\nu^* \otimes W_s) = (W_s^* \otimes W_s)^\dagger (W_\nu^* \otimes W_\nu) = O \tag{7.74a}$$

$$(W_s^* \otimes W_\nu)^\dagger (W_\nu^* \otimes W_s) = (W_s^* \otimes W_\nu)^\dagger (W_\nu^* \otimes W_\nu) = (W_\nu^* \otimes W_s)^\dagger (W_\nu^* \otimes W_\nu) = O \tag{7.74b}$$

Thus, if the "leakage" of signal power to the noise subspace is negligible, then

$$\Pi_\nu \bar{R}_k \simeq O \tag{7.75}$$

Eq.(7.75) is valid regardless of whether Π_ν is formed from using the eigenvectors of the extended quasi noise subspace or the eigenvectors of the entire extended pseudo noise subspace. Thus, using Eq.(7.75) in Eqs.(7.56a) and (7.64a), we can conclude that VEC-MUSIC is approximately unbiased. In addition, from Eqs.(7.56b) and (7.64b) the estimates are also approximately consistent.

7.4 Numerical Experiments

Effects of Distribution of Signals

So far, we have not assumed any particular distribution for the distributed signals. In the first two experiments, we examine via computer simulations the performance of VEC-MUSIC in the DOA estimation of distributed signals. In each experiment, a 4-sensor linear

array with uniform spacing of half-wavelength of the signal source is used. All sensors are omnidirectional and have unity gain.

Experiment 7.1

In this experiment, a distributed source with Gaussian power density distribution function with mean DOA $\bar{\phi} = -.5454 rad.$ and spread $\sigma = .3491 rad.$ on each side of mean DOA (corresponding to physical angle of arrival $[-10°, 6.7°]$) arrives at the array. The number of snapshots is taken to be 100. In the estimation of $\bar{\phi}$, we employ Eq.(7.16) to estimate the DOA. We use the extended quasi noise subspace and the extended pseudo noise subspace to form the corresponding noise projector. In order to examine the effects of the distribution of the signal, we assume that the source is uniformly distributed and compare the result of the DOA estimate with that under the correct assumption of Gaussian distribution. Fig.7.1a plots the root mean square (RMS) errors of mean DOA estimates of VEC-MUSIC under the assumptions of uniform distribution (plotted by solid line) and Gaussian distribution (plotted by dashed line) when the extended quasi noise subspace is used. As expected, the performance of VEC-MUSIC assuming the signal is uniformly distributed is very close to the performance with the correct assumption. This is because the criterion of VEC-MUSIC using the extended quasi noise subspace for the estimation of DOA of distributed signals is asymptotically equivalent to the criterion of DISPARE II, thus the VEC-MUSIC algorithm will have the same property of robustness against signal distributions if we make the assumption of uniform distribution for the distributed signals. Fig.7.1b plots the RMS errors of mean DOA estimates of VEC-MUSIC under the assumptions of uniform distribution (plotted by solid line) and the correct assumption of Gaussian distribution (plotted by dashed line) when the extended pseudo noise subspace is used. We observe that when using extended pseudo noise subspace VEC-MUSIC also yields similar performance under both assumptions of uniform and Gaussian distributions.

Experiment 7.2

In this experiment, a distributed source with uniform power density function of mean DOA $\bar{\phi} = -.5454rad.$ and spread $\sigma = .3491rad.$ on each side of mean DOA (corresponding to physical angle of arrival $[-10°, 6.7°]$) arrives at the array, The number of snapshots is taken to be 100. In the estimation of $\bar{\phi}$, we employ Eq.(7.16) to estimate the DOA. We again use both the extended pseudo noise subspace and the extended quasi noise subspace to form the corresponding noise projector. In order to examine the effects of distribution of the signal, we assume that the source is Gaussian distributed and compare the performance with that under the correct assumption of uniform distribution. Figs.7.2a and 7.2b show the corresponding RMS errors of mean DOA estimates of VEC-MUSIC under the assumptions of uniform distribution (plotted by solid line) and Gaussian distribution (plotted by dashed line) when either the extended quasi noise subspace or the extended pseudo noise subspace is used. We observe that, when assuming Gaussian distribution the performance of VEC-MUSIC deteriorate for either noise subspace being utilized, the failure of VEC-MUSIC when extended pseudo noise subspace being used is especially severe.

The results of Experiments 7.1 and 7.2 indicate that the VEC-MUSIC algorithm is robust to signal distribution by assuming uniform distribution. Equipped with the above observation, in the following experiments we assume that signals are all uniformly distributed. We will formulate the performance of VEC-MUSIC by assuming the distributed signals to be all uniformly distributed and then employ Eq.(7.23) to locate the estimates of the DOA using the expression of $\bar{R}_k$ corresponding to a uniformly distributed signal.

Further Experiments

To further demonstrate the performance of VEC-MUSIC, the following numerical experiments are conducted. Again, in each experiment, a 4-sensor linear array with uniform spacing of half-wavelength of the signal source is used. All sensors are omnidirectional and

have unity gain. Because of our interest in the ability of VEC-MUSIC to distinguish a point source from a distributed one, instead of concentrating on the accuracies of the estimates, we plot the spatial spectra of the algorithm for the respective searches as expressed by Eqs.(7.16) and (7.18). In each experiment the spread of the distributed signal is estimated by locating the peak of the 2-D search in Eq.(7.16) and the spatial spectrum for the distributed signal is the inverse of the profile of Eq.(7.16) along $\dot{\sigma}$. The spatial spectrum for the point source signal, on the other hand, is the inverse of the function in Eq.(7.18).

Experiment 7.3

In this experiment, a distributed source with uniform power density function of mean DOA $\bar{\phi} = -.5454rad.$ and spread $\sigma = .5236rad.$ on each side of mean DOA (corresponding to physical angle $[-10°, 9.9°]$) arrives at the array, and a point source impinges upon the array at $-1.5708rad.$ (corresponding to physical angle $-30°$). The powers of the point source and the distributed source are 6dB and 0dB respectively, and no noise. The number of snapshots is 100 and the results is plotted by one typical trial. From the discussion above, in the estimation of $\bar{\phi}$, we assume that the source is uniformly distributed and the spread σ is approximately known. In this experiment, we use the entire extended pseudo noise subspace to form the noise projector.

Fig.7.3a shows the spectrum of MUSIC assuming respectively the existence of 2 and 3 existing point signal sources. Fig.7.3b illustrates the spectrum of VEC-MUSIC for point source search and distributed source search. It can be observed that the MUSIC algorithm assuming 2 signals introduces biases to the estimates of both signals. Assuming 3 signals the MUSIC algorithm yields an accurate estimate for the point signal whereas the distributed signal is viewed upon as two point sources for which both peaks are far away from the true value of $\bar{\phi}$. On the other hand, in the search for a point source the VEC-MUSIC algorithm shows a dominant peak accurately situated at the true DOA of the point source, whereas

in the search for a distributed signal, a dominant peak is also located very close to the true $\bar{\phi}$. The simulation results also indicate that VEC-MUSIC can distinguish between a point source and a distributed source, since in each of the searches, the dominant peak corresponds to only the signal for which the algorithm seeks.

Experiment 7.4

In this experiment, the scenario is the same as the previous experiment except that the point source is within the uniformly distributed source. The point source arrives at the array at an angle $-.9185rad.$ which is within the spread of the distributed signal ($\bar{\phi} = -.5454rad.$, $\sigma = .5236rad.$). Again the number of snapshots is 100 and the results are plotted by one typical trial. The noise projector is formed by using the eigenvectors of the entire extended pseudo noise subspace.

Fig.7.4a shows the spectra of MUSIC assuming the existence of 2 and 3 existing point signal sources respectively denoted by dashed and solid line. Fig.7.4b plots the spectra of VEC-MUSIC for the point and distributed source searches, where solid and dashed lines denote the cases for point and distributed sources, respectively. Again, we can observe that the MUSIC algorithm fails to locate the distributed source accurately (introducing unacceptably large estimation error) in both cases of assuming 2 and 3 signals. However, the VEC-MUSIC locates both the point and distributed signals with high degrees of accuracy. Again, in either case of search, the VEC-MUSIC algorithm only yields a dominant peak which corresponds to only the signal for which is sought. We can see that VEC-MUSIC manifests its ability to distinguish point source from distributed source even when the point source merges into the distributed one.

Experiment 7.5

In this experiment, again, a distributed source with uniform power density function of

mean DOA $\bar{\phi} = -.5454rad.$ and spread $\sigma = .5236rad.$ on each side of the mean DOA (corresponding to physical angle $[-10°, 9.9°]$) arrives at the array, and a point source impinges upon the array at $-1.5708rad.$ (corresponding to physical angle $-30°$). The powers of the point source and the distributed source are 0dB and 3dB, respectively; and noise power is -26dB. The number of snapshots is taken to be 200. We employ VEC-MUSIC to estimate the DOA of the two signals and we utilize in one case the extended quasi noise subspace and in another the extended pseudo noise subspace for the formation of the noise projector.

Fig.7.5a illustrates the spectrum of VEC-MUSIC for the point and the distributed source search when extended quasi noise subspace is employed. We observe that, under this scenario, VEC-MUSIC for the distributed source search recognizes the signal being sought for but with a large error. VEC-MUSIC for point source search fails having mistakingly identified the distributed signal as also a point source and locating it with a even larger error. Fig.7.5b shows the spectra of VEC-MUSIC for the point and the distributed source search when the extended pseudo noise subspace is used. We observe that in each of the searches, a dominant peak is obtained corresponding to the signal being sought after. The unwanted signal in each of the searches is reduced to some minor humps in the spectrum. Furthermore, the dominant, peak in each case represents an accurate location of the desired signal. This experiment demonstrates an important advantage of employing extended pseudo noise subspace.

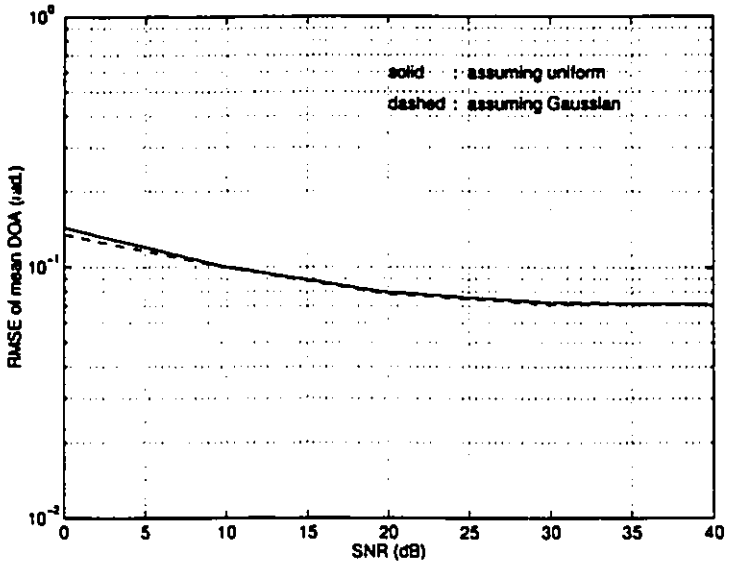
We notice that when the extended quasi noise subspace is used, VEC-MUSIC is rather vulnerable to background noise. By using the extended pseudo noise subspace noise effect can be reduced. This may be due to the characteristic of the extended pseudo noise subspace which contains the extended quasi noise subspace as well as the mixture of it with signal components onto which the projection can withhold other unwanted peaks in VEC-MUSIC spectrum. In other words, the use of the extended quasi noise subspace which has a smaller

dimensionality may allow spurious peaks to occur. This is because much of the signal energy is contained in the extended pseudo noise subspace, any vector there which is not in the extended quasi noise subspace will produce a peak when projected onto the extended quasi noise subspace.

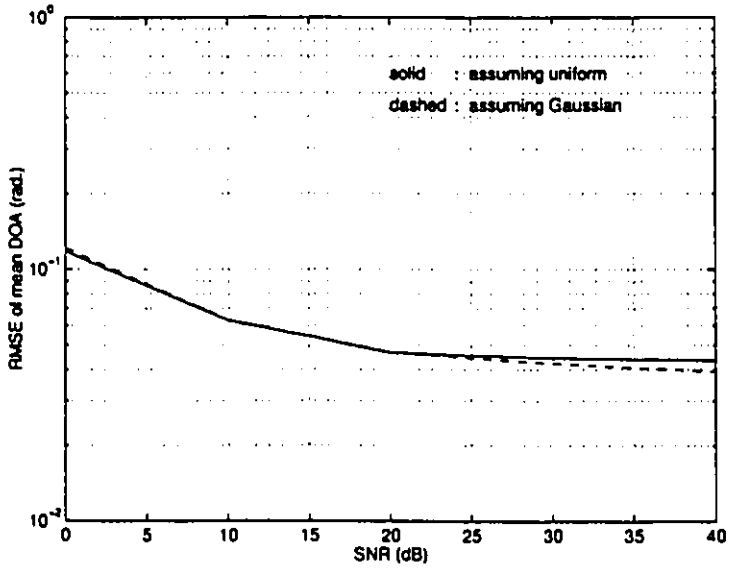
Experiment 7.6

The final experiment in this chapter is to examine the accuracy of the VEC-MUSIC estimates under different SNR via computer simulations and compare the simulated results to the theoretical evaluation of the performance of VEC-MUSIC. A single distributed source with uniform distribution of mean DOA $\bar{\phi} = -.5454rad.$ and spread $\sigma = .3491rad.$ on each side of mean DOA (corresponding to physical angle $[-10°, 6.7°]$) arrives at the array. The number of snapshots is 200. We apply the criterion given by Eq.(7.16) to estimate the DOA of the distributed source by jointly locating $\hat{\bar{\phi}}$ and $\hat{\sigma}$. We also employ Eqs.(7.56a) and (7.56b) to evaluate the RMS errors of the estimates of DOA for the distributed source. In both the computer simulation and the theoretical evaluation the extended pseudo noise subspace is used to form the noise projector. Since we are only interested in the mean DOA, Fig.7.6 plots the simulated RMS error of mean DOA estimate of VEC-MUSIC (crossed line) and the theoretical RMS error of mean DOA estimate of VEC-MUSIC (solid line) for the range of SNR being $0dB \sim 40dB$. We observe that the simulation result and the theoretical evaluation are in close agreement when SNR is above $20dB$; however, when SNR is under $20dB$, the discrepancy between the simulation result and the theoretical evaluation occurs. This is not surprising because the analysis of the asymptotic performance of VEC-MUSIC given in Section 7.2 is accurate under the condition when there is no noise. When noise occurs, the validity of $(\bar{S} + \Delta\bar{R}_x)\hat{\Pi}_s = \hat{\Pi}_s(\bar{S} + \Delta\bar{R}_x)$ is violated and the expression of Eqs.(7.33) in Theorem 7.2 no longer holds. However, this experiment suggests that when SNR is high enough (for instance, in this case SNR $\geq 20dB$), the theoretical evaluation

given by Eqs.(7.56) is close to the simulated performance of VEC-MUSIC.

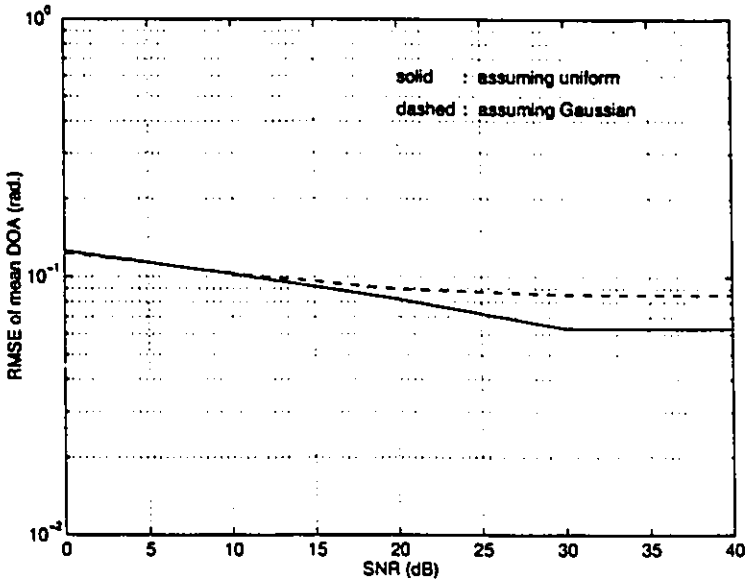


(a) The extended quasi noise subspace

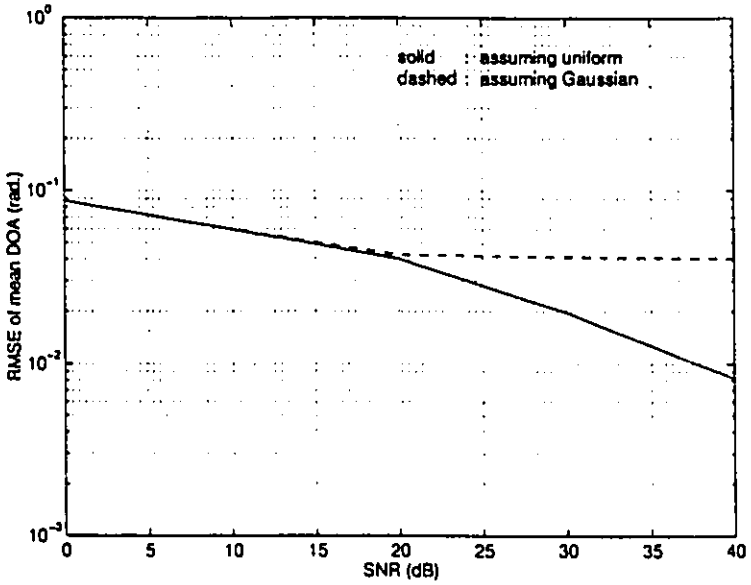


(b) The extended pseudo noise subspace

Figure 7.1: The estimation errors of mean DOA of a Gaussian distributed signal

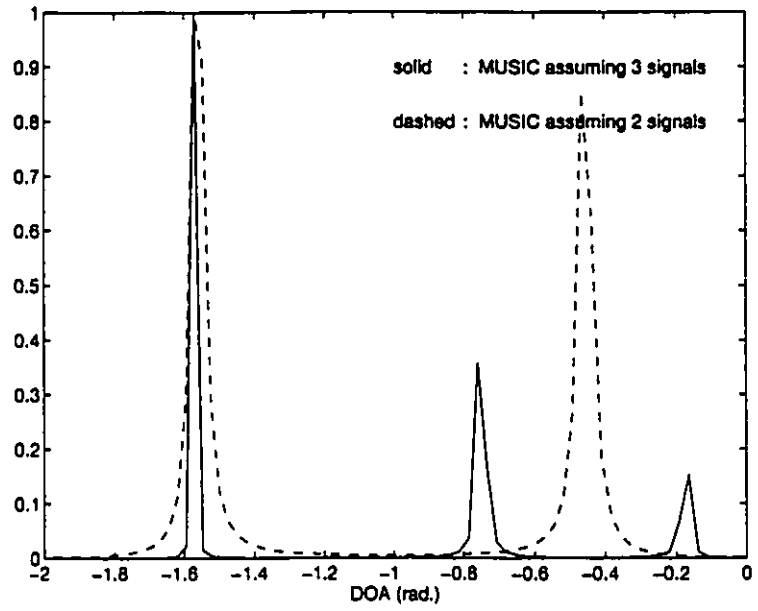


(a) The extended quasi noise subspace

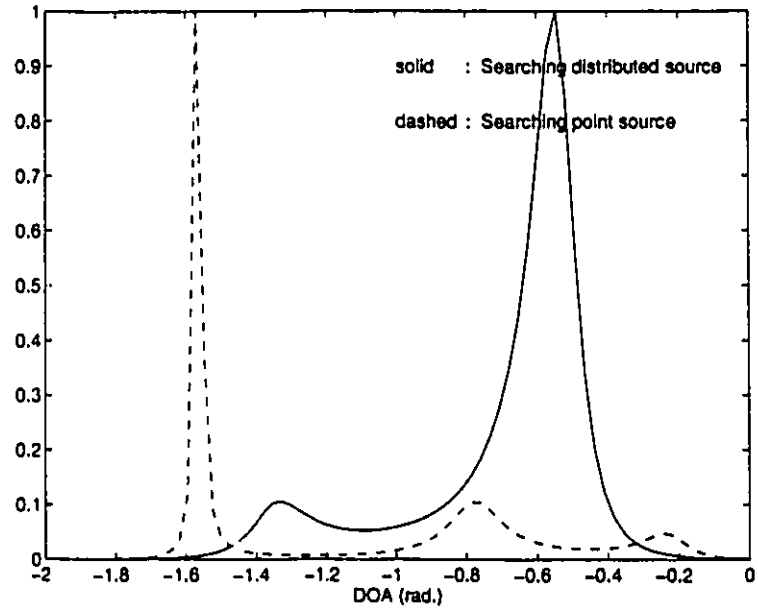


(b) The extended pseudo noise subspace

Figure 7.2: The estimation errors of mean DOA of a uniformly distributed signal

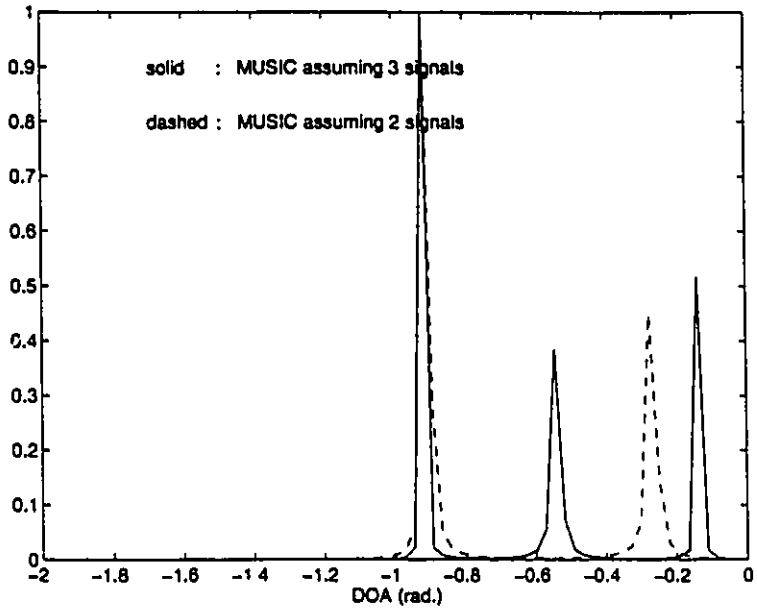


(a) Spectra of MUSIC

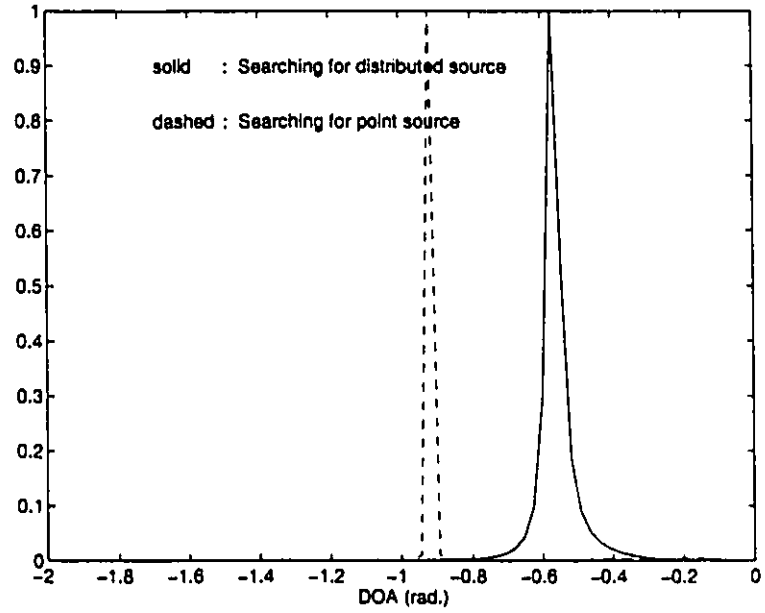


(b) Spectra of VEC-MUSIC

Figure 7.3: MUSIC and VEC-MUSIC spectra for 2 well-separated signals

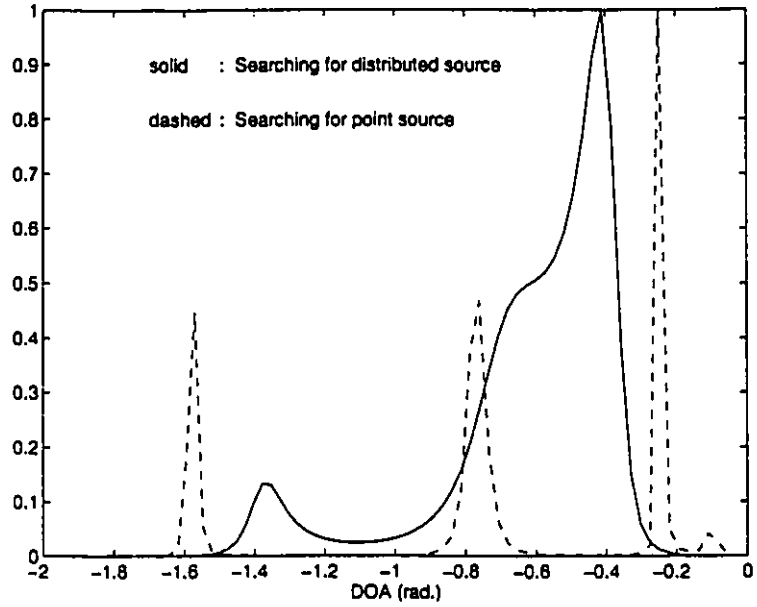


(a) Spectra of MUSIC

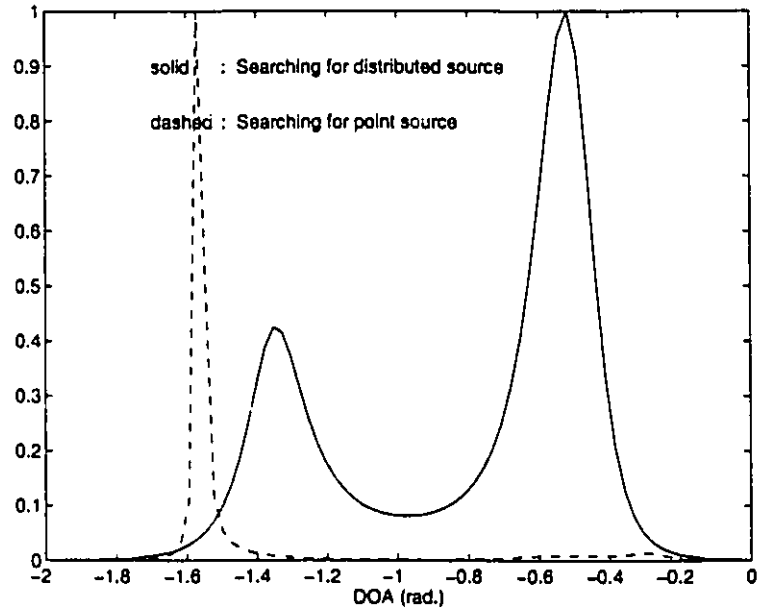


(b) Spectra of VEC-MUSIC

Figure 7.4: MUSIC and VEC-MUSIC spectra for 2 merged signals



(a) Extended quasi noise subspace



(b) Extended pseudo noise subspace

Figure 7.5: VEC-MUSIC spectra for two different noise subspaces

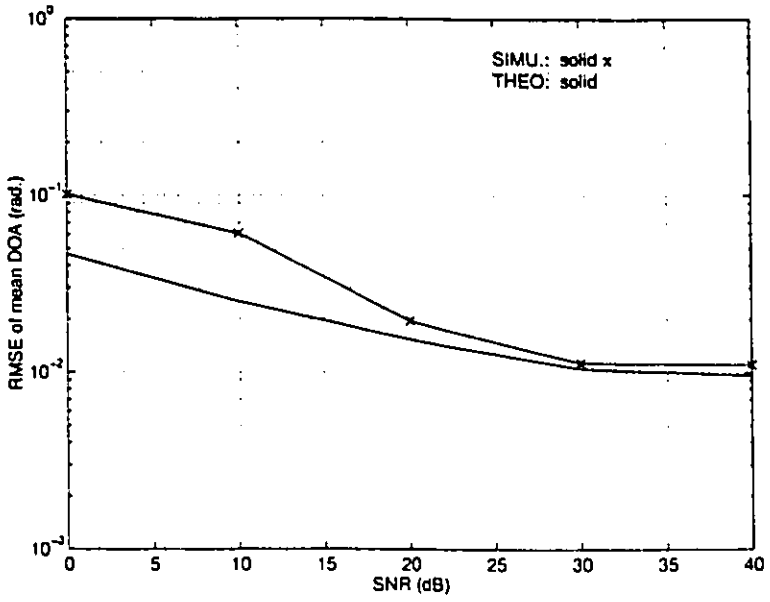


Figure 7.6: Theoretical evaluation and simulation of RMSE of mean DOA of a uniformly distributed signal with $(\bar{\phi}, \sigma) = (-.5454, .3491)$ arriving at a uniform linear array of 4 sensors; 200 snapshots are used for comparison.

Chapter 8

Conclusions

8.1 Summary

In this thesis, we have examined the DOA estimation problem of high resolution sensor array processing in the presence of spatially distributed sources.

The main assumption made is that a spatially distributed source consists of a large amount of statistically independent discrete sources (point sources) closely clustered together. This assumption is often (and justifiably) used to simplify calculations in many applications and is also supported by many experimental results. Based on this assumption, we have developed a mathematical model of spatially distributed sources for sensor array processing (and conceived point sources as a special case of distributed sources), particularly, in the formulation of the correlation matrix of the signals received by sensor array. The mathematical model has been validated by experiments conducted in Communications Research Center, Ottawa [1].

We have found that although the correlation matrix of spatially distributed signal is

[1] This statement refers to a private conversation with Dr. Robert W. Jenkins in CRC, Ottawa on March 18, 1994.

theoretically full-rank, which renders conventional high resolution array processing methods inapplicable in this case, the dimensionality of the signal subspace can be approximated by a number usually much smaller than the number of sensors; the remaining dimension of the matrix can be regarded as in the quasi noise subspace. Furthermore, by examining various spatial distributions of the sources, we arrived at the conclusion that without *a priori* knowledge of the distribution, the assumption of a spatially uniform distribution is robust to the variation of distributions. We then developed two methods of DISPARE I and DISPARE II based on different weighting projections of the eigenvectors of the matrix onto the estimated quasi noise subspace in a white ambient noise environment. Computer simulation results show that both of the methods are effective compared to MUSIC. The theoretical analysis of the asymptotical performance of the two methods are also presented. We further inferred that due to the existence of erroneous dominant peaks in the profile of the criterion of DISPARE I, DISPARE II is a more robust method for the estimation of the DOA of spatially distributed signals.

We have also explored higher-order statistics of spatially distributed signals in sensor array processing, especially the fourth-order cumulant matrix of the received signals by the sensor array. Due to the attributes of higher-order cumulants, Gaussian noise vanishes. This enables us to tackle the problem of DOA estimation of non-Gaussian signals in various noise environments, no matter whether it is white noise or coloured noise with known or unknown correlations as long as it is Gaussian. A DOA estimation method has been proposed and computer simulations have shown the merit of this method in suppressing Gaussian noise. Asymptotical performance of the method is also presented.

VEC-MUSIC belongs to another category. Statistically, it only involves second-order statistics; while in its extended data space, spatially distributed and point sources have their own geometric subspaces, which provides the distinction between them.

8.2 Comparison of the Various Methods

In this thesis, several DOA estimation methods have been proposed, each has its own merits. For different distribution of signals, all methods present robustness when assuming that the signals are uniformly distributed. For various noise environments, as long as noise is Gaussian, under moderate SNR the DOA estimation method based on the fourth-order cumulants of the received data evidences the advantage of suppressing Gaussian noise; under high SNR the DISPARE method yields better performance than the DOA estimation algorithm based on the fourth-order cumulants even when noise is spatially correlated. In aiming to distinguish point sources from distributed sources, the VEC-MUSIC method stands out against others. It also offers a unique feature of different choices of noise subspace, although the use of the entire extended pseudo noise subspace has been demonstrated to be less susceptible to spurious peaks when searching for the two different kinds of signals. However, VEC-MUSIC is highly sensitive to background noise. As for computational load, DISPARE requires $O(N \times M^2)$ flops [2], while the other two methods need $O(N \times M^4)$ flops, where N is the number of samples and M is the number of sensors. Therefore, DISPARE algorithms are much simpler than the other two algorithms, and the algorithm based on higher-order statistics bears the heaviest computation because much larger number of samples is required to estimate higher-order statistics than second-order statistics for the same estimation accuracy. In general, DISPARE is more applicable in the sense that (1) assuming signal sources are spatially in uniform distribution, DISPARE yields almost as good performance as that under the right assumption of distribution; (2) under high SNR, even under unknown correlated background noise, DISPARE yields fairly good estimation; (3) compared to the other two methods, DISPARE is computationally simpler.

[2] One "flop" is one floating point arithmetic operation.

8.3 Future Work

Future work in several aspects are suggested as follows. Firstly, in implementing the DOA estimation methods proposed in this thesis, a two-dimensional exhaustive search is employed which is computationally intensive. Instead, we can exploit the relaxation/alternating minimization type of technique which is conceptually simple for multidimensional optimization. The technique is iterative; at each iteration a minimization is performed with respect to a single parameter while all the other parameters are held fixed, thus rendering the search to one-dimensional. For example, to minimize the criterion of DISPARE II given by Eq.(4.19) following steps can be used:

1. Fix σ_k to the initial estimate *a priori* given and minimize Eq.(4.19) with respect to $\bar{\phi}_k$,
2. Fix $\bar{\phi}_k$ to the most recent estimate available and minimize Eq.(4.19) with respect to σ_k,
3. Iterate the above steps until converges.

This will likely lighten the computation requirement. Secondly, at present, DISPARE is a projection of signal covariance matrix $\boldsymbol{R}_k(\bar{\phi}, \sigma)$ onto quasi noise subspace with equal weighting. How to choose optimal weights is worth to seek. Thirdly, in exploiting the higher-order statistics, the cumulant matrix defined in Chapter 6 is used. We can further examine the other alternatives of using the higher-order cumulants. For example, the *diagonal slice* and the *contracted quadricovariance* [97] can be applied. We can compare the higher-order statistical information hidden in these matrices, and more importantly, the latter two are of reduced dimension thus lighten computational load. Finally, throughout this thesis we have assumed that the distributed signals are on the same plane as the array, the extension to a three-dimensional space is straightforward. We also can extend the study to arrays

with arbitrary geometry so that the DOA estimation may involve two angles, such as the simultaneous estimation of the angles of elevation and azimuth.

Inspired by radio communications via ionospheric scattering in the Arctic area, we initiate the study of distributed source localization and perceive the demands of this research subject in many other applications such as radar system performing low-angle tracking in a rough sea environment and mobile communications in an urban area. It will be of great interest to apply the mechanism presented in this thesis to these engineering problems.

In this thesis, a study on high resolution DOA estimation for spatially distributed sources is presented and some new methods in array processing for source localization are developed. The subject is still in its infancy. What is presented in this thesis is no more than the initial few steps of a long journey. Looking ahead, there still exist many exciting problems surrounding the topic.

Appendix A

Prolate Spheroidal Wave Functions

A.1 Prolate Spheroidal Wave Functions of Zero Order

The *prolate spheroidal wave functions* (PSWF) of zero order $\vartheta_m(c,x)$, $m = 1,2,\cdots$, are the continuous solutions of the following integral equation

$$\eta_m(c)\vartheta_m(c,x) = \int_{-1}^{1} \frac{\sin c(x-y)}{\pi(x-y)}\vartheta_m(c,y)dy \tag{A.1}$$

where $-\infty < x < \infty$, $\eta_1(c) < \eta_2(c) < \cdots$.

The PSWF $\vartheta_m(c,x)$ are doubly orthogonal, i.e.,

$$\int_{-\infty}^{\infty} \vartheta_m(c,x)\vartheta_n(c,x) = \delta_{mn} \tag{A.2}$$

$$\int_{-1}^{1} \vartheta_m(c,x)\vartheta_n(c,x) = \eta_m(c)\delta_{mn} \tag{A.3}$$

for $m,n = 1,2,\cdots$.

Numerical values of $\eta_m(c)$ for various values of c are given in [100]. It is observed that for a fix value of c, $\eta_m(c)$ fall off to zero rapidly with increasing m once m has exceeded c.

A small value of $\eta_m(c)$ implies that $\vartheta_m(c,x)$ will have most of its energy outside the interval $(-1,\ 1)$ whereas a large value of $\eta_m(c)$ implies that $\vartheta_m(c,x)$ will largely be concentrated in $(-1,\ 1)$.

A.2 Discrete Prolate Spheroidal Sequences

The *discrete prolate spheroidal sequences* (DPSS) $\{\zeta_m^{(i)}(\sigma,M)\}$ is defined as the real solution to the system of equations

$$\eta_m(\sigma,M)\zeta_m^{(i)}(\sigma,M)=\sum_{j=1}^{M}\frac{\sin\sigma(i-j)}{\pi(i-j)}\zeta_m^{(i)}(\sigma,M),\quad i=\pm1,\pm2,\cdots \tag{A.4}$$

for each $m=1,2,\cdots,M$, where $\zeta_m^{(i)}$ denotes the ith element of sequence ζ_m. It is understood here, of course, that when $i=j$ the expression $\frac{\sin\sigma(i-j)}{\pi(i-j)}$ has the value $\frac{\sigma}{\pi}$.

We see here $\eta_m(\sigma,M),\quad m=1,2,\cdots,M$ are the eigenvalues of an $M\times M$ matrix $\boldsymbol{R}(\sigma,M)$ with the ijth entry

$$\boldsymbol{R}(\sigma,M)[i;j]=\frac{\sin\sigma(i-j)}{\pi(i-j)},\quad i,j=1,2,\cdots,M \tag{A.5}$$

and M vectors obtained by index-limiting the DSPP $\{\zeta_m^{(i)}(\sigma,M)\},\ m=1,2,\cdots,M$, to the index set $(1,M)$ are the M corresponding eigenvectors of the matrix $\boldsymbol{R}(\sigma,M)$.

The DPSS's are doubly orthogonal, i.e.,

$$\sum_{i=1}^{M}\zeta_m^{(i)}(\sigma,M)\zeta_n^{(i)}(\sigma,M)=\eta_m(\sigma,M)\delta_{mn} \tag{A.6a}$$

$$\sum_{i=-\infty}^{\infty}\zeta_m^{(i)}(\sigma,M)\zeta_n^{(i)}(\sigma,M)=\delta_{mn} \tag{A.3b}$$

for $m,n=1,2,\cdots,M$. The DPSS $\{\zeta_m^{(i)}(\sigma,M)\},\ i=\pm1,\pm2,\cdots$, are band limited to $(-\sigma,\sigma)$ in frequency domain, the energy concentration of index-limited DPSS $\{\zeta_m^{(i)}(\sigma,M)\}$,$m=1,2,\cdots,M$, is given by

$$\eta_m(\sigma,M)=\frac{\sum_{i=1}^{M}\mid\zeta_m^{(i)}(\sigma,M)\mid^2}{\sum_{i=-\infty}^{\infty}\mid\zeta_m^{(i)}(\sigma,M)\mid^2}\quad m=1,\cdots,M \tag{A.4}$$

which is the ratio of the energy of index-limited DPSS to the energy of entire DPSS.

For fixed σ, $\eta_1(\sigma, M) \geq \eta_2(\sigma, M) \geq \cdots \geq \eta_M(\sigma, M)$; and when the length M of the index-limited DPSS increases, more eigenvalues $\eta_m(\sigma, M)$ have significant value because more energy is involved. It is also known [82] that as $M \to \infty$

$$\begin{aligned} \eta_m(\sigma, M) &\to 1 \quad \text{if} \quad k = \tfrac{\sigma M}{\pi}(1-\epsilon) \\ \eta_m(\sigma, M) &\to 0 \quad \text{if} \quad k = \tfrac{\sigma M}{\pi}(1+\epsilon) \end{aligned} \tag{A.5}$$

and this is true for any ϵ satisfying $0 < \epsilon < 1$. This fact states that for $M \to \infty$ the energy concentration is in the first $\frac{\sigma M}{\pi}$ eigenvalues.

In the next section, we will examine the energy concentration of the DPSS for not very large M by relating the DPSS to the PSWF.

A.3 Relationship between PSWF and DPSS

For not very large bandwidth σ and large M, let $c > 0$ and y is a given real number, if $M = [\frac{2c}{\sigma}]$ and $i = [\frac{M}{2}(1+y)]$ where $[x]$ denotes the nearest integer to x, then

$$\eta_m(\sigma, M) \simeq \eta_m(c) \tag{A.6}$$

$$\sqrt{\frac{M}{2}}\zeta_m^{(i)}(\sigma, M) \simeq \frac{\pm 1}{\sqrt{\eta_m(c)}}\vartheta_m(c, y) \tag{A.7}$$

In Eq.(A.7) when m is even the plus sign is taken when $\int_{-1}^{1} \vartheta_m(c, x)dx > 0$, when m is odd the plus sign is taken when $\int_{-1}^{1} x\vartheta_m(c, x)dx > 0$; otherwise, the negative sign is to be used.

From the observation of the values of $\eta_m(c)$ and the above relationship between PSWF and DPSS, we infer that for not very large M, after $m > \frac{\sigma M}{2}$, $\eta_m(\sigma, M)$ drop to zero rapidly such that the energy concentration is in the first $\frac{\sigma M}{2}$ eigenvalues.

Appendix B

Proofs of Theorems

B.1 Proof of Theorem 5.1

As we know, given N sample of observations $\boldsymbol{x}(n)$, $n = 1, \cdots, N$, the covariance matrix $\boldsymbol{R}_x$ is estimated by

$$\begin{aligned} N\hat{\boldsymbol{R}}_x &= \sum_{n=1}^{N}[\boldsymbol{x}(n) - \bar{\boldsymbol{x}}][\boldsymbol{x}(n) - \bar{\boldsymbol{x}}]^\dagger \\ &= \sum_{n=1}^{N}[\boldsymbol{x}(n) - \boldsymbol{\mu}][\boldsymbol{x}(n) - \boldsymbol{\mu}]^\dagger - N(\bar{\boldsymbol{x}} - \boldsymbol{\mu})(\bar{\boldsymbol{x}} - \boldsymbol{\mu})^\dagger \end{aligned} \tag{B.1}$$

where $\boldsymbol{\mu} = \mathrm{E}\{\boldsymbol{x}(n)\}$ and $\bar{\boldsymbol{x}} = \frac{1}{N}\sum_{n=1}^{N}\boldsymbol{x}(n)$.

Let

$$\boldsymbol{K}(n) = [\boldsymbol{x}(n) - \boldsymbol{\mu}][\boldsymbol{x}(n) - \boldsymbol{\mu}]^\dagger \tag{B.2a}$$

$$\boldsymbol{J}_N = (\bar{\boldsymbol{x}} - \boldsymbol{\mu})(\bar{\boldsymbol{x}} - \boldsymbol{\mu})^\dagger \tag{B.2b}$$

Therefore, we have

$$N\,\mathrm{vec}(\hat{\boldsymbol{R}}_x) = \sum_{n=1}^{N}\mathrm{vec}(\boldsymbol{K}(n)) - N\,\mathrm{vec}(\boldsymbol{J}_N) \tag{B.3}$$

We note that $\mathrm{vec}(\boldsymbol{K}(n))$ are i.i.d. random vectors with

$$\mathrm{E}\{\mathrm{vec}(\boldsymbol{K}(n))\} = \mathrm{vec}(\mathrm{E}\{\boldsymbol{K}(n)\}) = \mathrm{vec}(\mathrm{E}\{[\boldsymbol{x}(n) - \boldsymbol{\mu}][\boldsymbol{x}(n) - \boldsymbol{\mu}]^\dagger\}) = \mathrm{vec}(\boldsymbol{R}_x) \tag{B.4}$$

By the Central Limit Theorem, we have

$$N^{-\frac{1}{2}} \sum_{n=1}^{N} [\mathrm{vec}(\boldsymbol{K}(n)) - \mathrm{vec}(\boldsymbol{R}_x)] \rightarrow \mathcal{N}(\boldsymbol{O}, \mathrm{cov}\{\mathrm{vec}(\boldsymbol{K}(n))\}) \tag{B.5}$$

as $N \rightarrow \infty$. Here, $\mathcal{N}(\boldsymbol{O}, \mathrm{cov}\{\mathrm{vec}(\boldsymbol{K}(n))\})$ denotes a multivariate complex Gaussian distribution with zero mean and covariance $\mathrm{cov}\{\mathrm{vec}(\boldsymbol{K}(n))\}$.

From Eq.(B.3), we have

$$N^{\frac{1}{2}} \mathrm{vec}(\hat{\boldsymbol{R}}_x) = N^{-\frac{1}{2}} \sum_{n=1}^{N} \mathrm{vec}(\boldsymbol{K}(n)) - N^{\frac{1}{2}} \mathrm{vec}(\boldsymbol{J}_N) \tag{B.6}$$

Again, by using Central Limit Theorem, we have

$$N^{-\frac{1}{2}} \sum_{n=1}^{N} [\boldsymbol{x}(n) - \boldsymbol{\mu}] \rightarrow \mathcal{N}(\boldsymbol{O}, \mathrm{cov}\{\boldsymbol{x}(n)\}) = \mathcal{N}(\boldsymbol{O}, \boldsymbol{R}_x) \tag{B.7}$$

as $N \rightarrow \infty$,

i.e.,

$$N^{\frac{1}{2}} (\bar{\boldsymbol{x}} - \boldsymbol{\mu}) \rightarrow \mathcal{N}(\boldsymbol{O}, \boldsymbol{R}_x) \tag{B.8}$$

as $N \rightarrow \infty$.

Therefore,

$$N^{\frac{1}{2}} (\bar{\boldsymbol{x}} - \boldsymbol{\mu}) \rightarrow \boldsymbol{O}, \text{ as } N \rightarrow \infty \text{ in probability}^1 \tag{B.9}$$

so that

$$N^{\frac{1}{2}} (\bar{\boldsymbol{x}} - \boldsymbol{\mu})(\bar{\boldsymbol{x}} - \boldsymbol{\mu})^{\dagger} \rightarrow \boldsymbol{O}, \text{ in probability as } N \rightarrow \infty \tag{B.10}$$

As a consequence of Eqs.(B.6) and (B.10), we have

$$\begin{aligned} N^{\frac{1}{2}} [\mathrm{vec}(\hat{\boldsymbol{R}}_x) - \mathrm{vec}(\boldsymbol{R}_x)] &= N^{-\frac{1}{2}} \sum_{n=1}^{N} [\mathrm{vec}(\boldsymbol{K}(n)) - \mathrm{vec}(\boldsymbol{R}_x)] - N^{\frac{1}{2}} \mathrm{vec}(\boldsymbol{J}_N) \\ &\rightarrow \mathcal{N}(\boldsymbol{O}, \mathrm{cov}\{\mathrm{vec}(\boldsymbol{K}(n))\}) \end{aligned} \tag{B.11}$$

[1] This means each component of vector $\bar{\boldsymbol{x}} - \boldsymbol{\mu}$, say, v_m, $m = 1, 2, \cdots, M$, $\forall\, \epsilon > 0$, $\lim_{N \rightarrow \infty} P\{|N^{\frac{1}{2}} v_m| > \epsilon\} = 0$. A similar definition holds for random matrices.

as $N \rightarrow \infty$,

such that [2]

$$N^{\frac{1}{2}}[\text{vec}(\hat{\boldsymbol{R}}_x) - \text{vec}(\boldsymbol{R}_x)] = N^{\frac{1}{2}}\text{vec}(\Delta \boldsymbol{R}_x) \rightarrow \mathcal{N}(\boldsymbol{O}, \text{cov}\{\text{vec}(\boldsymbol{K}(n))\}) \tag{B.12}$$

as $N \rightarrow \infty$.

Now, we evaluate the asymptotic covariance matrix $\text{cov}\{\text{vec}(\boldsymbol{K}(n))\}$.

$$\begin{aligned}
& \text{cov}\{\text{vec}(\boldsymbol{K}(n))\} \\
= \; & \text{cov}\{\text{vec}([\boldsymbol{x}(n) - \boldsymbol{\mu}][\boldsymbol{x}(n) - \boldsymbol{\mu}]^\dagger)\} \\
= \; & \text{cov}\{[\boldsymbol{x}(n) - \boldsymbol{\mu}]^* \otimes [\boldsymbol{x}(n) - \boldsymbol{\mu}]\} \\
= \; & \text{E}\{[(\boldsymbol{x}(n) - \boldsymbol{\mu})^* \otimes (\boldsymbol{x}(n) - \boldsymbol{\mu})][(\boldsymbol{x}(n) - \boldsymbol{\mu})^* \otimes (\boldsymbol{x}(n) - \boldsymbol{\mu})]^\dagger\} \\
& -\text{E}\{(\boldsymbol{x}(n) - \boldsymbol{\mu})^* \otimes (\boldsymbol{x}(n) - \boldsymbol{\mu})\}\text{E}\{(\boldsymbol{x}(n) - \boldsymbol{\mu})^* \otimes (\boldsymbol{x}(n) - \boldsymbol{\mu})\}^\dagger
\end{aligned} \tag{B.13}$$

In general, given an underlying distribution for $\boldsymbol{x}(n)$, it is rather tedious (in terms of the algebraic manipulation involved) to find the elements of the asymptotic covariance matrix, since this involves finding all the fourth-order moments of the distribution. However, the calculation are fairly straightforward when sampling from a complex circular Gaussian distribution with mean $\boldsymbol{\mu}$ and covariance $\boldsymbol{R}_x$ denoted as $\mathcal{CN}(\boldsymbol{\mu}, \boldsymbol{R}_x)$. In this case the elements of the asymptotic covariance matrix are given by

$$\begin{aligned}
& \text{E}\{(x_k(n) - \mu_k)^*(x_l(n) - \mu_l)(x_s(n) - \mu_s)(x_t(n) - \mu_t)^*\} \\
& -\text{E}\{(x_k(n) - \mu_k)^*(x_l(n) - \mu_l)\}\text{E}\{(x_s(n) - \mu_s)^*(x_t(n) - \mu_t)\}^* \\
= \; & \text{E}\{(x_k(n) - \mu_k)^*(x_s(n) - \mu_s)\}\text{E}\{(x_l(n) - \mu_l)(x_t(n) - \mu_t)^*\} \\
= \; & \boldsymbol{R}_x^T[k;\, s]\boldsymbol{R}_x[l;\, t]
\end{aligned} \tag{B.14}$$

Therefore, the asymptotic covariance matrix is given by

$$\text{cov}\{\text{vec}(\boldsymbol{K}(n))\} = \boldsymbol{R}_x^T \otimes \boldsymbol{R}_x \tag{B.15}$$

[2] Here, we use the fact that if $A(N) \rightarrow A$ in distribution and $B_N \rightarrow 0$ in probability, then $A(N) + B_N \rightarrow A$ in distribution. See, e.g. [92], Section 2c.

Hence, Eq.(B.12) can be equivalently represented as

$$\mathrm{E}\{\mathrm{vec}(\Delta \boldsymbol{R}_x)\} = \boldsymbol{O}$$

$$\mathrm{E}\{\mathrm{vec}(\Delta \boldsymbol{R}_x)\mathrm{vec}^\dagger(\Delta \boldsymbol{R}_x)\} = \frac{1}{N}\boldsymbol{R}_x^T \otimes \boldsymbol{R}_x$$

Since $\Delta \boldsymbol{R}_x$ is Hermitian,

$$\mathrm{vec}^T(\Delta \boldsymbol{R}_x) = \mathrm{vec}^\dagger(\Delta \boldsymbol{R}_x^T) = [\boldsymbol{\Upsilon}\mathrm{vec}(\Delta \boldsymbol{R}_x)]^\dagger = \mathrm{vec}^\dagger(\Delta \boldsymbol{R}_x)\boldsymbol{\Upsilon}^\dagger \tag{B.16}$$

therefore,

$$\begin{aligned}
\mathrm{E}\{\mathrm{vec}(\Delta \boldsymbol{R}_x)\mathrm{vec}^T(\Delta \boldsymbol{R}_x)\} &= \mathrm{E}\{\mathrm{vec}(\Delta \boldsymbol{R}_x)\mathrm{vec}^\dagger(\Delta \boldsymbol{R}_x)\}\boldsymbol{\Upsilon}^\dagger \\
&= \frac{1}{N}(\boldsymbol{R}_x^T \otimes \boldsymbol{R}_x)\boldsymbol{\Upsilon}^\dagger \\
&= \frac{1}{N}(\boldsymbol{R}_x^T \otimes \boldsymbol{R}_x)\boldsymbol{\Upsilon}
\end{aligned}$$

where $\boldsymbol{\Upsilon}$ is the permutation matrix defined in Eq.(5.5). □

B.2 Proof of Theorem 5.2

We observe that the estimation errors of estimated signal projector $\hat{P}_s$ and quasi-noise projector $\hat{P}_\nu$ are results of the perturbation ΔR_x of the exact covariance matrix R_x. If we denote the nth-order variational of $\hat{P}_s$ and $\hat{P}_\nu$ by $\delta^n P$, then we may express $\hat{P}_s$ and $\hat{P}_\nu$ as [86]

$$\hat{P}_s = P_s + \delta P + \cdots + \delta^n P + \cdots = P_s + \Delta P_s \tag{B.17a}$$

$$\hat{P}_\nu = P_\nu - \delta P - \cdots - \delta^n P - \cdots = P_\nu - \Delta P_s \tag{B.17b}$$

where ΔP_s is given by Eq.(5.3b).

The expansion of the resulting error (perturbation) ΔP_s can be carried out to any desired order if a recurrence formular for $\delta^n P$ is derived.

Let the pseudo-inverse of S be defined as

$$S^{\#} = \sum_{i=1}^{M_0} \frac{w_i w_i^\dagger}{\mu_i - \sigma_\nu^2} \tag{B.18}$$

then, we have

$$SS^{\#} = S^{\#}S = P_s \tag{B.19}$$

We also know that orthonormal projectors are self-adjoint [67, 68], i.e.,

$$P_s^\dagger = P_s, \quad P_\nu^\dagger = P_\nu \tag{B.20a}$$

and idempotent, i.e.,

$$P_s P_s = P_s, \quad P_\nu P_\nu = P_\nu \tag{B.20b}$$

By replacing P_s and P_ν in Eqs.(B.20) by their expressions in Eqs.(B.17), we have

$$(P_s + \delta P + \cdots + \delta^n P + \cdots)(P_s + \delta P + \cdots + \delta^n P + \cdots) = P_s + \delta P + \cdots + \delta^n P + \cdots \tag{B.21a}$$

$$(P_\nu - \delta P - \cdots - \delta^n P - \cdots)(P_\nu - \delta P - \cdots - \delta^n P - \cdots) = P_\nu - \delta P - \cdots - \delta^n P - \cdots \tag{B.21b}$$

We equate the terms of corresponding order on both side of Eqs.(B.21) to obtain some identities. The first-order terms result in the following identities:

$$\delta P = (\delta P)P_s + P_s(\delta P) \tag{B.22a}$$

$$\delta P = (\delta P)P_\nu + P_\nu(\delta P) \tag{B.22b}$$

Premultiplying and postmultiplying Eq.(B.22a) by P_s and likewise for Eq.(B.22b) by P_ν, we conclude that

$$P_s(\delta P)P_s = O \tag{B.23a}$$

$$P_\nu(\delta P)P_\nu = O \tag{B.23b}$$

Note that matrix δP can be expressed as a sum of alternating projections, i.e.,

$$\begin{aligned} \delta P &= I(\delta P)I \\ &= (P_s + P_\nu)(\delta P)(P_s + P_\nu) \\ &= P_s(\delta P)P_s + P_s(\delta P)P_\nu + P_\nu(\delta P)P_s + P_\nu(\delta P)P_\nu \\ &= P_s(\delta P)P_\nu + P_\nu(\delta P)P_s \end{aligned} \tag{B.24}$$

where the identities in Eqs.(B.23) are used.

We know

$$\hat{P}_s\hat{R}_x = \sum_{i=1}^{M_0} \hat{\mu}_i\hat{w}_i\hat{w}_i^\dagger = \hat{R}_x\hat{P}_s \tag{B.25a}$$

thus,

$$(R_x + \triangle R_x)\hat{P}_s = \hat{P}_s(R_x + \triangle R_x) \tag{B.25b}$$

We know that, given the approximated signal subspace dimension M_0, R_x can be rewritten as

$$R_x \simeq S + \sigma_\nu^2 I \tag{B.26}$$

therefore,

$$(S + \sigma_\nu^2 I + \triangle R_x)\hat{P}_s = \hat{P}_s(S + \sigma_\nu^2 I + \triangle R_x) \tag{B.27}$$

which is equivalent to

$$(S + \triangle R_x)(P_s + \delta P) = (P_s + \delta P)(S + \triangle R_x) \tag{B.28}$$

with the first-order approximation of $\triangle P$ in $\hat{P}_s$.

The first-order terms of Eq.(B.28) yield an equation

$$S(\delta P) + \triangle R_x P_s = (\delta P)S + P_s \triangle R_x \tag{B.29}$$

Since $S^{\#}S = P_s$, we premultiply Eq.(B.29) by $S^{\#}$ and postmultiply by P_ν and obtain

$$P_s(\delta P)P_\nu = S^{\#}\triangle R_x P_\nu \tag{B.30}$$

Similarly, we derive $P_\nu(\delta P)P_s$ by premultiplying Eq.(B.29) with P_ν and postmultiplying with $S^{\#}$ to obtain

$$P_\nu(\delta P)P_s = P_\nu \triangle R_x S^{\#} \tag{B.31}$$

From Eq.(B.24), we have

$$\delta P = S^{\#}\triangle R_x P_\nu + P_\nu \triangle R_x S^{\#} \tag{B.32}$$

The higher-order terms $\delta^n P$ $(n > 1)$ can be derived in a similar manner. □

B.3 Proof of Theorem 7.1

As we know, the covariance matrix of $\mathrm{vec}(\boldsymbol{x}(n)\boldsymbol{x}^\dagger(n))$, $\bar{\boldsymbol{R}}_x$ is estimated using the equation

$$\begin{aligned} N\hat{\bar{\boldsymbol{R}}}_x &= \sum_{n=1}^{N}[\mathrm{vec}(\boldsymbol{x}(n)\boldsymbol{x}^\dagger(n)) - \mathrm{vec}(\hat{\bar{\boldsymbol{R}}}_x)][\mathrm{vec}(\boldsymbol{x}(n)\boldsymbol{x}^\dagger(n)) - \mathrm{vec}(\hat{\bar{\boldsymbol{R}}}_x)]^\dagger \\ &= \sum_{n=1}^{N}[\mathrm{vec}(\boldsymbol{x}(n)\boldsymbol{x}^\dagger(n)) - \mathrm{vec}(\boldsymbol{R}_x) - (\mathrm{vec}(\hat{\bar{\boldsymbol{R}}}_x) - \mathrm{vec}(\boldsymbol{R}_x))] \\ &\quad [\mathrm{vec}(\boldsymbol{x}(n)\boldsymbol{x}^\dagger(n)) - \mathrm{vec}(\boldsymbol{R}_x) - (\mathrm{vec}(\hat{\bar{\boldsymbol{R}}}_x) - \mathrm{vec}(\boldsymbol{R}_x))]^\dagger \\ &= \sum_{n=1}^{N}[\mathrm{vec}(\boldsymbol{x}(n)\boldsymbol{x}^\dagger(n)) - \mathrm{vec}(\boldsymbol{R}_x)][\mathrm{vec}(\boldsymbol{x}(n)\boldsymbol{x}^\dagger(n)) - \mathrm{vec}(\boldsymbol{R}_x)]^\dagger \\ &\quad -N[\mathrm{vec}(\hat{\bar{\boldsymbol{R}}}_x) - \mathrm{vec}(\boldsymbol{R}_x)][\mathrm{vec}(\hat{\bar{\boldsymbol{R}}}_x) - \mathrm{vec}(\boldsymbol{R}_x)]^\dagger \end{aligned} \tag{B.33}$$

Let

$$\boldsymbol{Z}(n) = [\mathrm{vec}(\boldsymbol{x}(n)\boldsymbol{x}^\dagger(n)) - \mathrm{vec}(\boldsymbol{R}_x)][\mathrm{vec}(\boldsymbol{x}(n)\boldsymbol{x}^\dagger(n)) - \mathrm{vec}(\boldsymbol{R}_x)]^\dagger \tag{B.34a}$$

$$\boldsymbol{B}_N = [\mathrm{vec}(\hat{\bar{\boldsymbol{R}}}_x) - \mathrm{vec}(\boldsymbol{R}_x)][\mathrm{vec}(\hat{\bar{\boldsymbol{R}}}_x) - \mathrm{vec}(\boldsymbol{R}_x)]^\dagger \tag{B.34b}$$

Therefore, we have

$$N\hat{\bar{\boldsymbol{R}}}_x = \sum_{n=1}^{N}\boldsymbol{Z}(n) - N\boldsymbol{B}_N \tag{B.35a}$$

i.e.,

$$N\,\mathrm{vec}(\hat{\bar{\boldsymbol{R}}}_x) = \sum_{n=1}^{N}\mathrm{vec}(\boldsymbol{Z}(n)) - N\,\mathrm{vec}(\boldsymbol{B}_N) \tag{B.35b}$$

We note that $\mathrm{vec}(\boldsymbol{Z}(n))$ are i.i.d. random vectors with

$$\mathrm{E}\{\mathrm{vec}(\boldsymbol{Z}(n))\} = \mathrm{vec}(\mathrm{E}\{\boldsymbol{Z}(n)\}) = \mathrm{vec}(\bar{\boldsymbol{R}}) \tag{B.36a}$$

$$\mathrm{cov}\{\mathrm{vec}(\boldsymbol{Z}(n))\} = \mathrm{E}\{[\mathrm{vec}(\boldsymbol{Z}(n)) - \mathrm{vec}(\bar{\boldsymbol{R}}_x)][\mathrm{vec}(\boldsymbol{Z}(n)) - \mathrm{vec}(\bar{\boldsymbol{R}}_x)]^\dagger\} \triangleq \boldsymbol{\Xi}_{\bar{R}} \tag{B.36b}$$

By Eq.(B.34a) and the Central Limit Theorem, we have

$$N^{-\frac{1}{2}}\sum_{n=1}^{N}[\mathrm{vec}(\boldsymbol{Z}(n)) - \mathrm{vec}(\bar{\boldsymbol{R}}_x)] \to \mathcal{N}(\boldsymbol{O}, \mathrm{cov}\{\mathrm{vec}(\boldsymbol{Z}(n))\}) = \mathcal{N}(\boldsymbol{O}, \boldsymbol{\Xi}_{\bar{R}}) \tag{B.37}$$

as $N \to \infty$.

From Eq.(B.11) as $N \to \infty$, we have

$$N^{\frac{1}{2}}[\text{vec}(\hat{\boldsymbol{R}}_x) - \text{vec}(\boldsymbol{R}_x)] \to \mathcal{N}(\boldsymbol{O}, \text{cov}\{\text{vec}(\boldsymbol{x}(n)\boldsymbol{x}^\dagger(n))\}) = \mathcal{N}(\boldsymbol{O}, \bar{\boldsymbol{R}}_x) \tag{B.38}$$

when $N \to \infty$. Therefore,

$$N^{\frac{1}{2}}[\text{vec}(\hat{\boldsymbol{R}}_x) - \text{vec}(\boldsymbol{R}_x)] \to \boldsymbol{O}, \text{ in probability as } N \to \infty \tag{B.39}$$

so that

$$N^{\frac{1}{2}}[\text{vec}(\hat{\boldsymbol{R}}_x) - \text{vec}(\boldsymbol{R}_x)][\text{vec}(\hat{\boldsymbol{R}}_x) - \text{vec}(\boldsymbol{R}_x)]^\dagger \to \boldsymbol{O}, \text{ in probability as } N \to \infty \tag{B.40}$$

i.e.,

$$N^{\frac{1}{2}}\boldsymbol{B}_N \to \boldsymbol{O}, \text{ in probability as } N \to \infty \tag{B.41}$$

Equivalently,

$$N^{\frac{1}{2}}\text{vec}(\boldsymbol{B}_N) \to \boldsymbol{O}, \text{ in probability as } N \to \infty \tag{B.42}$$

From Eq.(B.35b), we have

$$N^{\frac{1}{2}}\text{vec}(\hat{\bar{\boldsymbol{R}}}_x) = N^{-\frac{1}{2}} \sum_{n=1}^{N} \text{vec}(\boldsymbol{Z}(n)) - N^{\frac{1}{2}}\text{vec}(\boldsymbol{B}_N) \tag{B.43}$$

Substituting Eqs.(B.41) and (B.43) into Eq.(B.37), we have

$$N^{\frac{1}{2}}\text{vec}(\hat{\bar{\boldsymbol{R}}}_x) - N^{\frac{1}{2}}\text{vec}(\bar{\boldsymbol{R}}_x) = N^{-\frac{1}{2}} \sum_{n=1}^{N}[\text{vec}(\boldsymbol{Z}(n)) - \text{vec}(\bar{\boldsymbol{R}}_x)] - N^{\frac{1}{2}}\text{vec}(\boldsymbol{B}_N) \to \mathcal{N}(\boldsymbol{O}, \boldsymbol{\Xi}_{\bar{R}}) \tag{B.44}$$

as $N \to \infty$,

i.e.,

$$N^{\frac{1}{2}}[\text{vec}(\hat{\bar{\boldsymbol{R}}}_x - \bar{\boldsymbol{R}}_x)] = N^{1/2}[\text{vec}(\triangle\bar{\boldsymbol{R}}_x)] \to \mathcal{N}(\boldsymbol{O}, \boldsymbol{\Xi}_{\bar{R}}) \tag{B.45}$$

as $N \to \infty$.

Now, we evaluate the asymptotic covariance matrix $\boldsymbol{\Xi}_{\bar{R}}$. From Eq.(B.34a) and using

$\text{vec}(\boldsymbol{x}\boldsymbol{x}^\dagger) = \boldsymbol{x}^* \otimes \boldsymbol{x}$, we have

$$\begin{aligned}
& \text{vec}(\boldsymbol{Z}(n)) \\
= \; & \text{vec}([\text{vec}(\boldsymbol{x}(n)\boldsymbol{x}^\dagger(n)) - \text{vec}(\boldsymbol{R}_x)][\text{vec}(\boldsymbol{x}(n)\boldsymbol{x}^\dagger(n)) - \text{vec}(\boldsymbol{R}_x)]^\dagger) \\
= \; & [\text{vec}(\boldsymbol{x}(n)\boldsymbol{x}^\dagger(n)) - \text{vec}(\boldsymbol{R}_x)]^* \otimes [\text{vec}(\boldsymbol{x}(n)\boldsymbol{x}^\dagger(n)) - \text{vec}(\boldsymbol{R}_x] \\
= \; & [\boldsymbol{x}^*(n) \otimes \boldsymbol{x}(n) - \text{vec}(\boldsymbol{R}_x)]^* \otimes [\boldsymbol{x}^*(n) \otimes \boldsymbol{x}(n) - \text{vec}^*(\boldsymbol{R}_x)] \qquad \text{(B.46)}
\end{aligned}$$

and from Eq.(B.34a), we see that

$$\text{cov}\{\text{vec}(\boldsymbol{Z}(n))\} = \text{E}\{\text{vec}(\boldsymbol{Z}(n))\text{vec}^\dagger(\boldsymbol{Z}(n))\} - \text{vec}(\bar{\boldsymbol{R}}_x)\text{vec}^\dagger(\bar{\boldsymbol{R}}_x) \qquad \text{(B.47)}$$

Let us define

$$\boldsymbol{c}_x \triangleq \text{E}\{\text{vec}(\text{vec}(\boldsymbol{Z}(n))\text{vec}^\dagger(\boldsymbol{Z}(n)))\} \qquad \text{(B.48)}$$

Again, using $\text{vec}(\boldsymbol{x}\boldsymbol{x}^\dagger) = \boldsymbol{x}^* \otimes \boldsymbol{x}$ together with Eq.(B.46)

$$\begin{aligned}
\boldsymbol{c}_x \; = \; & \text{E}\{\text{vec}^*(\boldsymbol{Z}(n)) \otimes \text{vec}(\boldsymbol{Z}(n))\} \\
= \; & \text{E}\{[\boldsymbol{x}^*(n) \otimes \boldsymbol{x}(n) - \text{vec}(\boldsymbol{R}_x)] \otimes [\boldsymbol{x}(n) \otimes \boldsymbol{x}^*(n) - \text{vec}^*(\boldsymbol{R}_x)] \\
& \otimes [\boldsymbol{x}(n) \otimes \boldsymbol{x}^*(n) - \text{vec}^*(\boldsymbol{R}_x)] \otimes [\boldsymbol{x}^*(n) \otimes \boldsymbol{x}(n) - \text{vec}(\boldsymbol{R}_x)]\} \qquad \text{(B.49)}
\end{aligned}$$

Therefore, $\boldsymbol{\Xi}_R$ is given as

$$\text{cov}\{\text{vec}(\boldsymbol{Z}(n))\} = \text{unvec}(\boldsymbol{c}_x) - \text{vec}(\bar{\boldsymbol{R}}_x)\text{vec}^\dagger(\bar{\boldsymbol{R}}_x) \qquad \text{(B.50)}$$

Hence, Eq.(B.45) can be equivalently represented as

$$\begin{aligned}
& \text{E}\{\text{vec}(\Delta\bar{\boldsymbol{R}}_x)\} = \boldsymbol{O} \\
& \text{E}\{\text{vec}(\Delta\bar{\boldsymbol{R}}_x)\text{vec}^\dagger(\Delta\bar{\boldsymbol{R}}_x)\} = \frac{1}{N}\boldsymbol{\Xi}_R
\end{aligned}$$

Since $\Delta\bar{\boldsymbol{R}}_x$ is Hermitian,

$$\text{vec}^T(\Delta\bar{\boldsymbol{R}}_x) = \text{vec}^\dagger(\Delta\bar{\boldsymbol{R}}_x^T) = [\bar{\boldsymbol{\Upsilon}}\text{vec}(\Delta\bar{\boldsymbol{R}}_x)]^\dagger = \text{vec}^\dagger(\Delta\bar{\boldsymbol{R}}_x)\bar{\boldsymbol{\Upsilon}}^\dagger \qquad \text{(B.51)}$$

therefore,

$$\begin{aligned} \mathrm{E}\{\mathrm{vec}(\Delta\bar{R}_x)\mathrm{vec}^T(\Delta\bar{R}_x)\} &= \mathrm{E}\{\mathrm{vec}(\Delta\bar{R}_x)\mathrm{vec}^\dagger(\Delta\bar{R}_x)\}\tilde{\Upsilon}^\dagger \\ &= \frac{1}{N}\Xi_R\tilde{\Upsilon}^\dagger \\ &= \frac{1}{N}\Xi_R\tilde{\Upsilon} \end{aligned}$$

where $\tilde{\Upsilon}$ is the permutation matrix defined in Eq.(7.30). □

Appendix C

Covariance of the sample Cumulants

The covariance of the sample fourth-order cumulants of the array output may be expressed as a sum of homogeneous products of cumulants up to eighth-order. The list of all possible such terms can easily be adopted from pp.259-260 of [95]; even for zero-mean variables, it shows an impressive 569 terms (or partitions). Fortunately, since we deal only with circular complex random variables, many of these terms vanish. The remaining 50 terms sum up as follows:

$$
\begin{aligned}
& \Xi_Q[i,j,k,l;\ u,v,s,t] \\
= & \lim_{N\to\infty} N\mathrm{E}\{(\hat{Q}[i,j;\ k,l]-Q[i,j;\ k,l])(\hat{Q}[u,v;\ s,t]-Q[u,v;\ s,t])^*\} \\
= & \ Q_8[i,j,k,l,;\ u,v,s,t]+Q_6[i,j,k;\ s,u,v]R[t;\ l]+Q_6[i,j,k;\ v,t,s]R[u;\ l] \\
& +Q_6[i,j,t;\ l,u,v]R[k;\ s]+Q_6[i,j,t;\ l,u,s]R[k;\ v]+Q_6[i,k,t;\ l,u,v]R[j;\ s] \\
& +Q_6[i,k,t;\ l,u,s]R[j;\ v]+Q_6[s,j,k;\ l,u,v]R[t;\ i]+Q_6[l,j,k;\ s,t,v]R[u;\ i] \\
& +Q[i,j;\ k,v]Q[l,u;\ t,s]+Q[i,j;\ k,s]Q[u,l;\ t,v]+Q[i,j;\ u,l]Q[v,k;\ t,s]
\end{aligned}
$$

$$
\begin{aligned}
&+Q[i,j;\ t,l]Q[s,k;\ u,v] + Q[i,k;\ u,l]Q[j,v;\ t,s] + Q[i,k;\ t,l]Q[s,j;\ u,v] \\
&+Q[l,j;\ k,v]Q[i,u;\ t,s] + Q[l,j;\ k,s]Q[i,u;\ t,v] + Q[i,j;\ u,v]Q[l,k;\ t,s] \\
&+Q[i,j;\ t,s]Q[l,k;\ u,v] + Q[i,j;\ u,s]Q[l,k;\ t,v] + Q[l,j;\ t,v]Q[l,k;\ u,s] \\
&+Q[i,u;\ t,l]Q[v,j;\ k,s] + Q[i,k;\ u,v]Q[l,j;\ t,s] + Q[i,k;\ u,s]Q[l,j;\ t,v] \\
&+Q[i,k;\ t,v]Q[l,j;\ u,s] + Q[i,k;\ t,s]Q[l,j;\ u,v] \\
&+Q[i,u;\ t,l]R[j;\ v]R[k;\ s] + Q[i,u;\ t,l]R[j;\ s]R[k;\ v] \\
&+Q[s,j;\ k,v]R[u;\ i]R[t;\ l] + Q[s,j;\ k,v]R[t;\ i]R[u;\ l] \\
&+Q[i,k;\ u,v]R[j;\ s]R[t;\ l] + Q[l,j;\ t,s]R[u;\ i]R[k;\ v] \\
&+Q[i,k;\ u,s]R[j;\ v]R[t;\ l] + Q[l,j;\ t,v]R[u;\ i]R[k;\ s] \\
&+Q[i,k;\ t,v]R[j;\ s]R[u;\ l] + Q[l,j;\ u,s]R[t;\ i]R[k;\ v] \\
&+Q[i,k;\ t,s]R[j;\ v]R[u;\ l] + Q[l,j;\ u,v]R[t;\ i]R[k;\ s] \\
&+Q[i,j;\ u,v]R[k;\ s]R[t;\ l] + Q[l,k;\ t,s]R[u;\ i]R[j;\ v] \\
&+Q[i,j;\ t,s]R[k;\ v]R[u;\ l] + Q[l,k;\ u,v]R[t;\ i]R[j;\ s] \\
&+Q[i,j;\ u,s]R[k;\ v]R[t;\ l] + Q[l,k;\ t,v]R[u;\ i]R[j;\ s] \\
&+Q[i,j;\ t,v]R[k;\ s]R[u;\ l] + Q[l,k;\ u,s]R[t;\ i]R[j;\ v] \\
&+R[j;\ i]R[k;\ l]R[u;\ v]R[t;\ s] + R[j;\ i]R[k;\ l]R[u;\ s]R[t;\ v] \\
&+R[k;\ i]R[j;\ l]R[u;\ v]R[t;\ s] + R[k;\ i]R[j;\ l]R[u;\ s]R[t;\ v]
\end{aligned}
\tag{C.1}
$$

where $Q_6[i,j,k;\ u,v,s]$ and $Q_8[i,j,k,l;\ u,v,s,t]$ denote the sixth-order and eighth-order cumulants of the array output:

$$
Q_6[i,j,k;\ u,v,s] \triangleq \text{cum}_6\{x_i^*, x_j, x_k, x_u^*, x_v, x_s^*\} \tag{C.2a}
$$

$$
Q_8[i,j,k,l;\ u,v,s,t] \triangleq \text{cum}_8\{x_i^*, x_j, x_k, x_l^*, x_u, x_v^*, x_s^*, x_t\} \tag{C.2b}
$$

□

Bibliography

[1] J. R. Williams, "Fast beamforming algorithms", *J. Acoust. Soc. Amer.*, vol. 44, pp. 1454-1455, 1968.

[2] P. Rudnick, "Digital beamforming in the frequency domain", *J. Acoust. Soc. Amer.*, vol. 46, pp. 1089-1090, 1969.

[3] D. E. Dudgeon, "Fundamentals of digital array processing", *Proc. IEEE*, vol. 65, pp. 898-904, June, 1977.

[4] F. J. Harris, "On the Use of windows for harmonic analysis with the discrete Fourier transform", *Proc. IEEE*, vol. 66, pp. 51-83, Jan., 1978.

[5] R. G. Pridham and R. A. Mucci, "A novel approach to digital beamforming", *J. Acoust. Soc. Amer.*, vol. 63, pp. 425-434, Feb., 1978.

[6] H. Fan, E. I. El-Masry, and W. K. Jenkins, "Resolution enhancement of digital beamformers", *IEEE Trans. ASSP*, vol. 32, no. 5, pp. 1041-1051, Oct., 1984.

[7] H. Akaike, "A new look at statistical model identification", *IEEE Trans. AC*, vol. 19, pp. 716-723, Dec., 1974.

[8] J. Rissanen, "Modelling by shortest data description", *Automatica*, vol. 14, pp. 465-471, 1978.

[9] G. Schwartz, "Estimating the dimension of a model", *Ann. Stat.*, vol. 6, no. 2, pp. 461-464, 1978.

[10] M. Wax and T. Kailath, "Detection of signals by information theoretic criteria", *IEEE Trans. ASSP*, vol. 33, pp. 387-392, April, 1985.

[11] Y. Yin and P. Krishnaiah, "On some nonparametric methods for detection of the number of signals," *IEEE Trans. ASSP*, vol. 35, pp. 1533-1538, 1987.

[12] K. M. Wong, Q. T. Zhang, J. P. Reilly, and P. C. Yip, "On information theoretic criteria for the determination of the number of signals in high resolution array processing," *IEEE Trans. ASSP*, vol. 38, no. 11. pp. 1959-1971, Nov., 1990.

[13] W. G. Chen, K. M. Wong, and J. P. Reilly, "Detection of the number of signals — a predictive eigen-threshold approach," *IEEE Trans. SP*, vol. 20, no. 5, pp. 1088-1098, May, 1991.

[14] Q. Wu and D. R. Fuhrmann, "A parametric method for determining the number of signals in narrowband direction-finding," *IEEE Trans. SP*, vol. 39, no. 8, pp. 1848-1857, Aug., 1991.

[15] J. Capon, "High resolution frequency wavenumber spectrum analysis," *Proc. IEEE*, vol. 57, pp. 1408-1418, Aug., 1969.

[16] V. F. Pisarenko, "The retrieval of harmonics from a covariance function," *Geophys. J. Roy. Astron. Soc.*, vol. 33, pp. 347-366, 1973.

[17] D. H. Johnson and S. R. DeGraff, "Improving the resolution of bearing in passive sonar arrays by eigenvalue analysis," *IEEE Trans. ASSP*, vol. 30, no. 4, pp. 638-647, Aug., 1982.

[18] D. H. Johnson, "The application of spectral estimation methods to bearing estimation problems", *Proc. IEEE*, vol. 70, pp. 1018-1028, Sept., 1982.

[19] D. W. Tufts and R. Kumaresan, "Estimation of frequencies of multiple sinusoids: Making linear predictions perform like maximum likelihood," *Proc. IEEE*, vol. 70, no. 9, pp. 975-989, Sept., 1982.

[20] R. Kumaresan and D. W. Tufts, "Estimating the angles of arrival of multiple plane waves," *IEEE Trans. AES*, vol. 19, pp. 134-139, Jan., 1983.

[21] G. Bienvenue and L. Kopp, "Optimality of high resolution array processing using the eigensystem approach," *IEEE Trans. ASSP*, vol. 31, no. 5, pp. 1235-1247, Oct., 1983.

[22] G. Su and M. Morf, "Signal subspace approach for multiple wideband emitter location", *IEEE Trans. ASSP*, vol. 31, no. 6, pp. 1502-1522, Dec., 1983.

[23] H. Wang and M. Kaveh, "Coherent signal-subspace processing for the detection and estimation of angles of arrival of multiple wideband sources," *IEEE Trans. ASSP*, vol. 33, no. 4, pp. 823-831, Aug., 1985.

[24] R. O. Schmidt, "Multiple emitter location and signal parameter estimation," *IEEE Trans. AP*, vol. 34, pp. 276-280, March, 1986.

[25] Y. Bresler and A. Macovski, "Exact maximum-likelihood parameter estimation of superimposed exponential signals in noise," *IEEE Trans. ASSP*, vol. 34, no. 5, pp. 1081-1089, Oct., 1986.

[26] J. F. Böhme, "Estimation of spectral parameters of correlated signals in wavefields", *Signal Processing*, vol. 11, pp. 329-337, 1986.

[27] K. Sharman and T. S. Durrani, "A comparative study of modern eigenstructure methods for bearing estimation — A new high performance approach," *Proc. 25 IEEE Conf. Dec. Contr.*, pp.1737-1742, Athens, Greece, Dec., 1986.

[28] K. M. Buckley and L. J. Griffiths, "Broadband signal subspace spatial spectrum (BASS-ALE) estimation", *IEEE Trans. ASSP*, vol. 36, no. 7, pp. 953-964, July, 1988.

[29] H. Hung and M. Kaveh, "Focusing matrices for coherent signal subspace processing", *IEEE Trans. ASSP*, vol. 36, no. 8, pp. 1272-1281, Aug., 1988.

[30] I. Ziskind and M. Wax, "Maximum likelihood localization of multiple sources by alternating projection," *IEEE Trans. ASSP*, vol. 36, no. 10, pp.1553-1560, Oct., 1988.

[31] R. Roy and T. Kailath, "ESPRIT — Estimation of signal parameters via rotational invariance techniques," *IEEE Trans. ASSP*, vol. 37, no. 7, pp. 984-995, July, 1989.

[32] J. Krolik and D. N. Swingler, "Multiple broadband source location using steered covariance matrices", *IEEE Trans. ASSP*, vol. 37, no. 10, pp. 1481-1494, Oct., 1989.

[33] P. Stoica and K. C. Sharman, "Novel eigenanalysis method for direction estimation", *Proc. IEE, Pt. F*, pp. 19-26, February, 1990.

[34] J. Krolik and D. N. Swingler, "Focused wideband array processing via spatial resampling", *IEEE Trans. ASSP*, vol. 38, no. 2, pp. 356-360, Feb., 1990.

[35] P. Stoica and K. C. Sharman, "Maximum likelihood methods for direction-of-arrival estimation," *IEEE Trans ASSP*, vol. 38, no. 7, pp. 1132-1143, July, 1990.

[36] M. Miller and D. Fuhrmann, "Maximum likelihood narrow-band direction finding and the EM algorithm", *IEEE Trans. ASSP*, vol. 38, pp. 1560-1577, September, 1990.

[37] M. Viberg and B. Ottersten, "Sensor array processing based on subspace fitting," *IEEE Trans. SP*, vol. 39, no.5, pp. 1110-1121, May, 1991.

[38] M. Viberg, B. Ottersten, and T. Kailath, "Detection and estimation in sensor arrays using weighted subspace fitting," *IEEE Trans. SP*, vol. 39, no.11, pp. 2436-2449, Nov., 1991.

[39] M. Kaveh and A. J. Barabell, "The statistical performance of the MUSIC and the Minimum-Norm algorithms in resolving plane waves in noise," *IEEE Trans. ASSP*, vol. 34, pp. 331-341, April, 1986.

[40] H. Wang and M. Kaveh, "Performance of signal subspace processing methods, Pt. I: Narrowband systems", *IEEE Trans. ASSP*, vol. 34, no. 5, pp. 1201-1209, Oct., 1986.

[41] H. Wang and M. Kaveh, "Performance of signal subspace processing methods, Pt. II: Coherent wideband systems", *IEEE Trans. ASSP*, vol. 35, no. 11, pp. 1583-1591, Nov., 1987.

[42] M. Kaveh, H. Wang, and H. Hung, "On the theoretical performance of a class of estimators of the number of narrowband sources", *IEEE Trans. ASSP*, vol. 35, no. 9, pp. 1350-1352, Sept., 1987.

[43] B. Porat and B. Friedlander, "Analysis of the relative efficiency of the MUSIC algorithm", *IEEE Trans. ASSP*, vol. 36, no. 4, pp. 532-534, April, 1988.

[44] P. Stoica and A. Nehorai, "MUSIC, maximum likelihood, and Cramér-Rao Bound," *IEEE Trans. ASSP*, vol. 37, no. 5, pp. 720-741, May, 1989.

[45] Q. T. Zhang, K. M. Wong, P. C. Yip, and J. P. Reilly, "Statistical analysis of the performance of information theoretic criteria in the detection of the number of signals in array processing", *IEEE Trans. ASSP*, vol. 37, no. 10, pp. 1557-1567, Oct., 1989.

[46] P. Stoica and A. Nehorai, "MUSIC, maximum likelihood, and Cramér-Rao bound: Further results and comparisons", *IEEE Trans. ASSP*, vol. 38, no. 12, pp. 2140-2150, Dec., 1990.

[47] P. Stoica and T. Söderström, "Statistical analysis of MUSIC and subspace rotation estimates of sinusoidal frequencies", *IEEE Trans. SP*, vol. 39, no. 8, pp. 1836-1857, Aug., 1991.

[48] P. M. Schultheiss and J. P. Ianniello, "Optimum range and bearing estimation with randomly perturbed arrays", *J. Acoust. Soc. Am.*, vol. 68, pp. 167-173, 1980.

[49] K. M. Wong, R. S. Walker, and G. Niezgoda, "Effects of random sensor motion on bearing estimation by the MUSIC algorithm", *IEE Proc.*, vol. 135, Pt. F. no. 3, pp. 233-250, June, 1988.

[50] B. Friedlander, "A sensitivity analysis of the MUSIC algorithm", *IEEE Trans. ASSP*, vol. 38, no. 10, pp. 1740-1751, Oct., 1990.

[51] Y. Rockah and P. M. Schultheiss, "Array shape calibration using sources in unknown locations, Pt. I: Far-field sources", *IEEE Trans. ASSP*, vol. 35, no. 3, pp. 286-299, March, 1987.

[52] Y. Rockah and P. M. Schultheiss, "Array shape calibration using sources in unknown locations, Pt. II: Near-field sources and estimation implementation", *IEEE Trans. ASSP*, vol. 35, no. 6, pp. 724-735, June, 1987.

[53] B. Friedlander and A. J. Weiss, "Direction finding in the presence of mutual coupling", *22nd Asilomar Conf. on Sig., Systems, and Computers*, Oct. 31 - Nov. 2, 1987.

[54] B. Friedlander and A. J. Weiss, "Eigenstructure methods for direction finding with sensor gain and phase uncertainties", *ICASSP'88*, April, 1988.

[55] A. J. Weiss, A. S. Willsky, and B. C. Levy, "Eigenstructure approach for array processing with unknown intensity coefficients", *IEEE Trans. ASSP*, vol. 36, no. 10, pp. 1613-1617, Oct., 1988.

[56] G. M. Wenz, "Acoustic ambient noise in the ocean: Spectra and sources", *Acoust. Soc. Am.*, vol. 34, no. 12, pp. 1936-1956, Dec., 1962.

[57] R. J. Urick, *Principles of Underwater Sound*, 3rd edition, McGraw-Hill Book Co., 1983.

[58] A. Paulraj and T. Kailath, "Eigenstructure methods for direction of arrival estimation in the presence of unknown noise fields," *IEEE Trans. ASSP*, vol. 34, no. 1, pp. 13-20, Feb., 1986.

[59] J. Le Cadre, "Parametric methods for spatial signal processing in the presence of unknown colored noise fields," *IEEE Trans. ASSP*, vol. 37, pp. 965-983, July, 1989.

[60] A. H. Tewfik, "Harmonic retrieval in the presence of colored noise," *Proc. ICASSP'89*, pp. 2069-2072, 1989.

[61] J. P. Reilly, K. M. Wong, and P. M. Reilly, "Direction of arrival estimation in the presence of noise with unknown arbitrary covariance matrices," *Proc. ICASSP'89*, pp. 2609-2612, 1989.

[62] K. M. Wong, J. P. Reilly, Q. Wu, and S. Qiao, "Estimation of the direction of arrival of signals in unknown correlated noise, Pt. I: The MAP approach and its implementation," *IEEE Trans. SP*, vol. 40, no. 8, pp. 2007-2017, Aug., 1992.

[63] K. M. Wong and J. P. Reilly, "Estimation of the direction of arrival of signals in unknown correlated noise, Pt. II: Asymptotic behaviour and performance of the MAP approach," *IEEE Trans. SP*, vol. 40, no. 8, pp. 2018-2028, Aug., 1992.

[64] M. Wax, "Detection and localization of multiple sources in noise with unknown covariance," *IEEE Trans. SP*, vol. 40, no. 1, pp. 245-249, Jan., 1992.

[65] P. Stoica and T. Söderström, "On array signal processing in spatially correlated noise fields", *IEEE Trans. CAS - II: Analog and Digital Signal Processing*, vol. 39, no. 12, pp. 879-882, Dec., 1992.

[66] R. Bellman, *Introduction to Matrix Analysis*, 2nd ed., McGraw-Hill, 1972.

[67] P. Lancaster and M. Tismenetsky, *The Theory of Matrices*, 2nd ed., Academic Press, 1985.

[68] A. W. Naylor and G. R. Sell, *Linear Operator Theory*, Holt, Reinhart, and Winston, 1971.

[69] F. A. Graybill, *Matrices with Applications in Statistics*, 2nd ed., Wadsworth, 1983.

[70] D. R. Brillinger, *Time Series*, Holden-Day, Inc., 1981.

[71] K. M. Wong, Q. Wu, and P. Stoica, "Generalized correlation decomposition applied to array processing in unknown noise", *Advances in Spectrum Analysis and Array Processing*, Ch.6, vol.III, Edited by S. Haykin, Prentice Hall, 1995.

[72] Q. Wu and K. M. Wong, "UN-MUSIC and UN-CLE: An Application of Generalized Correlation Analysis to the Estimation of the Directions of Arrival", *IEEE Trans. SP*, Vol. 42, Sep., 1994.

[73] Q. Wu and K. M. Wong, "Estimation of DOA in unknown noise: performance analysis of UN-MUSIC and UN-CLE, and the optimality of CCD," *IEEE Trans. SP*, vol. 43, no. 2, pp. 454-468, Feb. 1995.

[74] Y. Meng, K. M. Wong and Q. Wu, "Estimation of th direction of arrival of spread sources in sensor array processing," *Proc. Int. Conf. on SP*, Bejing, China, Oct., 1993.

[75] T-P. Jantti, "The influence of extended sources on the theoretical performance of the MUSIC and ESPRIT methods: Narrow-band Sources," *Proc ICASSP'92*, vol.2, pp. 429-432, Mar., 1992.

[76] S. Anderson, M. Millnert and M. Viberg, "An Adaptive Array For Mobile Communication Systems", *IEEE Trans. Vehicular Technology*, Vol. 40, No. 1, Feb., 1991.

[77] H. L. Van Trees, *Detection, Estimation, and Modulation Theory*, Pt. III, John Wiley & Sons, Inc., 1971.

[78] P. Beckmann and A. Spizzichino, *The Scattering of Electromagnetic Waves from Rough Surfaces*, Pergamon Press, 1963.

[79] A. Ishimaru *Wave Propagation and Scattering in Random Media*, vol.1, Academic Press, 1978.

[80] R. W. Jenkins, "A field-aligned scattering model for high-latitude propagation", *Tech. Memo. DRL/TM095/92, Comm. Research Center*, Ottawa, Dec., 1992.

[81] R. W. Jenkins, "A simulation study of HF direction finding in the presence of F-region scattering and sporadic-E", *Report 93-004, Comm. Research Center*, Ottawa, 1993.

[82] D. Slepian, "Prolate spheroidal wave function, Fourier analysis and uncertainty - V: The discrete case", *The Bell System Technical Journal*, pp. 1371-1430, 1978.

[83] S. Valaee, P. Kabal, and B. Champagne, "Localization of distributed sources," *Proc. Fourteenth GRETSI Symposium on Signal and Image Peocessing*, Juan Les Pins, France, Sept.,1993.

[84] Y. Zhou and P. Yip, "Estimation of DOA based on array manifold space," *Circuits, Systems, and Signal Processing*, vol.14, no.2, Feb., 1995.

[85] Y. Meng and K. M. Wong, "Estimation of the directions of arrival of spatially dispersed signals in array processing," *Submitted to IEE Proc. Radar, Sonar, and Navigation.*

[86] H. Krim and P. Forster, and J. G. Proakis, "Operator approach to performance analysis of Root-MUSIC and Root-Min-Norm", *IEEE Trans. SP*, vol. 40, no. 7, pp. 1687-1696, July, 1992.

[87] J. F. Cardoso, "Blind identification of independent components with higher order moments," *Proc. Workshop Higher Order Spectral Anal.*, Vail, CO, June, 1989.

[88] B. Porat and B. Friedlander, "Performance analysis of parameter estimation algorithms based on high-order moments", *J. Adaptive Contro. Signal Processing*, vol. 3, pp. 191-229, 1989.

[89] C. Nikias and A. Petropulu, "Higher-Order Spectra Analysis: A Nonlinear Signal Processing Framework", *Prentice-Hall, Inc.*, 1993.

[90] B. Porat and B. Friedlander, "Direction finding algorithms based on high-order statistics", *IEEE Tran. SP*, vol. 39, No. 9, Sept., 1991.

[91] Y. Meng and K. M. Wong, "A direction finding algorithm based on higher-order statistics for distributed signals," *Proc. IEEE SP / ATHOS Worksh. Higher-Order Stat.*, Begur, Spain, June, 1995.

[92] C. R. Rao, "Linear Statistical Inference and Its Applications", 2nd ed. *John Wiley & Sons, New York*, 1973.

[93] R. Pan and C. Nikias, "Harmonic decomposition in cumulant domains," *Proc. ICASSP*, 1988, pp. 2356-2359.

[94] E. Moulines and J. F. Cardoso, "Second-order versus fourth-order MUSIC algorithms. An asymptotical statistical performance analysis", *Proc. Int. Worksh. Higher Order Spect. Anal.*, Chamrousse, France, 1991, pp. 121-130.

[95] D. Shin and J. Mendel, "Comparison between correlation based and cumulant based approaches to the harmonic retrieval and related problems", *Tech. Rep. 177, Univ. of Sourthern California, SIPI*, 1991.

[96] A. Graham, "Kronecker Products and Matrix Calculus with Applications", *John Wiley & Sons, New York*, 1981.

[97] J. F. Cardoso and E. Moulines, "Asymptotic performance analysis of direction-finding algorithms based on fourth-order cumulants", *IEEE Tran. SP*, vol. 43, pp. 214-224, Jan., 1995.

[98] Q. Wu, K. M. Wong, Y. Meng and W. Read, "DOA estimation of point and scattered sources - VEC-MUSIC," *Proc. IEEE 7th Worksh. on Stat. Sig. & Array Proc.*, Quebec City, Canada, June, 1994.

[99] M. G. Kendall and A. Stuart, "The Advanced Theory of Statistics," vol. 2, *Charles Griffin Co. Ltd., London*, 1961.

[100] D. Slepian and E. Sonnenblick, "Eigenvalues associated with prolate spheroidal wave functions of zero order", *The Bell System Technical Journal*, vol. 44, pp. 1745-1759, 1965.

[101] D. K. Barton, "Low-angle radar tracking," *Proc. IEEE*, vol. 62, no. 6, pp. 687-704, 1974.