MODIFIED ZIV-ZAKAI LOWER BOUND ON THE ERRORS OF THE ESTIMATION OF DOA

By

XIAODONG GU, M. Eng.

A Thesis

Submitted to the School of Graduate Studies

in Partial Fulfilment of the Requirements

for the Degree

Doctor of Philosophy

McMaster University

MODIFIED ZIV-ZAKAI LOWER BOUND ON THE ERRORS OF THE ESTIMATION OF DOA

TO MY PARENTS

DOCTOR OF PHILOSOPHY — McMASTER UNIVERSITY

(Electrical and Computer Engineering) — Hamilton, Ontario

TITLE: **Modified Ziv-Zakai Lower Bound on the Errors of the Estimation of DOA**

AUTHOR: Xiaodong Gu

M. Eng. (East China Institute of Chemical Technology)

B. Eng. (Anhui Institute of Mechanical and Electrical Engineering)

SUPERVISOR: Dr. K.M. Wong

Professor, Department of Electrical and Computer Engineering

B.Sc.(Eng.), Ph.D., (University of London)

D.I.C. (University of London)

Fellow, I.E.E.

Fellow, Royal Statistical Society

Fellow, Institute of Physics

NUMBER OF PAGES: xiv, 154

ABSTRACT

This thesis has been directed toward the problem of deriving a computable tight lower bound on the error of DOA estimation with the array processing. This work is developed based on the logic implied in Ziv-Zakai's idea and the work for Cramer-Rao lower bound (CRLB).

A profound understanding of Ziv-Zakai's idea is presented. The lower bound on the variance of DOA estimate with one incoming signal is derived applying the logic of Ziv and Zakai. Then the modified Ziv-Zakai lower bound (MZLB) on the covariance matrix of the multiple DOA estimates is developed. The theoretic analysis and the simulation results show that MZLB is a tight lower bound over a wide range of signal-noise ratio. It follows the SNR-threshold phenomenon occurring in the performance of the DOA estimation well, and it is easily computable.

It is proved that, the maximum-likelihood estimation of DOA parameters based on Data Model(2) discussed in this thesis is asymptotically efficient.

Acknowledgement

The author wishes to express sincere gratitude to Dr. K.M. Wong for his constant encouragement, continued assistant and expert guidance and supervision throughout the course of this work. Thanks are also due to Drs. P.C. Yip, J.P. Reilly and S. Qiao, the members of the Supervisory Committee, for their continued interest and useful suggestions.

It is the author's pleasure to acknowledge the inspiring discussions with her colleagues Q.T. Zhang, Q. Jin, X. Huang, H. Dai and W. Chen.

Many thanks go to other friends for their valuable help.

I would also like to acknowledge the financial support provided by Telecommunication Research Institute of Ontario and Department of Electrical and Computer Engineering (McMaster University).

Contents

List of Figures

Chapter 1

Introduction

1.1 INTRODUCTION

Bounds on error is an important aspect of estimation theory. In Section 1.2, we introduce the general knowledge about the lower bound on estimation error. In Section 1.3, we review previous works on the performance analysis of DOA (directions of arrival) estimation in array signal processing, and we also discuss the problems of error measure and error bound under low signal-noise ratio. In Section 1.4, we briefly introduce Ziv-Zakai's idea to derive a tight lower bound on mean-square error. In Section 1.5, we describe the scope and the structure of this thesis.

1.2 LOWER BOUNDS ON ESTIMATION ERRORS

In estimation theory, three issues are of interest:

1. Estimation procedures.

2. Measure of errors.

3. Bounds on errors.

Designing estimation procedures to extract the information from the received data which contain signal and noise, is a well developed and still very active research topic in both theory and application. Because there is noise in the received data and the number of the received data is finite, there must be some errors in the estimation results. This is the fundamental limitation in nature.

Since there are errors in the estimates, it is necessary to measure them. Performance analysis is very important in estimation theory. An estimate without performance analysis is incomplete. Performance analysis includes two aspects: the measure of errors and their error bounds. The most commonly used error measures are the bias and the variance of an estimate. Their bounds reveal the limitation of the accuracy of estimation. The error bounds provide an important approach to study the estimator performance. Through the closeness of the estimate performance to its lower bound, the evaluation of the estimator can be obtained.

The main known results on the lower bound problem are:

(1). The Cramer-Rao lower bound (CRLB) [78] [81].

(2). The Barankin lower bound (BLB) [1] [81].

(3). The Ziv-Zakai lower bound (ZZLB) [100].

Seidman [71] summarized the available lower bounds and compared them for a special case, which gave the nature and the effectiveness of all the lower bounds mentioned in the above.

CRLB is an achievable lower bound under some conditions. Also, it is simple to understand and easily computable. Thus it is used popularly. For linear estimations, performance analysis is relatively straightforward and for these estimators, the CRLB is achievable. However, for nonlinear estimations, in general, it is impossible to calculate precisely the estimation errors. With a finite number of samples, under high signal-noise ratio (SNR), CRLB can be approximately achieved by the variance of the nonlinear estimation. But under low SNR, most of the error measures of the nonlinear estimation are obtained only by computer simulation results, i.e., by evaluating the errors numerically. Also, CRLB is not achievable. The performance analysis of nonlinear estimation under low SNR seems to be a difficult problem.

1.3 REVIEW

Recently, various novel estimators have been developed for the directions of arrival estimation in array signal processing, e.g., MUSIC (multiple-signal-classification), MLE (maximum-likelihood estimation) (see, e.g., [6] [69] and [80]). Much work has been conducted to evaluate the performances of these estimators. There are two important aspects of performance evaluation of these estimators, *viz.*, the error measure, and the lower bound on these estimation errors.

A series of systematic works on the performance analysis of DOA estimation have been done by Stoica and Nehorai [73]-[77]. In their papers, the performances of MUSIC and MLE were derived and evaluated under asymptotic conditions, the Cramer-Rao lower bounds on the covariance matrix of DOA estimates were developed. Many other works

have been developed on the performances of DOA estimation under different conditions. For example, Rife and Boorstyn [66][67] derived CRLB for single-tone and multi-tone parameter estimation from discrete-time observations and gave the evaluation of ML estimate; Wong *et als.* [92] derived the CRLB on the covariance matrix for the DOA estimates incorporating random motion of sensors; Porat and Friedlander [61] derived the asymptotic formulas for the variance of the DOA estimates obtained by the MUSIC algorithm for uncorrelated sources, and compared them with CRLB; Gorman and Hero [23] derived CRLB for parameter estimation with constraints.

As we will see in Chapter 2, DOA estimation is nonlinear. With the finite samples and high SNR, or with the large samples and finite SNR, CRLB on DOA estimate error is achievable. The works mentioned in the above mainly study and compare the performance of DOA estimation with CRLB under these conditions, i.e. in the asymptotic sense.

However, with finite samples, under low SNR, CRLB on DOA estimate error does not follow closely to the estimator performance and is a loose bound. It does not follow the SNR-threshold phenomenon which is common to all nonlinear estimations.

For conciseness, in the following, when we mention SNR without mentioning the number of samples particularly, it means that the number of samples is finite.

Although, performance analysis of DOA estimation in a wide range of SNR seems to be a difficult problem, some work has been developed. In 1986, Kaveh and Barabell [36] derived a formula to find out at which SNR the threshold will occur for two closely spaced signals with the eigen-type estimation method; in 1990, Lee and Wengrovitz [45] extended the threshold expression in [36] to a more general class of problems.

Also, some lower bounds on the DOA estimate errors which are tight over a wide range of SNR, have been developed. In 1986, Weiss and Weinstein [86] proposed an error bound containing two types of free parameters. This W-W lower bound was reported to be

significantly tighter than the CRLB by choosing the optimal values of the free parameters. In 1987, Nohara [56] applied this W-W lower bound to derive the error bound on DOA estimates. Their results have shown that, W-W lower bound is valid in the SNR-threshold region. However the complexity in calculating these lower bounds far exceeds that of the CRLB. Actually, W-W lower bound and Barankin lower bound use a similar idea that, both of them apply Schwarz inequality to derive the lower bound, and obtain a tight lower bound by choosing an optimal free parameters. Therefore they share the common disadvantage that the complexity of the computation makes them lose practical significance.

1.4 ZIV-ZAKAI'S IDEA

In 1969, Ziv and Zakai [100] presented a new lower bound for the time-delay estimation. The main idea is that, by imagining a binary detection procedure based on an estimate, from which the error measure in detection is transferred into the error measure in estimation, and making use of the Chernoff formula which is an error measure approximation of the binary detection problem under some conditions, a tight lower bound on the mean-square error (MSE) of the time-delay estimate is obtained. This is a totally new idea to derive the lower bound without applying Schwarz inequality (CRLB, BLB and W-W lower bound are derived with Schwarz inequality). Since the Chernoff formula gives a lower bound on error probability which is effective no matter the SNR is high or low, the resulting Ziv-Zakai lower bound (ZZLB) is expected to be a tight lower bound under high SNR as well as low SNR. It follows the threshold phenomenon occurring in the nonlinear time-delay estimation [10] [87]. ZZLB is easily computable.

Since the binary detection is detecting one signal identified by one parameter, naturally, Ziv-Zakai's idea works for the single parameter estimation.

The question that arises is if Ziv-Zakai's idea can be applied to DOA estimation

which is a multiple parameter estimation.

1.5 SCOPE AND STRUCTURE OF THE THESIS

This thesis works on *applying Ziv-Zakai's idea to derive the lower bound on the covariance matrix of DOA estimates which contain multiple signals with multiple types of unknown parameters.*

The contributions of this thesis are:

1. Establishing the relationship between the performance of DOA estimates based on Data Model(2) which is a simplified model, and that based on Data Model(1) which is a practical model.

2. Proving that MLE of DOA parameters based on Data Model(2) is asymptotically efficient.

3. Giving a profound interpretation of the logic implied in Ziv-Zakai's idea. Applying this logic to derive a lower bound on the variance of DOA estimate based on Data Model(2) for one incoming signal. This is a basic step to show how Ziv-Zakai's idea works for DOA estimation.

4. Developing the modified Ziv-Zakai lower bound (MZLB) on the covariance matrix of DOA estimates in white noise based on Data Model(1). MZLB is a tight lower bound in a wide range of SNR and easily computable.

5. Examining the SNR-threshold and $\Delta\phi$-threshold phenomena of CRLB and MZLB for DOA estimation.

6. Comparing the performances of the MLE of DOA with MZLB and CRLB by simulations. The simulation results show that MZLB is much tighter than CRLB under low

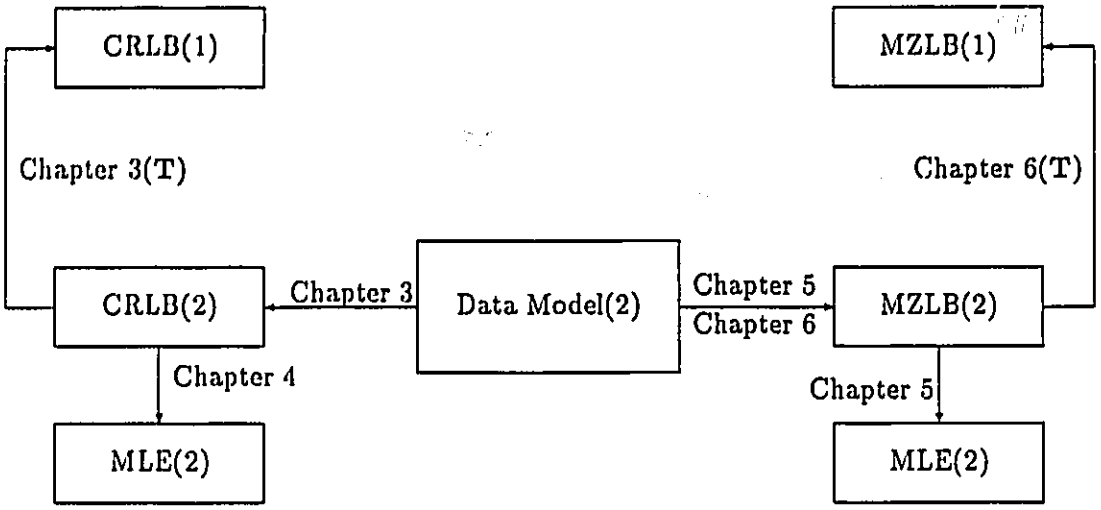


Figure 1.1: **The main structure of the thesis.**

SNR, and it follows the threshold phenomenon well.

Since we are interested in the lower bound on the covariance matrix of DOA estimates, some models used in the DOA estimation with array processing are introduced in Chapter 2.

The relationship among the main parts of this thesis is shown by Fig.1.1.

We start our research work from studying CRLB. In Chapter 3, we introduce CRLB(1), the CRLB based on Data Model(1), and discuss its properties. It is found that CRLB(1) shows the threshold phenomenon against the separations of DOA ($\Delta\phi$-threshold), but it cannot follow the threshold phenomenon against SNR (SNR-threshold). We establish Data Model(2), with which the DOA estimate errors is decoupled with the signal amplitude estimate errors such that the resulting CRLB(2), the CRLB based on Data Model(2), is independent of $\Delta\phi$. Also, we establish the relationship between CRLB(2) and CRLB(1) with a transformation T.

In Chapter 4, we prove that, in the asymptotic sense, the maximum-likelihood estimation of DOA parameters based on Data Model(2) (MLE(2)) is efficient, the errors of the MLE(2) is independent of each other and is independent of the locations of DOA. The discussion about the achievablity of CRLB(2) shows that, Data Model(2) is a basic data model to study the achievablity of a lower bound. Therefore, the work of MZLB is started with Data Model(2). With the simulations, we show that, the variance of MLE(1), the maximum-likelihood estimation of DOA based on Data Model(1) is a shift of the variance of MLE(2) approximately.

In Chapter 5, we give a profound interpretation of Ziv-Zakai's idea, modify the Ziv-Zakai's lower bound which is for the random parameter estimation, to the lower bound for the deterministic parameter estimation and develop the lower bound on variance of DOA estimate based on Data Model(2) for one incoming signal. An understanding of this case is fundamental to the understanding of the multiple incoming signal case.

In Chapter 6, first, we derive MZLB(2), the MZLB based on Data Model(2), for multiple incoming signals; then, we develop MZLB(1), the MZLB based on Data Model(1), on the covariance matrix of DOA estimates for multiple incoming signals with multiple types of unknown parameters from MZLB(2) with a transformation T. We discuss its properties. MZLB(1) is a tight lower bound over a wide range of SNR and is easily computable.

Chapter 7 gives the simulation results under the different conditions and some discussions.

Chapter 8 gives the summary, conclusions and the directions of some future research work.

Chapter 2

Preliminaries

2.1 INTRODUCTION

In this chapter, we briefly introduce the estimation of the directions of arrival in array processing, and define some models in Section 2.2. The notations used in this thesis are summarized in Section 2.3.

2.2 DOA ESTIMATION IN ARRAY PROCESSING

If we have signals coming from different directions, it is of interest, in many engineering areas, to find out the directions from which the signals arrive. To receive the incoming signals, we use sensors. A set of sensing elements placed in a known spatial pattern, in general, is referred to as an array. The task of array signal processing is to extract the information of interest from the received data which is the output of the array as shown in Fig. 2.1 in which a linear array of M uniformly spaced sensors is depicted. The sensors are all assumed to be identical isotropic elements. The distance between any two adjacent sensors is assumed to be $d_0 = \lambda/2$, where λ is the wavelength. This will be the array pattern used throughout this thesis.

Array signal processing has been applied in different disciplines such as radar, sonar, seismology, and biomedical engineering *etc.*. In general, the information to be extracted may be different in different contexts.

In this research work, we concentrate on the bearing estimation, i.e., the estimation of the directions of arrival of the signals. For simplicity, it is commonly referred to as DOA estimation.

In our study, we assume that K narrow band signals are generated from K distinct far field sources in space. The signals travel as plane waves from the sources toward the sensor array through a linear medium whose only effect on the signals is that of pure time delay. The wave front is shown by the lines l_i in Fig. 2.1. The signals received from the other sensors, are merely a delayed version of those received at the reference sensor (the origin 0).

Let $s_k(n)$ denote the k^{th} incoming signal from the direction α_k at the n^{th} snapshot as shown in Fig. 2.1, (where, $k = 1, ..., K, n = 1, ..., N$). The k^{th} signal arriving at the m^{th} sensor is,

$$\begin{aligned} s_{km}(n) &= a_k(n)e^{j(\frac{2\pi(m-1)d_0\sin(\alpha_k)}{\lambda})} \\ &= a_k(n)e^{j(m-1)\phi_k} \end{aligned} \tag{2.1}$$

where, $a_k(n)$ is the complex amplitude of the signal, and

$$\phi_k = \frac{2\pi d_0 \sin(\alpha_k)}{\lambda} = \pi \sin(\alpha_k) \tag{2.2}$$

since $d_0 = \lambda/2$.

The output of the m^{th} sensor is assumed to be a superposition of the K signals,

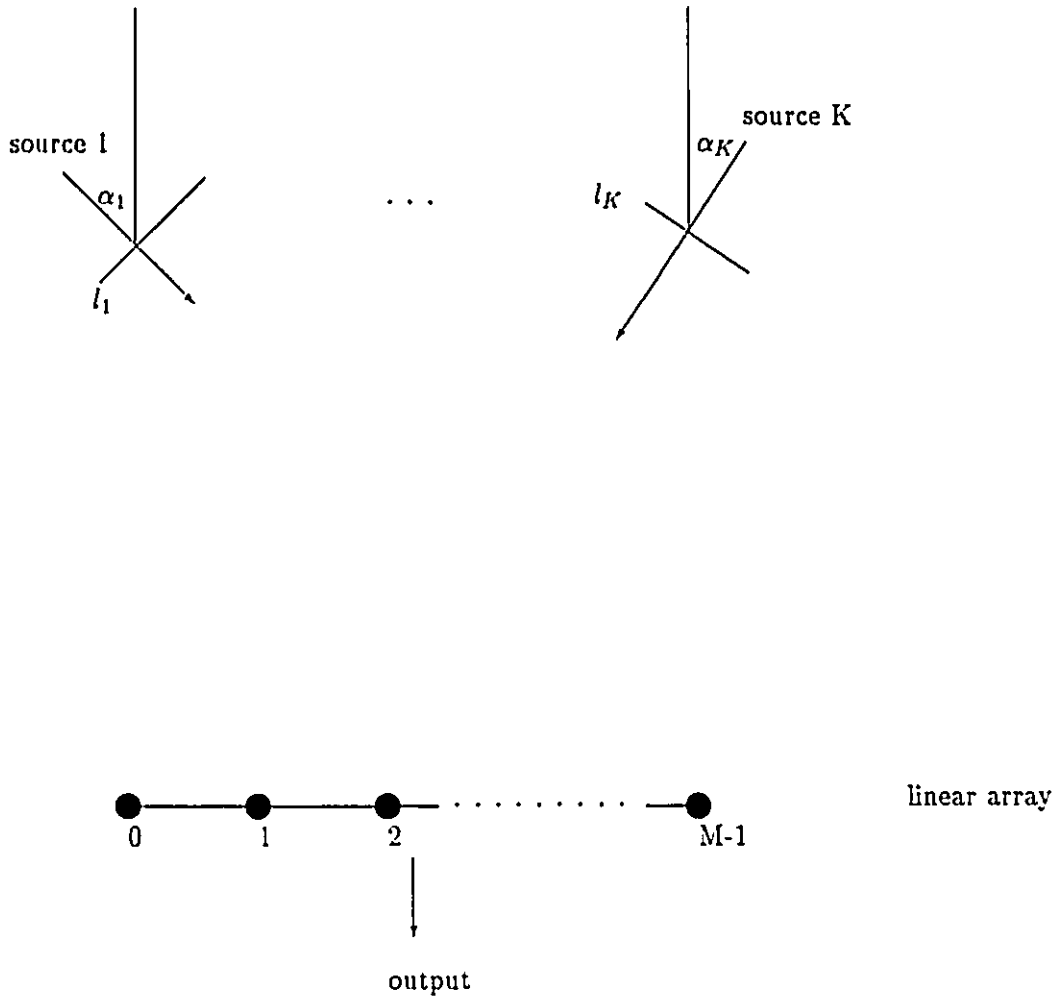


Figure 2.1: DOA estimation with array processing.

corrupted by the additive sensor noise,

$$r_m(n) = \sum_{k=1}^{K} s_{km}(n) + e_m(n), \tag{2.3}$$

where, $e_m(n)$ is the noise of the m^{th} sensor at the n^{th} snapshot.

Using the matrix notation, we have,

$$\begin{aligned} \mathrm{r}(n) &= \mathrm{s}(n) + \mathrm{e}(n) \\ &= \mathrm{D}(\vec{\phi})\mathrm{a}(n) + \mathrm{e}(n), \end{aligned} \tag{2.4}$$

where,

$$\mathrm{r}(n) = \begin{bmatrix} r_1(n) \\ r_2(n) \\ \vdots \\ r_M(n) \end{bmatrix}, \tag{2.5}$$

$$\vec{\phi} = \begin{bmatrix} \phi_1 \\ \phi_2 \\ \vdots \\ \phi_K \end{bmatrix}, \tag{2.6}$$

$$\mathrm{D}(\vec{\phi}) = [\mathrm{d}_1, ..., \mathrm{d}_K], \tag{2.7}$$

$$\mathrm{d}_k = \begin{bmatrix} 1 \\ e^{j\phi_k} \\ e^{j2\phi_k} \\ \vdots \\ e^{j(M-1)\phi_k} \end{bmatrix}, \tag{2.8}$$

$$\mathbf{a}(n) = \begin{bmatrix} a_1(n) \\ a_2(n) \\ \vdots \\ a_K(n) \end{bmatrix}, \tag{2.9}$$

and

$$\mathbf{e}(n) = \begin{bmatrix} e_1(n) \\ e_2(n) \\ \vdots \\ e_M(n) \end{bmatrix}. \tag{2.10}$$

Eq.(2.4) is the model of the observed data used in this thesis. There are three main problems associated with fitting models of Eq.(2.4) to the data $\mathbf{r}(n)$ [77].

1. Detection of the number of the signals K. Methods for this problem are well documented in the literature (see, e.g., [84] , [82], and [96]). In this thesis, we assume that the number of the signals K is known.

2. Estimation of the signal amplitude $\mathbf{a}(n)$. Once an estimate of $\vec{\phi}$ is available, the estimation of $\mathbf{a}(n)$ reduces to a simple linear estimation. We will not discuss estimating $\mathbf{a}(n)$ explicitly. However, the estimation of $\mathbf{a}(n)$ will implicitly appear in the following analysis. Note that, since it is required to estimate $\{\mathbf{a}(n)\}_{n=1}^{N}$ (not their " average characteristics," such as their covariance matrix), we will consider these variables as being deterministic. This assumption does not exclude that $\mathbf{a}(n)$ are samples from a random process. Thus, the distributional results derived in this thesis should be interpreted as being conditional on $\{\mathbf{a}(n)\}_{n=1}^{N}$.

3. Estimation of DOA $\vec{\phi}$ from the output data $\mathbf{r}(n)$ of array sensors. From Eq.(2.8), we see that the signal is a nonlinear function of the DOA parameter ϕ_k. DOA estimation

(e.g. MLE in Chapter 7) is nonlinear.

This thesis deals with the performance analysis of DOA estimates.

In addition, we make the following assumptions on the received data:

1. The signals and the noise are independent of each other. They are ergodic.
2. The noise from the sensors are Gaussian and white spatially and temporally. They have equal power (σ_e^2).
3. DOA estimate is unbiased. Thus, in this thesis, the mean-square error (MSE) of an estimate and the variance (VAR) of an estimate are the same.

2.3 NOTATIONAL CONVENTIONS

The following notational conventions will be used in this thesis.

M: the number of the sensors.

N: the number of the snapshots.

K: the number of the signals.

σ_{sk}^2: the power of the k^{th} signal.

σ_e^2: the power of noise.

$\rho_k = \frac{\sigma_{sk}^2}{\sigma_e^2}$: signal-noise ratio (SNR) for the k^{th} received signal.

bold lower-case character or character under arrow: vector.

bold upper-case character: matrix.

R_x: covariance matrix or correlation matrix of x.

I_K: K dimensions identity matrix.

$\Re(\mathrm{R})$: the real part of the complex matrix R.

$\Im(\mathrm{R})$: the image part of the complex matrix R.

$C^{M\times K}$: (MxK) dimension space.

a_i: the i^{th} vector in matrix A.

$A[i,j]$: the $(i,j)^{th}$ element in matrix A.

$\dot{\mu}$: the derivative of μ.

$\ddot{\mu}$: the second derivative of μ.

$|x|$: the absolute value of x.

det: determinant.

diag: diagonal matrix.

H: conjugate transpose.

$*$: conjugate.

Tr: trace.

E: expectation.

$p(x)$: probability density of random variable x.

$P(x > h)$: probability of $x > h$.

$\mathrm{A} \odot \mathrm{B}$: the Hadamard product of $\mathrm{A} \in C^{M\times K}$ and $\mathrm{B} \in C^{M\times K}$, i.e.,

$$[\mathrm{A} \odot \mathrm{B}][i,j] = A[i,j]B[i,j]. \tag{2.11}$$

complement error function:

$$\text{erfc\_}\{x\} = \frac{1}{\sqrt{2\pi}} \int_x^\infty e^{-t^2/2} dt. \tag{2.12}$$

Delta function:

$$\delta_{il} = \begin{cases} 1 & i = l \\ 0 & i \neq l. \end{cases} \tag{2.13}$$

Chapter 3

CRLB on the Covariance Matrix of DOA Estimates with Different Data Models

3.1 INTRODUCTION

Cramer-Rao lower bound(CRLB) on the covariance matrix of estimates is the most commonly used lower bound. For DOA estimation, CRLB has received significant attention and has been evaluated under different signal models and different environments. (see, e.g., [61][73] and [92] *etc.*)

Two data models for DOA estimation are considered in this thesis. These models introduce some assumptions on the data, which are helpful in studying lower bounds. Different data models result in different lower bounds.

$B_{CR}(1)$, the CRLB on the covariance matrix of DOA estimates based on Data Model(1), is derived in [77], in which SNR is coupled with the separations of DOA($\Delta\phi$). This complicates the performance analysis. We establish Data Model(2) in this chapter in order to decouple SNR with $\Delta\phi$ in CRLB performance. The resulting $B_{CR}(2)$, the CRLB

on the covariance matrix of DOA estimates based on Data Model(2), is independent of $\Delta\phi$. Data Model(2) provides a basic data model to study the achievablity of CRLB. We also show that, $B_{CR}(1)$ can be obtained from $B_{CR}(2)$ with a transformation T which is the function of $\Delta\phi$ and is independent of SNR.

CRLB and Fisher's information matrix are introduced in Section 3.2. Data Model(1) and $B_{CR}(1)$ are described in Section 3.3. It is indicated that $B_{CR}(1)$ is a function of SNR as well as the separations of DOA, and $B_{CR}(1)$ shows the $\Delta\phi$-threshold phenomenon. In Section 3.4, $B_{CR}(2)$ is derived based on Data Model(2). It is indicated that $B_{CR}(2)$ is independent of the separations of DOA. This will simplify the analysis of the achievablity of CRLB in Chapter 4. In Section 3.5, the relationship between $B_{CR}(2)$ and $B_{CR}(1)$ is discussed, and a transformation T between them is established. It is shown that T is dominated by the separations of DOA, the number of sensors, and is independent of the signal-noise ratio.

3.2 CRAMER-RAO LOWER BOUND

Cramer-Rao bound is a lower bound on the covariance matrix R_θ of estimates. It is quite well known and easy to apply.

Let $\vec{\theta}$ be an unknown parameter vector, θ_i be the i^{th} element in $\vec{\theta}$, and $\hat{\theta}$ be the estimate of $\vec{\theta}$. The covariance matrix of $\hat{\theta}$ is

$$R_\theta = E[(\hat{\theta} - \vec{\theta})(\hat{\theta} - \vec{\theta})^T]. \tag{3.1}$$

CRLB on the covariance matrix R_θ is given by [81]

$$B_{CR}^{\theta} = J^{-1}, \tag{3.2}$$

where, J is the commonly called Fisher's information matrix:

$$\mathrm{J} = E(\vec{\psi}\vec{\psi}^T), \tag{3.3}$$

in which,

$$\vec{\psi} \triangleq \frac{\partial \ln L}{\partial \vec{\theta}}. \tag{3.4}$$

Here, ln L is the log-likelihood function of the observation data.

Equivalently, the $(i,l)^{th}$ element in J can be obtained by

$$J[i,l] = -E(\frac{\partial^2}{\partial\theta_i\partial\theta_l} \ln L), \tag{3.5}$$

where, $E(x)$ is the expectation of x.

The following conditions are assumed to be satisfied for Eq.(3.2):

1. $\hat{\theta}$ is any unbiased estimate of $\vec{\theta}$;

2.
$$\frac{\partial^2 p(\mathrm{r}(n)/\vec{\theta})}{\partial\theta_i\partial\theta_l} \tag{3.6}$$

and
$$\frac{\partial p(\mathrm{r}(n)/\vec{\theta})}{\partial\theta_i} \tag{3.7}$$

exist and are absolutely integrable.

In DOA estimation, the likelihood function of the observed data described in Chapter 2 is given by

$$\begin{aligned} L &= \prod_{n=1}^{N} p(\mathrm{r}(n)/\vec{\phi}) \\ &= \frac{1}{(\det[\pi\mathrm{R}_e])^N} \exp\{-\sum_{n=1}^{N}[\mathrm{r}(n)-\mathrm{s}(n)]^H\mathrm{R}_e^{-1}[\mathrm{r}(n)-\mathrm{s}(n)]\} \end{aligned}$$

$$= \frac{1}{(\pi\sigma_e^2)^{NM}} \exp\{-\frac{1}{\sigma_e^2}\sum_{n=1}^{N}[\mathrm{r}(n)-\mathrm{s}(n)]^H[\mathrm{r}(n)-\mathrm{s}(n)]\}, \tag{3.8}$$

since the noise covariance matrix is given by

$$\mathrm{R}_e = E\{\mathrm{e}(n)\mathrm{e}^H(n)\} = \sigma_e^2\mathrm{I}.$$

Taking logarithms,

$$\ln L = -NM\ln(\pi\sigma_e^2) - \frac{1}{\sigma_e^2}\sum_{n=1}^{N}[\mathrm{r}(n)-\mathrm{s}(n)]^H[\mathrm{r}(n)-\mathrm{s}(n)], \tag{3.9}$$

where,

$$\mathrm{s}(n) = \mathrm{D}(\vec{\phi})\mathrm{a}(n).$$

This log-likelihood function involves three types of parameters: the DOA ($\vec{\phi}$), the signal amplitudes (a), and the power of the noise (σ_e^2). Since there are K arriving signals, there are K DOA parameters to be estimated.

Different assumptions may be made on the number of the types of unknown parameters and the number of the incoming signals, which result in the different data models. The CRLB so derived are different with the different data models.

3.3 CRLB(1)

3.3.1 Data Model(1)

We make the following assumptions on the received data described in Chapter 2:

1. The sequence $\{\mathrm{a}(n)\}_{n=1}^{N}$ is frozen in all realizations of the random data $\{\mathrm{r}(n)\}_{n=1}^{N}$; $\{\mathrm{e}(n)\}_{n=1}^{N}$ varies from realization to realization. The received data $\{\mathrm{r}(n)\}$, which is conditional on $\{\mathrm{a}(n)\}$, have a Gaussian distribution:

$$\{\mathrm{r}(n)\} \sim N(\mathrm{D}(\vec{\phi})\mathrm{a}(n), \sigma_e^2\mathrm{I}). \tag{3.10}$$

2. The unknown parameter vector is

$$\vec{\theta} = \{\vec{\phi}^T, \Re[\mathrm{a}^T(1)], \Im[\mathrm{a}^T(1)], ..., \Re[\mathrm{a}^T(N)], \Im[\mathrm{a}^T(N)], \sigma_e^2\}^T, \tag{3.11}$$

in which, $\vec{\phi}$ is the DOA parameter vector

$$\vec{\phi} = [\phi_1, ..., \phi_K]^T,$$

$\Re[\mathrm{a}(n)]$ is the real part of the K dimension vector $\mathrm{a}(n)$. $\Im[\mathrm{a}(n)]$ is the image part of the K dimension vector $\mathrm{a}(n)$. σ_e^2 is the power of the white noise. Therefore $\vec{\theta}$ is a (K+2NK+1)-dimension vector.

Note that, here, the number of the arriving signals is K, and there are three types of unknown parameters. This is referred to as the *m*ultiple incoming *s*ignals with *m*ultiple *t*ypes of unknown parameters (MSMT) in this thesis.

3. The incoming signals are uncorrelated, and for large N,

$$\sum_{n=1}^{N} a_k^*(n) a_l(n) \simeq \begin{cases} N\sigma_{sk}^2 & l = k \\ 0 & l \neq k. \end{cases} \tag{3.12}$$

The data model with these assumptions is called Data Model(1). This data model is referred to as the "conditional model" introduced in [77]. We denote CRLB for DOA estimation based on Data Model(1) as CRLB(1).

3.3.2 CRLB(1)

The work introduced in this section is developed by Stoica and Nehorai [77].

Since there are K unknown DOA, CRLB(1) on the covariance matrix of DOA estimates based on Data Model(1) is a K×K matrix. It is obtained from the Fisher's information matrix J which is a (K+2NK+1)×(K+2NK+1) matrix.

Substituting Eq.(3.11) into Eq.(3.4) yields

$$\vec{\psi}^T = \frac{\partial \ln L}{\partial\{\vec{\phi}^T, \Re[\mathbf{a}^T(1)], \Im[\mathbf{a}^T(1)], ..., \Re[\mathbf{a}^T(N)], \Im[\mathbf{a}^T(N)], \sigma_e^2\}}. \tag{3.13}$$

Then the Fisher's information matrix can be obtained such that

$$\mathrm{J} = E(\vec{\psi}\vec{\psi}^T), \tag{3.14}$$

in which, (in the following equations, $n1$ and $n2$ mean snapshots)

$$E[\frac{\partial \ln L}{\partial \vec{\phi}} \frac{\partial \ln L}{\partial \vec{\phi}^T}] = \frac{2}{\sigma_e^2} \sum_{n=1}^{N} \Re[\mathbf{X}^H(n)\dot{\mathbf{D}}^H \dot{\mathbf{D}}\mathbf{X}(n)] \triangleq \Gamma, \tag{3.15}$$

$$E[\frac{\partial \ln L}{\partial \Re[\mathbf{a}(n1)]} \frac{\partial \ln L}{\partial \Re[\mathbf{a}(n2)]^T}] = \frac{2}{\sigma_e^2} \Re[\mathbf{D}^H \mathbf{D}]\delta_{n1,n2} \triangleq \Re[\mathrm{H}], \tag{3.16}$$

$$E[\frac{\partial \ln L}{\partial \Re[\mathbf{a}(n1)]} \frac{\partial \ln L}{\partial \Im[\mathbf{a}(n2)]^T}] = -\frac{2}{\sigma_e^2} \Im[\mathbf{D}^H \mathbf{D}]\delta_{n1,n2} \triangleq -\Im[\mathrm{H}], \tag{3.17}$$

$$E[\frac{\partial \ln L}{\partial \Im[\mathbf{a}(n1)]} \frac{\partial \ln L}{\partial \Im[\mathbf{a}(n2)]^T}] = \frac{2}{\sigma_e^2} \Re[\mathbf{D}^H \mathbf{D}]\delta_{n1,n2} \triangleq \Re[\mathrm{H}], \tag{3.18}$$

$$E[\frac{\partial \ln L}{\partial \Re[\mathbf{a}(n)]} \frac{\partial \ln L}{\partial \vec{\phi}^T}] = \frac{2}{\sigma_e^2} \Re[\mathbf{D}^H \dot{\mathbf{D}}\mathbf{X}(n)] \triangleq \Re[\mathbf{G}(n)], \tag{3.19}$$

$$E[\frac{\partial \ln L}{\partial \Im[\mathbf{a}(n)]} \frac{\partial \ln L}{\partial \vec{\phi}^T}] = \frac{2}{\sigma_e^2} \Im[\mathbf{D}^H \dot{\mathbf{D}}\mathbf{X}(n)] \triangleq \Im[\mathbf{G}(n)], \tag{3.20}$$

and

$$E[\frac{\partial \ln L}{\partial \sigma_e^2} \frac{\partial \ln L}{\partial \theta_i}] = \begin{cases} \frac{MN}{\sigma_e^4} & \text{if } \theta_i = \sigma_e^2 \\ 0 & \text{otherwise,} \end{cases} \tag{3.21}$$

where,

$$\dot{\mathrm{D}} = [\dot{\mathrm{d}}_1, ..., \dot{\mathrm{d}}_K], \tag{3.22}$$

$$\dot{\mathrm{d}}_i = \frac{\partial \mathrm{d}(\phi_i)}{\partial \phi_i}, \tag{3.23}$$

$$\mathrm{X}(n) = \mathrm{diag}(a_1(n), ..., a_K(n)). \tag{3.24}$$

Therefore, the CRLB on the covariance matrix of the DOA estimates is given by

$$\mathrm{B}_{CR} = [\Gamma - \sum_{n=1}^{N} \Re\{\mathrm{G}^H(n)\mathrm{H}^{-1}\mathrm{G}(n)\}]^{-1} \tag{3.25}$$

$$= [\frac{2}{\sigma_e^2} \sum_{n=1}^{N} \Re\{\mathrm{X}^H(n)\dot{\mathrm{D}}^H[\mathrm{I} - \mathrm{D}(\mathrm{D}^H\mathrm{D})^{-1}\mathrm{D}^H]\dot{\mathrm{D}}\mathrm{X}(n)\}]^{-1}. \tag{3.26}$$

Let $\mathrm{R}_a = E(\mathrm{a}\mathrm{a}^H)$ denote the signal correlation matrix. Under the assumption Eq.(3.12), $\sum_{n=1}^{N} \mathrm{X}(n)\mathrm{X}(n) \simeq N\mathrm{diag}(\sigma_{s1}^2, ..., \sigma_{sK}^2) = N\mathrm{R}_a$. Thus Eq.(3.26) becomes

$$\mathrm{B}_{CR} = \frac{\sigma_e^2}{2N}[\Re\{(\dot{\mathrm{D}}^H[\mathrm{I} - \mathrm{D}(\mathrm{D}^H\mathrm{D})^{-1}\mathrm{D}^H]\dot{\mathrm{D}}) \odot \mathrm{R}_a^H\}]^{-1}$$

$$= \mathrm{diag}(\frac{1}{2\rho_1 N\dot{\mathrm{d}}_1^H\mathrm{P}_e\dot{\mathrm{d}}_1}, ..., \frac{1}{2\rho_K N\dot{\mathrm{d}}_K^H\mathrm{P}_e\dot{\mathrm{d}}_K}), \tag{3.27}$$

where, $\odot$ is the Hadamard product which is defined in Chapter 2,

$$\rho_k = \frac{\sigma_{sk}^2}{\sigma_e^2} \tag{3.28}$$

is the signal-noise ratio (SNR) of the k^{th} received signal, and

$$\mathrm{P}_e = \mathrm{I} - \mathrm{D}(\mathrm{D}^H\mathrm{D})^{-1}\mathrm{D}^H \tag{3.29}$$

is the projection matrix on the noise space.

We call Eq.(3.27) $B_{CR}(1)$, since it is based on Data Model(1).

In Eq.(3.27), the term $\dot{d}_k^H P_e \dot{d}_k$ is a function of the separations of DOA. Therefore $B_{CR}(1)$ is a function of SNR as well as the separations of DOA.

For $M \rightarrow \infty$, it has been proved in [77] that, Eq.(3.27) becomes

$$B_{CR} \simeq \text{diag}(\frac{6}{\rho_1 N M^3}, ..., \frac{6}{\rho_K N M^3}). \tag{3.30}$$

For one incoming signal (i.e. K=1), Eq.(3.27) becomes

$$B_{CR} = \frac{6}{\rho_1 N M(M^2-1)}. \tag{3.31}$$

Note that, in Eq.(3.31), the number of the incoming signal is one, but the types of the unknown parameters are multiple. This is referred to as the *s*ingle incoming *s*ignal with *m*ultiple *t*ypes of unknown parameters (SSMT) in this thesis.

3.3.3 CRLB(1) and Threshold Phenomena

In the DOA estimation, the performance against SNR and the performance against the DOA separations ($\Delta\phi$) are of interest. There exist two types of threshold phenomenon. In a graph of the variance against $\Delta\phi$, when the DOA separations decreases, if there exists a threshold γ, such that when $\Delta\phi \leq \gamma$, the variance of DOA estimate suddenly increases very fast, then we call this the $\Delta\phi$-threshold phenomenon. Similarly, in a graph of the variance against SNR, if there exists a threshold ρ, such that when $SNR < \rho$, the variance of DOA estimate suddenly increases very fast, then we call this the SNR-threshold phenomenon. A tight lower bound on the variance is expected to show these threshold phenomena. In the following, we examine CRLB(1).

$B_{CR}^{11}(1)$, the CRLB(1) on the variance of the first DOA estimate, against $\Delta\phi$ and against SNR are plotted in Fig.3.1 and Fig.3.2, where, the number of sensors M=8, the

number of snapshots N=50, the number of incoming signals K=2.

In Fig.3.1(a), $B^{11}_{CR}(1)$ is plotted against $\Delta\phi$ with SNR=-10 dB. It is shown that, when $\Delta\phi > \gamma = 2\pi/M$, $B^{11}_{CR}(1)$ is almost unchanged with the different $\Delta\phi$; when $\Delta\phi \leq \gamma$, $B^{11}_{CR}(1)$ increases very fast with decreasing $\Delta\phi$. In Fig.3.1(b), $B^{11}_{CR}(1)$ is plotted against $\Delta\phi$ with SNR=+10 dB. The same observations persist. Therefore, CRLB(1) shows the $\Delta\phi$-threshold phenomenon. Comparing (a) and (b) in Fig.3.1, we see that, the value of the threshold γ is independent of the value of SNR, since it is caused by the term $\dot{\mathbf{d}}_k^H \mathbf{P}_e \dot{\mathbf{d}}_k$ in Eq.(3.27) which is independent of SNR. But CRLB(1) against $\Delta\phi$ with SNR=10 dB is a parallel shift of that with SNR=-10 dB.

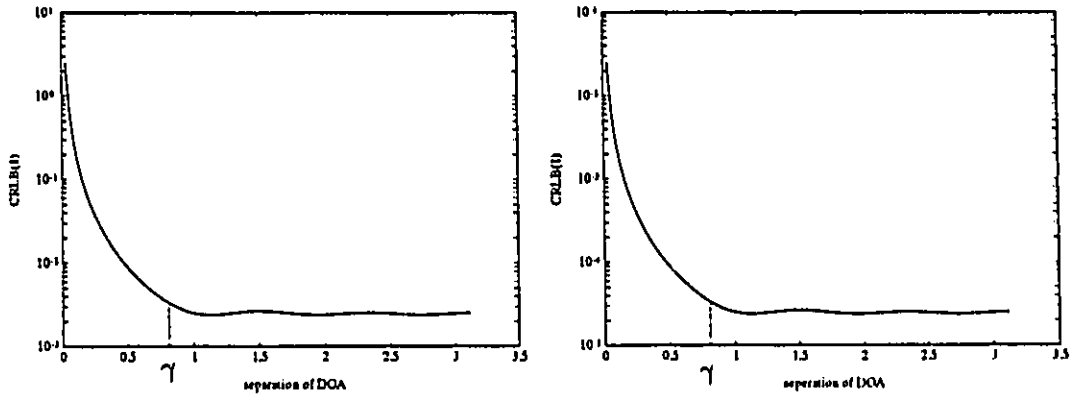
In Fig.3.2(a), $B^{11}_{CR}(1)$ is plotted against SNR with $\Delta\phi = 0.5$ radian. It is shown that, $B^{11}_{CR}(1)$ is a linear function of SNR. In Fig.3.2(b), $B^{11}_{CR}(1)$ is plotted against SNR with $\Delta\phi = 2.0$ radian. The same observations persist. Obviously, CRLB(1) does not show the SNR-threshold phenomenon. Comparing (a) and (b) in Fig.3.2, we see that, the CRLB(1) against SNR with $\Delta\phi = 2.$ radian is a parallel shift of that with $\Delta\phi = 0.5$ radian.

Therefore, we have,

Observation 3.1 *CRLB(1) shows the $\Delta\phi$-threshold phenomenon. The value of the threshold γ is independent of SNR.*

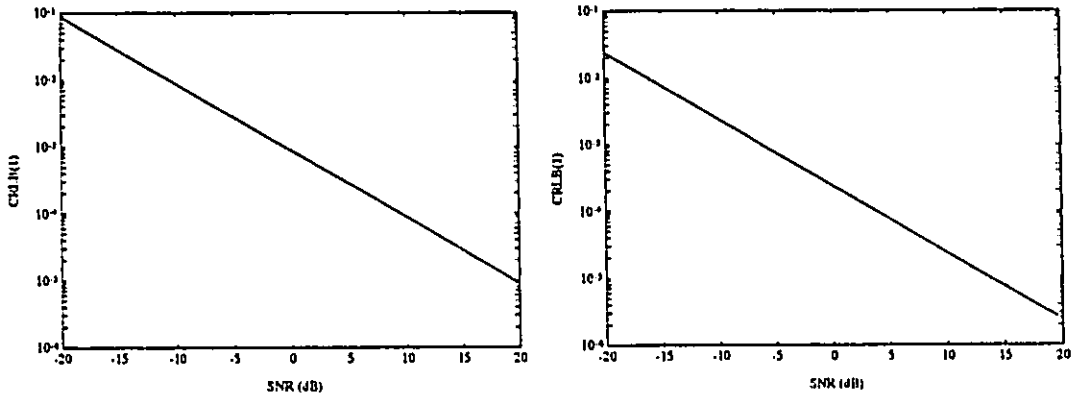
Observation 3.2 *CRLB(1) does not show the SNR-threshold phenomenon. It is a linear function of SNR for a particular value of $\Delta\phi$.*

Observation 3.3 *In CRLB(1), $\Delta\phi$ and SNR are coupled with each other in the way that, CRLB(1) against $\Delta\phi$ with different SNR are parallel to each other, or, CRLB(1) against SNR with different $\Delta\phi$ are parallel to each other.*



(a). SNR=-10 dB. (b). SNR=+10 dB.

Figure 3.1: $B_{CR}^{11}(1)$ against $\Delta\phi$.



(a). $\Delta\phi = 0.5$ radian (b). $\Delta\phi = 2.$ radian

Figure 3.2: $B_{CR}^{11}(1)$ against SNR.

3.4 CRLB(2)

3.4.1 Data Model(2)

We make the following assumptions on the received data

$$\mathbf{r}(n) = \mathbf{D}(\vec{\phi})\mathbf{a}(n) + \mathbf{e}(n),$$

1. The sequence $\{\mathbf{a}(n)\}_{n=1}^{N}$ is frozen in all realizations of the random data $\{\mathbf{r}(n)\}_{n=1}^{N}$; $\{\mathbf{e}(n)\}_{n=1}^{N}$ varies from realization to realization. The received data $\{\mathbf{r}(n)\}$, which is conditional on $\{\mathbf{a}(n)\}$, belongs to Gaussian distribution:

$$\{\mathbf{r}(n)\} \sim N(\mathbf{D}(\vec{\phi})\mathbf{a}(n), \sigma_e^2\mathbf{I}). \tag{3.32}$$

2. The unknown parameter vector is

$$\vec{\theta} = \{\vec{\phi}\}, \tag{3.33}$$

where, $\vec{\phi} = [\phi_1, ..., \phi_K]^T$ is a K-dimensional DOA parameter vector, i.e., only the DOA of the signals are unknown.

Note that, here, the number of incoming signals is K, and the type of the unknown parameter is one. This is referred to as the *m*ultiple incoming *s*ignals with the *s*ingle *t*ype of unknown parameters (MSST) in this thesis.

3. The incoming signals are uncorrelated, and for large N,

$$\sum_{n=1}^{N} a_k^*(n)a_l(n) \simeq \begin{cases} N\sigma_{sk}^2 & l = k \\ 0 & l \neq k. \end{cases} \tag{3.34}$$

The data model with these assumptions is called Data Model(2). Data Model(2) is different from Data Model(1) on the number of the types of the unknown parameters, i.e.,

apart from the DOA of the signals, the parameters of signal amplitudes and noise power are assumed to be known. We denote CRLB for DOA estimation based on Data Model(2) as CRLB(2).

3.4.2 CRLB(2)

Substituting Eq.(3.33) into Eq.(3.4) yields

$$\vec{\psi} = \frac{\partial \ln L}{\partial \vec{\phi}}. \tag{3.35}$$

The Fisher's information matrix with Data Model(2)

$$\begin{aligned} \mathrm{J} &= E(\vec{\psi}\vec{\psi}^T) \\ &= E[\frac{\partial \ln L}{\partial \vec{\phi}} \frac{\partial \ln L}{\partial \vec{\phi}^T}] \\ &= \Gamma. \end{aligned} \tag{3.36}$$

The last step in Eq.(3.36) is from Eq.(3.15), and from that,

$$\Gamma = \frac{2}{\sigma_e^2} \sum_{n=1}^{N} \Re[\mathrm{X}^H(n)\dot{\mathrm{D}}^H\dot{\mathrm{D}}\mathrm{X}(n)] \tag{3.37}$$

$$= \frac{2N}{\sigma_e^2} \Re[\dot{\mathrm{D}}^H\dot{\mathrm{D}} \odot \mathrm{R}_a] \tag{3.38}$$

$$= \frac{2N}{\sigma_e^2} \frac{M(M-1)(2M-1)}{6} \mathrm{diag}(\sigma_{s1}^2, ..., \sigma_{sK}^2) \tag{3.39}$$

$$= \mathrm{diag}(\frac{N\rho_1 M(M-1)(2M-1)}{3}, ..., \frac{N\rho_K M(M-1)(2M-1)}{3}), \tag{3.40}$$

where, Eq.(3.34) has been used from Eq.(3.37) to Eq.(3.38), $\mathrm{R}_a = \mathrm{diag}(\sigma_{s1}^2, ..., \sigma_{sK}^2)$ is the signal correlation matrix, and the $(k,k)^{th}$ element of $\dot{\mathrm{D}}^H\dot{\mathrm{D}}$ is

$$\dot{\mathrm{d}}_k^H\dot{\mathrm{d}}_k = \sum_{m=1}^{M} [j(m-1)e^{j(m-1)\phi_k}]^*[j(m-1)e^{j(m-1)\phi_k}]$$

$$= \sum_{m=1}^{M}(m-1)^2$$

$$= \frac{M(2M-1)(M-1)}{6}.$$

CRLB on the covariance matrix of the estimated $\hat{\phi}$ based on Data Model(2) is the inverse of the Fisher's information matrix directly,

$$\mathrm{B}_{CR}(2) = \Gamma^{-1} = \mathrm{diag}\left(\frac{3}{N\rho_1 M(M-1)(2M-1)}, \cdots, \frac{3}{N\rho_K M(M-1)(2M-1)}\right). \quad (3.41)$$

We call Eq.(3.41) $\mathrm{B}_{CR}(2)$ since it is based on Data Model(2).

$\mathrm{B}_{CR}(2)$ is independent of the separations of DOA. Data Model(2) provides a basic model to study the achievablity of CRLB. This will be examined in Chapter 4.

For one incoming signal, Eq.(3.41) becomes

$$B_{CR}(2) = \frac{3}{N\rho_1 M(M-1)(2M-1)}. \quad (3.42)$$

Note that, here, there is only one incoming signal, and one type of the unknown parameter. This is referred to as the *s*ingle incoming *s*ignal with *s*ingle *t*ype of unknown parameter (SSST) in this thesis.

3.4.3 CRLB(2) and Threshold Phenomena

With the same examples as used in Section 3.3.3, $B_{CR}^{11}(2)$, the CRLB(2) on the variance of the first DOA estimate, against $\Delta\phi$ and against SNR are plotted in Fig.3.3 and Fig.3.4.

In Fig.3.3, $B_{CR}^{11}(2)$ is plotted against $\Delta\phi$ with SNR=-10 dB. It is shown that, $B_{CR}^{11}(2)$ does not vary with $\Delta\phi$. This is obvious since $B_{CR}^{11}(2)$ in Eq.(3.41) is independent of $\Delta\phi$.

In Fig.3.4, $B_{CR}^{11}(2)$ is plotted against SNR with $\Delta\phi = 0.5$ radian. It is shown that,

$B^{11}_{CR}(2)$ is a linear function of SNR. Obviously, CRLB(2) does not show the SNR-threshold phenomenon.

Therefore, we have,

Observation 3.4 *CRLB(2) does not show the $\Delta\phi$-threshold phenomenon.*

Observation 3.5 *CRLB(2) does not show the SNR-threshold phenomenon. It is a linear function of SNR.*

Observation 3.6 *CRLB(2) against SNR is independent of $\Delta\phi$.*

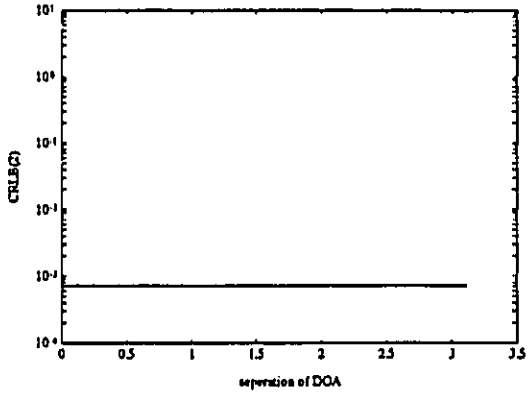


Figure 3.3: $B_{CR}^{11}(2)$ against $\Delta\phi$ (SNR=-10 dB).

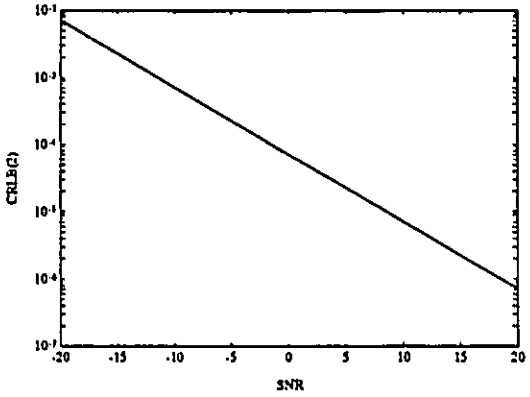


Figure 3.4: $B_{CR}^{11}(2)$ against SNR ($\Delta\phi = 0.5$ radian).

3.5 THE RELATIONSHIP BETWEEN $B_{CR}(2)$ AND $B_{CR}(1)$

The main difference between Data Model(1) and Data Model(2) is the number of the types of the unknown parameters.

$B_{CR}(1)$ is based on Data Model(1), the unknown parameter space includes the DOA as well as the signal amplitudes. Since the error of the noise power estimate is not coupled with these parameters, see Eq.(3.21), it is ignored here. In $B_{CR}(1)$ (Eq.(3.25)), G and H are related to the differential of the log-likelihood function with respect to the amplitude $a(n)$, and Γ is from the differential of the log-likelihood function with respect to the DOA $\vec{\phi}$. $B_{CR}(2)$ is based on data model(2), the unknown parameter space only includes the DOA. In the Fisher's matrix corresponding to Data Model(2) (Eq.(3.36)), Γ is obtained from the differential of the log-likelihood function with respect to the DOA $\vec{\phi}$, and does not involve G and H, i.e., not involving the signal amplitudes $a(n)$ since the signal amplitudes are assumed to be known.

Because Γ is independent of the separations of DOA,

$$B_{CR}(2) = \Gamma^{-1} \tag{3.43}$$

is independent of the separations of the DOA. Because G, H are related to the separations of DOA,

$$B_{CR}(1) = [\Gamma - \sum_{n=1}^{N} \Re\{G^H(n)H^{-1}G(n)\}]^{-1} \tag{3.44}$$

is a function of the separations of the DOA.

From Eqs.(3.27) and (3.41), we see that, both $B_{CR}(1)$ and $B_{CR}(2)$ are diagonal matrices, their diagonal elements are positive. So they are full rank. The transformation matrix between them is given by

$$T = B_{CR}(1)B_{CR}^{-1}(2)$$

$$= \text{diag}\left(\frac{M(M-1)(2M-1)}{6\dot{\mathbf{d}}_1^H \mathbf{P}_e \dot{\mathbf{d}}_1}, \ldots, \frac{M(M-1)(2M-1)}{6\dot{\mathbf{d}}_K^H \mathbf{P}_e \dot{\mathbf{d}}_K}\right). \tag{3.45}$$

For a given set of signals with the fixed DOA and the fixed number of sensors, the elements of T are constants independent of SNR and the number of snapshots.

Consequently, $\mathrm{B}_{CR}(1)$ can be viewed in another way: it is obtained from $\mathrm{B}_{CR}(2)$ with the transformation T. The usefulness of this transformation will be more apparent later on when the Ziv-Zakai lower bound for DOA estimation is derived.

The following graphs show the relationship between $\mathrm{B}_{CR}(1)$ and $\mathrm{B}_{CR}(2)$, which depends on the transformation T.

As examples, with K=2, M=8, and N=50, and the equal signal power, $B_{CR}^{11}(1)$, the lower bound on the variance of the first DOA estimate $\hat{\phi}_{01}$ with Data Model(1), and $B_{CR}^{11}(2)$, the lower bound on the variance of $\hat{\phi}_{01}$ with Data Model(2), against SNR under the different $\Delta\phi$ are plotted in Fig.3.5. The solid lines represent $B_{CR}^{11}(1)$, and the dash lines represent $B_{CR}^{11}(2)$. From which we see that:

1. $B_{CR}^{11}(1)$ is a parallel shift of $B_{CR}^{11}(2)$.

   This is because of that the transformation T is independent of SNR.

2. The distance of this parallel shift depends on $\Delta\phi$. When $\Delta\phi > 2\pi/M$, the distance of the parallel shift are almost the same for the different $\Delta\phi$, as shown in (a) and (b). When $\Delta\phi < 2\pi/M$, the distance of this parallel shift varies with $\Delta\phi$, the smaller the $\Delta\phi$, the larger the parallel shift, as shown in (c) and (d). This is because of that, the transformation T is dominated by $\Delta\phi$ and M.

Similar observations persist when $B_{CR}^{22}(1)$ and $B_{CR}^{22}(2)$, which are the lower bounds on the variances of the second DOA estimate based on Data Model(2), are examined. Therefore, we have

Observation 3.7 *As a function of SNR, CRLB(1) is a parallel shift of CRLB(2).*

With K=2, N=50, SNR=0 dB under different M, $B_{CR}^{11}(1)$ and $B_{CR}^{11}(2)$ against $\Delta\phi$ are plotted in Fig.3.6. The solid lines represent CRLB with Data Model(1), and the dash lines represent CRLB with Data Model(2). Here,

1. When $\Delta\phi < \gamma$, $B_{CR}^{11}(1)$ at point γ suddenly increases with $\Delta\phi$ decreasing. $B_{CR}^{11}(2)$ dose not.

2. The $\Delta\phi$-threshold γ in $B_{CR}^{11}(1)$ is different with the different M as shown in plots (a), (b), and (c).

3. When $\Delta\phi > \gamma$, $B_{CR}^{11}(1)$ is approximately equal to $B_{CR}^{11}(2)$ times a constant. The dash-dotted lines are obtained from multiplying $B_{CR}^{11}(2)$ by a constant τ. τ is given by the ratio of $B_{CR}(1)$ and $B_{CR}(2)$,

$$\begin{aligned}\tau &= \frac{B_{CR}(1)}{B_{CR}(2)} \\ &= \frac{6M(M-1)(2M-1)}{3M(M-1)(M+1)} \\ &= \frac{2(2M-1)}{M+1},\end{aligned}$$

where, $B_{CR}(1)$ and $B_{CR}(2)$ are from Eqs.(3.31) and Eq.(3.42). The plots show that, for $\Delta\phi > \gamma$,

$$B_{CR}^{11}(1) \simeq \tau B_{CR}^{11}(2).$$

These are because of that, T is dominated by $\Delta\phi$ and M. And we have

Observation 3.8 *CRLB(2) does not show the $\Delta\phi$-threshold phenomenon. But CRLB(1), which may be obtained from CRLB(2) with a transformation* T, *shows the $\Delta\phi$-threshold phenomenon.*

With K=2, N=50, M=8 and different SNR, $B_{CR}^{11}(1)$ and $B_{CR}^{11}(2)$ against $\Delta\phi$ are plotted in Fig.3.7, where,

1. The $\Delta\phi$-threshold γ in $B_{CR}^{11}(1)$ is the same with the different SNR in plots (a), (b) and (c).

2. Plots (a), (b) and (c) are the same except the coordinate values on the vertical axis.

These are because of that, T is independent of SNR.

Therefore, combining Figs.3.5, 3.6 and 3.7, we have

Observation 3.9 *The transformation* T *is independent of SNR, and is dominated by* $\Delta\phi$ *and* M.

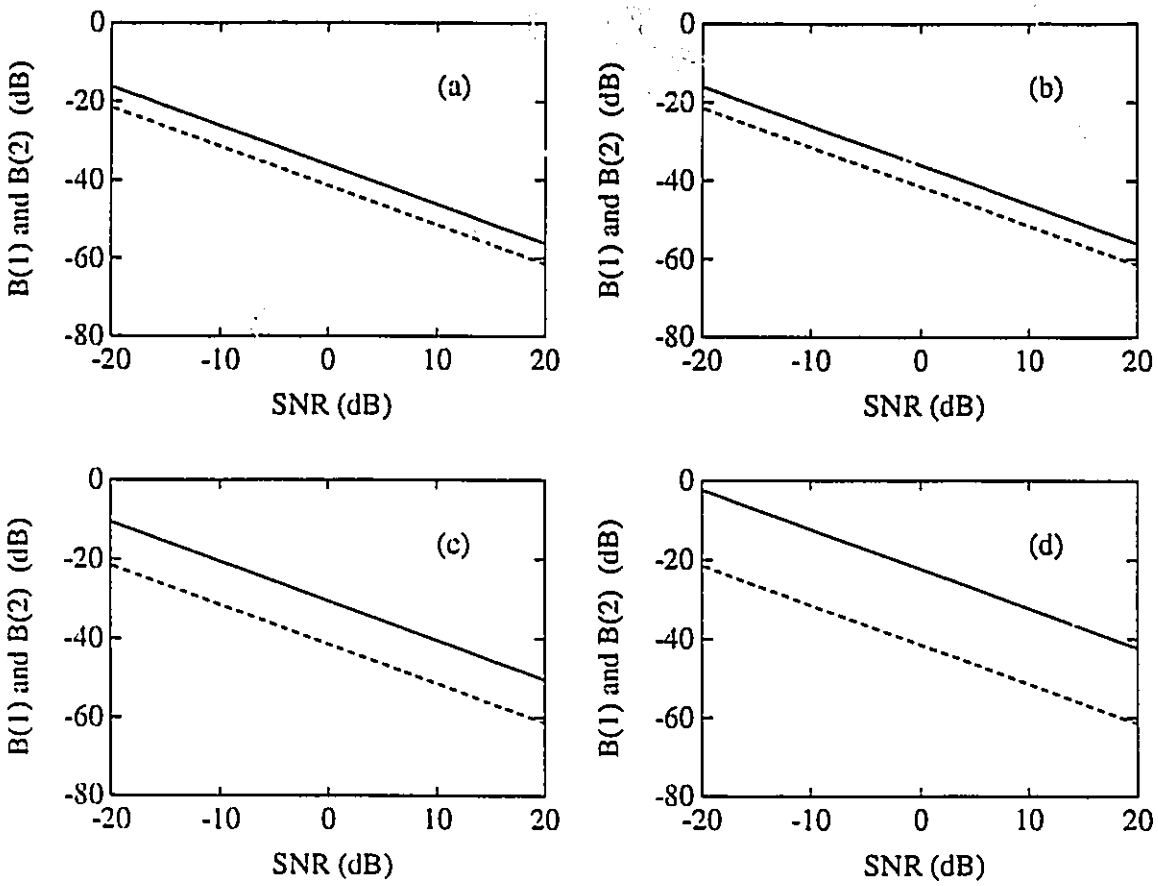
3.6 SUMMARY

In this chapter we have the following results:

1. In CRLB(1), $\Delta\phi$ and SNR are coupled with each other. CRLB(1) shows the $\Delta\phi$-threshold phenomenon, and does not show the SNR-threshold phenomenon.

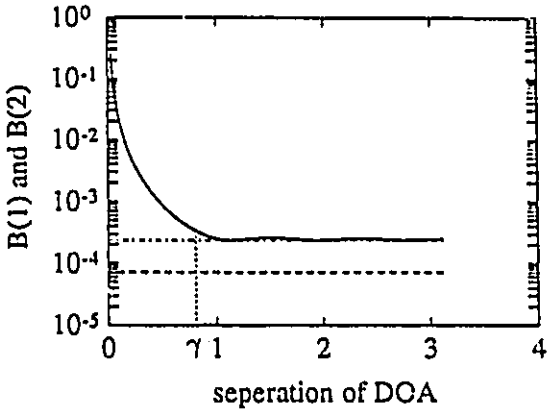
2. CRLB(2) is independent of $\Delta\phi$, and does not show the the SNR-threshold phenomenon.

3. CRLB(1) can be obtained from $B_{CR}(2)$ with a transformation T. This transformation is dominated by the separations of DOA and the number of sensors, and is independent of SNR.

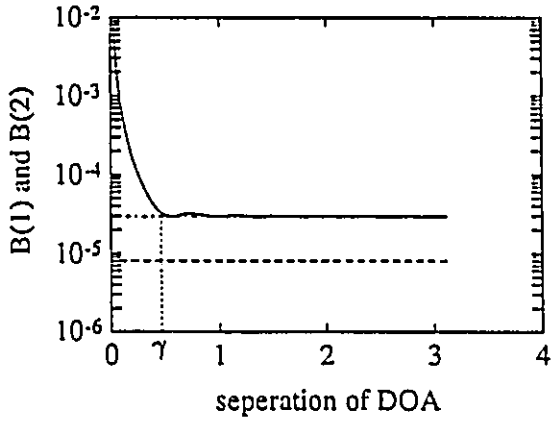


- - $B^{11}_{CR}(2)$ — $B^{11}_{CR}(1)$
M=8 N=50
(a). $\Delta\phi = 2.0$ (b). $\Delta\phi = 1.0$
(c). $\Delta\phi = 0.5$ (d). $\Delta\phi = 0.2$

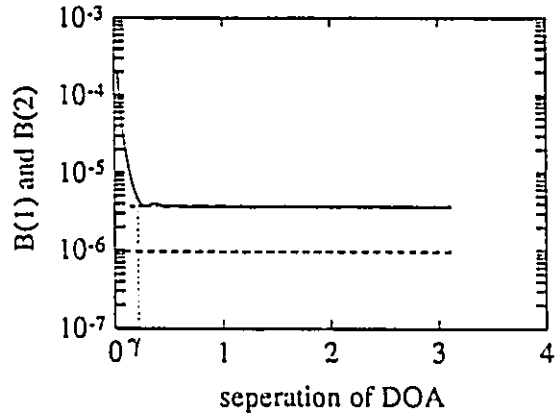
Figure 3.5: $B_{CR}(1)$ and $B_{CR}(2)$ against SNR with different $\Delta\phi$ (radian).



M=8
N=50
SNR=0dB (a)
— $B_{CR}(2)$
— $B_{CR}(1)$

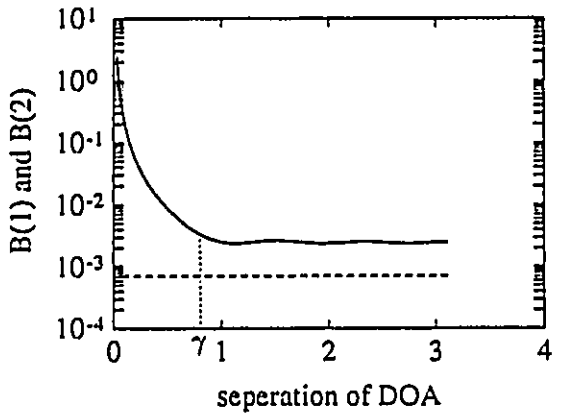


M=16
N=50
SNR=0dB (b)
— $B_{CR}(2)$
— $B_{CR}(1)$

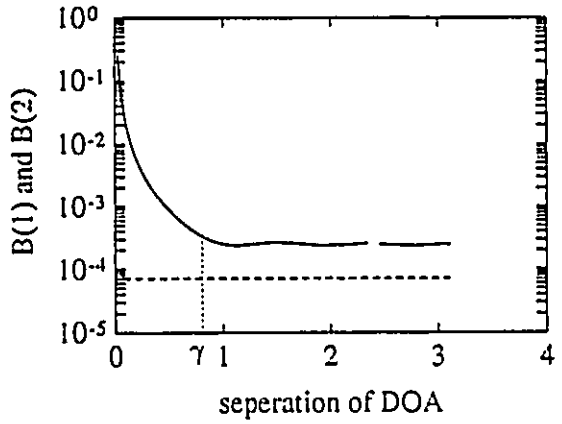


M=32
N=50
SNR=0dB (c)
— $B_{CR}(2)$
— $B_{CR}(1)$

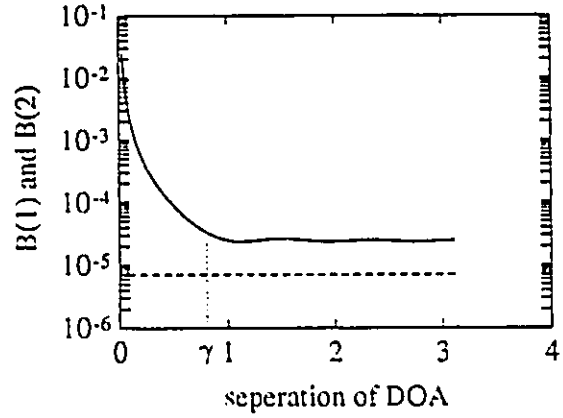
Fig. 3.6: $B_{CR}(1)$ and $B_{CR}(2)$ against $\Delta\phi$ under different M



(a)



(b)



(c)

Fig. 3.7: $B_{CR}(1)$ and $B_{CR}(2)$ against $\Delta\phi$ with different SNR

Chapter 4

The Efficiency of the MLE of DOA Based on Data Model(2)

4.1 INTRODUCTION

For DOA estimation, it has been shown in [77] that the maximum-likelihood estimation(MLE) based on Data Model(1) is not statistically efficient if the number of sensors M is small, even if the number of snapshots N is large. In this chapter, we derive the covariance matrix of MLE(2), the MLE of DOA based on Data Model(2), to indicate that the MLE of DOA based on Data Model(2) is asymptotically efficient.

In Section 4.2, the equiprobable contours and the forms of the covariance matrices are introduced to give the geometric interpretations of the error spread and the covariance matrix. In Section 4.3, the covariance matrix of the MLE(2) is derived. It is shown that CRLB(2) is achievable by the variance of the MLE(2) as a linearized approximation, and the MLE(2) is asymptotically efficient. It is also indicated that the errors of the MLE(2) of DOA are independent of each other and are independent of the locations of DOA asymptotically. In Section 4.4, the SNR-threshold phenomenon of the MLE(2) is discussed. It is shown that CRLB(2) fails to follow the SNR-threshold phenomenon. In Section 4.5, the simulation

results show that, the variance of MLE(1), the maximum-likelihood estimation of DOA based on Data Model(1), is a shift of the variance of MLE(2) approximately.

4.2 GEOMETRIC INTERPRETATION OF COVARIANCE MATRIX

For conciseness, we give the analysis for K=2 in detail. The results can be readily extended to K> 2 case.

Suppose there are two incoming signals with the unknown deterministic DOA (ϕ_{01}, ϕ_{02}). The parameter vector to be estimated is

$$\vec{\phi}_0 = \begin{bmatrix} \phi_{01} \\ \phi_{02} \end{bmatrix}.$$

The unbiased estimate of $\vec{\phi}_0$ is denoted as $\hat{\phi}$. Let the error vector be

$$\vec{\phi}_\epsilon = \hat{\phi} - \vec{\phi}_0 = \begin{bmatrix} \hat{\phi}_1 - \phi_{01} \\ \hat{\phi}_2 - \phi_{02} \end{bmatrix} = \begin{bmatrix} \phi_{\epsilon 1} \\ \phi_{\epsilon 2} \end{bmatrix}. \tag{4.1}$$

The covariance matrix of estimate $\hat{\phi}$ is

$$\begin{aligned} \mathrm{R}_\phi &= E[\vec{\phi}_\epsilon \vec{\phi}_\epsilon^T] \\ &= \begin{bmatrix} \sigma_1 & \sigma_{12} \\ \sigma_{21} & \sigma_2 \end{bmatrix}, \end{aligned} \tag{4.2}$$

where, $\sigma_1 = E(\phi_{\epsilon 1}^2)$ and $\sigma_2 = E(\phi_{\epsilon 2}^2)$ are the variances of the estimates $\hat{\phi}_1$ and $\hat{\phi}_2$, $\sigma_{12} = E(\phi_{\epsilon 1}\phi_{\epsilon 2})$ is the covariance of these two estimates, $\sigma_{21} = \sigma_{12}$.

The best way to determine how the covariance matrix provides a measure of error spread is to consider the special case in which the $\vec{\phi}_\epsilon$ are jointly Gaussian [81]. For algebraic

simplicity we let $E(\vec{\phi}_\varepsilon) = \vec{0}$. The joint probability density for two Gaussian variables is

$$p(\vec{\phi}_\varepsilon) = \frac{1}{(\det[2\pi \mathrm{R}_\phi])^{1/2}} \exp(-\frac{1}{2}\vec{\phi}_\varepsilon^T \mathrm{R}_\phi^{-1} \vec{\phi}_\varepsilon). \tag{4.3}$$

A typical density for two variables is shown in Fig. 4.1(a) geometrically.

From Eq.(4.3) we observe that the equiprobable contours are defined by the relation

$$\vec{\phi}_\varepsilon^T \mathrm{R}_\phi^{-1} \vec{\phi}_\varepsilon = C^2, \tag{4.4}$$

which is the equation of an ellipse for K=2. In Fig.4.1(b), we show the equiprobable contours.

Equiprobable contours can be viewed as the geometric interpretation of the covariance matrix R_ϕ. R_ϕ and the respective equiprobable contours are of three forms:

F-I. In Eq.(4.2), if $\sigma_1 \neq \sigma_2$ and $\sigma_{12} = \sigma_{21} \neq 0$, i.e., the variances are not equal to each other and the errors are correlated, then R_ϕ is a symmetric matrix and the respective equiprobable contours are shown in Fig. 4.2.

F-II. In Eq.(4.2), if $\sigma_1 \neq \sigma_2$ but $\sigma_{12} = \sigma_{21} = 0$, i.e., the variances are not equal to each other and the errors are uncorrelated, then

$$\mathrm{R}_\phi = \mathrm{diag}(\sigma_1, \sigma_2) \tag{4.5}$$

is a diagonal matrix and the respective equiprobable contours are shown in Fig. 4.3.

F-III. In Eq.(4.2), if $\sigma_1 = \sigma_2 = \sigma$ and $\sigma_{12} = \sigma_{21} = 0$, i.e., the variances are equal to each other and the errors are uncorrelated, then

$$\mathrm{R}_\phi = \sigma \mathrm{I} \tag{4.6}$$

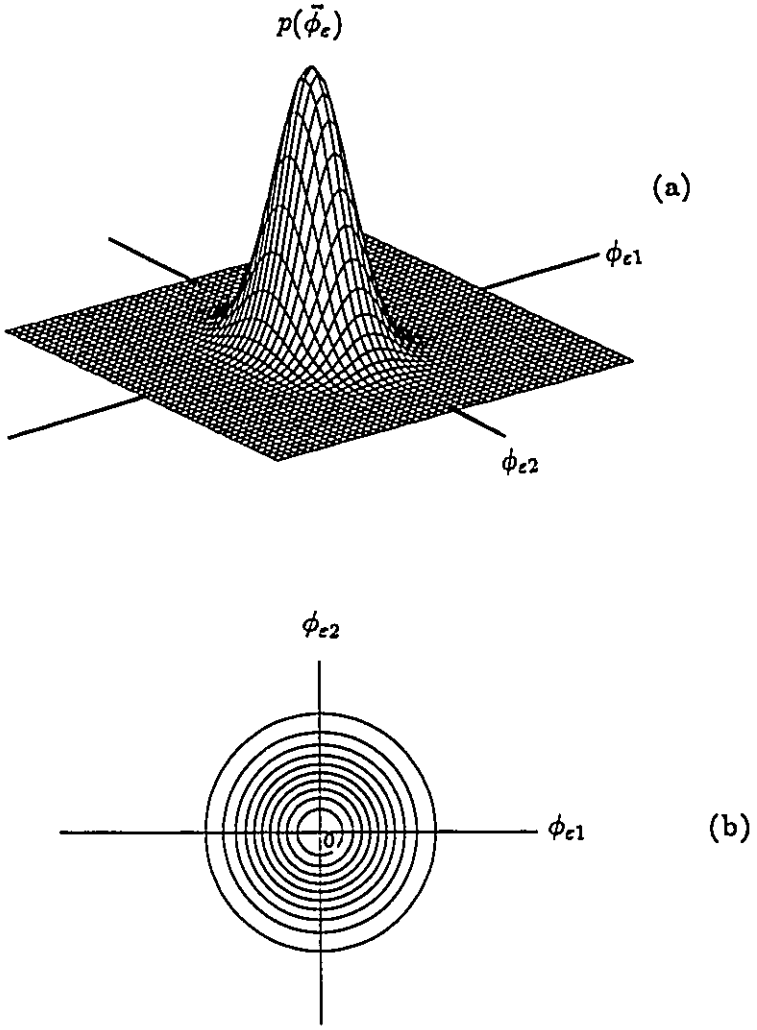


Figure 4.1: (a). Probability density function $p(\vec{\phi}_e)$. (b). Equiprobable contours of $p(\vec{\phi}_e)$.

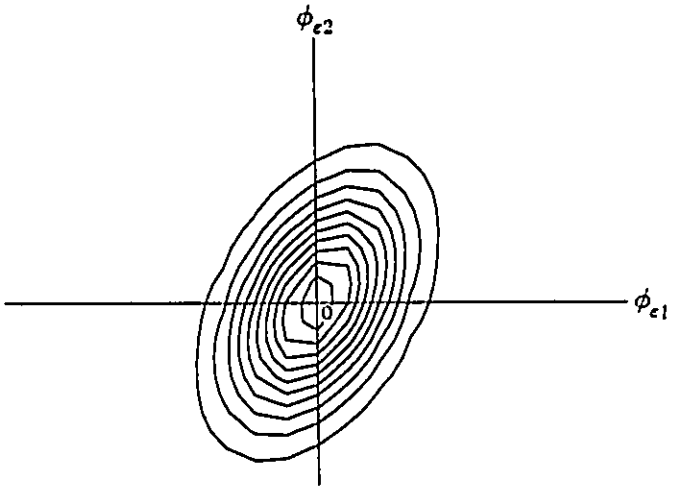


Figure 4.2: Equiprobable contours w.r.t. correlated variables.

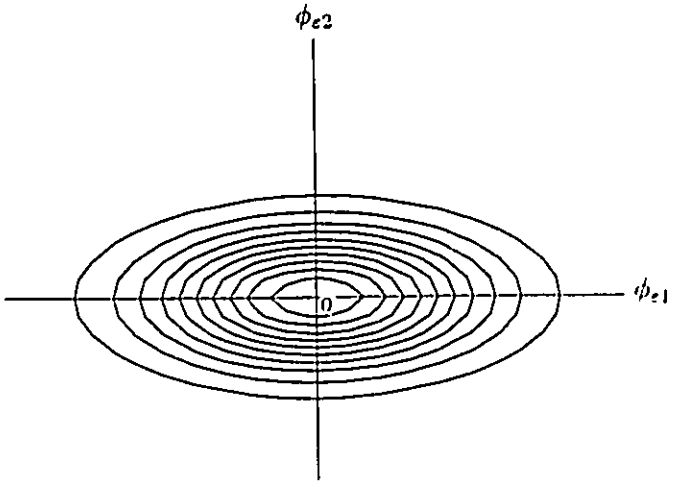


Figure 4.3: Equiprobable contours w.r.t. uncorrelated variables.

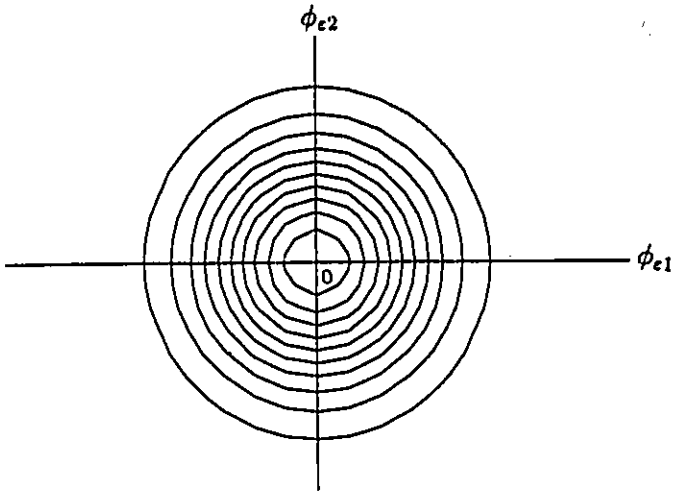


Figure 4.4: **Equiprobable contours w.r.t. uncorrelated variables and equal variances.**

is an identity matrix times the variance, and the respective equiprobable contours are shown in Fig. 4.4.

The equiprobable contours considered here have the following properties:

B-I. These ellipses increase in both major and minor diameters monotonically with increasing C.

B-II. These ellipses are concentric with fixed major and minor axes for increasing C.

B-III. The equiprobable contours concentrating around the centre represent the small errors in error distribution. The equiprobable contours far away from the centre represent the large errors in error distribution.

4.3 THE EFFICIENCY OF MLE(2)

According to [81], any estimate that satisfies the CRLB with an equality is called an efficient estimate. Also if an efficient estimate exists, it is the maximum-likelihood estimate. MLE is used to study the achievablity of CRLB. Let MLE(1) denote the MLE of DOA based on Data Model(1). The covariance matrix of MLE(1) has been derived in [77]. In that paper, a rather unusual result is obtained such that, MLE(1) is statistically inefficient for $M < \infty$, even though $N \to \infty$, where M is the numb :r of sensors and N is the number of snapshots. The reason, as discussed by [77], is that the inconsistency of the estimates of the signal amplitudes a and the noise power σ_e^2 in the case of $M < \infty$ implies that the errors of the MLE(1) cannot not achieve the CRLB for large N if M is small.

Data Model(2) introduced in Chapter 3 assumes that the signal amplitudes and the noise power are known. Since there are no estimations of the signal amplitudes and the noise power, CRLB(2) is then achievable by the error of MLE(2). We have the following theorem.

Theorem 4.1 *For the uncorrelated signals in the Gaussian white noise, the MLE of DOA based on Data Model(2) is asymptotically efficient.*

proof:

For convenience, we rewrite Eq.(3.41) as:

$$\mathrm{B}_{CR}(2) = \mathrm{diag}\left(\frac{3}{N\rho_1 M(M-1)(2M-1)}, \cdots, \frac{3}{N\rho_K M(M-1)(2M-1)}\right). \tag{4.7}$$

With the log-likelihood function shown in Eq.(3.9), the realization of the MLE of DOA based on Data Model(2) is given by the minimizer of the following function :

$$f(\vec{\phi}) = \sum_{n=1}^{N} [\mathbf{r}(n) - \mathbf{D}(\vec{\phi})\mathbf{a}(n)]^H [\mathbf{r}(n) - \mathbf{D}(\vec{\phi})\mathbf{a}(n)], \tag{4.8}$$

where, r(n), a(n) and $\mathrm{D}(\vec{\phi})$ are defined in Eqs.(2.5) to (2.10).

Expanding $\dot{f}$, the derivative of $f(\vec{\phi})$, in a Taylor series and ignoring the 2nd and the higher order terms, equating $\dot{f}$ to be zero for MLE, we have

$$0 = \dot{f}(\hat{\phi}) \simeq \dot{f}(\vec{\phi})|_{\vec{\phi}=\vec{\phi}_0} + \ddot{f}(\vec{\phi})|_{\vec{\phi}=\vec{\phi}_0}(\hat{\phi} - \vec{\phi}_0), \tag{4.9}$$

where, the k^{th} element of $\dot{f}(\vec{\phi})$ is

$$\frac{\partial f(\vec{\phi})}{\partial \phi_k} = \sum_{n=1}^{N}\{-\mathrm{a}^H(n)\frac{\partial \mathrm{D}^H(\vec{\phi})}{\partial \phi_k}[\mathrm{r}(n)-\mathrm{D}(\vec{\phi})\mathrm{a}(n)]-[\mathrm{r}(n)-\mathrm{D}(\vec{\phi})\mathrm{a}(n)]^H\frac{\partial \mathrm{D}(\vec{\phi})}{\partial \phi_k}\mathrm{a}(n)\}, \tag{4.10}$$

and the $(k,l)^{th}$ element of the Hessian matrix $\ddot{f}(\vec{\phi})$ is

$$\begin{aligned}\frac{\partial^2 f(\vec{\phi})}{\partial \phi_k \partial \phi_l} &= \sum_{n=1}^{N}\{-\mathrm{a}^H(n)\frac{\partial^2 \mathrm{D}^H(\vec{\phi})}{\partial \phi_k \partial \phi_l}\mathrm{e}(n) + \mathrm{a}^H(n)\frac{\partial \mathrm{D}^H(\vec{\phi})}{\partial \phi_k}\frac{\partial \mathrm{D}(\vec{\phi})}{\partial \phi_l}\mathrm{a}(n) \\ &\quad -\mathrm{e}^H(n)\frac{\partial^2 \mathrm{D}(\vec{\phi})}{\partial \phi_k \partial \phi_l}\mathrm{a}(n) + \mathrm{a}^H(n)\frac{\partial \mathrm{D}^H(\vec{\phi})}{\partial \phi_l}\frac{\partial \mathrm{D}(\vec{\phi})}{\partial \phi_k}\mathrm{a}(n)\},\end{aligned} \tag{4.11}$$

in which, $\mathrm{e}(n) = \mathrm{r}(n) - \mathrm{D}(\vec{\phi})\mathrm{a}(n)$ is the noise vector.

Note that

$$\frac{\partial \mathrm{D}(\vec{\phi})}{\partial \phi_k}\mathrm{a}(n) = a_k(n)[0, je^{j\phi_k}, ..., (M-1)je^{j(M-1)\phi_k}]^T. \tag{4.12}$$

Now we consider two cases: $l \neq k$ and $l = k$. First, under the case $l \neq k$, it is obvious that

$$\frac{\partial^2 \mathrm{D}^H(\vec{\phi})}{\partial \phi_k \partial \phi_l} = \mathbf{0}. \tag{4.13}$$

Also, because of Eq.(4.12), we have

$$\begin{aligned}\sum_{n=1}^{N}\mathrm{a}^H(n)\frac{\partial \mathrm{D}^H}{\partial \phi_k}\frac{\partial \mathrm{D}}{\partial \phi_l}\mathrm{a}(n) &= \sum_{n=1}^{N}a_k^*(n)a_l(n)\sum_{m=1}^{M}(m-1)^2e^{j(\phi_l-\phi_k)} \\ &= 0.\end{aligned} \tag{4.14}$$

The last step on Eq.(4.14) is due to Eq.(3.34) in Section 3.4.1, i.e.,

$$\sum_{n=1}^{N} a_k^*(n)a_l(n) \simeq \begin{cases} 0 & l \neq k, \\ N\sigma_{sk}^2 & l = k. \end{cases} \tag{4.15}$$

Therefore, the off-diagonal elements of the Hessian matrix in Eq.(4.11) are zeros.

Secondly, in the case $l = k$, let us consider the case that the estimate $\hat{\phi}$ is close enough to the true DOA $\vec{\phi}_0$. Under this condition, the change of the array manifold with respect to $\vec{\phi}$ within this region can be approximated by a linear function, i.e., the slopes $\frac{\partial D[m,k]}{\partial \phi_k}$ (m=1, ..., M) are constants, where $D[m,k]$ is the $(m,k)^{th}$ element in matrix D. The meaning of this can be seen in Fig.4.5. Here, we show the real part of $D[m,k]$, i.e.,

$$u(\phi_k) \triangleq \Re\{D[m,k]\} = \cos[(m-1)\phi_k].$$

By Taylor series expansion around the DOA ϕ_{0k}, we have

$$\frac{\partial^2 u(\phi_k)}{\partial \phi_k^2} \simeq 0.$$

Therefore, under the condition that the estimate $\hat{\phi}$ is close enough to the true DOA $\vec{\phi}_0$, i.e., in the region of small estimate errors,

$$\frac{\partial^2 \mathbf{D}}{\partial \phi_k^2} = 0. \tag{4.16}$$

Substituting Eqs.(4.13), (4.14) and (4.16) into Eq.(4.11) yields

$$\frac{\partial^2 f(\vec{\phi})}{\partial \phi_l \partial \phi_k} = \begin{cases} 0 & l \neq k \\ 2\sum_{n=1}^{N} \mathbf{a}^H(n)\frac{\partial \mathbf{D}^H}{\partial \phi_k}\frac{\partial \mathbf{D}}{\partial \phi_k}\mathbf{a}(n) & l = k. \end{cases} \tag{4.17}$$

By Eq.(4.12),

$$\begin{aligned} \sum_{n=1}^{N} \mathbf{a}^H(n)\frac{\partial \mathbf{D}^H}{\partial \phi_k}\frac{\partial \mathbf{D}}{\partial \phi_k}\mathbf{a}(n) &= \sum_{n=1}^{N} a_k^*(n)a_k(n)\sum_{m=1}^{M}(m-1)^2 \\ &= \frac{N\sigma_{sk}^2 M(M-1)(2M-1)}{6}, \end{aligned} \tag{4.18}$$

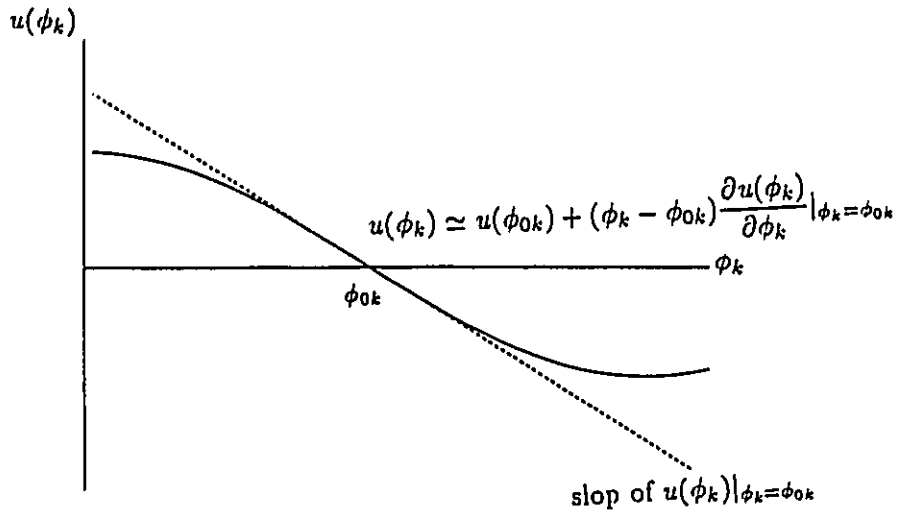


Figure 4.5: **The behaviour of error variance in the presence of small errors ϕ_e.**

where, Eq.(4.15) is used.

Substituting Eq.(4.18) into Eq.(4.17) results in the Hessian matrix

$$\begin{aligned} \mathbf{Q} &\triangleq \frac{\partial^2 f(\vec{\phi})}{\partial \vec{\phi}^2} \\ &= \text{diag}\left(\frac{N\sigma_{s1}^2 M(M-1)(2M-1)}{3}, ..., \frac{N\sigma_{sK}^2 M(M-1)(2M-1)}{3}\right). \end{aligned} \tag{4.19}$$

Note that the $(k,l)^{th}$ element in matrix $E\{[\frac{\partial f(\vec{\phi})}{\partial \vec{\phi}}][\frac{\partial f(\vec{\phi})}{\partial \vec{\phi}}]^H\}$ is

$$E[(\frac{\partial f(\vec{\phi})}{\partial \phi_k})(\frac{\partial f(\vec{\phi})}{\partial \phi_l})^*] = \begin{cases} 0, & l \neq k \\ \frac{N\sigma_{s_k}^2\sigma_e^2 M(M-1)(2M-1)}{3} & l = k. \end{cases} \tag{4.20}$$

See Appendix A for the proof of Eq.(4.20). Thus,

$$E\{[\frac{\partial f(\vec{\phi})}{\partial \vec{\phi}}][\frac{\partial f(\vec{\phi})}{\partial \vec{\phi}}]^H\} = \text{diag}\left(\frac{N\sigma_{s1}^2\sigma_e^2 M(M-1)(2M-1)}{3}, ..., \frac{N\sigma_{sK}^2\sigma_e^2 M(M-1)(2M-1)}{3}\right). \tag{4.21}$$

Now, we are ready to obtain the covariance matrix of the MLE(2) of DOA. From

Eq.(4.9), we have

$$\hat{\vec{\phi}} - \vec{\phi} \simeq -\mathbf{Q}^{-1}\frac{\partial f(\vec{\phi})}{\partial \vec{\phi}}. \tag{4.22}$$

Let $\mathbf{R}_{\phi}(2)$ be the covariance matrix of the MLE of DOA based on Data Model(2). With Eqs.(4.22), (4.19) and (4.21), we have

$$\begin{aligned}
\lim_{\hat{\phi}\to\vec{\phi}_0} \mathbf{R}_{\phi}(2) &\triangleq \mathbf{R}_{\phi l}(2) && (4.23)\\
&= E[(\hat{\vec{\phi}} - \vec{\phi}_0)(\hat{\vec{\phi}} - \vec{\phi}_0)^T]|_{\hat{\vec{\phi}}\to\vec{\phi}_0} \\
&\simeq \mathbf{Q}^{-1}E\{[\frac{\partial f(\vec{\phi})}{\partial \vec{\phi}}][\frac{\partial f(\vec{\phi})}{\partial \vec{\phi}}]^H\}\mathbf{Q}^{-1} \\
&= \mathrm{diag}(\frac{3}{N\rho_1 M(M-1)(2M-1)}, \ldots, \frac{3}{N\rho_K M(M-1)(2M-1)}). && (4.24)
\end{aligned}$$

When N is large, the estimates $\hat{\vec{\phi}}$ will be close to the true DOA $\vec{\phi}_0$ enough such that Eq.(4.24) is the covariance matrix of MLE(2) in the asymptotic sense (i.e., the large N). The comparison of Eq.(4.24) with Eq.(4.7) shows that

$$\mathbf{R}_{\phi l}(2) \simeq \mathbf{B}_{CR}(2). \tag{4.25}$$

So, $\mathbf{B}_{CR}(2)$ can be achieved by the covariance matrix of MLE(2) asymptotically. Theorem 4.1 follows.

#.

Eq.(4.24) shows the following properties of the covariance matrix $\mathbf{R}_{\phi}$ of MLE of DOA based on Data Model(2) for $\hat{\vec{\phi}} - \vec{\phi}_0$:

1. The covariance matrix $\mathbf{R}_{\phi l}$ is a diagonal matrix. This means that the errors of the estimates using MLE(2) are uncorrelated with each other. It can be proved [81] that the MLE(2) is Gaussian distributed when $N \to \infty$, so the errors $\phi_{ek} = \hat{\phi}_k - \phi_{0k} (k = 1, \ldots, K)$ from MLE are independent of each other for large N.

2. The variances of the MLE(2) of DOA are independent of the locations of DOA. When the signal power are equal, i.e., $\sigma_{s1}^2 = ... = \sigma_{sK}^2 = \sigma_s^2$, the covariance matrix $R_{\phi I}$ is a identity matrix times a constant, i.e., the form F-III discussed in Section 4.2.

Therefore we can conclude that

Lemma 4.1 *Under the condition that $\hat{\phi} \to \vec{\phi}_0$, with MLE based on Data Model(2), the errors of DOA estimates are independent of each other, and the errors of DOA estimates are independent of the locations of DOA.*

The achievablity of CRLB(2) by the variance of MLE(2) can be interpreted geometrically as shown in Fig.4.6 and Fig.4.7 ($\sigma_{s1}^2 = \sigma_{s2}^2 = \sigma_s^2$). Here, the dotted lines are the equiprobable contours by the equation

$$\vec{\phi}_\epsilon^T R_\phi^{-1}(2)\vec{\phi}_\epsilon = C^2,$$

and the solid lines are the Cramer-Rao bound ellipses by the equation

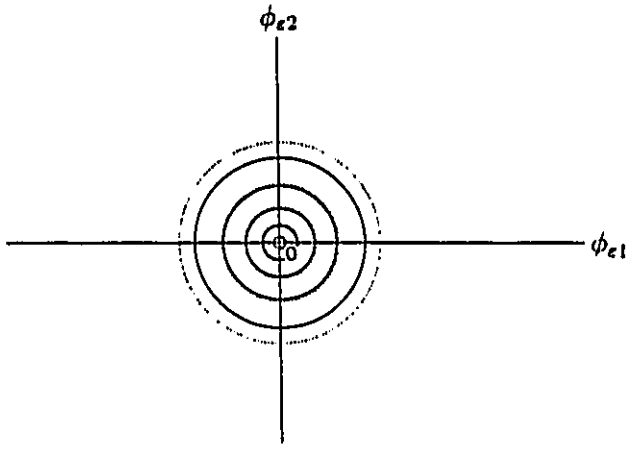
$$\vec{\phi}_\epsilon^T B_{CR}^{-1}(2)\vec{\phi}_\epsilon = C^2.$$

Around the centre of the equiprobable contours, i.e., the small $|\vec{\phi}_\epsilon|$, we have Eq.(4.25), i.e.,

$$\lim_{\vec{\phi}\to\vec{\phi}_0} R_\phi(2) \triangleq R_{\phi I}(2) \simeq B_{CR}(2), \tag{4.26}$$

so the dotted lines almost lie on the solid lines. For large $|\vec{\phi}_\epsilon|$, Eq.(4.26) does not hold due to the fact that the higher order terms in the Taylor expansion of Eq.(4.9) cannot be ignored. Thus the dotted lines lie outside of the solid lines.

The common saying that CRLB is achievable under high SNR for nonlinear estimation can be explained as the following:

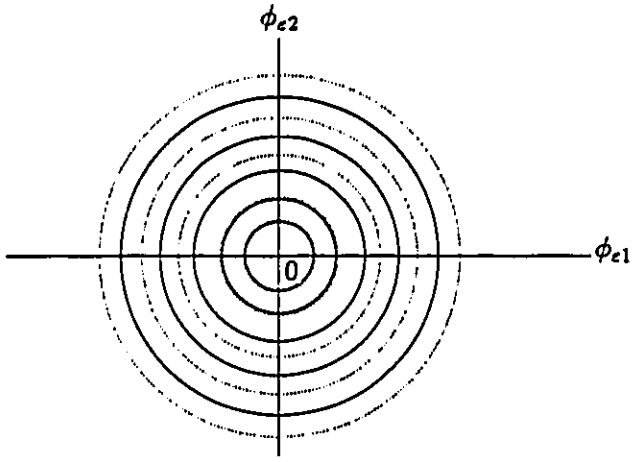


$$-\ \vec{\phi}_\epsilon^T \mathbf{B}_{CR}^{-1}(2)\vec{\phi}_\epsilon = C^2, \qquad \ldots\ \vec{\phi}_\epsilon^T \mathbf{R}_{\phi}^{-1}(2)\vec{\phi}_\epsilon = C^2.$$

Figure 4.6: **Equiprobable contours and CRLB(2) ellipses under high SNR.**

Under high SNR, the estimates with small estimation errors statistically dominate the variance. Most of the relevant equiprobable contours concentrate around the centre as shown by the dotted lines in Fig.4.6. The errors belonging to the equiprobable contours away from the centre can be ignored, thus CRLB is achievable.

Under low SNR, a large number of the estimation errors lie away from the centre of the equiprobable contours. This condition is exemplified in Fig.4.7. Under this condition the errors which lie away from the centre dominate the variance. As being seen from Fig.4.7, the ellipses for the CRLB and that for the variance of MLE do not agree closely for large error, thus CRLB is not achievable in this case.



$$\cdots \vec{\phi}_e^T \mathrm{B}_{CR}^{-1}(2)\vec{\phi}_e = C^2, \qquad \cdots \vec{\phi}_e^T \mathrm{R}_{\phi}^{-1}(2)\vec{\phi}_e = C^2.$$

Figure 4.7: **Equiprobable contours and CRLB(2) ellipses under low SNR.**

4.4 THE SNR-THRESHOLD PHENOMENON

Performance analysis against SNR is of interest in practice. Fig.4.8 shows the comparison between the variance of MLE and CRLB with an example based on Data Model(2). The number of sensors is 8, the number of snapshots is 50, and the number of the incoming signal is one.

The solid line represents the simulation performance of the maximum-likelihood estimate. The performance is divided into three segments. When $SNR > \rho^1$, MLE performs like a linear estimator, we call this area the linear region. When $\rho^2 < SNR < \rho^1$, the estimation error of MLE abruptly increases. The reason for this phenomenon is that, the noise level involved in the received data is high enough to make the signal observations from the received data be subject to ambiguities. The SNR at which the variance rises very rapidly as SNR decreases is called the threshold region. This SNR-threshold phenomenon occurs during $\rho^2 < SNR < \rho^1$. When $SNR < \rho^2$, the signal observations are completely

dominated by noise and are essentially useless for DOA estimation. This area is called the ambiguity region, and usually is ignored in the performance analysis. The dash line is $B_{CR}(2)$. We see that, under high SNR($SNR > \rho^l$), CRLB(2) is achievable since under this condition the errors around zero dominate the performance. This is the ambiguity-free region. But when $SNR < \rho^l$, CRLB(2) becomes very loose. Under low SNR (SNR$< \rho^l$), the variance of the estimate is affected by the large error ϕ_ϵ, but CRLB(2) is a linearized approximation only for the small error as discussed in Section 4.3. Therefore CRLB(2) is not a close approximation to the performance in the SNR-threshold region. It cannot show the SNR-threshold phenomenon.

Recalling the analysis in Chapter 3.5, CRLB(1) is obtained from CRLB(2) by the transformation T, and T is independent of SNR, which results in CRLB(1) against SNR being a parallel shift of CRLB(2) (Fig.3.5). Therefore, we have,

Observation 4.1 *CRLB(1) cannot show the SNR-threshold phenomenon due to the fact that CRLB(2) cannot show the SNR-threshol phenomenon.*

So Data Model(2) is a basic data model to study the performance of DOA estimate against SNR.

4.5 MLE(1) AND MLE(2)

In the above sections, the performance of the MLE of DOA based on Data Model(2) is discussed. Since the practical model is Data Model(1), we need to see the variance of the MLE(1), the maximum-likelihood estimation of DOA based on Data Model(1), to find the relationship of the performances between MLE(1) and MLE(2).

In Fig.4.9 to Fig.4.12, the variances of MLE(1) and MLE(2), the CRLB(1) and

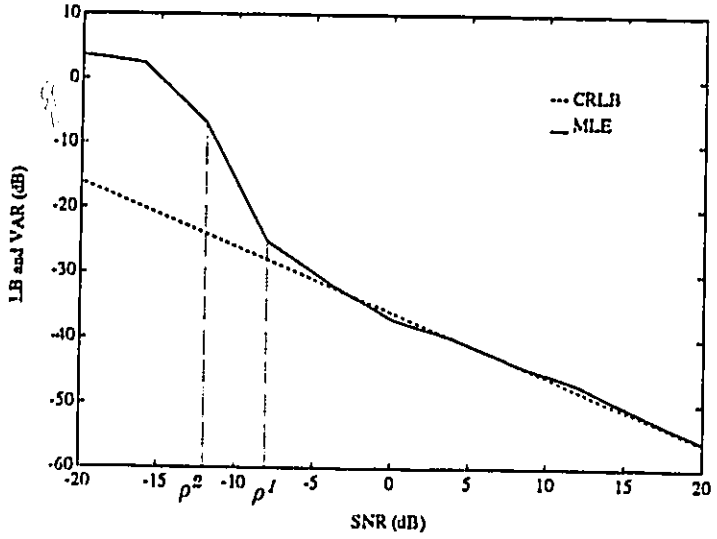


Figure 4.8: **CRLB(2) and the SNR-threshold phenomenon in MLE(2).**

CRLB(2) for different number of sensors, different number of snapshots, and different number of incoming signals are plotted. These simulations show that, for the ambiguity free region and the threshold region,

1. The variance of MLE(1) is a shift of the variance of MLE(2) approximately.
2. The amount of the shift of variances of MLE(2) to MLE(1) is the same as the amount of shift of CRLB(2) to CRLB(1).

These are very useful observations.

4.6 SUMMARY

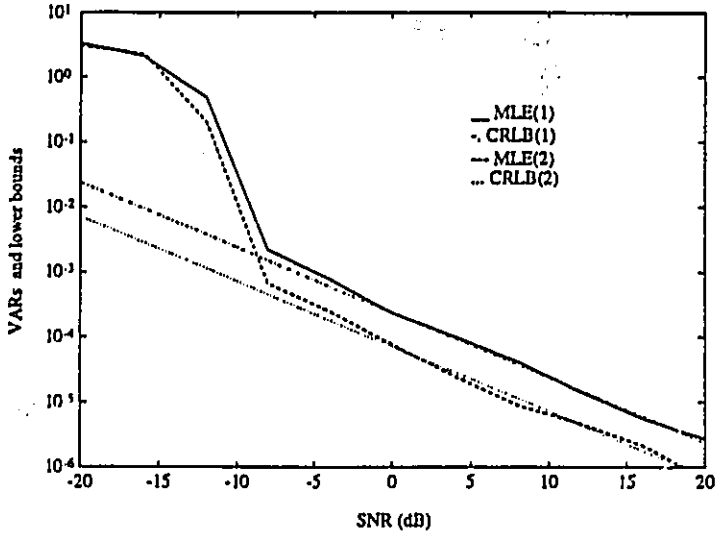
In this chapter, we have the following results:

1. The MLE of DOA based on Data Model(2) is asymptotically efficient.

2. Lemma 4.1 shows that under the condition $\hat{\phi} \to \vec{\phi}_0$, the errors of the MLE(2) of DOA are independent of each other, and are independent of the locations of DOA.

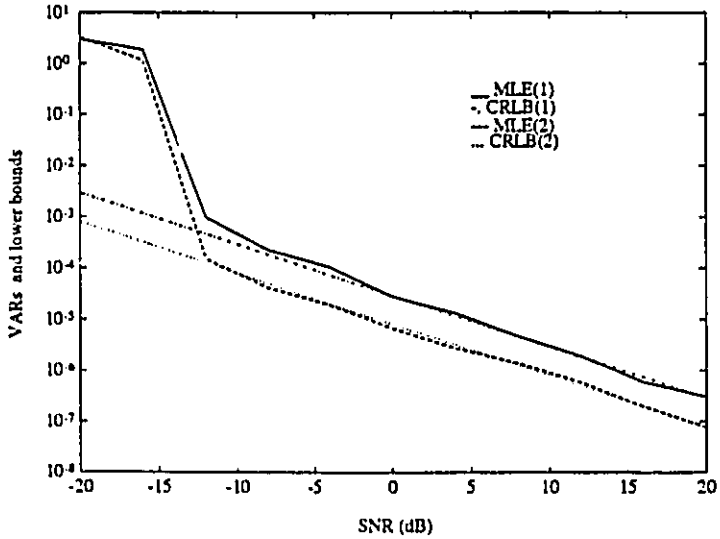
3. Data Model(2) is a basic data model to study the achievablity of a lower bound.

4. The variance of the MLE(1) is a shift of the variance of the MLE(2) for the ambiguity free region and the threshold region, and the amount of this shift is the same as the shift of CRLB(2) to CRLB(1).



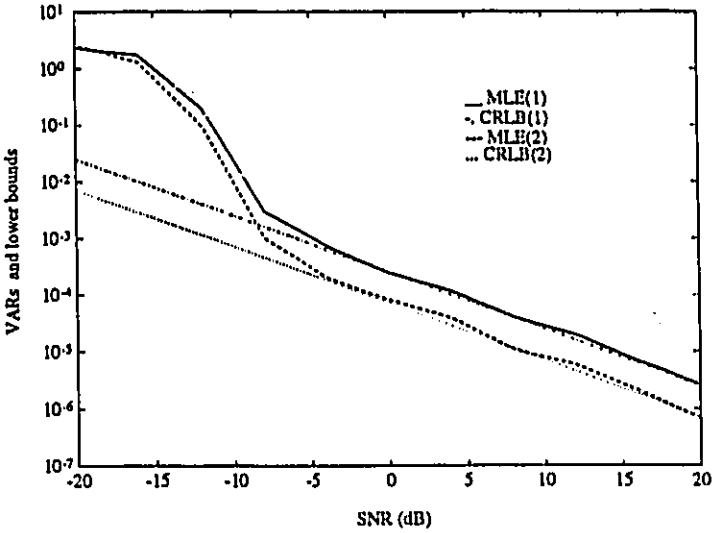
N=50, M=8, K=1.

Figure 4.9: The relationship of the variances of MLE(1) and MLE(2).



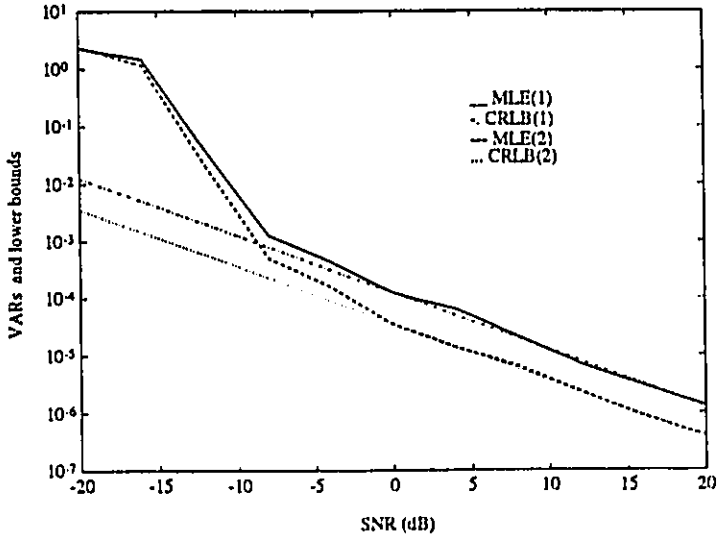
N=50, M=16, K=1.

Figure 4.10: The relationship of the variances of MLE(1) and MLE(2).



N=50, M=8, K=2[-.5 .5]

Figure 4.11: The relationship of the variances of MLE(1) and MLE(2).



N=100, M=8, K=2[-.5 .5]

Figure 4.12: The relationship of the variances of MLE(1) and MLE(2).

Chapter 5

Modified Ziv-Zakai Lower Bound on Variance of DOA Estimate for One Signal

5.1 INTRODUCTION

In this chapter, using Ziv-Zakai's idea, a lower bound on the variance of the DOA estimate for one signal in Gaussian white noise is derived. An understanding of this case is fundamental to the understanding of the general case which will be discussed in the following chapters. Section 5.2 presents an exposition of Ziv-Zakai's idea. Section 5.3 shows how this idea works for the DOA estimation in array processing, and derives the modified Ziv-Zakai lower bound (MZLB) on the variance of DOA estimate. In Section 5.4 the evaluation of this lower bound with Data Model(2) is developed. In Section 5.5, the properties of this lower bound are discussed, and MZLB with Data Model(1) is obtained.

5.2 ZIV-ZAKAI'S IDEA

Let ϕ_0 be the true parameter, and $\hat{\phi}$ be its unbiased estimate. The variance of this estimate is defined as

$$\begin{aligned} E[(\hat{\phi}-\phi_0)^2] &= \int_{-\infty}^{\infty} x^2 dP(\hat{\phi}-\phi_0 \le x) \\ &= \int_{0}^{\infty} x^2 dP(|\hat{\phi}-\phi_0| \le x). \end{aligned} \tag{5.1}$$

In Eq.(5.1), we consider the probability density function of the estimate error $(\hat{\phi}-\phi_0)$ being symmetric about zero.

A lower bound on this variance can be obtained from the lower bound on the error distribution $P(|\hat{\phi}-\phi_0| \le x)$, and the lower bound on $P(|\hat{\phi}-\phi_0| \le x)$ can be developed from a binary detection procedure. This idea was first presented by Ziv and Zakai in 1969 [100]. In order to have a better understanding of Ziv-Zakai's idea, the logic employed needs to be reviewed.

Ziv-Zakai's idea is to link the estimation error to the binary detection error. By imagining a binary detection procedure, the error probability of which is bounded, an expression can be established between the estimation error in Eq.(5.1) and the detection error bound. Because this detection error bound takes into account both the high and low SNR (Appendix B), it is expected that the resulting estimation bound will be a tight one.

Thus, Ziv-Zakai's idea can be applied for the general single parameter estimation in four steps as shown in Fig. 5.1. Among these four steps, the third one is the critical step from which the error distribution in Eq.(5.1) is obtained from the error probability in binary detection procedure.

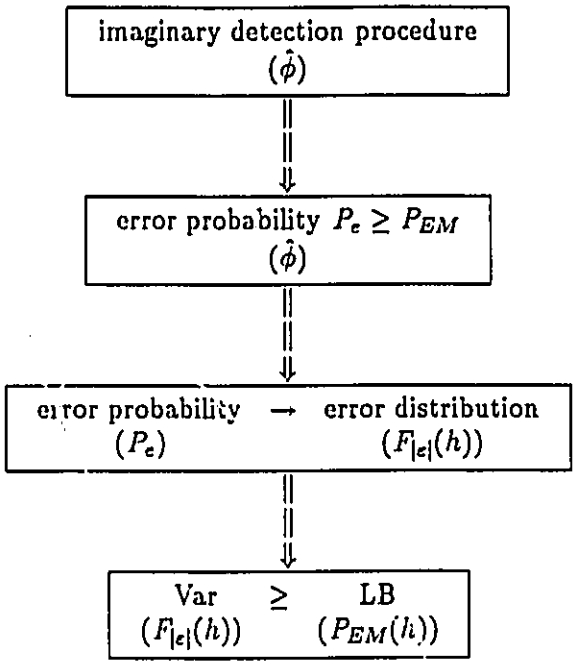


Figure 5.1: Diagram of the idea for deriving the error bound.

5.3 FROM DETECTION ERROR BOUND TO ESTIMATION ERROR BOUND

5.3.1 Imaginary Detection Procedure Based on Estimate

With the received data as described in Chapter 2, we imagine a binary hypotheses. Suppose the incoming DOA is of two possible values, ϕ_0 and ϕ_1. It is desired to determine which one is true. Then we establish the following hypotheses such that:

$$H_0 : \qquad \phi = \phi_0,$$

$$H_1 : \qquad \phi = \phi_1.$$

Based on some arbitrary estimate $\hat{\phi}$ of the incoming DOA, the decision criterion is

$$|\hat{\phi} - \phi_0| \underset{H_1}{\overset{H_0}{\lesseqgtr}} |\hat{\phi} - \phi_1|. \tag{5.2}$$

Eq.(5.2) is a binary detection procedure based on an estimate $\hat{\phi}$. It connects the estimate from the real received data to an imaginary detection procedure.

5.3.2 Error Probability and Its Lower Bound

Let us consider, for the time being, $\phi_1 - \phi_0 > 0$. First, we develop the error measure for the detection procedure described in Eq.(5.2).

Let $p(\hat{\phi}|H_i), i = 0, 1$, be the probability density function (PDF) of the estimate $\hat{\phi}$ under the hypothesis H_i. The situation for the decision procedure is depicted in Fig.5.2, where,

$$\Phi_D = \frac{\phi_0 + \phi_1}{2}. \tag{5.3}$$

Note that the decision region is a periodic function of ϕ. In Fig.5.2, $\pi + \Phi_D$ is equivalent to $-\pi + \Phi_D$.

We make the following assumptions:

A-I. $p(\hat{\phi}|\phi_i)$ is symmetric about ϕ_i, $p(\hat{\phi}|\phi_0)$ and $p(\hat{\phi}|\phi_1)$ are identical except for the translation of the axis, i.e., the error probability is independent of the location of the DOA;

A-II.

$$P(H_1) = P(H_0) = \frac{1}{2}; \tag{5.4}$$

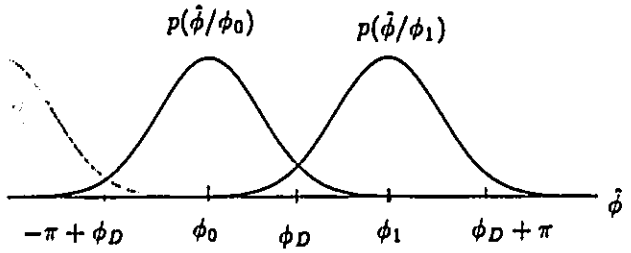


Figure 5.2: **Error probabilities with the decision scheme Eq.(5.2).**

A–III. The cost of making erroneous decisions is equal;

Under assumptions A-I to A-III, the error probability of the detection by the decision rule of Eq.(5.2) is given by (see Fig.5.2)

$$\begin{aligned} P_{e1} &= P(H_0)P(\hat{\phi} > \Phi_D|H_0) + P(H_1)P(\hat{\phi} \leq \Phi_D|H_1) \\ &= P(\hat{\phi} > \Phi_D|H_0), \end{aligned} \tag{5.5}$$

where, we have ignored the negligible probabilities of $P\{\hat{\phi} > (\Phi_D + \pi)|H_0\}$ and $P\{\hat{\phi} \leq (\Phi_D - \pi)|H_1\}$.

If we denote the estimation error by ε such that

$$\varepsilon = \hat{\phi} - \phi_0 \tag{5.6}$$

and let

$$h = \phi_1 - \phi_0, \tag{5.7}$$

then, using Eq.(5.3), Eq.(5.5) can be written as

$$\begin{aligned} P_{e1} &= P\{(\hat{\phi} - \phi_0) > (\frac{\phi_0 + \phi_1}{2} - \phi_0)\} \\ &= P\{\varepsilon > h/2\}. \end{aligned} \tag{5.8}$$

Eq.(5.2) is not an optimum decision rule, hence P_{e1} is not the minimum probability of error. However, the Bayes criterion with equal error cost leads to an optimum likelihood ratio test (LRT) yielding minimum decision error probability P_{EM} (described in Appendix B), therefore,

$$P_{e1} \geq P_{EM}(h). \tag{5.9}$$

The RHS of Eq.(5.9) is the lower bound on the error probability of the imaginary detection procedure based on the estimate $\hat{\phi}$.

5.3.3 Definition of the Variance of DOA Estimate Based on P_{e1}

If $\phi_1 - \phi_0 < 0$, similarly we obtain

$$\begin{aligned} P_{e2} &= P(\hat{\phi} \leq \Phi_D | H_0) \\ &= P(\varepsilon \leq h/2) \\ &\geq P_{EM}(h). \end{aligned} \tag{5.10}$$

Then, combining Eq.(5.8) and Eq.(5.10) results in

$$P(|\varepsilon| > |h/2|) = P_{e1} + P_{e2}. \tag{5.11}$$

The RHS of Eq.(5.11) is the error probability of the imaginary binary detection procedure under the hypotheses H_0 (ϕ_0) and H_1 (ϕ_1).

Under the condition that the error measure is independent of the location of DOA, both sides of Eq.(5.11) only depend on the distance $h = |\phi_1 - \phi_0|$ between two DOA. Thus, $P(|\varepsilon| > |h/2|)$ becomes the error distribution of the estimate $\hat{\phi}$, since $\varepsilon = \hat{\phi} - \phi_0$. This is the relationship between the error of the imaginary binary detection and the error of the real estimation problem. Using this relationship, we can develop the lower bound of the estimate error.

Let

$$F_c(h/2) \triangleq P(|\varepsilon| > |h/2|), \tag{5.12}$$

where we define

$$F_c(h) = P(|\varepsilon| > |h|). \tag{5.13}$$

Note that $F_c(h)$ is a decreasing function of h. In order to define the variance, let

$$\begin{aligned} F_{|\varepsilon|}(h) &= 1 - F_c(h) \\ &= P(|\varepsilon| \leq |h|), \end{aligned} \tag{5.14}$$

then, $F_{|\varepsilon|}(h)$ can be interpreted as the cumulative distribution of the estimate error.

Note that $|\varepsilon| \in [0, \pi]$. Therefore, the consideration of Eq.(5.13) can be concentrated in the range $0 \leq h \leq \pi$.

From Eq.(5.1), the variance of DOA estimate can therefore be expressed in terms of $F_{|\varepsilon|}(h)$ such that

$$E[\varepsilon^2] = \int_0^{\pi} h^2 dF_{|\varepsilon|}(h). \tag{5.15}$$

5.3.4 Error Bound on the Variance of the Estimate in Terms of P_{EM}

From Eqs.(5.9), (5.10), (5.11) and (5.12), we have

$$\frac{1}{2}F_c(h/2) \geq P_{EM}(h). \tag{5.16}$$

Multiplying both sides of Eq.(5.16) by h and integrating from 0 to 2π, Eq.(5.16) becomes

$$\begin{aligned}
&\int_0^{2\pi} hP_{EM}(h)dh \\
&\leq \frac{1}{2}\int_0^{2\pi} hF_c(\frac{h}{2})dh \\
&= 2\int_0^{\pi} hF_c(h)dh \\
&= h^2F_c(h)\,|_0^{\pi} - \int_0^{\pi} h^2dF_c(h).
\end{aligned} \tag{5.17}$$

Note that $|\varepsilon| \in [0, \pi]$, from Eq.(5.13), we have

$$F_c(h)|_{h=\pi} = P(|\varepsilon| > \pi) = 0. \tag{5.18}$$

Using Eqs.(5.14), (5.15) and (5.18) in Eq.(5.17), we can write

$$E[\varepsilon^2] \geq \int_0^{2\pi} hP_{EM}(h)dh \triangleq B_Z. \tag{5.19}$$

Eq.(5.19) interprets the lower bound on the variance of the DOA estimate. The RHS of Eq.(5.19) is referred to as the modified Ziv-Zakai lower bound (MZLB).

5.4 EVALUATION OF THE ERROR BOUND

To evaluate the bound in Eq.(5.19), first, we need to find an interpretation for $P_{EM}(h)$. We recall that [81], if the received data $\bar{r}$ is real and satisfies the assumptions A-I to A-III, then

the optimum LRT yields the minimum decision error probability given by (see Appendix B)

$$P_{EM} = \exp[\mu(q_M) + q_M^2\ddot{\mu}(q_M)/2]\mathrm{erfc}_*\{q_M\sqrt{\ddot{\mu}(q_M)}\}/2 + \exp[\mu(q_M) + (1-q_M)^2\ddot{\mu}(q_M)/2]\mathrm{erfc}_*\{(1-q_M)\sqrt{\ddot{\mu}(q_M)}\}/2, \quad (5.20)$$

where,

$$\mathrm{erfc}_*\{\eta\} \triangleq \frac{1}{\sqrt{2\pi}}\int_\eta^\infty e^{-t^2/2}dt, \quad (5.21)$$

$$\mu(q) \triangleq \ln\int_{-\infty}^{\infty}[p(\vec{r}|H_1)]^q[p(\vec{r}|H_0)]^{1-q}d\vec{r}, \quad (5.22)$$

with $p(\vec{r}|H_i)$ being the PDF of the received data under hypothesis H_i, and q_M satisfies the condition that

$$\dot{\mu}(q_M) = 0, \quad (5.23)$$

where, $0 \le q_M \le 1$.

It is shown in Appendix C that with Data Model(2), for K=1, the minimum error probability in Eq.(5.20) in our problem is

$$P_{EM}(h) = \mathrm{erfc}_*\{[\rho N\sum_{m=1}^{M}(1-cos(h(m-1)))]^{1/2}\} = \mathrm{erfc}_*\{f(h)\}, \quad (5.24)$$

where we define the signal to noise ratio (SNR) as

$$\rho = \sigma_e^{-2}\sigma_s^2, \quad (5.25)$$

and

$$f(h) = [\rho N\sum_{m=1}^{M}(1-cos(h(m-1)))]^{1/2}. \quad (5.26)$$

Secondly, we need to perform the integration in Eq.(5.19). The limits of the integral is from 0 to 2π, but h is the distance between two distinct possible DOA. The correct decision cannot be made if the two DOA overlap. Therefore, we should leave a guard separation between the two possible DOA, and modify the range to $h \in [0, H]$, where $H = 2\pi - \delta$, with δ being a small but finite angle.

Let $B_Z(2)$ denote the MZLB based on Data Model(2). Now, by substituting Eq.(5.24) into Eq.(5.19), we have

$$B_Z(2) = \int_0^H \text{erfc}_*\{f(h)\}h\,dh \tag{5.27}$$

$$= \frac{H^2}{2}\text{erfc}_*\{f(H)\} + \frac{1}{2\sqrt{2\pi}}\int_0^H h^2 e^{-f^2/2}\frac{df}{dh}dh$$

$$\triangleq I_1 + I. \tag{5.28}$$

The integral I on the RHS of Eq.(5.28) can be broken up into two parts such that

$$B_Z(2) = \frac{H^2}{2}\text{erfc}_*\{f(H)\} + \frac{1}{2\sqrt{2\pi}}\int_0^{\Delta} h^2 e^{-f^2/2}(\frac{df}{dh})dh + \frac{1}{2\sqrt{2\pi}}\int_{\Delta}^H h^2 e^{-f^2/2}(\frac{df}{dh})dh$$

$$= I_1 + I_0 + I_2, \tag{5.29}$$

where,

$$I_1 = \frac{H^2}{2}\text{erfc}_*\{f(H)\}, \tag{5.30}$$

$$I_0 = \frac{1}{2\sqrt{2\pi}}\int_0^{\Delta} h^2 e^{-f^2/2}(\frac{df}{dh})dh, \tag{5.31}$$

$$I_2 = \frac{1}{2\sqrt{2\pi}}\int_{\Delta}^H h^2 e^{-f^2/2}(\frac{df}{dh})dh. \tag{5.32}$$

Now we evaluate I_0 and give the values of Δ and H.

From the definition of $f(h)$ in Eq.(5.26) we have

$$\frac{df}{dh} = \frac{1}{2}\sqrt{\rho N}[\sum_{m=1}^{M}(m-1)\sin((m-1)h)][2\sum_{m=1}^{M}\sin^2((m-1)h/2)]^{-\frac{1}{2}}. \tag{5.33}$$

Thus, substituting Eq.(5.33) into I_0 in Eq.(5.31), we have

$$I_0 = \frac{\sqrt{\rho N}}{2\sqrt{2\pi}}\int_0^{\Delta}\frac{h^2}{2\sqrt{2}}[\sum_{m=1}^{M}\sin^2((m-1)h/2)]^{-\frac{1}{2}}[\sum_{m=1}^{M}(m-1)\sin((m-1)h)]e^{-f^2/2}dh. \tag{5.34}$$

We choose Δ such that, for $h \in [0, \Delta]$, every part of the integrand in Eq.(5.34) is positive. Furthermore, Δ is chosen such that the value of the largest component $\sin((M-1)\Delta/2)$ in the summations is still on the first rising part of the sine curve, so that the following approximations are closely satisfied:

1.
$$\sum_{m=1}^{M}\sin^2((m-1)h/2) \simeq \sum_{m=1}^{M}(m-1)^2h^2/4, \tag{5.35}$$

therefore,

$$\begin{aligned}[\sum_{m=1}^{M}\sin^2((m-1)h/2)]^{-\frac{1}{2}} &\simeq [\sum_{m=1}^{M}(m-1)^2h^2/4]^{-\frac{1}{2}} \\ &= \frac{2\sqrt{6}}{h\sqrt{M(M-1)(2M-1)}};\end{aligned} \tag{5.36}$$

2. Each of the terms under the summation within the second set of brackets in Eq.(5.34) can be approximated by a straight line joining the extreme values 0 and Δ such that

$$(m-1)\sin((m-1)h) \simeq \frac{(m-1)\sin((m-1)\Delta)}{\Delta}h. \tag{5.37}$$

Therefore,

$$\sum_{m-1}^{M}(m-1)\sin((m-1)h) \simeq \frac{h}{\Delta}\sum_{m=1}^{M}(m-1)\sin((m-1)\Delta); \tag{5.38}$$

3. Similar to 2., with the approximation

$$\sin\frac{(m-1)h}{2} \simeq \frac{h}{\Delta}\sin(\frac{(m-1)\Delta}{2}), \tag{5.39}$$

the exponential term in Eq.(5.34) becomes

$$\begin{aligned}
\exp(-f^2/2) &= \exp\{\frac{-\rho N}{2}\sum_{m=1}^{M} 2\sin^2\frac{(m-1)h}{2}\} \\
&\simeq \exp\{\frac{-\rho N h^2}{2\Delta^2}\sum_{m=1}^{M} 2\sin^2((m-1)\Delta/2)\} \\
&= \exp\{\frac{-\rho N h^2}{2\Delta^2}\sum_{m=1}^{M} 1-\cos((m-1)\Delta)\}.
\end{aligned} \tag{5.40}$$

The choice of Δ is arbitrary as long as Eqs.(5.35), (5.37) and (5.39) are closely satisfied. We choose

$$\Delta = \frac{\pi}{2M}. \tag{5.41}$$

With this choice, the sum of the RHS of Eq.(5.38) can be evaluated quite simply as the following:

$$\begin{aligned}
\sum_{m=1}^{M}(m-1)\sin((m-1)h) &\simeq \frac{h}{\Delta}\sum_{m=1}^{M}(m-1)\sin((m-1)\Delta) \\
&= -[\frac{d}{d\Delta}\sum_{m=1}^{M}\cos((m-1)\Delta)]\frac{h}{\Delta} \\
&= -[\frac{d}{d\Delta}\Re\{\sum_{m=1}^{M}e^{j(m-1)\Delta}\}]\frac{h}{\Delta} \\
&= -\frac{1}{2}[\frac{d}{d\Delta}\{1-\frac{\cos(M\Delta)-\cos((M-1)\Delta)}{1-\cos(\Delta)}\}]\frac{h}{\Delta} \\
&\simeq \frac{4M^2}{\pi^2}\frac{h}{\Delta} \\
&= \frac{8M^3h}{\pi^3}
\end{aligned} \tag{5.42}$$

The details of Eq.(5.42) are shown in Appendix E. The RHS of Eq.(5.40) can be evaluated as (Appendix F):

$$\exp(-f^2/2) \simeq \exp(-0.07\rho N M^3 h^2) = \exp(-\eta_1^2 h^2/2), \tag{5.43}$$

where,

$$\eta_1^2/2 = 0.07\rho N M^3. \tag{5.44}$$

Substituting Eqs.(5.36), (5.42) and (5.43) into Eq.(5.34), integrating, simplifying, and defining

$$\eta_0^2 = \rho N M(M-1)(2M-1)/12, \tag{5.45}$$

we obtain (see Appendix G)

$$I_0 \simeq \frac{1}{2\eta_0^2}\sqrt{\frac{M(M-1)(2M-1)}{2M^3}}[\frac{1}{2} - \text{erfc}_*\{\eta_1\Delta\} - \frac{\eta_1\Delta}{\sqrt{2\pi}}e^{-\eta_1^2\Delta^2/2}]. \tag{5.46}$$

In Eq.(5.46), for a reasonably large number of sensors M, the term under the square root sign can be approximated to unity. Also, the last two terms in the brackets are negligible compared with the first term. Therefore, we have

$$I_0 \simeq \frac{1}{4\eta_0^2} = \frac{3}{\rho N M(M-1)(2M-1)}. \tag{5.47}$$

At high SNR, this value of I_0 plays the dominant role on the estimation error bound as expressed in Eq.(5.29). It is reasonable then, to argue that the guard separation δ should be chosen to be the same value as Δ, i.e., $H = 2\pi - (\pi/2M)$. With this choice of H and Δ, it can be seen I_0 and I_1 in Eqs.(5.31) and (5.30) are always positive. In general, I_1 rises steeply to significance under low SNR while dropping fast to negligible values under high SNR. On the other hand, I_0 is a linear function of SNR. It decreases rather slowly with the increase in SNR and is the dominant part under high SNR. The behaviour of the two functions are exemplified in Fig.5.3 for M=8 and N=50. In contrast, I_2 is always negative.

This can be easily seen from the fact that I_2 has exactly the same integrand as that in I_0 in Eq.(5.34), where, while the other parts are all positive, the average value of the part in the second set of brackets is negative for $h \in (\Delta, H]$. The magnitude of the value of I_2, in general, quickly diminishes to a negligible value as the SNR increases and is significant only under extremely low SNR. An example of this behaviour is shown in Fig.5.4 for M=8 and N=50.

The bound $B_Z(2)$ is a combination of these three components, and can be partitioned into three separate sections as shown in Fig.5.5. For the SNR to be higher than ρ^1, the ambiguity in DOA estimation can essentially be resolved and $B_Z(2)$ is dominated by the value of I_0. This is a linear region since I_0 is a linear function. The transition from the ambiguity-free region to the region dominated by ambiguity occurs when the SNR falls below ρ^1 until it reaches ρ^2. Here, the value of I_1 rises to significance and becomes the dominant part. This is the SNR-threshold phenomenon in the estimation of DOA, and the error bound is essentially the combination of the terms I_0 and I_1. When the SNR falls below ρ^2, DOA observations are subject to ambiguities. The value of I_2 becomes significant too and the resulting bound is essentially the combination of the terms of I_0, I_1 and I_2. The bound $B_Z(2)$ shown in Fig.5.5 is for M=8 and N=50. Fig.5.6 shows the bound $B_Z(2)$ for M=16 and M=32 respectively. In each of these cases, the three regions can be distinctly observed.

Since we mainly consider the performance of the lower bound in the ambiguity-free region and the threshold region, ignoring the complicated integral I_2 will not significantly affect the performance of $B_Z(2)$ in these two regions as shown in Fig.5.7, so that we can write

$$\begin{aligned} B_Z(2) &\simeq I_0 + I_1 \\ &= \frac{3}{\rho N M(M-1)(2M-1)} + \frac{H^2}{2}\mathrm{erfc}_*\{f(H)\} \end{aligned} \tag{5.48}$$

for the ambiguity-free region and the threshold region. In the following, we will concentrate our studies on these two regions.

Eq.(5.48) is the MZLB(2) for the SSST case.

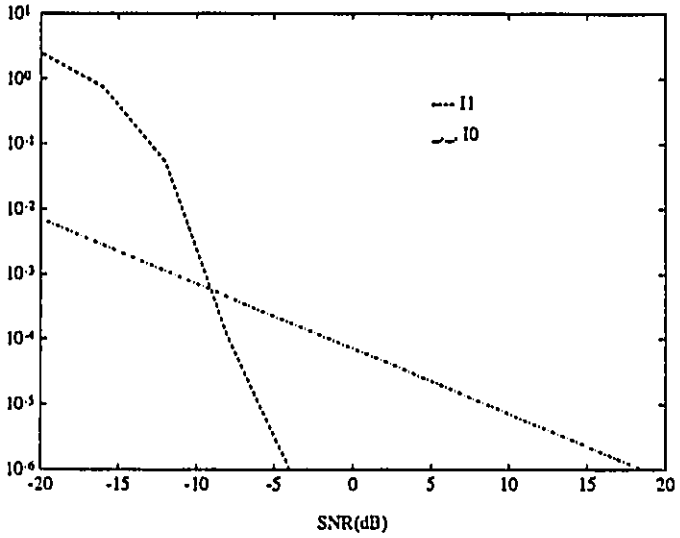


Figure 5.3: I_0 and I_1 for M=8 N=50.

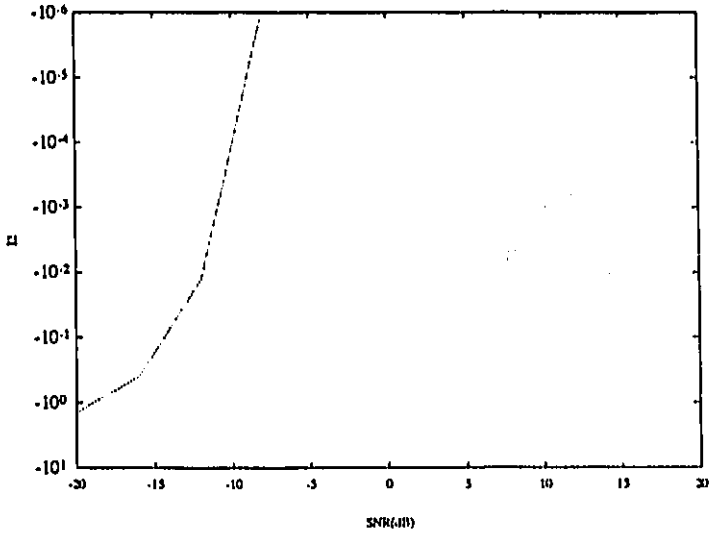


Figure 5.4: I_2 for M=8 N=50.

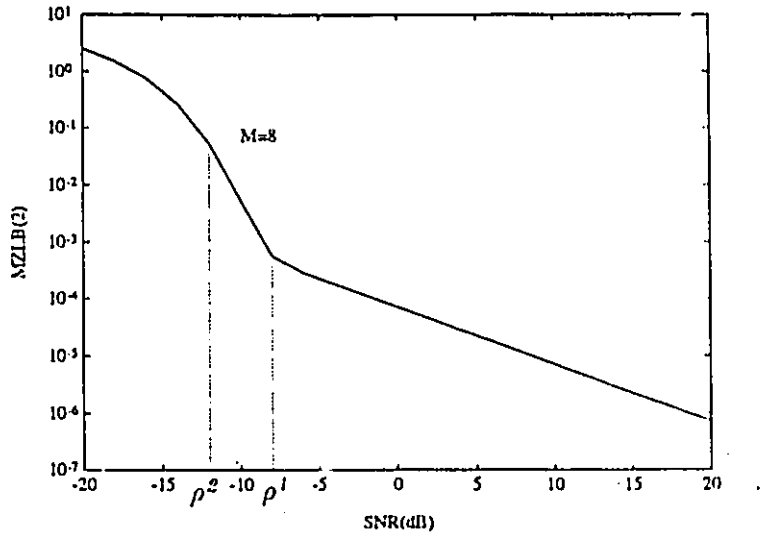


Figure 5.5: $B_Z(2)$ for M=8 N=50.

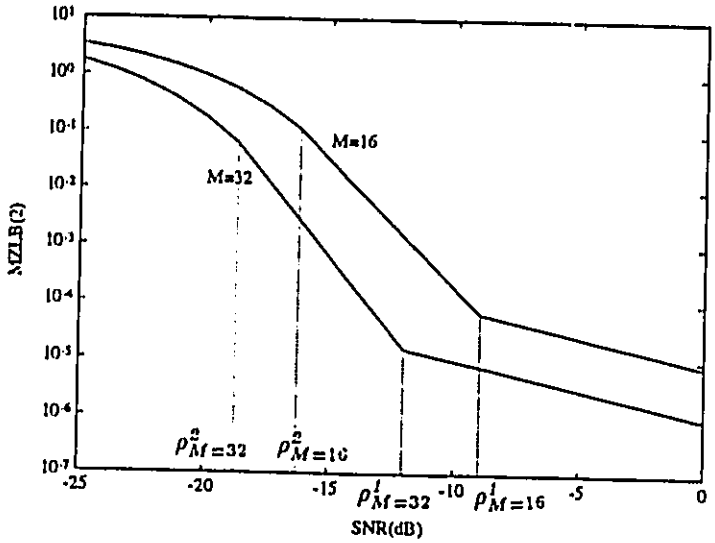


Figure 5.6: $B_Z(2)$ for M=16 and M=32 (N=50).

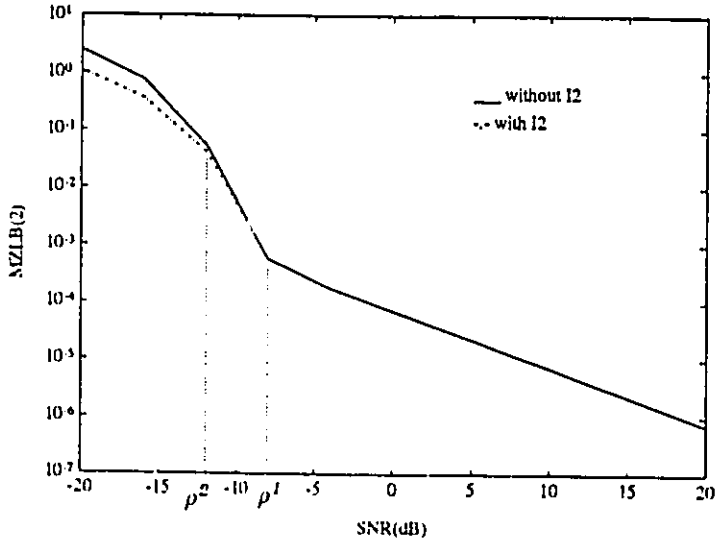


Figure 5.7: $B_Z(2)$ with I_2 and without I_2 (M=8, N=50).

5.5 PROPERTIES OF MZLB

5.5.1 $B_Z(2)$ and $B_{CR}(2)$

It is interesting to compare the bound $B_Z(2)$ in Eq.(5.48) with the well-known Cramer-Rao bound. We have derived CRLB based on Data Model(2) in Chapter 3.4 that,

$$B_{CR}(2) = \frac{3}{\rho N M(M-1)(2M-1)}. \tag{5.49}$$

Comparing Eqs.(5.48) and (5.49), we note that $B_{CR}(2)$ is the same as I_0. Thus, we can conclude that the MZLB includes the CRLB as a part which dominates when the SNR is high, i.e.

$$B_Z(2) \simeq B_{CR}(2) + I_1. \tag{5.50}$$

When SNR is low, it is the term I_1 in MZLB that shows the threshold. We call this term the *threshold term*. Fig.5.8 shows the variance of the MLE of DOA of a single signal based on Data Model(2). The estimation was carried out with M=8 sensors and N=50 snapshots. The performance is evaluated over 200 trials. It can be observed that, when the SNR is high ($\geq -8dB$), the MLE performance reaches both the CRLB and the MZLB. However, for lower SNR, the MZLB follows the threshold region of the MLE performance very much more closely than the CRLB due to the additional terms of I_1 in Eq.(5.48).

5.5.2 MZLB with Data Model(1)

In practice, the amplitude of the arriving signal is often unknown, so Data Model(2) is not a practical model while Data Model(1) is more realistic. In Chapter 4, Fig.4.9 to Fig.4.12 show that, the variance of the MLE(1) almost is a shift of the variance of the MLE(2). Also, these plots show that, the relationship between the variance of MLE(2) and MLE(1) for the ambiguity free region and the threshold region is the same as the relationship between

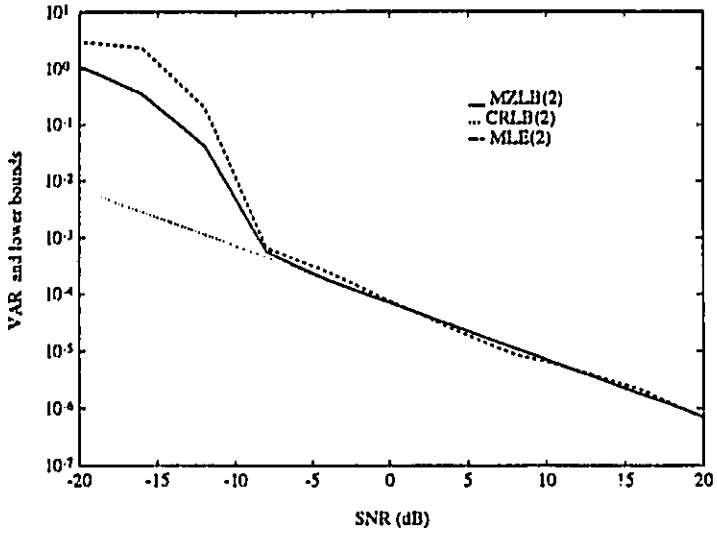


Figure 5.8: **MZLB, CRLB and MLE with Data Model(2).**

CRLB(1) and CRLB(2). Since the MZLB for Data Model(1) is not easy to derive due to the fact that there will be multiple types of parameters involved, these observations make us conjecture that the MZLB(1), the MZLB based on Data Model(1), is a shift of the MZLB(2), the MZLB based on Data Model(2), and the relationship between MZLB(1) and MZLB(2) can be found with the relationship between CRLB(1) and CRLB(2). Since it is discussed in Chapter 3, CRLB(1) is related with CRLB(2) by a transformation T, we have

$$B_Z(1) = TB_Z(2), \tag{5.51}$$

where, (see Eqs.(3.31) and (5.49))

$$T = \frac{B_{CR}(1)}{B_{CR}(2)} = \frac{2(2M-1)}{M+1}. \tag{5.52}$$

In Eq.(5.51), $B_Z(1)$ is the modified ZZLB for the estimation of the DOA of a signal when its amplitude is deterministic and unknown. T is the difference for the lower bounds with the different data models. Fig.5.9 shows the MZLB as interpreted in Eq.(5.51) together with the performance of the MLE of one DOA based on Data Model(1). The agreement

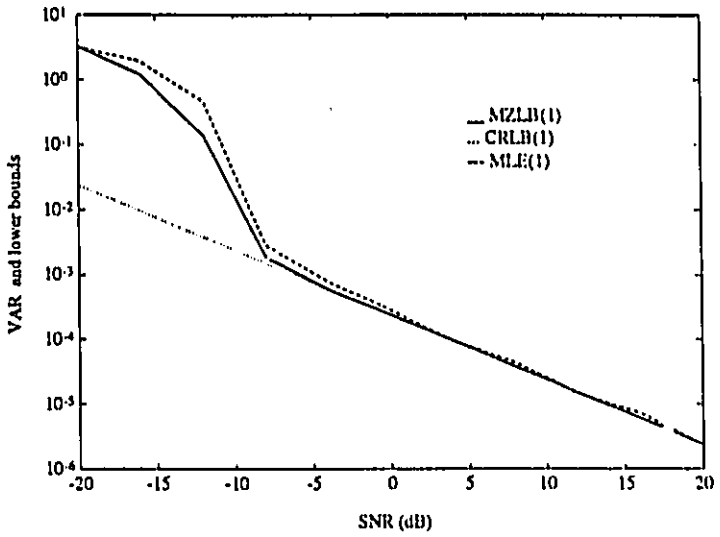


Figure 5.9: MZLB, CRLB and MLE with Data Model(1).

of the MLE performance and the $B_Z(1)$ is very much similar to that shown in Fig.5.8 in which the amplitude of the signal is known.

T in Eq.(5.52) is a special case of the transformation T in Eq.(3.45) discussed in Chapter 3.

It is worthwhile pointing out that the idea we used here to solve the complicated practical problem involving multiple parameters is, firstly, to simplify the problem with some assumptions (Data Model(2)) in order to obtain the bound $B_Z(2)$, and then to use the transformation T on it to obtain the more realistic bound $B_Z(1)$.

5.6 SUMMARY

In this Chapter, we have the following results:

1. With Ziv-Zakai's idea, MZLB for one incoming signal case is developed.

2. MZLB shows the SNR-threshold phenomenon.

Chapter 6

MZLB on the Covariance Matrix of DOA Estimates for Multiple Uncorrelated Signals

6.1 INTRODUCTION

In chapter 5, we have derived the lower bound for the SSST case using Ziv-Zakai's idea. The practical DOA estimation involves multiple signals with multiple types of parameters. So in this chapter, we develop MZLB for the MSMT case based on the work in Chapter 4 and Chapter 5.

First, in Section 6.2, MZLB(2) under the MSST case is derived with Assumption 6.1 which is an extension of Lemma 4.1 in Chapter 4.3. The properties of MZLB(2) are discussed in Section 6.3. Then, in Section 6.4, MZLB(1), the modified Ziv-Zakai lower bound on the covariance matrix based on Data Model(1) (the MSMT case), is developed from MZLB(2) with a transformation T. The geometric interpretation of this transformation is given in Section 6.5. We propose Theorem 6.1, the final result, in Section 6.6. The properties of MZLB(1) are discussed in Section 6.7, it is indicated that, MZLB(1) shows the $\Delta\phi$-threshold

phenomenon as well as the SNR-threshold phenomenon.

6.2 MZLB(2) FOR MSST CASE

6.2.1 Preliminary

The original Ziv-Zakai's idea is to derive a lower bound on the variance of an estimate by comparing a binary detection procedure based on an arbitrary estimate, with the optimal binary detection procedure for which a tight lower bound is known. We have remarked in Appendix B that the binary detection procedure is detecting one incoming signal identified by one parameter. Therefore, ZZLB is developed for one parameter estimation (the SSST case).

The practical DOA estimation not only involves the estimation of the multiple types of parameters, but also involves the estimation of the multiple DOA parameters. Therefore, in DOA estimation, not only the errors of the different type estimates are coupled with each other, but also the errors of the same DOA type estimates are coupled with each other. The performance analysis under this case is more complicated than that for one DOA estimate developed in Chapter 5. It is very hard to use Ziv-Zakai's idea to derive the lower bound on these coupled errors.

Under some reasonable assumptions, the complicated problem can be decomposed into several sub-problems, such that these sub-problems are independent of each other and are easier to solve. In Chapter 3 and Chapter 4, this method has been applied to decouple the DOA estimation errors with the errors from the signal amplitude estimation and the noise power estimation, i.e., to decouple the estimation errors of the different type of parameters, which results in the performance analysis of the DOA estimation against SNR being independent of that against $\Delta\phi$. In the following, this method will be applied to

decouple the DOA estimation errors, i.e., the estimation errors of the same type parameters, such that the performance analysis of K DOA estimates is decomposed into K independent sub-problems. Then Ziv-Zakai's idea is applied to solve these sub-problems.

Let

$$\vec{\phi}_0 = \begin{bmatrix} \phi_{01} \\ \phi_{02} \\ \vdots \\ \phi_{0K} \end{bmatrix}$$

be the DOA parameter vector to be estimated on K-dimension parameter space $\{\phi_1, \phi_2, ..., \phi_K\}$.

Let the estimate of $\vec{\phi}_0$ be $\hat{\phi}$, and the error vector be

$$\vec{\phi}_e = \hat{\phi} - \vec{\phi}_0 = \begin{bmatrix} \hat{\phi}_1 - \phi_{01} \\ \hat{\phi}_2 - \phi_{02} \\ \vdots \\ \hat{\phi}_K - \phi_{0K} \end{bmatrix} = \begin{bmatrix} \phi_{e1} \\ \phi_{e2} \\ \vdots \\ \phi_{eK} \end{bmatrix}. \tag{6.1}$$

Lemma 4.1 developed in Chapter 4.3 indicates that, under the condition that $\hat{\phi} \to \vec{\phi}_0$ (i.e. the small error $|\vec{\phi}_e|$), the errors of the MLE of DOA based on Data Model(2) are independent of each other and are independent of the locations of DOA. Now, we idealize the behaviour of the MLE(2) of DOA under the large error $|\vec{\phi}_e|$, such that Lemma 4.1 holds also for the large $|\vec{\phi}_e|$ (without the condition $\hat{\phi} \to \vec{\phi}_0$), i.e., Lemma 4.1 holds for all sizes of errors in the error distribution. Therefore, we have

Assumption 6.1 *With the MLE based on Data Model(2), the errors of DOA estimates are independent of each other, and the errors are independent of the locations of DOA.*

The following simulation results show that, when the seperation of DOA is larger than the beam width, ρ_{12}, the correlation coefficient of the errors of two DOA estimates, is

is less than 0.1 and thus the correlation between the errors of DOA estimates is ignorable. Therefore, the Assumption 6.1 is reasonable.

Example (1): K=2[-.5 .5], M=8, N=50

$SNR(db)$	ρ_{12}
20	.0828
16	.0217
12	.0628
8	.0496
4	.0596
0	.0178
−4	.0519
−8	−.0181
−12	.0220
−16	.0793
−20	.0806

Example (2): K=2[-.5 .5], M=16, N=50

$SNR(db)$	ρ_{12}
−20	.0508
−16	.0487
−12	−.0274
−8	−.0840
−4	−.0915
0	.0203
4	−.0672
8	.0200
12	−.0562
16	.0168
20	.0910

Then, with $\mathbf{R}_\phi(2)$ being such an idealized covariance matrix of the DOA estimates, we may write:

$$\mathbf{R}_\phi(2) = E(\vec{\phi}_\epsilon \vec{\phi}_\epsilon^T) = \text{diag}(\sigma_1, ..., \sigma_K), \tag{6.2}$$

where, the variances of DOA estimates σ_k $(k = 1, ..., K)$ are independent of each other and are independent of the locations of DOA. From Eq.(4.24), we see that, when the incoming

signals have equal powers, σ_k are equal to each other.

For conciseness, in the following, all the discussions are developed for K=2, where, K is the number of the incoming signals. The results can be extended to $K > 2$ cases directly.

6.2.2 MZLB(2)

In this section, we derive the lower bound on the covariance matrix $\mathrm{R}_\phi(2)$ with Assumption 6.1.

The received data vector at the n^{th} snapshot from a M-elements array is

$$\mathrm{r}(n) = \mathrm{s}(n) + \mathrm{e}(n), \tag{6.3}$$

in which, signal

$$\mathrm{s}(n) = [\mathrm{d}_{01}, \mathrm{d}_{02}] \begin{bmatrix} a_1(n) \\ a_2(n) \end{bmatrix} = \mathrm{s}_{01} + \mathrm{s}_{02}, \tag{6.4}$$

where, d_{0k} is the k^{th} steering vector, $a_k(n)$ is the amplitude of the k^{th} incoming signal, s_{01} arrives from the direction ϕ_{01}, and s_{02} arrives from the direction ϕ_{02}.

Since the errors are independent of each other, we may derive the variance of $\hat{\phi}_1$ setting $\phi_{\varepsilon 2} = \hat{\phi}_2 - \phi_{02} = 0$, where, $\hat{\phi}_1$ is the estimate of ϕ_{01}, and $\phi_{\varepsilon 2}$ is the error of the estimate $\hat{\phi}_2$. To find the $E(\phi_{\varepsilon 1}^2)$, where E denotes the expectation, we imagine a binary detection procedure. Suppose the incoming DOA vector is of two possibilities: $\vec{\phi}_0$ and $\vec{\phi}_1$, as shown in Fig. 6.1,

where,

$$\vec{\phi}_0 = \begin{bmatrix} \phi_{01} \\ \phi_{02} \end{bmatrix},$$

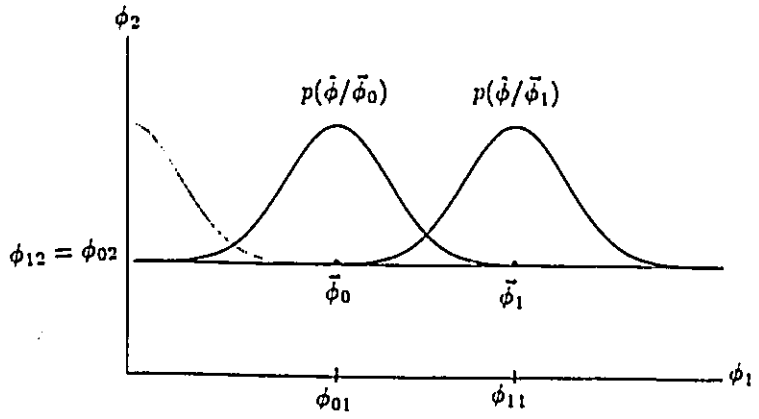


Figure 6.1: Imaginary binary detection procedure for K=2.

$$\vec{\phi}_1 = \begin{bmatrix} \phi_{11} \\ \phi_{12} \end{bmatrix},$$

and

$$\phi_{12} = \phi_{02}. \tag{6.5}$$

We establish the following hypotheses such that

H_0 : $DOA = \vec{\phi}_0$,

H_1 : $DOA = \vec{\phi}_1$.

Based on some arbitrary estimate $\hat{\phi}$ of the incoming DOA, which is

$$\hat{\phi} = \begin{bmatrix} \hat{\phi}_1 \\ \hat{\phi}_2 \end{bmatrix},$$

the decision rule is

$$|\hat{\phi} - \vec{\phi}_0| \underset{H_1}{\overset{H_0}{\lesssim}} |\hat{\phi} - \vec{\phi}_1|. \tag{6.6}$$

Note that $\phi_{\epsilon 2} = 0$ and Eq.(6.5) result in $\hat{\phi}_2 = \phi_{02} = \phi_{12}$, therefore,

$$|\hat{\phi} - \vec{\phi}_0| = \sqrt{(\hat{\phi}_1 - \phi_{01})^2 + (\hat{\phi}_2 - \phi_{02})^2} = |\hat{\phi}_1 - \phi_{01}| \tag{6.7}$$

and

$$|\hat{\phi} - \vec{\phi}_1| = \sqrt{(\hat{\phi}_1 - \phi_{11})^2 + (\hat{\phi}_2 - \phi_{12})^2} = |\hat{\phi}_1 - \phi_{11}|. \tag{6.8}$$

Substituting Eqs.(6.7) and (6.8) into Eq.(6.6), then comparing it with Eq.(5.2), we see that Eq.(6.6) is the same as Eq.(5.2). Thus following the four steps described in Chapter 5.3, we have

$$E(\phi_{\epsilon 1}^2) = \int_0^{\pi} h^2 dP(\varepsilon 1 \le h)|_{\phi_{\epsilon 2}=0} \ge \int_0^{2\pi} h P_{EM1}(h) dh|_{\phi_{\epsilon 2}=0}. \tag{6.9}$$

Now, the same procedure of setting $\phi_{\epsilon 1} = 0$ for the variance of $\hat{\phi}_2$ can be applied so that,

$$E(\phi_{\epsilon 2}^2) = \int_0^{\pi} h^2 dP(\varepsilon 2 \le h)|_{\phi_{\epsilon 1}=0} \ge \int_0^{2\pi} h P_{EM2}(h) dh|_{\phi_{\epsilon 1}=0}. \tag{6.10}$$

The RHS of Eqs.(6.9) and (6.10) are the lower bounds on the variances of the estimates $\hat{\phi}_1$ and $\hat{\phi}_2$.

6.2.3 Evaluation of MZLB(2)

To evaluate MZLB, first, we need to find $P_{EM1}(h)$ in Eq.(6.9) since the PDF of the received data for two incoming signals is different from that for one incoming signal. The Gaussian conditional PDF of the received data is given by

$$p(\mathrm{r}/\vec{\phi}_i) = \frac{1}{(\pi\sigma_e^2)^{MN/2}} \exp[-\frac{1}{\sigma_e^2} \sum_{n=1}^{N} (\mathrm{r} - \mathrm{s}_{i1} - \mathrm{s}_{i2})^H (\mathrm{r} - \mathrm{s}_{i1} - \mathrm{s}_{i2})], \tag{6.11}$$

where, $i = 0, 1$ depends on the hypothesis. With Eq.(6.11), $P_{EM1}(h)|_{\epsilon 2=0}$ in Eq.(6.9) is derived in Appendix H, which is (see Eqs.(H.10) and (H.11))

$$P_{EM1}(h)|_{\phi_{\epsilon 2}=0} = \text{erfc}_*\{f_1(h)\}, \tag{6.12}$$

where,

$$f_1(h) = \{N\rho_1 \sum_{m=1}^{M}[1 - \cos((m-1)h)]\}^{.5}. \tag{6.13}$$

Let $B_Z^{il}(2)$ denote the $(i,l)^{th}$ element in the lower bound matrix $\mathrm{B}_Z(2)$, Eq.(6.9) becomes

$$\begin{aligned} E(\phi_{\epsilon 1}^2) &\geq \int_0^{2\pi} \text{erfc}_*\{f_1(h)\}hdh \\ &\triangleq B_Z^{11}(2). \end{aligned} \tag{6.14}$$

Similarly, Eq.(6.10) becomes

$$\begin{aligned} E(\phi_{\epsilon 2}^2) &\geq \int_0^{2\pi} hP_{EM2}(h)dh|_{\phi_{\epsilon 1}=0} \\ &= \int_0^{2\pi} \text{erfc}_*\{f_2(h)\}hdh \\ &\triangleq B_Z^{22}(2), \end{aligned} \tag{6.15}$$

where,

$$f_2(h) = \{N\rho_2 \sum_{n=1}^{M}[1 - \cos((m-1)h)]\}^{.5}.$$

Comparing Eqs.(6.14) and (6.15) with Eq.(5.27), we see that the lower bounds $B_Z^{11}(2)$ and $B_Z^{22}(2)$ on the variances of $\hat{\phi}_1$ and $\hat{\phi}_2$ are as the same as that for one incoming signal. Thus using the results developed in Chapter 5 (see Eq.(5.48)), we can evaluate the integrals in Eqs.(6.14) and (6.15):

$$B_Z^{11}(2) \simeq \frac{H^2}{2}\text{erfc}_*\{f_1(H)\} + \frac{3}{\rho_1 N M(M-1)(2M-1)}, \tag{6.16}$$

$$B_Z^{22}(2) \simeq \frac{H^2}{2}\text{erfc}_*\{f_2(H)\} + \frac{3}{\rho_2 NM(M-1)(2M-1)}, \tag{6.17}$$

where, $H = 2\pi - \pi/2M$ as developed in Chapter 5.4.

Since the estimate errors are assumed to be uncorrelated, i.e., $\sigma_{12} = \sigma_{21} = 0$, their lower bounds

$$B_Z^{12}(2) = B_Z^{21}(2) = 0. \tag{6.18}$$

Combing Eqs.(6.16) to (6.18), MZLB(2), the lower bound on the covariance matrix $\mathrm{R}_\phi(2)$ is

$$\begin{aligned}
&\mathrm{B}_Z(2) \\
&\simeq \text{diag}\Big(\frac{H^2}{2}\text{erfc}_*\{f_1(H)\} + \frac{3}{\rho_1 NM(M-1)(2M-1)}, \\
&\quad \frac{H^2}{2}\text{erfc}_*\{f_2(H)\} + \frac{3}{\rho_2 NM(M-1)(2M-1)}\Big) \\
&= \text{diag}\Big(\frac{H^2}{2}\text{erfc}_*\{f_1(H)\}, \frac{H^2}{2}\text{erfc}_*\{f_2(H)\}\Big) + \mathrm{B}_{CR}(2),
\end{aligned} \tag{6.19}$$

where, the first term is the threshold term, and $\mathrm{B}_{CR}(2)$ is the CRLB on the covariance matrix of DOA estimates with Data Model(2) as defined in Eq.(3.41),

$$\mathrm{B}_{CR}(2) \simeq \text{diag}\Big(\frac{3}{\rho_1 NM(M-1)(2M-1)}, \frac{3}{\rho_2 NM(M-1)(2M-1)}\Big). \tag{6.20}$$

6.3 MZLB(2) AND THRESHOLD PHENOMENA

With two examples $B_Z^{11}(2)$ and $B_{CR}^{11}(2)$ against SNR are plotted in Fig.6.2 (a) and (b), The dash line in Fig.6.2 is the variance of the MLE of ϕ_{01} obtained from the simulation based on Data Model(2), which is denoted as $V_{ML}^{11}(2)$.

It can be seen that, because of the threshold term in Eq.(6.19), $B_Z^{11}(2)$ (the solid line) shows the SNR-threshold phenomenon, and is much tighter than $B_{CR}^{11}(2)$ (the dotted line) when the noise level is high. It is obvious that Eq.(6.19) is independent of $\Delta\phi$, so MZLB(2) cannot show the $\Delta\phi$-threshold phenomenon.

Therefore, we have,

Observation 6.1 *MZLB(2) shows the SNR-threshold phenomenon.*

Observation 6.2 *MZLB(2) does not show the $\Delta\phi$-threshold phenomenon.*

Observation 6.3 *MZLB(2) against SNR is independent of $\Delta\phi$.*

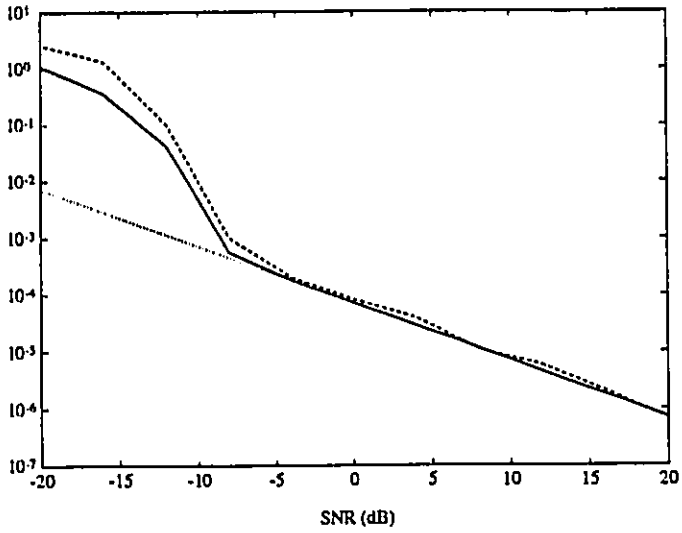
The geometric interpretation of MZLB(2) is shown in Fig.6.3 ($\sigma_{s1}^2 = \sigma_{s2}^2$). Here, the dotted lines show the equiprobable contours given by

$$\vec{\phi}_\epsilon^T [\mathrm{R}_\phi(2)]^{-1} \vec{\phi}_\epsilon = C^2,$$

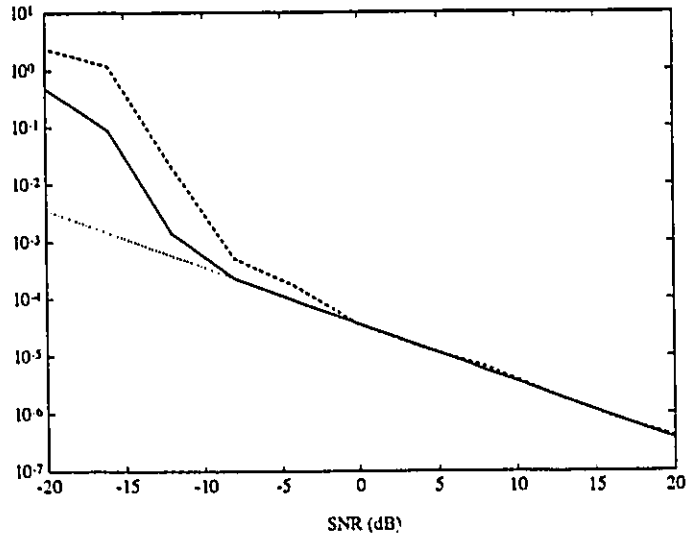
the solid lines show the bound ellipses given by

$$\vec{\phi}_\epsilon^T [\mathrm{B}_Z(2)]^{-1} \vec{\phi}_\epsilon = C^2.$$

Comparing Fig.6.3 and Fig.4.7, the difference between MZLB(2) and CRLB(2) in the geometric interpretation is that MZLB(2) bound ellipses are much closer to the equiprobable contours than CRLB(2) bound ellipses are when the errors spread out from the zero.



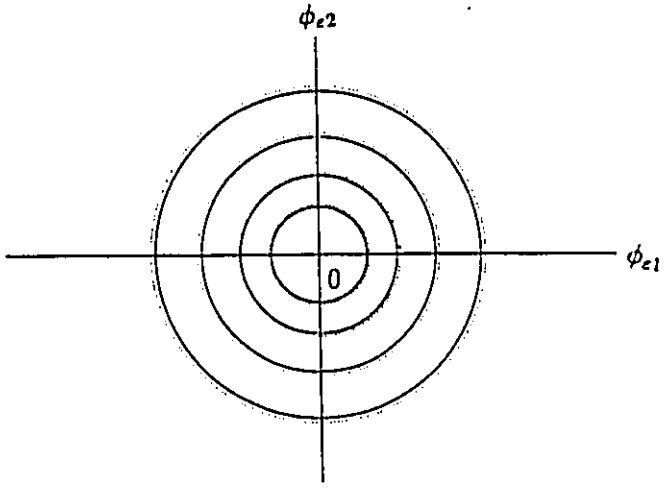
(a). M=8 N=50 K=2[-.5 .5]



(b). M=16 N=50 K=2[-.5 .5]

— $B_Z^{11}(2)$ ... $B_{CR}^{11}(2)$ $---V_{ML}^{11}(2)$

Figure 6.2: Lower bounds and variance of MLE based on Data Model(2).



$$\_\vec{\phi}_\epsilon^T \mathbf{B}_Z^{-1}(2)\vec{\phi}_\epsilon = C^2, \qquad \ldots\vec{\phi}_\epsilon^T \mathbf{R}_\phi^{-1}(2)\vec{\phi}_\epsilon = C^2.$$

Figure 6.3: Equiprobable contours and MZLB ellipses based on Data Model(2).

6.4 MZLB(1)

In Chapter 4 and Section 6.3, Data Model(2) was used in order to apply the Ziv-Zakai's idea to derive the error bound. The resulting MZLB(2) for MSST case is independent of the DOA separations $\Delta\phi$. Since Data Model(2) is not a practical model as it assumes the knowledge of the signal amplitudes and the noise power, we need to employ Data Model(1) and find MZLB based on Data Model(1). Now, we make use of the results for Data Model(2) so that the bounds for Data Model(1) can be developed.

Let $\mathbf{R}_\phi(1)$ be the covariance matrix of the DOA estimates based on Data Model(1). Since $\mathbf{R}_\phi(2)$ is a diagonal covariance matrix (see Eq.(6.2)), it is full rank. Thus, we can find a transformation T such that

$$\mathbf{R}_\phi(1) = \mathbf{T}\mathbf{R}_\phi(2). \tag{6.21}$$

By the coordinate transformation of the error vector $\vec{\phi}_\varepsilon$,

$$\phi_{\varepsilon 1} = h\cos\beta$$

and

$$\phi_{\varepsilon 2} = h\sin\beta,$$

we define the error vector in the polar coordinate

$$\vec{\phi}_{\varepsilon h} = \begin{bmatrix} h\cos\beta \\ h\sin\beta \end{bmatrix}, \tag{6.22}$$

where,

$$h = \sqrt{\phi_{\varepsilon 1}^2 + \phi_{\varepsilon 2}^2} = |\vec{\phi}_\varepsilon| \in [0, \lambda(\beta)],$$

$$\beta = \arctan(\frac{\phi_{\varepsilon 2}}{\phi_{\varepsilon 1}}) \in [0, 2\pi],$$

in which, $\lambda(\beta)$ is the maximum length of h and is therefore a function of the angle β. For example, when $\beta = 0$, $\lambda(0) = \pi$ and $h \in [0, \pi]$; when $\beta = \pi/4$, $\lambda(\pi/4) = \sqrt{2}\pi$ and $h \in [0, \sqrt{2}\pi]$, as shown in Fig.6.4. Therefore the covariance matrix of the estimate $\hat{\phi}$ becomes

$$\begin{aligned} \mathrm{R}_\phi(l) &= \int_0^\pi \int_0^\pi \vec{\phi}_\varepsilon \vec{\phi}_\varepsilon^T p_l(\vec{\phi}_\varepsilon) d\phi_{\varepsilon 1} d\phi_{\varepsilon 2} \\ &= \int_0^{2\pi} \int_0^{\lambda(\beta)} \vec{\phi}_{\varepsilon h} \vec{\phi}_{\varepsilon h}^T p_l(\vec{\phi}_{\varepsilon h}) dh d\beta, \end{aligned} \tag{6.23}$$

where, $p_l(\vec{\phi}_\varepsilon)$ is the PDF of the vector $\vec{\phi}_\varepsilon$, $p_l(\vec{\phi}_{\varepsilon h})$ is the PDF of the vector $\vec{\phi}_{\varepsilon h}$, $l = 1, 2$ depends on the data model.

We examine this covariance matrix in two regions. The first one is with $h \leq \Delta$, Δ is a small value, $h = |\vec{\phi}_\varepsilon| \in [0, \Delta]$ means that the errors distribute around the zero. The

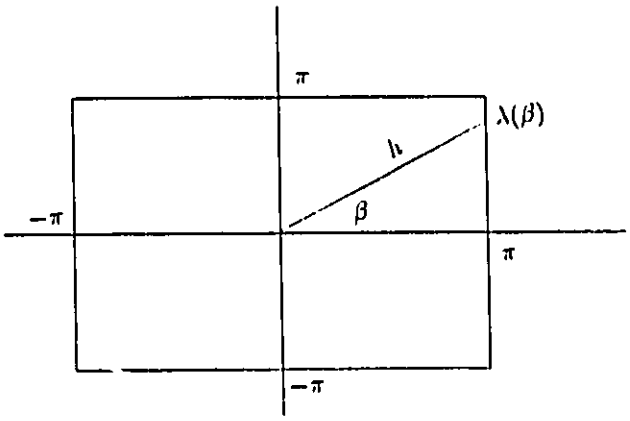


Figure 6.4: $h \in [0, \lambda(\beta)]$

second one is with $\Delta < h \leq \lambda(\beta)$, $h = |\vec{\phi}_e| \in [\Delta, \lambda(\beta)]$ means that the errors spread away from the zero. Then we have, based on Data Model(1),

$$\begin{aligned} \mathrm{R}_{\phi}(1) &= \int_0^{2\pi} \int_0^{\lambda(\beta)} \vec{\phi}_{eh} \vec{\phi}_{eh}^T p_1(\vec{\phi}_{eh}) dh d\beta \\ &= \int_0^{2\pi} \int_0^{\Delta} \vec{\phi}_{eh} \vec{\phi}_{eh}^T p_1(\vec{\phi}_{eh}) dh d\beta + \int_0^{2\pi} \int_{\Delta}^{\lambda(\beta)} \vec{\phi}_{eh} \vec{\phi}_{eh}^T p_1(\vec{\phi}_{eh}) dh d\beta \\ &\triangleq \mathrm{R}_{\phi l}(1) + \mathrm{R}_{\phi g}(1), \end{aligned} \tag{6.24}$$

and, based on Data Model(2),

$$\begin{aligned} \mathrm{R}_{\phi}(2) &= \int_0^{2\pi} \int_0^{\lambda(\beta)} \vec{\phi}_{eh} \vec{\phi}_{eh}^T p_2(\vec{\phi}_{eh}) dh d\beta \\ &= \int_0^{2\pi} \int_0^{\Delta} \vec{\phi}_{eh} \vec{\phi}_{eh}^T p_2(\vec{\phi}_{eh}) dh d\beta + \int_0^{2\pi} \int_{\Delta}^{\lambda(\beta)} \vec{\phi}_{eh} \vec{\phi}_{eh}^T p_2(\vec{\phi}_{eh}) dh d\beta \\ &\triangleq \mathrm{R}_{\phi l}(2) + \mathrm{R}_{\phi g}(2). \end{aligned} \tag{6.25}$$

Substituting Eqs.(6.24) and (6.25) into Eq.(6.21) yields

$$\mathrm{R}_{\phi l}(1) + \mathrm{R}_{\phi g}(1) = \mathrm{T}[\mathrm{R}_{\phi l}(2) + \mathrm{R}_{\phi g}(2)]. \tag{6.26}$$

As proposed in Chapter 5.5.2, now, we conjecture that

Assumption 6.2 *T is a linear transformation such that the transformation between* $\mathrm{R}_{\phi g}(1)$ *and* $\mathrm{R}_{\phi g}(2)$ *is the same as that between* $\mathrm{R}_{\phi l}(1)$ *and* $\mathrm{R}_{\phi l}(2)$.

Using Assumption 6.2 in Eq.(6.26), we have

$$\mathrm{R}_{\phi l}(1) = \mathrm{T}\mathrm{R}_{\phi l}(2), \tag{6.27}$$

$$\mathrm{R}_{\phi g}(1) = \mathrm{T}\mathrm{R}_{\phi g}(2). \tag{6.28}$$

To obtain the transformation T, we note that, because CRLB is achievable by the covariance matrix for $|\vec{\phi}_e| \in [0, \Delta]$ as discussed in Chapter 4, we have

$$\mathrm{B}_{CR}(2) \simeq \mathrm{R}_{\phi l}(2), \tag{6.29}$$

and

$$\mathrm{B}_{CR}(1) \simeq \mathrm{R}_{\phi l}(1) \tag{6.30}$$

for large M [77]. Substituting Eqs.(6.29) and (6.30) into Eq.(6.27), we have

$$\mathrm{B}_{CR}(1) \simeq \mathrm{T}\mathrm{B}_{CR}(2). \tag{6.31}$$

Therefore

$$\mathrm{T} \simeq \mathrm{B}_{CR}(1)\mathrm{B}_{CR}^{-1}(2), \tag{6.32}$$

where, $\mathrm{B}_{CR}(2)$ is a positive definite matrix (see Eq.(3.41), it is a diagonal matrix and all the elements are positive).

T transforms the covariance matrix based on Data Model(2) to the covariance matrix based on Data Model(1). Since under high SNR, the covariance matrices reach the lower bounds, therefore it is reasonable to assume that the same linear transformation T applies

to the bounds of Data Model(1) and Data Model(2). Thus, the MZLB based on Data Model(1) can be obtained from MZLB(2), i.e., (see Eqs.(6.21), (6.19), and (6.31))

$$\begin{aligned} \mathrm{B}_Z(1) &= \mathrm{T}\mathrm{B}_Z(2) \\ &\simeq \mathrm{T}[0.5H^2\mathrm{diag}(\mathrm{erfc}_*\{f_1(H)\},\mathrm{erfc}_*\{f_2(H)\}) + \mathrm{B}_{CR}(2)] \\ &= 0.5H^2\mathrm{T}\mathrm{diag}(\mathrm{erfc}_*\{f_1(H)\},\mathrm{erfc}_*\{f_2(H)\}) + \mathrm{B}_{CR}(1). \end{aligned} \tag{6.33}$$

The RHS of Eq.(6.33) is the MZLB on the covariance matrix of the DOA estimates with two uncorrelated signals in the white noise based on Data Model(1).

Comparing Eq.(6.32) and Eq.(3.45), we see that, this is the same transformation T,

$$\begin{aligned} \mathrm{T} &= \mathrm{B}_{CR}(1)\mathrm{B}_{CR}^{-1}(2) \\ &= \mathrm{diag}(\frac{M(M-1)(2M-1)}{6\dot{\mathrm{d}}_1^H\mathrm{P}_e\dot{\mathrm{d}}_1}, \frac{M(M-1)(2M-1)}{6\dot{\mathrm{d}}_2^H\mathrm{P}_e\dot{\mathrm{d}}_2}). \end{aligned} \tag{6.34}$$

6.5 GEOMETRIC INTERPRETATION OF T

For conciseness, here we assume the equal signal powers, i.e., $E[a_1^2] = E[a_2^2] = \sigma_s^2$. Then, Eq.(6.2) becomes

$$\mathrm{R}_\phi(2) = \sigma\mathrm{I}, \tag{6.35}$$

$\mathrm{R}_\phi(2)$ is of the form F-III discussed in Chapter 4.2. The equiprobable contours given by

$$\vec{\phi}_e^T\mathrm{R}_\phi^{-1}(2)\vec{\phi}_e = C^2 \tag{6.36}$$

are plotted in Fig.6.5, where, the ellipses 1' and 2' represent the equiprobable contours for the smaller errors $|\vec{\phi}_e|$, the ellipses 3' and 4' represent the equiprobable contours for the larger errors $|\vec{\phi}_e|$.

Since $B_{CR}(1)$ in Eq.(3.27) is of form F-II in Chapter 4.2, $R_{\phi l}(1) \simeq B_{CR}(1)$ (see Eq.(6.30)) is of form F-II. The equiprobable contours for $R_{\phi l}(1)$ are shown by the ellipses 1 and 2 in Fig. 6.6. The equiprobable contours 1 and 2 in Fig. 6.6 can be mapped from 1' and 2' in Fig.6.5 with a transformation T. T can be obtained by Eq.(6.32) as discussed in Section 6.4. Then, with Assumption 6.2, the same transformation T maps the equiprobable contours 3' and 4' in Fig.6.5 to the equiprobable contours 3 and 4 in Fig.6.6 which are corresponding to the large $|\vec{\phi}_e|$.

Thus the equiprobable contours for $R_\phi(1)$ given by

$$\vec{\phi}_e^T R_\phi^{-1}(1)\vec{\phi}_e = C^2 \tag{6.37}$$

shown in Fig.6.6 is of the form F-II.

Geometrically, T maps a family of ellipses (circles) into another family of ellipses.

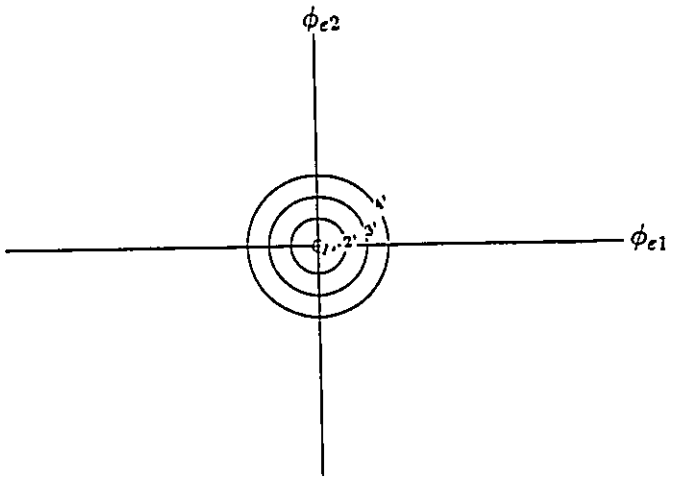


Figure 6.5: Equiprobable contours of $p(\vec{\phi}_e)$ and MZLB ellipses based on Data Model(2).

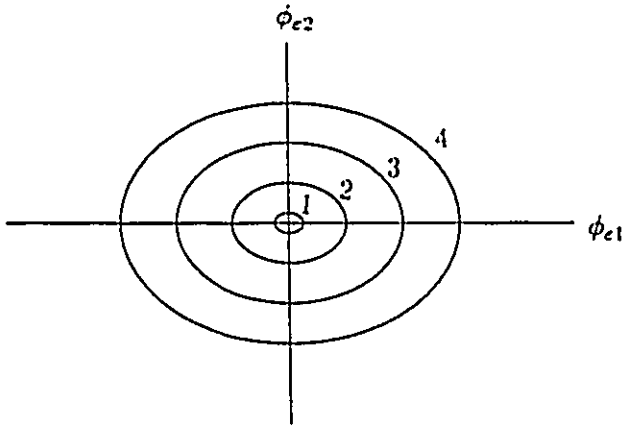


Figure 6.6: Equiprobable contours of $p(\vec{\phi}_e)$ based on Data Model(1).

6.6 MZLB ON THE COVARIANCE MATRIX OF DOA ESTIMATES FOR MULTIPLE UNCORRELATED SIGNALS

MZLB(1) derived in Section 6.4 is based on the Assumptions 6.1 and 6.2. These assumptions are true for the small estimation errors. Assumptions 6.1 and 6.2 hypothesize that these assumptions hold for the larger errors and therefore idealize the behaviour of the covariance matrix of DOA estimates for the large $|\vec{\phi}_e|$. We call the errors under these considerations the idealized errors.

Using these assumptions and extending the case for K=2 to the general case, we therefore can propose the following :

Theorem 6.1 *The modified Ziv-Zakai lower bound (MZLB) on the covariance matrix of the DOA estimates for multiple uncorrelated signals in the white noise is*

$$\mathrm{B}_Z(1) = \mathrm{T}\mathrm{B}_Z(2) \tag{6.38}$$

$$\simeq .5H^2\mathrm{T}\mathrm{diag}(\mathrm{erfc}_*\{f_1(H)\},...,\mathrm{erfc}_*\{f_K(H)\}) + \mathrm{B}_{CR}(1), \tag{6.39}$$

in which,

$$\mathrm{T} = \mathrm{B}_{CR}(1)\mathrm{B}_{CR}^{-1}(2)$$

$$= \mathrm{diag}(\frac{M(M-1)(2M-1)}{6\dot{\mathrm{d}}_1^H\mathrm{P}_e\dot{\mathrm{d}}_1},...,\frac{M(M-1)(2M-1)}{6\dot{\mathrm{d}}_K^H\mathrm{P}_e\dot{\mathrm{d}}_K}), \tag{6.40}$$

$$H = 2\pi - \frac{\pi}{2M}, \tag{6.41}$$

$$f_k(H) = [\rho_k N \sum_{m=1}^{M}(1-\cos((m-1)H))]^{.5}, \tag{6.42}$$

$$\mathrm{erfc}_*\{x\} = \frac{1}{\sqrt{2\pi}}\int_x^\infty e^{-t^2/2}dt, \tag{6.43}$$

$$\mathrm{B}_{CR}(1) = \mathrm{diag}(\frac{1}{2\rho_1 N\dot{\mathrm{d}}_1^H P_e\dot{\mathrm{d}}_1},...,\frac{1}{2\rho_K N\dot{\mathrm{d}}_K^H P_e\dot{\mathrm{d}}_K}). \tag{6.44}$$

Since there is a threshold term in $\mathrm{B}_Z(1)$, and the computation for this term is simple, MZLB is of the following advantages:

1. It is tight in a wide range of SNR. It follows the SNR-threshold phenomenon as well as the $\Delta\phi$-threshold phenomenon occurring in the DOA estimation.

2. It is easily computable.

6.7 MZLB(1) AND THRESHOLD PHENOMENA

The difference between MZLB and CRLB is the threshold term (the first term in Eq.(6.39)). First, we discuss the properties of the SNR-threshold phenomenon of MZLB(1) in detail.

Property 6.1 *MZLB(1) shows the SNR-threshold phenomenon.*

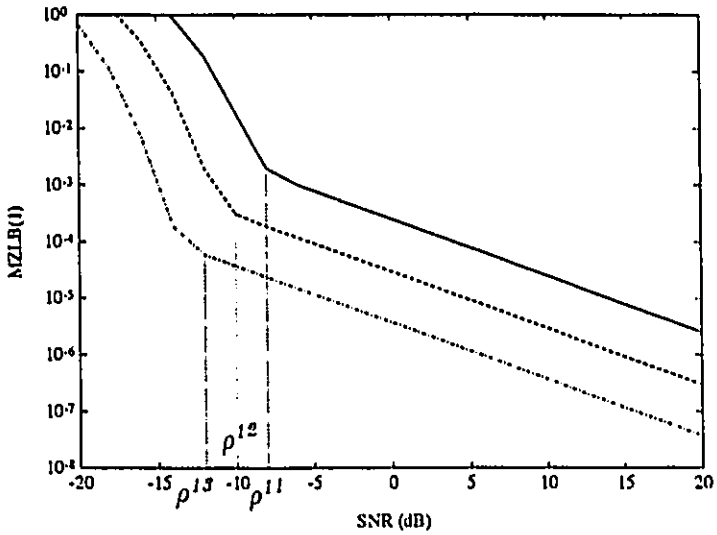
1. *The larger the number of sensors M, the lower the SNR at which the threshold ρ^l of MZLB(1) occurs.*
2. *The larger the number of snapshots N , the lower the SNR at which the threshold ρ^l of MZLB(1) occurs.*
3. *The SNR at which the threshold ρ^l of MZLB(1) occurs does not vary with the separations of DOA ($\Delta\phi$), but the whole lower bound varies with the $\Delta\phi$.*

   proof: Let $B_Z^{il}(1)$ denote the $(i,l)^{th}$ element in $\mathrm{B}_Z(1)$.

1. From Eqs.(6.39), (6.42) and (6.43), we see that when M increases, $f_k(II)$ increases since the number of the terms, which are positive, in the summation of Eq.(6.42) increases. This results in that the complement error function $\mathrm{erfc}_*\{f_k(II)\}$, therefore the threshold term, decreases. This decreasing is much faster than the decreasing of Eq.(6.44) which results from the same increasing of M. Therefore, the threshold ρ^l of $B_Z^{11}(1)$ shifts toward the left with M increasing as shown in Fig.6.7. Here, the number of signals K=2, $\phi_{01} = -.5$ radian, $\phi_{02} = .5$ radian, the number of snapshots N=50; the solid line represents the behaviour of $B_Z^{11}(1)$ with the threshold ρ^{l1} for the number of sensors $M_1 = 8$, the dash line represents the behaviour of $B_Z^{11}(1)$ with the threshold ρ^{l2} for $M_2 = 16$, the dash-dotted line represents the behaviour of $B_Z^{11}(1)$ with the threshold ρ^{l3} for $M_3 = 32$. It is shown that, $\rho^{l1} > \rho^{l2} > \rho^{l3}$ for $M_1 < M_2 < M_3$. The same property is observed with $B_Z^{22}(1)$.

2. When N in Eq.(6.42) increases, $f_k(H)$ increases. This results in erfc$_*\{f_k(H)\}$ in the threshold term in Eq.(6.39) decreases. This decreasing is much faster than the decreasing of Eq.(6.44) which results from the same increasing of N. Therefore, the threshold ρ^l shifts toward the left with N increasing as shown in Fig.6.8. Here, K=2, $\phi_{01} = -.5$ radian, $\phi_{02} = .5$ radian, M=8; the solid line represents the behaviour of $B_Z^{11}(1)$ with the threshold ρ^{l1} for $N_1 = 50$, the dash line represents the behaviour of $B_Z^{11}(1)$ with the threshold ρ^{l2} for $N_2 = 100$, the dash-dotted line represents the behaviour of $B_Z^{11}(1)$ with the threshold ρ^{l3} for $N_3 = 200$. It is shown that, $\rho^{l1} > \rho^{l2} > \rho^{l3}$ for $N_1 < N_2 < N_3$. The same property is observed with $B_Z^{22}(1)$.

3. In Eq.(6.38), $\mathrm{B}_Z(2)$ is a function of SNR and is independent of $\Delta\phi$ (see Eq.(6.19)). T is a function of $\Delta\phi$ and is independent of SNR (see Eq.(6.40)). So, $\mathrm{B}_Z(1) = \mathrm{T}\mathrm{B}_Z(2)$ against SNR with different $\Delta\phi$ are parallel to each other. Therefore, the threshold ρ^l does not vary with $\Delta\phi$ as shown in Fig.6.9. Here, K=2, M=8, and N=50; the solid line represents the behaviour of $B_Z^{11}(1)$ with the threshold ρ^{l1} for $\Delta\phi_1 = 0.2$ radian, the dash line represents the behaviour of $B_Z^{11}(1)$ with the threshold ρ^{l2} for $\Delta\phi_2 = 0.5$ radian, the dash-dotted line represents the behaviour of $B_Z^{11}(1)$ with the threshold ρ^{l3} for $\Delta\phi_3 = 1$ radian. It is shown that, $\rho^{l1} = \rho^{l2} = \rho^{l3}$ for $\Delta\phi_1 < \Delta\phi_2 < \Delta\phi_3$. The same property is observed with $B_Z^{22}(1)$. This may be explained as the following: With Data Model(2), i.e., the signal amplitude a is assumed to be known, MZLB(2) is independent of the separations of DOA ($\Delta\phi$). With the Data Model(1), i.e., the parameters to be estimated include the signal amplitude a, MZLB(1) is a function of $\Delta\phi$. Therefore it can be concluded that the dependence of MZLB(1) on $\Delta\phi$ is due to the necessity of the estimation of a. The MLE of a is linear (see Chapter 7, Eq.(7.2)). The resulting influence on the DOA estimation is linear, i.e. MZLB(1) against SNR with the different $\Delta\phi$ are parallel to each other; therefore, the SNR at which the SNR-threshold ρ^l occurs does not vary with $\Delta\phi$.



N=50 K=2[-.5 .5]
— M_1=8 - - - M_2=16 ··· M_3=32

Figure 6.7: SNR-threshold of MZLB(1) under different M.

#.

Second, we will discuss the $\Delta\phi$-threshold phenomenon of MZLB(1).

Property 6.2 *MZLB(1) shows the $\Delta\phi$-threshold phenomenon.*

1. *The larger the sensor number M, the smaller the threshold γ.*
2. *The $\Delta\phi$-threshold γ of MZLB(1) is independent of SNR.*

Properties 6.5 can be seen with an example as shown in Fig.6.10 and Fig.6.11. In Fig. 6.10, with SNR=0 dB, N=50, K=2, the dash-dotted line represents the behaviour of $B_Z^{11}(1)$ with the threshold γ_1 for $M_1 = 8$; the dash line represents the behaviour of $B_Z^{11}(1)$ with the threshold γ_2 for $M_2 = 16$; the solid line represents the behaviour of $B_Z^{11}(1)$ with the threshold γ_3 for $M_3 = 32$. It is shown that, $\gamma_1 > \gamma_2 > \gamma_3$ for $M_3 > M_2 > M_1$. The same property is observed with $B_Z^{22}(1)$.

In Fig. 6.11, with SNR=0 dB, N=50, K=2, the dash-dotted line represents the behaviour of $B_Z^{11}(1)$ with the threshold γ_1 for $SNR_1 = -10dB$; the dash line represents the behaviour of $B_Z^{11}(1)$ with the threshold γ_2 for $SNR_2 = 0dB$; the solid line represents the behaviour of $B_Z^{11}(1)$ with the threshold γ_3 for $SNR_3 = 10dB$. It is shown that, $\gamma_3 = \gamma_2 = \gamma_3$ for $SNR_3 > SNR_2 > SNR_1$. The same property is observed with $B_Z^{22}(1)$.

Finally, we discuss how SNR and $\Delta\phi$ couple with each other in MZLB(1).

Property 6.3 *In MZLB(1), $\Delta\phi$ and SNR are coupled with each other in the way that, MZLB(1) against $\Delta\phi$ with the different SNR are parallel to each other, or, MZLB(1) against SNR with the different $\Delta\phi$ are parallel to each other.*

Property 6.6 can be seen obviously in Fig.6.9 and Fig.6.11.

6.8 SUMMARY

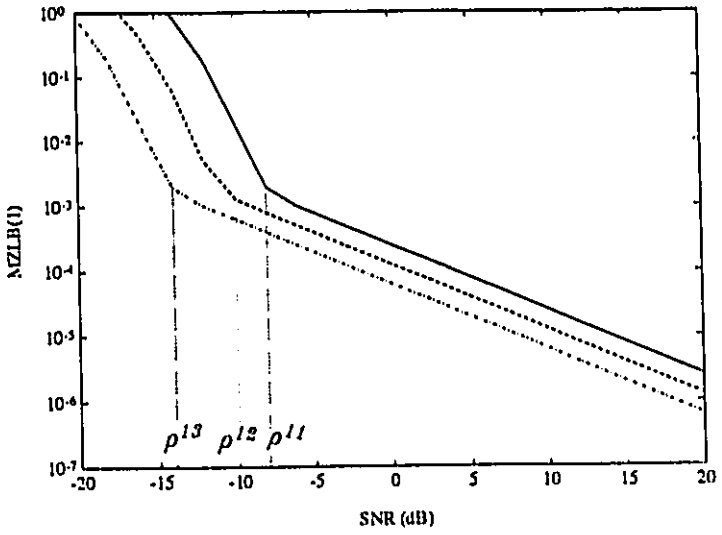
In this Chapter, we have the following results:

1. With Assumption 6.1, which is based on Lemma 4.1 derived in Chapter 4.3, the MZLB(2) for the MSST case is developed based on the MZLB(2) for the SSST case conducted in Chapter 5.

2. With Assumption 6.2, which idealizes the behaviour of the large estimation errors, MZLB(1) is developed from MZLB(2) with a transformation T.

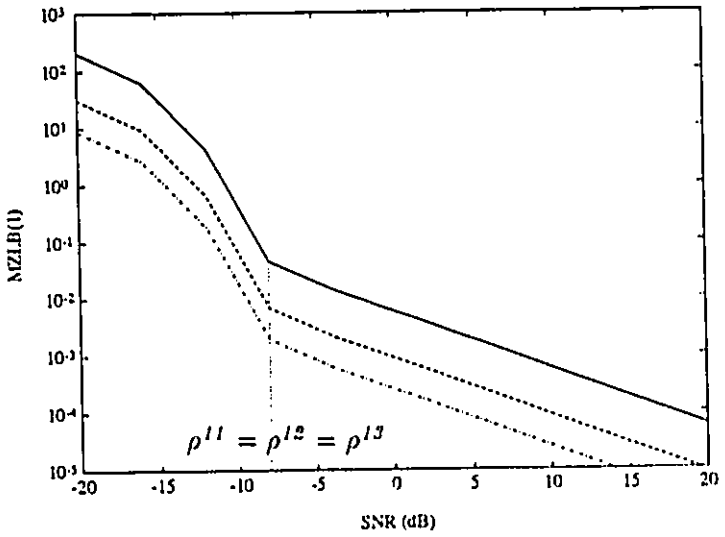
3. MZLB(1), the modified Ziv-Zakai lower bound for the MSMT case, shows the $\Delta\phi$-threshold phenomenon as well as the SNR-threshold phenomenon.

4. MZLB(1) is much tighter than CRLB(1) under low SNR. It is easily computable.



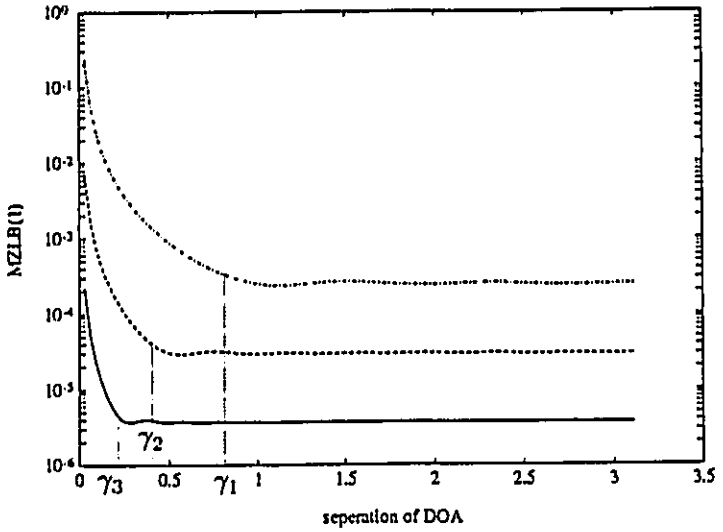
M=8 K=2[-.5 .5]
— N_1=50 - - - N_2=100 ··· N_3=200

Figure 6.8: SNR-threshold of MZLB(1) under different N.



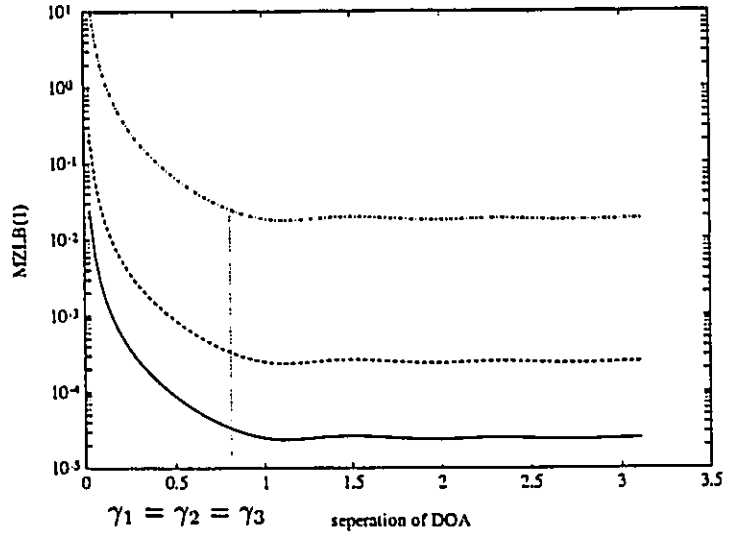
M=8 N=50 K=2
— $\Delta\phi_1$ = 0.2[−.1.1] - - - $\Delta\phi_2$ = 0.5[−.25.25] ··· $\Delta\phi_3$ = 1[−.5.5]

Figure 6.9: SNR-threshold of MZLB(1) under different $\Delta\phi$.



SNR=0 dB N=50 K=2
--- M_1=8 - - - M_2=16 — M_3=32

Figure 6.10: $\Delta\phi$-threshold of MZLB(1) with different M.



M=8 N=50 K=2
--- SNR_1=-10 dB - - - SNR_2=0 dB — SNR_3=10 dB

Figure 6.11: $\Delta\phi$-threshold of MZLB(1) with different SNR.

Chapter 7

Simulations and Discussions

7.1 INTRODUCTION

In Section 7.2, MZLB(1) and CRLB(1) are compared with the performance of the MLE of DOA parameters with simulation results. It is shown that MZLB(1) is a tight lower bound over a wide range of SNR. In Section 7.3, we discuss the MZLB for correlated signals, and give simulation results. In Section 7.4, we discuss the achievablity of MZLB by the variance of MLE.

7.2 SIMULATIONS BASED ON DATA MODEL(1)

With the log-likelihood function shown in Eq.(3.9), the realization of the MLE is given by the minimizer of the following function:

$$f(\vec{\phi}) = \sum_{n=1}^{N} [\mathbf{r}(n) - \mathbf{D}(\vec{\phi})\mathbf{a}(n)]^{H} [\mathbf{r}(n) - \mathbf{D}(\vec{\phi})\mathbf{a}(n)]. \tag{7.1}$$

Based on Data Model(1), the MLE of a(n) is obtained from Eq. (7.1) by

$$\hat{\mathbf{a}}(n) = (\mathbf{D}^{H}\mathbf{D})^{-1}\mathbf{D}^{H}\mathbf{r}(n). \tag{7.2}$$

By substituting Eq.(7.2) in Eq.(7.1), the MLE of DOA becomes the following optimization problem,

$$\hat{\phi} = \arg\min_{\vec{\phi}} \mathrm{Tr}\{\mathbf{P}_e(\vec{\phi})\hat{\mathbf{R}}_r\}, \tag{7.3}$$

where,

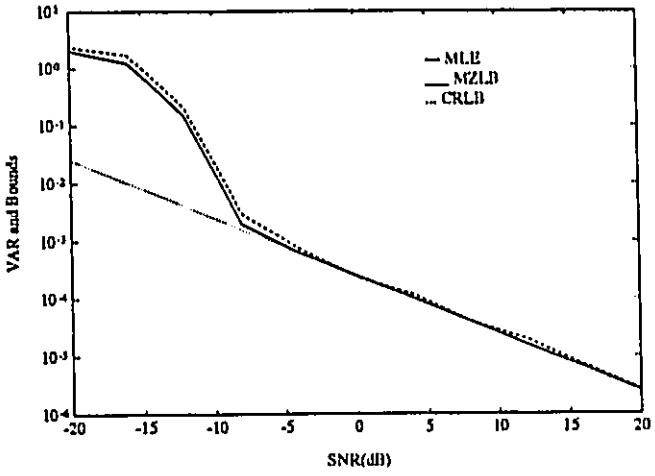
$$\hat{\mathbf{R}}_r = \frac{1}{N}\sum_{n=1}^{N} \mathbf{r}(n)\mathbf{r}^H(n)$$

is the correlation matrix of the received data, $\mathbf{P}_e(\vec{\phi})$ is a projector defined in Eq.(3.29).

With various number of snapshots, various number of sensors and various number of signals, the MLE of DOA parameters (i.e., Eq.(7.3)) is implemented on the computer. The variance of MLE is compared with MZLB(1) and CRLB(1) in Fig. 7.1 to Fig. 7.7, where, MZLB(1) is given by Eq.(6.39) and is shown by the solid lines, CRLB(1) is given by Eq.(3.27) and is shown by the dotted lines, and the variance of MLE is given by the dash lines.

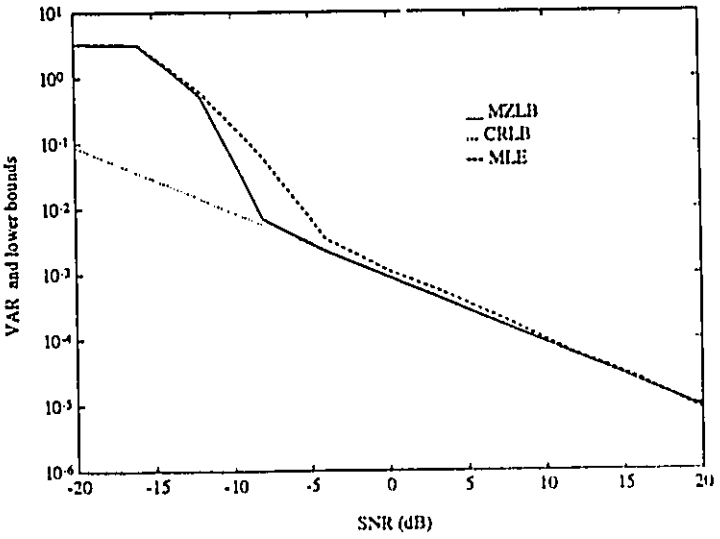
All simulations shown in these figures are for $\hat{\phi}_1$ which is the estimate of the first DOA ϕ_{01}. Similar observations persist when $\hat{\phi}_k$, which is the estimate of the k^{th} DOA $\phi_{0k}(k = 1, ..., K)$, is examined.

The simulation results show that, CRLB is tight only when SNR is high; MZLB is a tight lower bound over a wide range of SNR, it follows the SNR-threshold phenomenon.



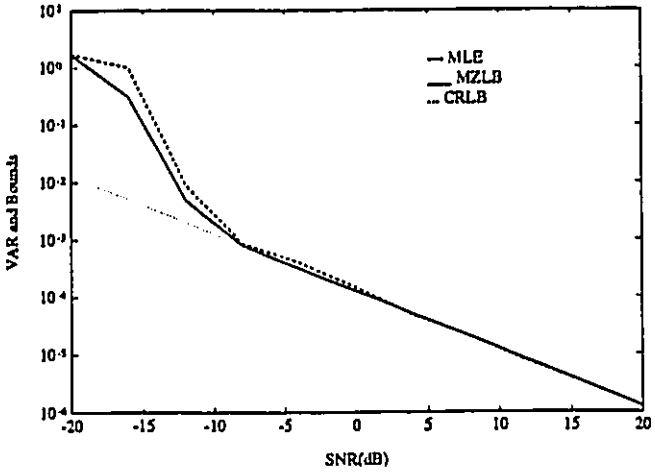
K=2[-.5 .5], M=8, N=50

Figure 7.1: MZLB, CRLB and MLE (K=2, M=8, N=50).



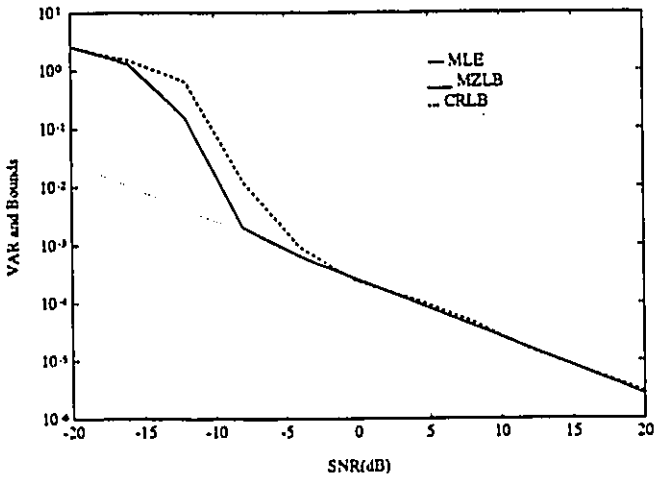
K=2[-.25 .25], M=8, N=50

Figure 7.2: MZLB, CRLB and MLE (closer seperation)(K=2, M=8, N=50).



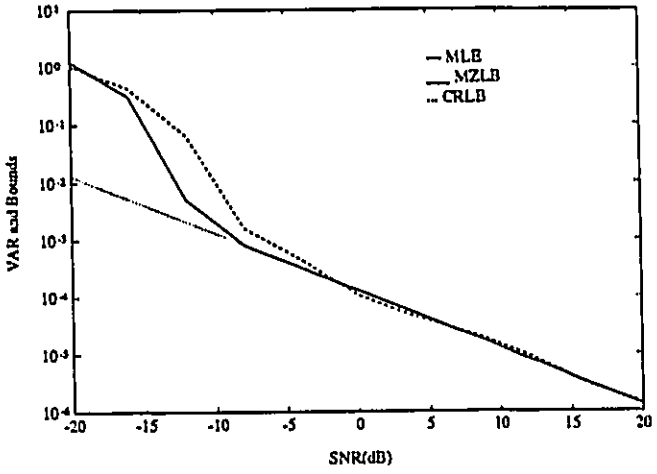
K=2[-.5 .5], M=8, N=100

Figure 7.3: MZLB, CRLB and MLE (K=2, M=8, N=100).



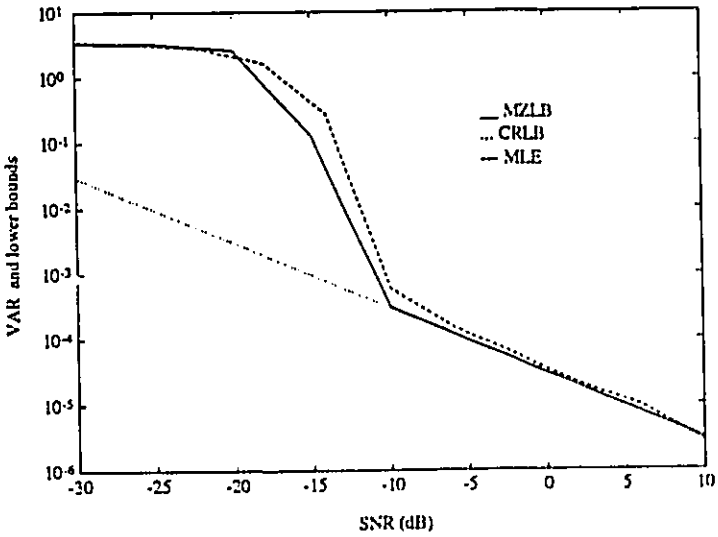
K=3[-.5 .5 1.57], M=8, N=50

Figure 7.4: MZLB, CRLB and MLE (K=3, M=8, N=50).



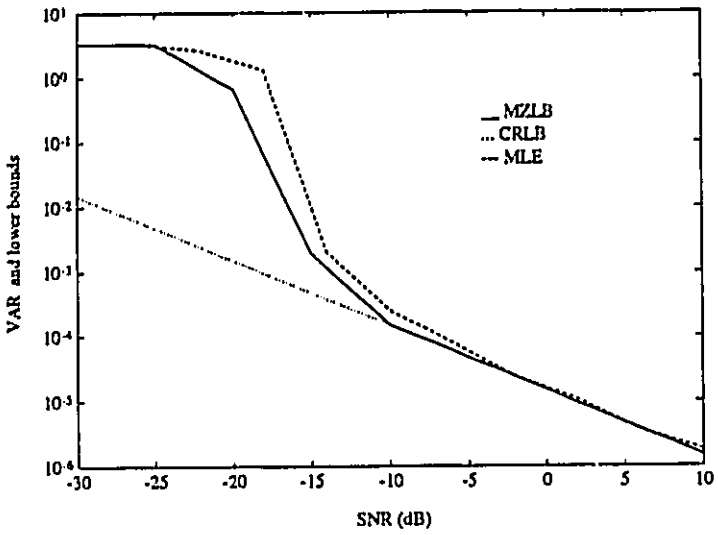
K=3[-.5 .5 1.57], M=8, N=100

Figure 7.5: MZLB, CRLB and MLE (K=3, M=8, N=100).



K=3[-.25 .25 1.57], M=16, N=50

Figure 7.6: MZLB, CRLB and MLE (K=3, M=16, N=50).



K=3[-.25 .25 1.57], M=16, N=100

Figure 7.7: MZLB, CRLB and MLE (K=3, M=16, N=100).

7.3 MZLB FOR THE CORRELATED SIGNALS CASE

Let B_Z^c and B_{CR}^c denote the MZLB and CRLB for the correlated signals case respectively. For this case, Eq.(3.26) becomes,

$$\begin{aligned} B_{CR}^c &= [\frac{2}{\sigma_e^2}\Re\{\mathbf{X}^H(n)\dot{\mathbf{D}}^H\mathbf{P}_e\dot{\mathbf{D}}\mathbf{X}(n)\}]^{-1} \\ &= [\frac{2N}{\sigma_e^2}\Re\{\dot{\mathbf{D}}^H\mathbf{P}_e\dot{\mathbf{D}}\odot\mathbf{R}_a^H\}]^{-1} \\ &= [\frac{2N}{\sigma_e^2}\Re\{\dot{\mathbf{D}}^H\mathbf{P}_e\dot{\mathbf{D}}\odot\mathbf{R}_{a0}^H\sigma_s^2\}]^{-1}, \end{aligned} \tag{7.4}$$

where,

$$\mathbf{P}_e = \mathbf{I} - \mathbf{D}(\mathbf{D}^H\mathbf{D})^{-1}\mathbf{D}^H, \tag{7.5}$$

and $\mathbf{R}_a = \sigma_s^2\mathbf{R}_{a0}$ is the signal correlation matrix with the assumption that all signals have equal power.

The off-diagonal elements of B^c_{CR} are not zeros because R_a is not a diagonal matrix for correlated signals. B^c_{CR} is of the form F-I as shown in Fig.4.2. Since CRLB is achievable for small $|\vec{\phi}_c|$ as discussed before, with the analogous idea shown in Chapter 6, we think the covariance matrix for the correlated signals can be obtained from the covariance matrix based on Data Model(2) with a transformation T_c. Then we propose that the MZLB for correlated signals may be obtained from MZLB(2) with this transformation T_c. Therefore,

$$B^c_Z = T_c B_Z(2), \tag{7.6}$$

in which

$$T_c = B^c_{CR} B^{-1}_{CR}(2). \tag{7.7}$$

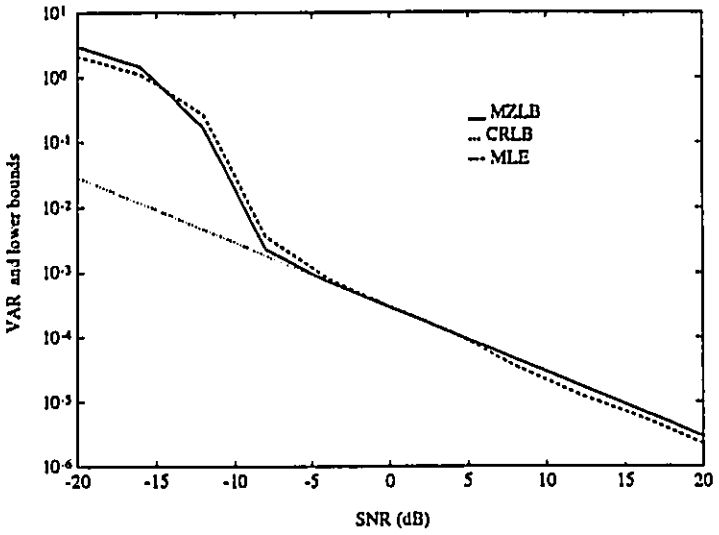
From Eq.(3.41),

$$B^{-1}_{CR}(2) = \frac{N\rho M(M-1)(2M-1)}{3} I. \tag{7.8}$$

Substituting Eqs.(7.4) and (7.8) into Eq.(7.7) yields

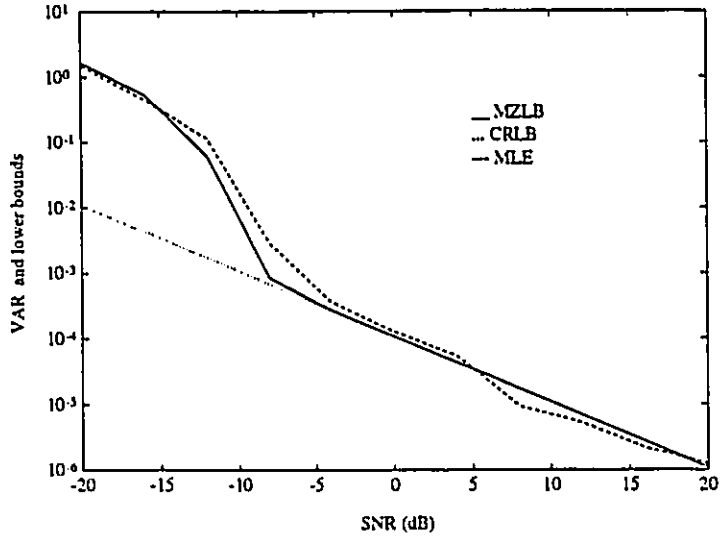
$$T_c = \frac{M(M-1)(2M-1)}{6} [\Re\{\dot{D}^H P_c \dot{D} \odot R^H_{a0}\}]^{-1}. \tag{7.9}$$

With the simulations, both the variances and the covariances (absolute values) from MLE are compared with B^c_Z and B^c_{CR}. The examples shown in Fig.7.8 to Fig.7.11 are for K=2 case, in which $\rho_{12} = \rho_{21}$ are the correlation coefficient. From these simulation results, it seems Eq.(7.6) works even though the idea presented here is only a conjecture.



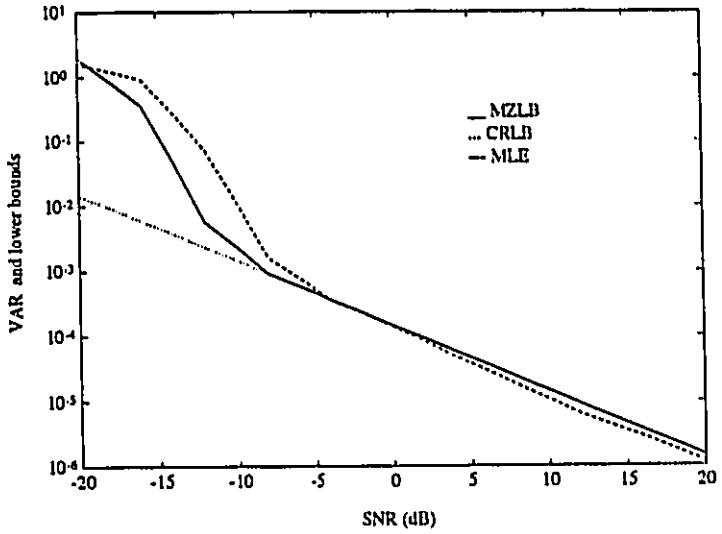
K=2[-.5 .5] $\rho_{12} = .5$ M=8 N=50

Figure 7.8: MZLB, CRLB and MLE (variance).



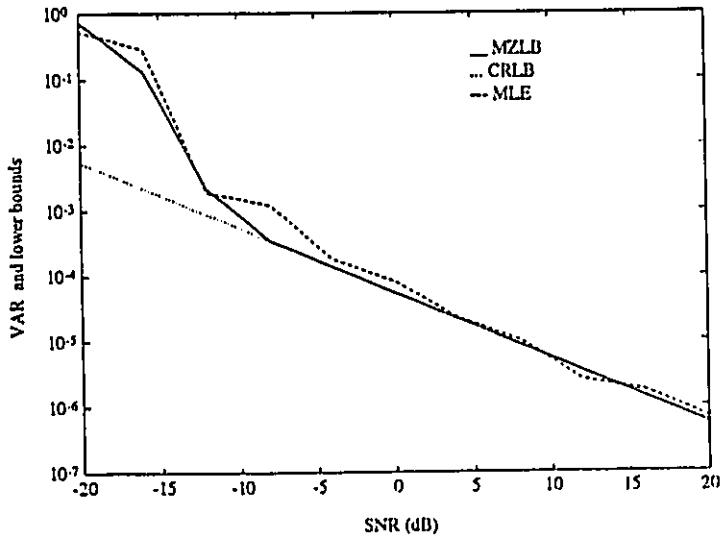
K=2[-.5 .5] $\rho_{12} = .5$ M=8 N=50

Figure 7.9: MZLB, CRLB and MLE (|covariance|).



K=2[-.5 .5] $\rho_{12} = .5$ M=8 N=100

Figure 7.10: MZLB, CRLB and MLE (variance).



K=2[-.5 .5] $\rho_{12} = .5$ M=8 N=100

Figure 7.11: MZLB, CRLB and MLE (|covariance|).

7.4 THE ACHIEVABLITY OF MZLB

The simulations in the above show that MZLB follows the SNR-threshold of MLE very well. But whether MZLB is achievable is still an open question [71]. We discuss this under Assumption 6.1 and Assumption 6.2.

During the derivation of the MZLB, the only inequality we have used is:

$$P_e \geq P_{EM}.$$

Recalling that, in Chapter 5, the variance of DOA estimate $E(|\varepsilon|^2)$ is derived from P_e, and its lower bound B_Z is derived from P_{EM}, therefore, the following relationship can be established:

$$\boxed{P_e \simeq P_{EM}}$$

$$\Uparrow$$

$$\Downarrow$$

$$\boxed{E(|\varepsilon|^2) \simeq B_Z}$$

i.e., if one of the approximates in the boxes holds, the other one holds.

For showing that B_Z is achievable, we need to prove that the second box holds. Since it has been proved that $B_Z = B_{CR}$ is achievable when SNR is high, i.e., the second box holds under high SNR, we are going to prove that it holds under low SNR with the description in Fig.7.12.

In Fig.7.12 (a) and (b), $p(\hat{\phi}/\phi_i)$ is the probability density function of the estimated parameter. It is assumed to be Gaussian if $\hat{\phi}$ is obtained from the maximum likelihood

estimator. The error probability P_e is represented by the shaded area (see Chapter 5.3). In Fig. 7.12 (c) and (d), $p(l/H_i)$ is the probability density function of the sufficient statistic l in binary detection. It is Gaussian distributed as shown in Appendix D. P_{EM} is the shaded area (see Appendix B).

MZLB is achievable by the variance of the maximum likelihood estimation under high SNR, since it includs CRLB and CRLB is achievable under high SNR. This results in $P_e \simeq P_{EM}$ when SNR is high, i.e. the shaded area in (a) approximates that in (c). Since both $p(\hat{\phi})$ and $p(l)$ are Gaussian, the approximation of the shaded areas in (b) and (d) will not be changed when SNR is low. Therefore, $P_e \simeq P_{EM}$ under low SNR.

The conclusion from this analysis is that, with our considerations, MZLB is achievable by the maximum likelihood estimator under both high SNR as well as low SNR.

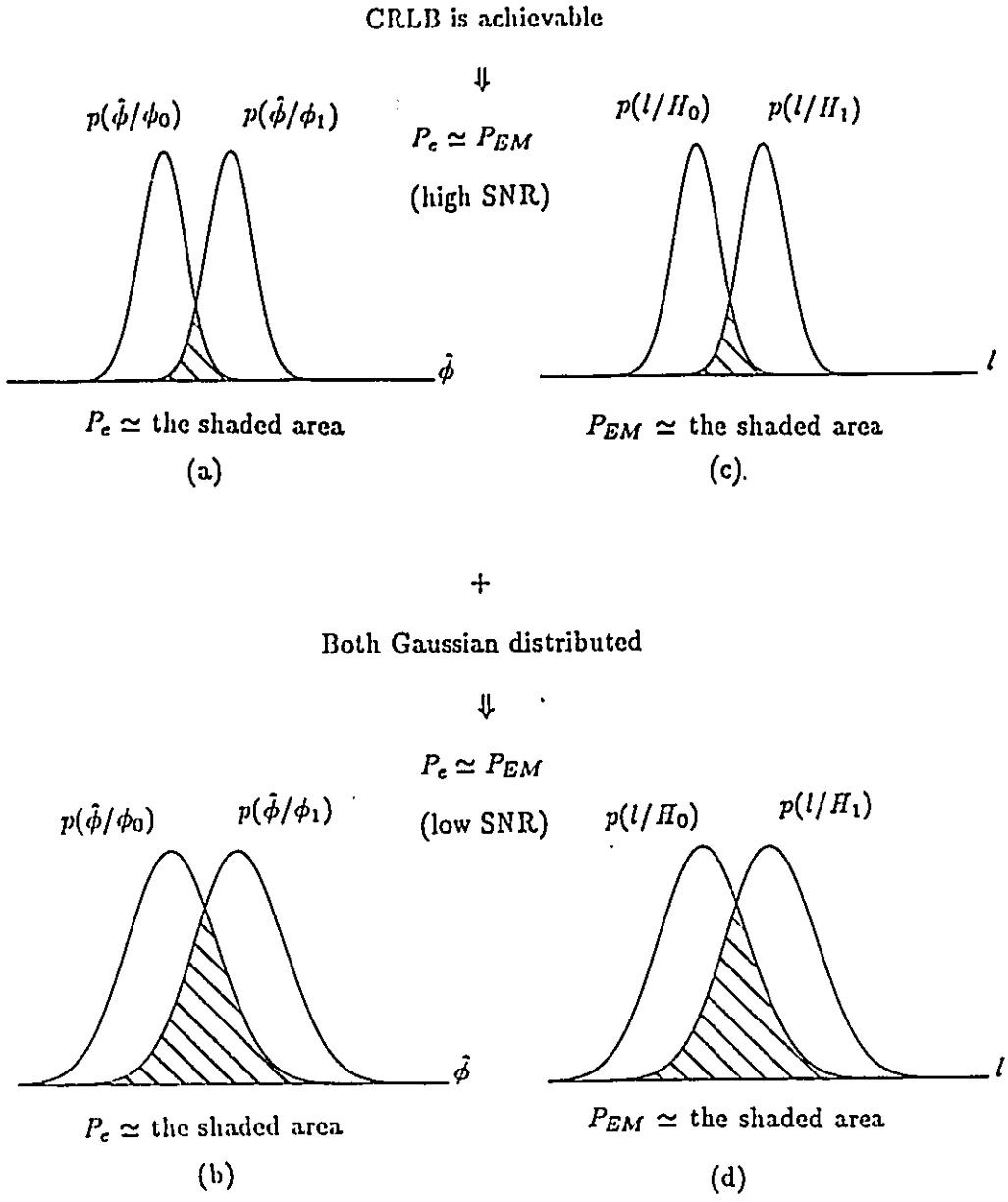


Figure 7.12: The achievablity of MZLB.

Chapter 8

Summary

Performance analysis for DOA estimation over a wide range of signal-noise ratio is a difficult problem. Some work has been conducted toward this direction. The work developed in this thesis addresses part of the problem.

Using a new idea, Ziv and Zakai developed a lower bound on MSE of the time delay estimate, which works for one incoming signal with one unknown parameter. The major contributions in this thesis are, by modifying the idea from the Ziv-Zakai lower bound on the error of one random parameter estimation, a lower bound on the variance of one deterministic DOA estimate has been derived; this has been extended to obtain the lower bound of the covariance matrix of DOA estimates which contains multiple incoming signals having multiple types of unknown parameters. The resulting MZLB is a computable tight lower bound.

The work developed in this thesis is summarized as the following:

1. Establishing Data Model(2) to study the performance of DOA estimate against SNR independently. Establishing the relationship between the performance of DOA with Data Model(2) and the performance of DOA with Data Model(1).

2. Proving that the MLE of DOA parameters based on Data Model(2) is asymptotically efficient.

3. Giving a profound interpretation of Ziv-Zakai's idea, and applying the logic in Ziv-Zakai's idea to derive a lower bound on the variance of one DOA estimate.

4. Developing MZLB(2) on the covariance matrix of multiple DOA estimation. The modified Ziv-Zakai lower bound is obtained from MZLB(2) with a transformation T.

5. Studying the SNR-threshold phenomenon as well as the $\Delta\phi$-threshold phenomenon in CRLB and MZLB.

6. Showing the tightness of MZLB by simulations.

The work developed in this thesis addresses the performance analysis of DOA estimation. The idea used here may be extended to the general performance analysis for nonlinear estimation.

The idea presented by Ziv and Zakai in [100] is excellent. Since there is not much work that applies Ziv-Zakai's idea to derive lower bounds on estimation errors, the research work in this area should be interested.

Appendix A

Proof of Eq.(4.20)

$\mathbf{e}(n)$ is the white Gaussian noise at the n^{th} snapshot. It is a M-dimension vector. $e_m(n)$ is the m^{th} element of $\mathbf{e}(n)$. The following relationships hold [77]:

1.

$$E\{\mathbf{e}(n)\mathbf{e}^T(n1)\} = \mathbf{0} \tag{A.1}$$

for $n = n1$ and $n \neq n1$.

2.

$$E\{e_m(n)e^*_{m1}(n1)\} = \begin{cases} 0, & n \neq n1, \text{ or } m1 \neq m \\ \sigma_e^2 & n = n1 \text{ and } m1 = m \end{cases} \tag{A.2}$$

since the noise are white spatially and temporally.

Eq.(4.10) results in,

$$E\{\frac{\partial f}{\partial \phi_k}\frac{\partial f^*}{\partial \phi_l}\}$$

$$= E\{[\sum_{n=1}^{N} -\mathbf{a}^H(n)\frac{\partial \mathbf{D}^H}{\partial \phi_k}\mathbf{e}(n) - \mathbf{e}^H(n)\frac{\partial \mathbf{D}}{\partial \phi_k}\mathbf{a}(n)]$$

$$[\sum_{n1=1}^{N} -\mathbf{a}^H(n1)\frac{\partial \mathbf{D}^H}{\partial \phi_l}\mathbf{e}(n1) - \mathbf{e}^H(n1)\frac{\partial \mathbf{D}}{\partial \phi_l}\mathbf{a}(n1)]^*\}$$

$$= \sum_{n=1}^{N}\sum_{n1=1}^{N}[E\{\mathbf{a}^H(n)\frac{\partial \mathbf{D}^H}{\partial \phi_k}\mathbf{e}(n)\mathbf{a}^T(n1)\frac{\partial \mathbf{D}^T}{\partial \phi_l}\mathbf{e}^*(n1)\}$$

$$+\mathbf{a}^H(n)\frac{\partial \mathbf{D}^H}{\partial \phi_k}E\{\mathbf{e}(n)\mathbf{e}^T(n1)\}\frac{\partial \mathbf{D}^*}{\partial \phi_l}\mathbf{a}^*(n1)$$

$$+E\{\mathbf{e}^H(n)\frac{\partial \mathbf{D}}{\partial \phi_k}\mathbf{a}(n)\mathbf{a}^T(n1)\frac{\partial \mathbf{D}^T}{\partial \phi_l}\mathbf{e}^*(n1)\}$$

$$+E\{\mathbf{e}^H(n)\frac{\partial \mathbf{D}}{\partial \phi_k}\mathbf{a}(n)\mathbf{e}^T(n1)\frac{\partial \mathbf{D}^*}{\partial \phi_l}\mathbf{a}^*(n1)\}], \tag{A.3}$$

where, the second term is equal to zero for both $l = k$ and $l \neq k$ because of Eq.(A.1). Similarly, the third term is equal to zero for both $l = k$ and $l \neq k$ since

$$E\{\mathbf{e}^H(n)\frac{\partial \mathbf{D}}{\partial \phi_k}\mathbf{a}(n)\mathbf{a}^T(n1)\frac{\partial \mathbf{D}^T}{\partial \phi_l}\mathbf{e}^*(n1)\} = \mathbf{a}^T(n1)\frac{\partial \mathbf{D}^T}{\partial \phi_l}E\{\mathbf{e}^*(n1)\mathbf{e}^H(n)\}\frac{\partial \mathbf{D}}{\partial \phi_k}\mathbf{a}(n).$$

Now we consider the first term in Eq.(A.3). For $l \neq k$, recalling Eq.(4.12), we have,

$$\sum_{n=1}^{N}\sum_{n1=1}^{N}E\{\mathbf{a}^H(n)\frac{\partial \mathbf{D}^H}{\partial \phi_k}\mathbf{e}(n)\mathbf{a}^T(n1)\frac{\partial \mathbf{D}^T}{\partial \phi_l}\mathbf{e}^*(n1)\}$$

$$= \sum_{n=1}^{N}\sum_{n1=1}^{N}a_k^*(n)a_l(n1)E\{[\sum_{m=1}^{M}(m-1)j\exp(j(m-1)\phi_k)e_m^*(n)]^*$$

$$[\sum_{m1=1}^{M}(m1-1)j\exp(j(m1-1)\phi_l)e_{m1}^*(n1)]\}$$

$$= 0, \tag{A.4}$$

where, Eq.(4.15) has been applied. Similarly, the fourth term in Eq.(A.3) becomes zero when $l \neq k$.

For $l = k$, using Eq.(A.2), the first term in Eq.(A.3) becomes

$$\sum_{n=1}^{N}a_k^*(n)a_k(n)E\{\sum_{m=1}^{M}[(m-1)j\exp(j(m-1)\phi_k)e_m^*(n)]^*[(m-1)j\exp(j(m-1)\phi_k)e_m^*(n)]\}$$

$$= \sum_{n=1}^{N}a_k^*(n)a_k(n)[\sum_{m=1}^{M}(m-1)^2E\{e_m^*(n)e_m(n)\}]$$

$$= \frac{N\sigma_{gk}^2\sigma_e^2 M(M-1)(2M-1)}{6}. \qquad (A.5)$$

Similarly, the fourth term in Eq.(A.3) equals to the RHS of Eq.(A.5).

Thus Eq.(4.20) follows.

Appendix B

Derivation of P_{EM} in Eq.(5.20)

Let us consider a binary detection problem with sensor array. The DOA of the incoming signal is of two possible values, ϕ_0 and ϕ_1. We form the following hypotheses:

H_0: $DOA = \phi_0$,

H_1: $DOA = \phi_1$.

The task of detection is using some procedure to decide which hypothesis is true.

Bayes criterion leads a likelihood-ratio test (LRT):

$$\ln \frac{p(\mathbf{r}|H_1)}{p(\mathbf{r}|H_0)} \triangleq l \underset{H_1}{\overset{H_0}{\lessgtr}} \ln \frac{P(H_0)}{P(H_1)} \triangleq \gamma, \tag{B.1}$$

where, $p(\mathbf{r}|H_i)$ is the PDF of the received data when H_i is true, l is a sufficient statistic, $P(H_i)$ is the probability of H_i occuring, and γ is the test threshold. We assume that the costs of making erroneous decisions are equal.

When the two hypotheses are equally likely, i.e. $P(H_0) = P(H_1) = 0.5$, $\gamma = 0$ and Eq.(B.1) results in a minimum error probability receiver.

The error probability of LRT is shown in Fig.B.1, where, $p(l/H_0)$ is the probability

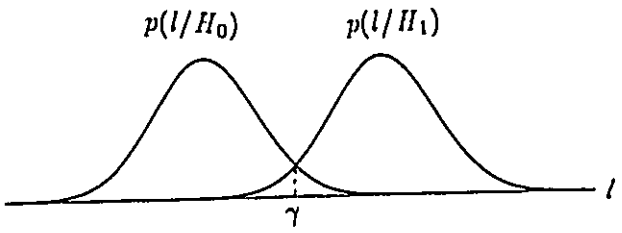


Figure B.1: **Error probabilities with LRT.**

density function of l under hypothesis H_0, $p(l/H_1)$ is the PDF of l under hypothesis H_1.

The probability of false alarm P_F (i.e. we say H_1 is true when H_0 is true) is

$$P_F = \int_{\gamma}^{\infty} p(l/H_0)dl.$$

The probability of a miss P_M (i.e. we say H_0 is true when H_1 is true) is

$$P_M = \int_{-\infty}^{\gamma} p(l/H_1)dl.$$

The total error probability is

$$P_e = P(H_0)P_F + P(H_1)P_M = (P_F + P_M)/2. \tag{B.2}$$

Now, we follow the way introduced in [81] to obtain an approximation of this minimum error probability P_e.

Under H_0, the moment-generating function of l is

$$\psi(q) \triangleq E(e^{ql}|H_0) \tag{B.3}$$

Define a function,

$$\begin{aligned}
\mu(q) &\triangleq \ln \psi(q) \\
&= \ln E(e^{ql(\mathbf{r})}|H_0) \\
&= \ln \int_{-\infty}^{\infty} e^{ql} p(l/H_0)dl \\
&= \ln \int_{-\infty}^{\infty} e^{ql(\mathbf{r})} p(\mathbf{r}/H_0)d\mathbf{r} \\
&= \ln \int_{-\infty}^{\infty} [\frac{p(\mathbf{r}/H_1)}{p(\mathbf{r}/H_0)}]^q p(\mathbf{r}/H_0)d\mathbf{r} \\
&= \ln \int_{-\infty}^{\infty} [p(\mathbf{r}/H_1)]^q [p(\mathbf{r})/H_0]^{1-q} d\mathbf{r}. \qquad \text{(B.4)}
\end{aligned}$$

Now, constructing a new probability density function

$$p_q(x) = \frac{e^{qx} p(x/H_0)}{\int_{-\infty}^{\infty} e^{ql} p(l/H_0)dl} = \frac{e^{qx} p(x/H_0)}{\exp(\mu(q))}, \qquad \text{(B.5)}$$

we have,

$$E(x) = \dot{\mu}(q) \qquad \text{(B.6)}$$

since

$$\begin{aligned}
E(x) &= \int_{-\infty}^{\infty} x p_q(x)dx \\
&= \frac{\int_{-\infty}^{\infty} x e^{qx} p(x/H_0)dx}{\exp(\mu(q))},
\end{aligned}$$

and

$$\begin{aligned}
\dot{\mu}(q) &= \frac{d}{dq}[\ln \int_{-\infty}^{\infty} e^{ql} p(l/H_0)dl] \\
&= \frac{\int_{-\infty}^{\infty} \frac{d}{dq} e^{ql} p(l/H_0)dl}{\int_{-\infty}^{\infty} e^{ql} p(l/H_0)dl}
\end{aligned}$$

$$= \frac{\int_{-\infty}^{\infty} l e^{ql} p(l/H_0)dl}{\exp(\mu(q))};$$

$$Var(x) = \ddot{\mu}(q) \tag{B.7}$$

since

$$\begin{aligned} Var(x) &= E(x^2) - [E(x)]^2 \\ &= \int_{-\infty}^{\infty} x^2 p_q(x)dx - [\frac{\int_{-\infty}^{\infty} x e^{qx} p(x/H_0)dx}{\exp(\mu(q))}]^2 \\ &= \frac{\int_{-\infty}^{\infty} x^2 e^{qx} p(x|H_0)dx}{\exp(\mu(q))} - [\frac{\int_{-\infty}^{\infty} x e^{qx} p(x/H_0)dx}{\exp(\mu(q))}]^2 \end{aligned}$$

and

$$\begin{aligned} \ddot{\mu}(q) &= \frac{d}{dq}\frac{\int_{-\infty}^{\infty} l e^{ql} p(l/H_0)dl}{\exp(\mu(q))} \\ &= \frac{[\int_{-\infty}^{\infty} l^2 e^{ql} p(l/H_0)dl]\exp(\mu(q)) - [\int_{-\infty}^{\infty} l e^{ql} p(l/H_0)dl][\int_{-\infty}^{\infty} l e^{ql} p(l/H_0)dl]}{[\exp(\mu(q))]^2} \\ &= \frac{[\int_{-\infty}^{\infty} l^2 e^{ql} p(l/H_0)dl]}{\exp(\mu(q))} - [\frac{\int_{-\infty}^{\infty} l e^{ql} p(l/H_0)dl}{\exp(\mu(q))}]^2. \end{aligned}$$

We choose $q = q_M$ such that

$$\dot{\mu}(q_M) = \gamma = 0 \tag{B.8}$$

for the minimum error probability receiver.

Now, using Eqs.(B.5) and (B.8), P_F can be written as

$$\begin{aligned} P_F &= \int_{\gamma}^{\infty} p(l/H_0)dl \\ &= \int_{\dot{\mu}}^{\infty} \exp(\mu(q) - qx) p_q(x)dx|_{q=q_M} \end{aligned}$$

$$= \exp(\mu(q) - q\dot{\mu}(q)) \int_{\dot{\mu}}^{\infty} \exp(q\dot{\mu}(q) - qx) p_q(x) dx|_{q=q_M}. \quad (B.9)$$

Note that, in our problem, l is Gaussion distributed (Appendix D), so, $p_q(x)$ in Eq.(B.5) is Gaussian. Let

$$y \triangleq \frac{x - E(x)}{[Var(x)]^{1/2}} = \frac{x - \dot{\mu}(q_M)}{\sqrt{\ddot{\mu}(q_M)}},$$

then $y \sim N(0,1)$, and

$$\begin{aligned} P_F &= \frac{1}{\sqrt{2\pi}} \exp(\mu(q_M) - q_M\dot{\mu}(q_M)) \int_0^{\infty} \exp(-q_M\sqrt{\ddot{\mu}(q_M)}y - y^2/2)dy \\ &= \exp\{\mu(q_M) + \frac{q_M^2}{2}\ddot{\mu}(q_M)\}\text{erfc}_*\{q_M\sqrt{\ddot{\mu}(q_M)}\}. \end{aligned} \quad (B.10)$$

In Similar way,

$$P_M = \exp\{\mu(q_M) + \frac{(1-q_M)^2}{2}\ddot{\mu}(q_M)\}\text{erfc}_*\{(1-q_M)\sqrt{\ddot{\mu}(q_M)}\}. \quad (B.11)$$

Therefore,

$$\begin{aligned} P_e &= (P_F + P_M)/2 \\ &= exp[\mu(q_M) + q_M^2\ddot{\mu}(q_M)/2]\text{erfc}_*\{q_M\sqrt{\ddot{\mu}(q_M)}\}/2 \\ &\quad + exp[\mu(q_M) + (1-q_M)^2\ddot{\mu}(q_M)/2]\text{erfc}_*\{(1-q_M)\sqrt{\ddot{\mu}(q_M)}\}/2 \\ &\triangleq P_{EM}. \end{aligned} \quad (B.12)$$

The approximation in the above exists under the following conditions:

(C1). The hypotheses H_0 and H_1 are equally likely,

(C2). The costs of making erroneous decisions are equal,

Remark

The following points are important to understand the nature of the lower bound.

• Under the condition (C1) and (C2), P_{EM} is the minimum error probability in binary detection, no matter the signal-noise ratio is high or low.

• In nature, the binary detection procedure is detecting one signal identified by one parameter which is of two possibile values.

Appendix C

Proof of Eq.(5.24)

Eq.(5.22) is valid for the real data as shown in Appendix B. However, the received data in our problem is complex. To overcome this, we re-formulate the data vector r(n) at the n^{th} snapshot into a 2M-dimensional vector such that

$$\vec{r}(n) = [Re(r_1(n)), ..., Re(r_M(n)), Im(r_1(n)), ..., Im(r_M(n))]^T, \tag{C.1}$$

and

$$\vec{s}_i(n) = [Re(s_{i1}(n)), ..., Re(s_{iM}(n)), Im(s_{i1}(n)), ..., Im(s_{iM}(n))]^T \tag{C.2}$$

$$= a(n)[1, \cos(\phi_i), ..., \cos(\phi_i(M-1)), 0, \sin(\phi_i), ..., \sin(\phi_i(M-1))]^T, \tag{C.3}$$

where, $r_m(n)$ is the complex data received by the m^{th} sensor at the n^{th} snapshot, $s_{im}(n)$ and ϕ_i are the signal and the angle of arrival when H_i is true. The quantity $a(n)$ is, in general, complex, having a random amplitude and a random phase at each snapshot. However, since we assume the signal is narrowband such that it does not vary while traversing from sensor 1 to sensor M during the same snapshot, we can apply an appropriate phase shift at each snapshot so that $a(n) = |a(n)|$ in Eq.(C.3) rendering the 2M-dimensional vector $\vec{s}(n)$ real.

The PDF of the received data can now be re-written as [22]

$$p(\vec{r}/H_i) = \frac{1}{\det(2\pi\vec{\mathrm{R}}_e)^{N/2}} \exp[-\frac{1}{2}\sum_{n=1}^{N}(\vec{r}(n) - \vec{s}_i(n))^T \vec{\mathrm{R}}_e^{-1}(\vec{r}(n) - \vec{s}_i(n))], \tag{C.4}$$

where,

$$\vec{\mathrm{R}}_e = \frac{1}{2}\begin{bmatrix} Re(\mathrm{R}_e) & -Im(\mathrm{R}_e) \\ Im(\mathrm{R}_e) & Re(\mathrm{R}_e) \end{bmatrix} = \frac{\sigma_e^2}{2}\mathrm{I}. \tag{C.5}$$

Then Eq.(5.22) can be written as

$$\begin{aligned} \mu(q) &= \ln \int_{-\infty}^{\infty} \frac{1}{(2\pi\sigma_e^2)^{MN/2}} \exp[-\frac{1}{\sigma_e^2}(\sum_{n=1}^{N} q(\vec{r} - \vec{s}_1)^T(\vec{r} - \vec{s}_1) \\ &\quad +(1-q)(\vec{r} - \vec{s}_0)^T(\vec{r} - \vec{s}_0))]d\vec{r}, \end{aligned} \tag{C.6}$$

where, for convenience, we have omitted the obvious dependences of $\vec{r}$ and $\vec{s}_i$ on n. Then

$$\begin{aligned} \mu(q) &= \ln \frac{1}{\det(2\pi\vec{\mathrm{R}}_e)^N} \int_{-\infty}^{\infty} \exp[-\frac{1}{\sigma_e^2}\sum_{n=1}^{N}(q\vec{r}^T\vec{r} - q\vec{r}^T\vec{s}_1 - q\vec{s}_1^T\vec{r} + q\vec{s}_1^T\vec{s}_1 \\ &\quad +\vec{r}^T\vec{r} - \vec{r}^T\vec{s}_0 - \vec{s}_0^T\vec{r} + \vec{s}_0^T\vec{s}_0 - q\vec{r}^T\vec{r} + q\vec{r}^T\vec{s}_0 + q\vec{s}_0^T\vec{r} - q\vec{s}_0^T\vec{s}_0)]d\vec{r} \\ &= \ln \frac{1}{[\det(2\pi\vec{\mathrm{R}}_e)]^{N/2}} \int_{-\infty}^{\infty} exp[-\frac{1}{\sigma_e^2}\sum_{n=1}^{N}(\vec{r}^T\vec{r} - q\vec{r}^T\vec{s}_1 - q\vec{s}_1^T\vec{r} \\ &\quad -(1-q)\vec{r}^T\vec{s}_0 - (1-q)\vec{s}_0^T\vec{r} + \vec{s}_0^T\vec{s}_0)]d\vec{r} \\ &= \ln \exp[-\frac{1}{\sigma_e^2}\sum_{n=1}^{N}(-q^2\vec{s}_1^T\vec{s}_1 - q(1-q)\vec{s}_1^T\vec{s}_0 - q(1-q)\vec{s}_0^T\vec{s}_1 \\ &\quad -(1-q)^2\vec{s}_0^T\vec{s}_0 + \vec{s}_0^T\vec{s}_0)]\frac{1}{[\det(2\pi\vec{\mathrm{R}}_e)]^{N/2}} \\ &\quad \int_{-\infty}^{\infty} \exp[-\frac{1}{2}\sum_{n=1}^{N}(\vec{r} - (q\vec{s}_1 + (1-q)\vec{s}_0))^T\vec{\mathrm{R}}_e^{-1}(\vec{r} - q(\vec{s}_1 + (1-q)\vec{s}_0))]d\vec{r} \\ &= \frac{1}{\sigma_e^2}\sum_{n=1}^{N}(q^2\vec{s}_1^T\vec{s}_1 + q(1-q)\vec{s}_1^T\vec{s}_0 + q(1-q)\vec{s}_0^T\vec{s}_1 + (1-q)^2\vec{s}_0^T\vec{s}_0 - \vec{s}_0^T\vec{s}_0) \end{aligned}$$

$$= \frac{1}{\sigma_e^2}\sum_{n=1}^{N}[2\sum_{m=1}^{M} a^2 q(-1+q) + q(1-q)(\vec{s}_0^T\vec{s}_1 + \vec{s}_1^T\vec{s}_0)].$$

Note that, $\phi_1 - \phi_0 = h$, with Eq.(3.34) in Data Model(2) and Eq.(C.3), we have,

$$\sum_{n=1}^{N}\vec{s}_1^T\vec{s}_0 = \sum_{n=1}^{N}\vec{s}_0^T\vec{s}_1$$

$$= \sum_{n=1}^{N} a^2(n)\sum_{m=1}^{M}\cos((m-1)\phi_0)\cos((m-1)\phi_1) + \sin((m-1)\phi_0)\sin((m-1)\phi_1)$$

$$= N\sigma_s^2\sum_{m=1}^{M} cos(h(m-1)).$$

So,

$$\mu(q) = 2\sigma_e^{-2}\sigma_s^2 N\sum_{m=1}^{M}[-q(1-q) + q(1-q)cos(h(m-1))]$$

$$= -2\rho N q(1-q)\sum_{m=1}^{M}[1 - cos(h(m-1))]. \qquad (C.7)$$

Eq.(B.8) results in

$$q_M = 1/2, \qquad (C.8)$$

for which Eq.(C.7) becomes

$$\mu(1/2) = -\frac{1}{2}\rho N\sum_{m=1}^{M}[1 - cos(h(m-1))], \qquad (C.9)$$

and

$$\ddot{\mu}(1/2) = 4\rho N\sum_{m=1}^{M}[1 - cos(h(m-1))]. \qquad (C.10)$$

Substituting Eqs.(C.9) and (C.10) into Eq.(5.20). we obtain

$$P_{EM}(h) \simeq \exp[\mu(1/2) + \ddot{\mu}(1/2)/8]\mathrm{erfc}_*\{\sqrt{\ddot{\mu}(1/2)/2}\}/2$$

$$+ \exp[\mu(1/2) + \ddot{\mu}(1/2)/8]\mathrm{erfc}_*\{\sqrt{\ddot{\mu}(1/2)/2}\}/2$$

$$= \operatorname{erfc}_*\{\sqrt{\ddot{\mu}(1/2)}/2\}$$

$$= \operatorname{erfc}_*\{[\rho N \sum_{m=1}^{M}(1-\cos(h(m-1)))]^{1/2}\}. \qquad \text{(C.11)}$$

Appendix D

Distribution of the Statistic l

In our problem,

$$p(\mathbf{r}|H_0) = (\pi\sigma_e^2)^{-MN}\exp[-\sigma_e^{-2}\sum_{n=1}^{N}(\mathbf{r}(n)-\mathbf{s}_0(n))^H(\mathbf{r}(n)-\mathbf{s}_0(n))],$$

$$p(\mathbf{r}|H_1) = (\pi\sigma_e^2)^{-MN}\exp[-\sigma_e^{2}\sum_{n=1}^{N}(\mathbf{r}(n)-\mathbf{s}_1(n))^H(\mathbf{r}(n)-\mathbf{s}_1(n))],$$

so,

$$\begin{aligned} l &= \ln[\frac{p(\mathbf{r}|H_1)}{p(\mathbf{r}|H_0)}] \\ &= \ln\{\exp[\frac{1}{\sigma_e^2}\sum_{n=1}^{N}(\mathbf{r}^H(n)(\mathbf{s}_1(n)-\mathbf{s}_0(n))+(\mathbf{s}_1(n)-\mathbf{s}_0(n))^H\mathbf{r}(n) \\ &\quad -\mathbf{s}_1^H(n)\mathbf{s}_1(n)+\mathbf{s}_0^H(n)\mathbf{s}_0(n)]\} \\ &= \frac{1}{2\sigma_e^2}\sum_{n=1}^{N}Re[\mathbf{r}^H(n)(\mathbf{s}_1(n)-\mathbf{s}_0(n))], \end{aligned}$$

where, we have used Eq.(3.34) in Data Model(2) such that

$$\sum_{n=1}^{N}\mathbf{s}_1^H(n)\mathbf{s}_1(n) = \sum_{n=1}^{N}\mathbf{s}_0^H(n)\mathbf{s}_0(n) \simeq MN\sigma_s^2.$$

So, l is a linear combination of Gaussian random data $\mathbf{r}(n)$ and is therefore of Gaussian distribution.

Appendix E

Proof of Eq.(5.42)

Let

$$C = 1 + \cos(x) + \cos(2x) + ... + \cos((M-1)x),$$

and

$$S = 1 + \sin(x) + \sin(2x) + ... + \sin((M-1)x),$$

then

$$\begin{aligned} C + jS &= 1 + e^{jx} + e^{j2x} + ... + e^{j(M-1)x} \\ &= \frac{1 - e^{jMx}}{1 - e^{jx}} \\ &= \frac{(1 - e^{jMx})(1 - e^{-jx})}{(1 - e^{jx})(1 - e^{-jx})} \\ &= \frac{1 - e^{-jx} - e^{jMx} + e^{j(M-1)x}}{1 - e^{jx} - e^{-jx} + 1} \\ &= \frac{1 - e^{-jx} + e^{j(M-1)x} - e^{jMx}}{2(1 - \cos(x))}. \end{aligned}$$

Equating the real part,

$$C = \sum_{m=1}^{M} \cos((m-1)x)$$

$$= \Re(C + jS)$$

$$= \frac{1 - \cos(x) + \cos((M-1)x) - \cos(Mx)}{2(1 - \cos(x))}. \tag{E.1}$$

With $x = \Delta = \frac{\pi}{2M}$, and the approximation (M is reasonable large)

$$\sin((M-1)\Delta) = \sin\frac{(M-1)\pi}{2M} \simeq \sin\frac{\pi}{2} = 1,$$

$$\sin\frac{(2M-1)\Delta}{2} = \sin\frac{(2M-1)\pi}{4M} \simeq \sin\frac{\pi}{2} = 1,$$

$$\sin\frac{\Delta}{2} = \sin\frac{\pi}{4M} \simeq \frac{\pi}{4M},$$

$$\cos(\Delta/2) = \cos(\pi/4M) \simeq 1,$$

we have

$$\frac{dC|_{x=\Delta}}{d\Delta}$$

$$= \frac{1}{2}\frac{d}{d\Delta}[1 - \frac{\cos(M\Delta) - \cos((M-1)\Delta)}{1 - \cos(\Delta)}]$$

$$= -\frac{1}{2}[\frac{(1 - \cos(\Delta))\{-M\sin(M\Delta) + (M-1)\sin((M-1)\Delta)\}}{(1 - \cos(\Delta))^2}$$

$$-\frac{\{\cos(M\Delta) - \cos((M-1)\Delta)\}\sin(\Delta)}{(1 - \cos(\Delta))^2}]$$

$$\simeq -\frac{1}{2}[\frac{(M-1)\{-M\sin(M\Delta)/(M-1) + 1\}}{2\sin^2(\Delta/2)}$$

$$+\frac{2\sin((2M-1)\Delta/2)\sin(\Delta/2)2\sin(\Delta/2)\cos(\Delta/2)}{4\sin^4(\Delta/2)}]$$

$$\simeq -\frac{1}{2}[\frac{-1}{2(\pi/4M)^2} + \frac{1}{(\pi/4M)^2}]$$

$$= -\frac{4M^2}{\pi^2}. \tag{E.2}$$

Appendix F

Proof of Eq.(5.43)

With Eq.(E.1), and the approximations,

$$\cos\Delta = \cos\frac{\pi}{2M} \simeq 1,$$

$$\sin(M\Delta) = 1,$$

$$\cos(M\Delta) = 0,$$

$$\cos(\Delta/2) = \cos\frac{\pi}{4M} \simeq 1,$$

$$\sin(\Delta/2) = \sin(\pi/4M) \simeq \pi/4M,$$

the summation term in Eq.(5.40) can be written as

$$\begin{aligned}
\sum_{m=1}^{M} 1-\cos((m-1)\Delta) &= M[1-\frac{1-\cos(\Delta)+\cos((M-1)\Delta)-\cos(M\Delta)}{2M(1-\cos(\Delta))}] \\
&\simeq M[1-\frac{\cos(M\Delta)\cos(\Delta)+\sin(M\Delta)\sin(\Delta)}{4M\sin^2\frac{\Delta}{2}}] \\
&\simeq M[1-\frac{2\sin\frac{\Delta}{2}\cos\frac{\Delta}{2}}{4M\sin^2\frac{\Delta}{2}}] \\
&\simeq M[1-1/M\Delta] \\
&= M(1-2/\pi)
\end{aligned}$$

$$= .363M. \tag{F.1}$$

Then Eq.(5.40) becomes

$$\begin{aligned} \exp(-f^2/2) &= \exp\{\frac{-\rho N h^2}{2\Delta^2} \sum_{m=1}^{M} (1 - \cos((m-1)\Delta/2))\} \\ &= \exp(\frac{-.363 M N \rho h^2}{2\Delta^2}) \\ &= \exp(-.07\rho N M^3 h^2), \end{aligned} \tag{F.2}$$

where, $\Delta = \pi/2M$ (See Eq.(5.41)).

Appendix G

Proof of Eq.(5.46)

Substituting Eqs.(5.36), (5.42) and (5.43) into Eq.(5.34) yields

$$\begin{aligned} I_0 &= \frac{\sqrt{\rho N}}{2\sqrt{2\pi}} \int_0^{\Delta} \frac{h^2}{2\sqrt{2}\, h\sqrt{M(M-1)(2M-1)}} \frac{2\sqrt{6}}{} \frac{8M^3 h}{\pi} \exp(-\eta_1^2 h^2/2) dh \\ &= \frac{\sqrt{\rho N}}{2\sqrt{2\pi}} \frac{1}{2\sqrt{2}} \frac{2\sqrt{6}}{\sqrt{M(M-1)(2M-1)}} \frac{8M^3}{\pi^3} \int_0^{\Delta} h^2 \exp(-\eta_1^2 h^2/2) dh \\ &= [\frac{4\sqrt{3}M^3\sqrt{\rho N}}{\pi^3\sqrt{M(M-1)(2M-1)}}][\frac{1}{\sqrt{2\pi}} \int_0^{\Delta} h^2 \exp(-\eta_1^2 h^2/2) dh]. \end{aligned} \tag{G.1}$$

The term in the second bracket in Eq.(G.1) is

$$\begin{aligned} \frac{1}{\sqrt{2\pi}} \int_0^{\Delta} h^2 e^{-\eta_1^2 h^2/2} dh &= \frac{1}{\sqrt{2\pi}} \int_0^{\Delta} -\frac{1}{\eta_1^2} h de^{-\eta_1^2 h^2/2} \\ &= -\frac{1}{\sqrt{2\pi}\eta_1^2} [h e^{-\eta_1^2 h^2/2}|_0^{\Delta} - \int_0^{\Delta} e^{-\eta_1^2 h^2/2} dh] \\ &= -\frac{1}{\eta_1^2} [\frac{\Delta e^{-\eta_1^2 \Delta^2/2}}{\sqrt{2\pi}} - \frac{1}{\eta_1} (\frac{1}{2} - \mathrm{erfc.}\{\eta_1 \Delta\})] \\ &= -\frac{1}{\eta_1^3} [\frac{\eta_1 \Delta e^{-\eta_1^2 \Delta^2/2}}{\sqrt{2\pi}} - \frac{1}{2} + \mathrm{erfc.}\{\eta_1 \Delta\}]. \end{aligned} \tag{G.2}$$

Then, using Eqs.(5.44) and (5.45), the term in the first bracket in Eq.(G.1) times $1/\eta_1^3$ is

$$\begin{aligned}
&\frac{4\sqrt{3}M^3\sqrt{\rho N}}{\pi^3\sqrt{M(M-1)(2M-1)}\eta_1^3} \\
&= \frac{4\sqrt{3}M^3\sqrt{\rho N}}{\pi^3\sqrt{M(M-1)(2M-1)}0.14\rho l l M^3\sqrt{0.14\rho N M^3}} \\
&= \frac{4\sqrt{3}}{\pi^3(0.14)^{1.5}}\frac{1}{\rho N}\frac{\sqrt{M(M-1)(2M-1)/M^3}}{M(M-1)(2M-1)} \\
&\simeq \frac{1}{2\eta_0^2}\sqrt{\frac{M(M-1)(2M-1)}{2M^3}}. \qquad \text{(G.3)}
\end{aligned}$$

So Eq.(5.46).

Appendix H

Proof of Eq.(6.12)

For the binary detection, the minimal error probability is given by [81]

$$
\begin{aligned}
& P_{EM}(h) \\
&= \exp[\mu(q_M) + q_M^2 \ddot{\mu}(q_M)/2]\mathrm{erfc}_*\{q_M\sqrt{\ddot{\mu}(q_M)}\}/2 \\
&\quad + \exp[\mu(q_M) + (1-q_M)^2 \ddot{\mu}(q_M)/2]\mathrm{erfc}_*\{(1-q_M)\sqrt{\ddot{\mu}(q_M)}\}/2, \qquad \text{(H.1)}
\end{aligned}
$$

in which,

$$
\mathrm{erfc}_*\{x\} = \frac{1}{\sqrt{2\pi}} \int_x^\infty e^{-t^2/2} dt, \qquad \text{(H.2)}
$$

$$
\mu(q) = \ln \int_{-\infty}^{\infty} [p(\vec{r}/H_1)]^q [p(\vec{r}/H_0)]^{1-q} d\vec{r}. \qquad \text{(H.3)}
$$

As in Appendix C, from Eq.(6.11), we have

$$
p(\vec{r}/H_i) = \frac{1}{(2\pi\sigma_e^2)^{MN/2}} \exp[-\frac{1}{\sigma_e^2} \sum_{n=1}^{N} (\vec{r} - \vec{s}_{i1} - \vec{s}_{i2})^T (\vec{r} - \vec{s}_{i1} - \vec{s}_{i2})], \qquad \text{(H.4)}
$$

where,

$$
\vec{s}_{01} = a_1(n) \begin{bmatrix} \Re(\mathrm{d}_{01}) \\ \Im(\mathrm{d}_{01}) \end{bmatrix}
$$

$$= a_1(n)[1, \cos(\phi_{01}), ..., \cos((M-1)\phi_{01}), 0, \sin(\phi_{01}), ..., \sin((M-1)\phi_{01})]^T,$$

$$\vec{s}_{02} = a_2(n) \begin{bmatrix} \Re(\mathrm{d}_{02}) \\ \Im(\mathrm{d}_{02}) \end{bmatrix}$$

$$= a_2(n)[1, \cos(\phi_{02}), ..., \cos((M-1)\phi_{02}), 0, \sin(\phi_{02}), ..., \sin((M-1)\phi_{02})]^T,$$

$$\vec{s}_{11} = a_1(n) \begin{bmatrix} \Re(\mathrm{d}_{11}) \\ \Im(\mathrm{d}_{11}) \end{bmatrix}$$

$$= a_1(n)[1, \cos(\phi_{11}), ..., \cos((M-1)\phi_{11}), 0, \sin(\phi_{11}), ..., \sin((M-1)\phi_{11})]^T,$$

$$\vec{s}_{12} = a_2(n) \begin{bmatrix} \Re(\mathrm{d}_{12}) \\ \Im(\mathrm{d}_{12}) \end{bmatrix}$$

$$= a_2(n)[1, \cos(\phi_{12}), ..., \cos((M-1)\phi_{12}), 0, \sin(\phi_{12}), ..., \sin((M-1)\phi_{12})]^T,$$

$$\vec{r}(n) = \begin{bmatrix} \Re(\mathrm{r}(n)) \\ \Im(\mathrm{r}(n)) \end{bmatrix},$$

$$\vec{e}(n) = \begin{bmatrix} \Re(\mathrm{e}(n)) \\ \Im(\mathrm{e}(n)) \end{bmatrix}.$$

where,

$$\vec{s}_{12} = \vec{s}_{02} \tag{H.5}$$

since $\phi_{12} = \phi_{02}$, see Eq.(6.5). As discussed in Appendix C, we may apply an appropriate phase shift at each snapshot so that $a_i(n) = |a_i(n)|$ rendering the 2M-dimensional vector $\vec{s}_i$ real.

Substituting Eq.(H.4) into Eq.(H.3), we obtain

$$
\begin{aligned}
&\mu(q) \\
&= \ln[2\pi\sigma_e^2]^{-MN(q+1-q)/2} \int_{-\infty}^{\infty} \exp\{-\sum_{n=1}^{N} \frac{1}{\sigma_e^2}[q(\vec{r}-\vec{s}_{11}-\vec{s}_{12})^T(\vec{r}-\vec{s}_{11}-\vec{s}_{12}) \\
&\quad +(1-q)(\vec{r}-\vec{s}_{01}-\vec{s}_{02})^T(\vec{r}-\vec{s}_{01}-\vec{s}_{02})]\}d\vec{r} \\
&= \ln[2\pi\sigma_e^2]^{-MN/2} \int_{-\infty}^{\infty} \exp\{-\sum_{n=1}^{N} \frac{1}{\sigma_e^2}[\vec{r}-q(\vec{s}_{11}+\vec{s}_{12})-(1-q)(\vec{s}_{01}+\vec{s}_{02})]^T \\
&\quad [\vec{r}-q(\vec{s}_{11}+\vec{s}_{12})-(1-q)(\vec{s}_{01}+\vec{s}_{02})]\}d\vec{r} \\
&\quad \exp\{-\sum_{n=1}^{N} \frac{1}{\sigma_e^2}[-q^2(\vec{s}_{11}+\vec{s}_{12})^T(\vec{s}_{11}+\vec{s}_{12})-q(1-q)(\vec{s}_{11}+\vec{s}_{12})^T(\vec{s}_{01}+\vec{s}_{02}) \\
&\quad -q(1-q)(\vec{s}_{01}+\vec{s}_{02})^T(\vec{s}_{11}+\vec{s}_{12})-(1-q)^2(\vec{s}_{01}+\vec{s}_{02})^T(\vec{s}_{01}+\vec{s}_{02}) \\
&\quad +q(\vec{s}_{11}^T\vec{s}_{12}+\vec{s}_{12}^T\vec{s}_{11}+\vec{s}_{11}^T\vec{s}_{11}+\vec{s}_{12}^T\vec{s}_{12})+(1-q)(\vec{s}_{01}^T\vec{s}_{02}+\vec{s}_{02}^T\vec{s}_{01}+\vec{s}_{01}^T\vec{s}_{01}+\vec{s}_{02}^T\vec{s}_{02})]\} \\
&= \ln\exp\{-\sum_{n=1}^{N} \frac{1}{\sigma_e^2}q(1-q)[(\vec{s}_{11}+\vec{s}_{12})^T(\vec{s}_{11}+\vec{s}_{12})+(\vec{s}_{01}+\vec{s}_{02})^T(\vec{s}_{01}+\vec{s}_{02}) \\
&\quad -(\vec{s}_{11}+\vec{s}_{12})^T(\vec{s}_{01}+\vec{s}_{02})-(\vec{s}_{01}+\vec{s}_{02})^T(\vec{s}_{11}+\vec{s}_{12})]\} \\
&= -\sum_{n=1}^{N} \frac{1}{\sigma_e^2}q(1-q)[(\vec{s}_{11}-\vec{s}_{01})+(\vec{s}_{12}-\vec{s}_{02})]^T[(\vec{s}_{11}-\vec{s}_{01})+(\vec{s}_{12}-\vec{s}_{02})] \\
&= -q(1-q)\sum_{n=1}^{N} \frac{1}{\sigma_e^2}(\vec{s}_{11}-\vec{s}_{01})^T(\vec{s}_{11}-\vec{s}_{01}), \qquad \text{(H.6)}
\end{aligned}
$$

where $\vec{s}_{12} - \vec{s}_{02} = \vec{0}$ because of Eq.(H.5).

We can evaluate the summation part in Eq.(H.6) so that

$$\sum_{n=1}^{N} \frac{1}{\sigma_e^2}(\vec{s}_{11} - \vec{s}_{01})^T(\vec{s}_{11} - \vec{s}_{01})$$

$$= \sum_{n=1}^{N} \frac{|a_1(n)|^2}{\sigma_e^2}[1, \cos(\phi_{11}) - \cos(\phi_{01}), ..., \cos(\phi_{11}(M-1)) - \cos(\phi_{01}(M-1)),$$

$$0, \sin(\phi_{11}) - \sin(\phi_{01}), ..., \sin((M-1)\phi_{11}) - \sin((M-1)\phi_{01})]$$

$$[1, \cos(\phi_{11}) - \cos(\phi_{01}), ..., \cos(\phi_{11}(M-1)) - \cos(\phi_{01}(M-1)),$$

$$0, \sin(\phi_{11}) - \sin(\phi_{01}), ..., \sin((M-1)\phi_{11}) - \sin((M-1)\phi_{01})]^T$$

$$\simeq N\rho_1 \sum_{m=1}^{M} [\cos((m-1)\phi_{11}) - \cos((m-1)\phi_{01})]^2 + [\sin((m-1)\phi_{11}) - \sin((m-1)\phi_{01})]^2$$

$$= 2N\rho_1 \sum_{m=1}^{M} (1 - \cos[(m-1)(\phi_{11} - \phi_{01})])$$

$$= 2N\rho_1 \sum_{m=1}^{M} (1 - \cos[(m-1)h]), \tag{H.7}$$

where, $\sum_{n=1}^{N} |a_1(n)|^2 \simeq N\sigma_{s1}^2$ is from Eq.(3.34), and

$$\rho_1 = \frac{\sigma_{s1}^2}{\sigma_e^2}$$

is defined as the signal-noise-ratio (SNR) of the first received signal.

As discussed in Appendix B, we choose $q = q_M = 0.5$ so that $\dot{\mu}(q) = 0$. Then

$$\mu_1(0.5) = -0.25N\rho_1 \sum_{m=1}^{M} [1 - \cos((m-1)h)], \tag{H.8}$$

$$\ddot{\mu}_1(0.5) = 2N\rho_1 \sum_{m=1}^{M} [1 - \cos((m-1)h)]. \tag{H.9}$$

From Eq.(H.1), therefore,

$$P_{EM1}(h) = \text{erfc}_*\{\sqrt{\ddot{\mu}_1(1/2)/2}\} = \text{erfc}_*\{f_1(h)\}, \tag{H.10}$$

where, $\text{erfc}_*\{x\}$ is defined in Eq.(H.2), and

$$f_1(h) = \{N\rho_1 \sum_{m=1}^{M}[1 - \cos((m-1)h)]\}^{.5}. \tag{H.11}$$

Bibliography

[1] Barankin, E. W., "Locally best unbiased estimates," *Ann. Math. Statist.*, vol. 20, pp. 447-501, 1949.

[2] Billini, S. and Tartara, G., "Bounds on error in signal parameter estimation," *IEEE Trans. Communication* , vol. COM-22, No. 3, pp. 304-342, 1974.

[3] Bienvenu, G. and Kopp, L., "Optimality of high resolution array processing," *IEEE Trans. Acoust., Speech, Signal Processing,* vol. ASSP-31, No. 10, pp. 1235-1248, 1983.

[4] Bobrovsky, B. Z. and Zakai. M., "A lower bound on the estimation error for certain diffusion processes," *IEEE Trans. Information Theory,* vol. IT-22, No. 1, pp.45-52, 1976.

[5] Brennan, L. E., "Angular accuracy of a pulsed search radar," *IRE Trans. Antennas Propag.*, vol. AP-9, pp. 268-275, 1961.

[6] Bresler, Y. and Macovski, A., "Exact Maximum likelihood parameter estimation of superimposed exponential signals in noise," *IEEE Trans. Acoust., Speech, Signal Processing,* vol. ASSP-34, No. 10, pp. 1081-1089, 1986.

[7] Burdic, W. S., *Underwater Acoustic System Analysis.* New York: Prentice Hall, 1984.

[8] Carter, G. C., "Variance bounds for passively locating an acoustic source with a symmetric line array," *J. Acoust. Soc. Amer.*, vol. 62, No. 4, pp. 922-926, 1977.

[9] Chapman, D. G. and Robbins, N., "Minimum variance estimation without regularity assumptions," *Ann. Math. Statist.*, vol. 22, pp. 581-586, 1951.

[10] Chazan, D. Zakai, M. and Ziv. J. , "Improved lower bounds on signal parameter estimation," *IEEE Trans. Information Theory*, vol. IT-21, No. 1, pp. 90-93, January 1975.

[11] Clergeot, H., Tressens, S., and Ouamri, A., "Performance of high resolution frequencies estimation methods compared to the Cramer-Rao bounds," *IEEE Trans. Acoust., Speech, Signal Processing*, vol. ASSP-37, No. 11, pp. 1730-1719, 1989.

[12] Cox, H., "Resolving power and sensitivity to mismatch of optimum array processors," *J. Acoust. Soc. Amer.*, vol. 54, No. 3, pp. 771-788, 1973.

[13] Cramer, H., *Mathematical Methods of Statistics*, Princeton: Princeton University Press, 1946.

[14] Farrier, D.R., "High-performance signal subspace beamformer," *IEE Proceedings*, vol. 136, No. 6, pp.283-288, 1989.

[15] Farrier, D.R., Jeffries, D.J., and Mardani, R., "Theoretical performance prediction of the MUSIC algorithm," *IEE Proceedings*, vol. 135, No. 3, pp.216-224, 1988.

[16] Fariedlander, B. and Porat, B., "Bounds for ARMA spectral analysis based on sample covariances," in *Proc. IEEE Int. Conf. Acoust., Speech, Signal Processing*, pp. 616-619, 1985.

[17] Feder, M. and Weinstein, E., "Parameter estimation of superimposed signals using the EM algorithm," *IEEE Trans. Acoust., Speech, Signal Processing*, vol. ASSP-36, No. 4, pp. 477-489, 1988.

[18] Garbriel, W., " Spectral analysis and adaptive array superresolution techniques," *Proc. IEEE*, vol. 68, No. 6, pp. 654-666, June, 1980.

[19] Gatlin, B., "Superresolution of multiple noise sources in antenna beam," *IEEE Trans. Antennas Propag.*, vol. AP-31, No. 3, pp. 456-462, May, 1983.

[20] Giere, R. N., *Explaining Science* , Chicago, London: The University of Chicago Press, 1988.

[21] Golub, G. H. and Van Loan, C. F., *Matrix Computation.* Baltimore, Maryland: John Hopkins University Press, 1983.

[22] Goodman, N. R., " Statistical analysis based on a certain multivariate complex, Gaussian distribution (an introduction) ," *Ann. Math. Statist.* No. 34., pp. 152-177, 1963.

[23] Gorman, J. D. and Hero, A.O., "Lower bounds for parametric estimation with constraints," *IEEE Trans. Information Theory,* vol. IT-36, No. 6, pp. 1285-1301, November 1990.

[24] Gu, X. and Wong, K. M., "A modified Ziv-Zakai lower bound and its application in array processing," in *Proc. IEEE Int. Conf. Acoust., Speech,Signal Processing,* Toronto, Ontario, pp. 1477-1480, 1991.

[25] Hahn, W. R., "Optimum signal processing for passive sonar range and bearing estimation," *J. Acoust. Soc. Amer.,* vol. 58, pp. 201-207, 1981.

[26] Hahn, W. R. and Tretter, S. A., "Optimum processing for delay-vector estimation in passive signal arrays," *IEEE Trans. Information Theory,* vol. IT-19, No. 5, pp. 608-614, September 1973.

[27] Haykin, S., (editor), *Array Processing.* Englewood Cliffs, New Jersey: Prentice Hall, 1985.

[28] Haykin, S., *Nonlinear Methods of Spectral Analysis.* New York: Springer-Verlag, 1983.

[29] Helstrom, C. W., *Statistical Theory of Signal Detection.* New York: Pergamon Press, 1960

[30] Hero, A.O., "A Cramer-Rao type lower bound for essentially unbiased parameter estimation," *Lincoln Laboratory,* Technical Report 890, January 3, 1992.

[31] Hero, A. O., "Poisson models and mean-squared error for correlate estimators of time delay," *IEEE Trans. Information Theory,* vol. IT-34, No. 2, pp. 287-303, March 1988.

[32] Hudson, J. E., *Adaptive Array Principles.* Stevenage, U. K., and New York: Perter Peregrinus LTD., 1981.

[33] Ianniello, J. P., "Time delay estimation via cross-correlation in the presence of large estimation errors," *IEEE Trans. Acoust., Speech, Signal Processing,* vol. ASSP-30, pp. 998-1003, 1982.

[34] Johnson, D. H. and Degraff, S. R. "Improving the resolution of bearing in passive sonar array by eigenvalue analysis," *IEEE Trans. Acoust., Speech, Signal Processing,* vol. ASSP-30, No. 4, pp. 638-647, 1982.

[35] Johnson, D. H., "The application of spectral estimation methods to bearing estimation problems," *Proc. IEEE,* vol. 70, No. 9, pp. 1018-1028, 1982.

[36] Kaveh, M. and Barabell, A. J.,"The statistical performance of the MUSIC and the minimum-norm algorithms in resolving plane waves in noise," *IEEE Trans. Acoust., Speech, Signal Processing,* vol. ASSP-34, No. 2, pp. 331-341, 1986.

[37] Kay, S. M., "Accurate frequency estimation at low signal-to-noise ratio," *IEEE Trans. Acoust., Speech, Signal Processing,* vol. ASSP-32, No. 3, pp. 331-341, 1984.

[38] Kendall, M. G., *Advanced Theory of Statistics.* New York: Macmillan, 1977.

[39] Kiefer, J., "On minimum variance estimators," *Ann. Math. Statist.*, vol. 23, pp. 627-629, 1952.

[40] Ksienski, A. A. and McGhee, R. B., "A decision theoretic approach to the angular resolution and parameter estimation problem for multiple targets," *IEEE Trans. Aerospace and Electronic Systems*, vol. AES-4, No. 3, pp. 443-455, 1968.

[41] Kumaresan, R. and Tufts, D. W., "Estimating the angles of arrival of multiple plane waves," *IEEE Trans. Aerospace and Electronic Systems*, vol. AES-19, pp. 134-139, 1983.

[42] Kuruc, A. R., "Lower bounds on multiple-source direction finding in the presence of direction-dependent antenna-array- calibration errors," *MIT Lincoln Laboratory*, Technical Report 799, October 24, 1989.

[43] Lang, S. W., "Frequency estimation with maximum entropy spectral estimators," *IEEE Trans. Acoust., Speech, Signal Processing*, vol. ASSP-28, No. 6, pp. 716-724, 1980.

[44] Lee, H. B., " The Cramer-Rao bound on frequency estimates of signals closely spaced in frequency," *IEEE Trans., Signal Processing*, vol. 40, No. 6, pp. 1508-1517, 1992.

[45] Lee, H. B. and Wengrovitz M. S., " Resolution threshold of beamspace MUSIC for two closely spaced emitters," *IEEE Trans. Acoust., Speech, Signal Processing*, vol. ASSP-38, No. 9, pp. 1545-1559, 1990.

[46] Lee, Y. W., *Statistical theory of Communication.* New York: Wiley, 1960.

[47] Ligett, W. S., *Passive sonar: Fitting models to multiple time series.* New York: Academic Press, 1973.

[48] Magnus, J. R., *Matrix Differential Calculus with Applications in Statistics and Econometrics.* New York: Wiley, 1988.

[49] Mayrague, "ESPRIT and TAM (Toeplitz approximation method) are theoretically equivalent," in *Proc. IEEE Int. Conf. Acoust., Speech, Signal Processing* New York, N. J., pp.2456-2459, 1988.

[50] McAulay, R. J. and Hofstetter, E. M., "Barankin bounds on parameter estimation," *IEEE Trans. Information Theory,* vol. IT-17, No. 6, pp. 669-676, November 1971.

[51] McAulay, R. J. and Seidman, L. P., "A useful form of the Barankin lower bound and its application to PPM threshold analysis," *IEEE Trans. Information Theory,* vol. IT-15, No. 2, pp. 273-279, march 1969.

[52] Messer, H. and Adar, Y., "New lower bounds of frequency estimation of a multitone random signal in noise," *Signal Processing,* 18, pp.413-424, 1989.

[53] Miller, K. S., *Complex Stochastic Processes, an Introduction to Theory and Application.* Addison-Wesley Publishing Company, Inc. 1974.

[54] Miller, M. I., and Fuhrmann, D.R.. "Maximum-likelihood narrow-band direction finding and the EM algorithm," *IEEE Trans. Acoust., Speech, Signal Processing,* vol. ASSP-38, No. 9, pp. 1560-1577, 1990.

[55] Muirhead, R. J., *Aspects of Multivariate Statistical Theory.* New York: Wiley, 1982.

[56] Nohara, T. J., *Two-Dimensional Spectrum Estimation with the MUSIC Algorithm and Classical Beamforming,* M. Eng. Thesis, ECE Department, McMaster University, Canaca, 1987.

[57] Ouibrahim, D. D. Weiner and Sarkar, T. K., "A generalized approach to direction finding," *IEEE Trans. Acoust., Speech, Signal Processing,* vol. ASSP-36, No. 4, pp. 610-612, 1988.

[58] Parthasarathy, S. and Tufts, D. W., "Maximum-likelihood estimation of parameters of exponentially damped sinusoids," *Proc. IEEE,* vol. 73, pp.1528-1530, 1985.

[59] Pillai, S. U. and Kwon, B. H., "Performance analysis of MUSIC-type high resolution estimators for direction finding in correlated and coherent scenes," *IEEE Trans. Acoust., Speech, Signal processing,* vol. ASSP-37, No. 8, pp. 1176-1189, 1989.

[60] Pisarenko, V. F., "The retrieval of harmonics from a covariance function," *Geophys. J. R. Astron. Soc.,* 33, pp. 347-366, 1973.

[61] Porat, B. and Friedlander B., "Analysis of the asymptotic relative efficiency of the MUSIC algorithm," *IEEE Trans. Acoust., Speech, Signal Processing,* vol. ASSP-36, No. 4, pp. 532-543, 1988.

[62] Priestley, M. B., *Spectral Analysis and Time Series.* London, New York: Academic Press, 1981.

[63] Rao, B. D. and Hari, K. V. S., "Performance analysis of ESPRIT and TAM in determining the direction of arrival of plane waves in noise," *IEEE Trans. Acoust., Speech, Signal Processing,* vol. ASSP-37, No. 12, pp. 1990-1995, 1989.

[64] Rao, B. D. and Hari, K. V. S., "Statistical performance analysis of the minimum-norm method," *IEE Proceedings,* vol. 136, No. 3, pp.125-134, 1989.

[65] Rendas, M. J. D. and Moura, M. F., "Cramer-Rao bound for location svstems in multipath environments, " *IEEE Trans. Acoust., Speech, Signal Processing,* vol. ASSP-39, No. 12, pp. 2593-2610, 1991.

[66] Rife, D. C. and Boorstyn, R. R., "Multiple tone parameter estimation from discrete-time observations," *The Bell System Technical J.,* vol. 55, No. 9, pp. 1389-1410 , 1976.

[67] Rife, D. C. and Boorstyn, R. R., "Single-tone parameter estimation from discrete-time observations," *IEEE Trans. Information Theory,* vol. IT-20, No. 5, pp. 591-598, September 1974.

[68] Roy, R. and Kailath, T., "ESPRIT-Estimation of signal parameters via rotational invariance techniques," *IEEE Trans. Acoust., Speech, Signal Processing*, vol. ASSP-37, No. 7, pp. 984-995, 1989.

[69] Schmidt, R. O. and Frands, R. E., "Multiple emitter location and signal parameter estimation," *IEEE Trans. Antennas Propag.*, vol. AP-34, No. 3, pp. 276-280, 1986.

[70] Schweppe, F. C., "Sensor-array data processing for multiple-signal sources," *IEEE Trans. Information Theory*, vol. IT-14, No. 2, pp. 294-305, March 1968.

[71] Seidman, L. P., "Performance limitations and error calculations for parameter estimation," *Proc. IEEE*, vol. 58, No. 5, pp. 644-652, 1970.

[72] Starer, D. and Nehorai, A., "Maximum likelihood estimation of exponential signals in noise using a Newton algorithm," *Proc. 4th ASSP Workshop on Spectrum Estimation and Modelling*, pp. 240-245, Minneapolis, MN, August 1988.

[73] Stoica, P. and Nehorai, A., "Performance comparison of subspace rotation and MUSIC methods for direction estimation," *IEEE Trans. Acoust., Speech, Signal Proces[illegible]*, vol. ASSP-39, No. 2, pp. 446-453, 1991.

[74] Stoica, P. and Nehorai, A., "MUSIC, Maximum likelihood, and Cramer-Rao bound: further results and comparisons," *IEEE Trans. Acoust., Speech, Signal Processing*, vol. ASSP-38, No. 12, pp. 2140-2150, 1990.

[75] Stoica, P. and Nehorai, A., "Performance study of conditional and unconditional direction-of-arrival estimation," *IEEE Trans. Acoust., Speech, Signal Processing*, vol. ASSP-38, No. 10, pp. 1783-1795, 1990.

[76] Stoica, P. and Sharman, K., "Maximum likelihood methods for direction-of-arrival estimation," *IEEE Trans. Acoust., Speech, Signal Processing*, vol. ASSP-38, No. 7, pp. 1132-1143, 1990.

[77] Stoica, P. and Nehorai, A., "MUSIC, maximum likelihood, and Cramer-Rao bound," *IEEE Trans. Acoust., Speech, Signal Processing,* vol. ASSP-37, No. 5, pp. 720-741, 1989.

[78] Swerling, "Parameter estimation accuracy formulas," *IEEE Trans. Information Theory,* vol. IT-10, pp. 302-313, October 1964.

[79] Takeuchi, K., *Foundations of Multivariate Analysis: A Unified Approach by Means of Projection onto Linear Subspaces.* New Delhi: Wiley Eastern, 1982.

[80] Tufts, D. W. and Kumeresan, R., "Estimation of frequencies of multiple sinusoids: making linear prediction perform like maximum likelihood, " *Proc. IEEE,* vol. 70, No. 9, pp. 975-989, 1982.

[81] Van Trees, H. L., *Detection, Estimation and Modulation Theory: Part I.* New York: Wiley, 1968.

[82] Wang, H. and Kaveh, M., "On the performance of signal-subspace processing-Part I: narrow-band systems," *IEEE Trans. Acoust., Speech, Signal Processing.* vol. ASSP-34, No. 10, pp. 1201-1209, 1986.

[83] Wax, M. and Ziskand, I., "On unique localization of multiple sources by passive sensor arrays," *IEEE Trans. Acoust., Speech, Signal Processing,* vol. ASSP-37, No. 7, pp. 1553-1560, 1989.

[84] Wax, M. and Kailath, T., "Detection of signals by information theoretic criteria," *IEEE Trans. Acoust., Speech, Signal Processing,* vol. ASSP-33, No. 4, pp. 387-392, 1985.

[85] Wax, M. and Ziv, J., "Improved bounds on the local mean-square error and the bias of parameter estimations," *IEEE Trans. Information Theory,* vol. IT-23, pp. 529-530, July 1977.

[86] Weiss, A. J. and Weinstein, E., "Lower bounds in parameter estimation - summary of results," in *Proc. IEEE Int. Conf. Acoust., Speech, Signal Processing*, Tokyo, pp. 567-572, 1986.

[87] Weiss, A. J., "Fundamental limitations in passive time delay estimation - part I: narrow-band systems," *IEEE Trans. Acous., Speech, Signal Processing*, vol. ASSP-31, No. 2, pp. 472-486, 1983.

[88] Weinstein, E., and Wiss, A. J., "A general class of lower bounds in parameter estimation," *IEEE Trans. Information Theory*, vol. IT-34, No. 2, pp. 338-342, March 1988.

[89] Weinstein, E., "Relations between Belini-Tartara, Chazan-Zakai-Ziv, and Wax-Ziv lower bounds," *IEEE Trans. Information Theory*, vol. IT-34, No. 2, pp. 342-343, March 1988.

[90] Weinstein, E., "Decentralization of the Gaussian maximum likelihood estimator and its applications to passive array processing," *IEEE Trans. Acoust., Speech, Signal Processing*, vol. ASSP-29, No. 5, pp.945-951, 1981.

[91] Williams, R. T., Prasad, S., Mahalanabis, A. K., and Sibul, L. H., "An improved spatial smoothing technique for bearing estimation in a multipath environment," *IEEE Trans. Acoust., Speech, Signal Processing*, vol. ASSP-36, No. 4, pp.425-432, 1988.

[92] Wong, K. M., Walker, R. S. and Niezgoda, G., "Effects of random sensor motion on bearing estimation by the MUSIC algorithm," *IEE Proceedings*, vol. 135, No. 3, pp. 233-250, 1988.

[93] Yin, Y. Q. and Krishnaiah, P. R., "On some nonparametric methods for detection of the number of signals," *IEEE Trans. Acoust., Speech, Signal Processing*, vol. ASSP-35, No. 11, pp. 1533-1538, 1987.

[94] Zakai, M. and Ziv, J. "On the Threshold Effect in Radar Range Estimation," *IEEE Trans. Information Theory*, vol. IT-15, No. 1, pp. 167-170, January 1969.

[95] Zhang, Q. T., Wong, K. M., Yip, J. P. and Reilly, J. P., "Statistical analysis of the performance of information criteria in the detection of the number of signals in array processing," *IEEE Trans. Acoust., Speech, Signal Processing*, vol. ASSP-37, No. 10, pp. 1555-1567, 1989.

[96] Zhao, L. C., Krishnaiah, P. R. and Bai, Z. D., "Remarks on Certain Criteria for detection of number of signals," *IEEE Trans. Acoust., Speech, Signal Processing*, vol. ASSP-35, No. 2, pp. 129-132, 1987.

[97] Zhao, L. C., Krishnaiah, P. R. and Bai, Z. D., "On detection of the number of signals in presence of white noise," *J. Multivariate Anal.* 20, pp. 1-25, 1986.

[98] Zhou, C., Haber, F. and Jaggard, D. L., "A resolution measure for the MUSIC algorithm and its application to plane wave arrivals contaminated by coherent interference," *IEEE Trans. Acoust., Speech, Signal Processing*, vol. ASSP-39, No. 2, pp. 454-463, 1991.

[99] Ziskind, I. and Wax, M., "Maximum likelihood localization of multiple sources by alternating projection," *IEEE Trans. Acous., Speech, Signal Processing*, vol. ASSP-36, No.10, pp. 1553-1560, 1988.

[100] Ziv, J. and Zakai, M. "Some Lower Bounds on Signal Parameter Estimation," *IEEE Trans. Information Theory*, vol. IT-15, No. 3, pp. 386-391, May 1969.

[101] Zlotoxski, M. D. and Lee, T. S., "Maximum likelihood based sensor array signal processing in the beamspace domain for low angle radar tracking," *IEEE Trans. Acoust., Speech, Signal Processing*, vol. ASSP-39, No. 3, pp. 656-671, 1991.