TEL-AVIV UNIVERSITY

The Iby and Aladar Fleischman Faculty of Engineering

Direction of Arrival Estimation of Known Signals in Channels with Memory

A thesis submitted towards the degree of

Master of Science in Electrical Engineering

by

Shaul Shulman

August 2004

TEL-AVIV UNIVERSITY

The Iby and Aladar Fleischman Faculty of Engineering

Direction of Arrival Estimation of Known Signals in Channels with Memory

A thesis submitted towards the degree of
Master of Science in Electrical Engineering

by

Shaul Shulman

This research was carried out in the
Department of Electrical Engineering-Systems,
under the supervision of Prof. Anthony J. Weiss

August 2004

Acknowledgement

I would like to thank my supervisor Prof. Anthony Weiss for his devoted guidance and support throughout the work on this thesis.

This thesis is dedicated to my parents Boris and Maya Shulman.

Abstract

We consider the problem of estimating the Directions-of-Arrival (DOA) of signals with known waveforms using a sensor array. The case of known waveforms is encountered in digital communications (after detection) or when training and synchronization waveforms are used. We address multiple signals with Inter-Symbol Interference (ISI).

The exact Maximum Likelihood (ML) estimator of angles and channel coefficients is presented, as well as more computationally efficient, approximate estimators. We show that assuming that the signal waveforms are uncorrelated, the multidimensional ML minimization problem decouples to a set of independent minimizations which can be separately solved for each one of the signals. This property has been shown to hold for signals with known waveforms, and here it is expanded to a case of ISI and multipath. The decoupled minimization significantly simplifies the computational complexity of the algorithm and eliminates the resolution limitation of closely spaced angles. The number of separate minimizations is equal to the number of uncorrelated sources. The order of each minimization is equal to the number of multipath reflections.

Cramer-Rao bounds (CRB) are also developed in order to evaluate the performance of the algorithm. Compact expressions of CRB are presented for single path and multipath cases.

Monte Carlo simulations verify the theoretical results. All estimators are shown to approach the CRB for long enough observations. It is shown that the decoupled estimation of uncorrelated sources coincides with the single source CRB at low SNR. At high SNR the interference due to other sources is the main contributor to the error and then the RMS error deviates from the CRB. The threshold point depends on the correlation between the source waveform and the interfering waveforms.

Contents

List of Figures

List of abbreviations and notations

Abbreviations

AWGN	-	Additive White Gaussian Noise
DOA	-	Direction of Arrival
ML	-	Maximum Likelihood
CRB	-	Cramer-Rao Bound
MSE	-	Mean Square Error
RMSE	-	Root Mean Square Error
PDF	-	Probability Density Function
ULA	-	Uniform Linear Array
SNR	-	Signal to Noise Ratio
i.i.d	-	Independently Identically Distributed

Notation

In this thesis, matrices and vectors are set in boldface, with upper-case letters used for matrices and lower-case letters for vectors. If nothing else is explicitly stated, the meaning of the following is:

$\mathbf{A}^T$	-	Transpose
$\mathbf{A}^*$	-	Complex Conjugate Transpose
$\mathbf{A}^+$	-	Complex Conjugate
$\mathbf{A}^\dagger$	-	The Moore-Penrose pseudo inverse of $\mathbf{A}$
$Tr\{\mathbf{A}\}$	-	Trace operator
$Vec(\mathbf{A})$	-	The column vector that is obtained by stacking the columns of $\mathbf{A}$
$[\mathbf{A}]_{ij}$	-	The $(i,j)^{th}$ element of the matrix $\mathbf{A}$
$\|\mathbf{A}\|_F$	-	The Frobenious norm of a matrix, defined as $\|\mathbf{A}\|_F^2 = Tr\{\mathbf{A}\mathbf{A}^*\}$
$\mathbf{A} \otimes \mathbf{B}$	-	The Kronecker product.

For an $N \times M$ matrix $\mathbf{A}$ and $K \times L$ matrix $\mathbf{B}$, it is defined as the $NK \times ML$ matrix

$$\mathbf{A} \otimes \mathbf{B} \triangleq \begin{bmatrix} a_{11}\mathbf{B} & \dots & a_{1M}\mathbf{B} \\ \vdots & \ddots & \vdots \\ a_{N1}\mathbf{B} & \dots & a_{NM}\mathbf{B} \end{bmatrix}$$

$\mathbf{A} \odot \mathbf{B}$	-	The Hadamard or Schur product, i.i. elementwise multiplication
$diag(\mathbf{A})$	-	A column vector containing the diagonal elements of the matrix $\mathbf{A}$
$diag(a_1, \dots, a_K)$	-	$K \times K$ diagonal matrix with diagonal elements a_1 through a_K
$\mathbf{I}_m$	-	The identity $m \times m$ matrix. Often, subscript m is omitted.
$\hat{\theta}$	-	An estimate of the parameter θ
θ_0	-	True value of the parameter θ
$Re[\bullet]$	-	The real part of a complex quantity
$Im[\bullet]$	-	The imaginery part of a complex quantity
$E\{x\}$	-	Expectation (mean) of a random variable x
$Var\{x\}$	-	Variance of a random variable x
$\delta_{t,s}$	-	The Kronecker delta
$\arg\min_x f(x)$	-	The minimizing argument of the function $f(x)$
$\arg\max_x f(x)$	-	The maximizing argument of the function $f(x)$

Chapter 1

Introduction

This thesis deals with the problem of estimating the Directions of Arrival (DOA) of electromagnetic waves impinging on an array of antenna elements (sensors). DOA estimation using an array of sensors has been a topic of extensive research in the past few decades. Classical applications include source localization in RADAR and SONAR, location of communication transmitters, seismology and radio astronomy.

Until recently, most direction finding algorithms were intended to deal with analog signals. In such cases the waveform of the received signal is not known, therefore the conventional DOA estimations methods do not utilize any information on the temporal structure of the signal. (A basic summary of these methods in given in chapter 2 of this thesis).

In the past decade there has been an increasing interest in direction finding of digital signals. For example, in cellular communications, DOA estimation of a mobile subscriber can be used for various "location based services."

In the case of a digital signal, the transmitted waveform can often be known to the receiver. This happens when either a training sequence is used or when the transmitted symbols are being demodulated and fed back.

Several previous works addressed a case of known signals([7],[8],[4],[17],[18]).

Works by Li et.al. ([7],[8]) showed that knowledge of signal waveform improves estimation performance and allows to separately estimate the angles of known signals, under the condition that they are weakly correlated with each other. The waveforms were assumed known up to a complex scaling, which represents a slow (flat) fading of the signal.

Caderval and Moses ([4]) extended this algorithm to a case of multipath. The array was assumed to receive several multipath reflections of each source and the angle of each reflection was estimated.

However, the algorithms mentioned above presume that the transmitted waveforms arrive at the array without any relative time delay. Unfortunately, in digital communications this is not always the case.

It is more likely that each reflection will have a different time delay when it impinges on the array. In a case of a single path, where each signal arrives from one angle only, the symbols can also suffer relative time delays if there are scatterers in the vicinity of the transmitter. This happens when the distance from the transmitter to the receiver is much larger than the distance to the scatterer, so all reflections arrive at the receiver from one angle but at different times and with different scalings.

This scenario is suitable for a rural environment. For example, imagine that a cellular phone is located inside a house and is received by an antenna positioned on a high mast at a base station few kilometers away. The only reflections that the signal may contain when it reaches the base station can origin from the walls of the house. When received at the base station, the relative angles of reflections are practically zero, but the relative time delays are not.

In an urban environment, in addition to local scattering, there will also be reflections from surrounding buildings, multiplying the number of DOAs from each source.

When the array response is periodically sampled, these relative time delays cause the data symbols of the reflections to sum up with different scalings. This phenomena is called Intersymbol Interference (ISI), and can be viewed as if the signal propagates through a channel that has memory, a channel which "remembers" and adds up previously transmitted symbols to each new received symbol.

It is therefore desirable to propose algorithms for a model of multiple known waveforms, possibly propagating through multiple paths, and all having a temporal parametrization in a from of ISI. This is the main purpose of this thesis.

It is important to note, that although we treat a situation where the transmitted signal is scattered in the near area, (which causes ISI but still preserves a single DOA) we assume the transmitter to be a single point. This is in contrast to technique known as distributed source modelling (see for example [9] and [2]), where the source location is assumed to be random and its mean is of interest.

In this work we state the data model and present Maximum Likelihood estimation algorithms for the problem at hand. Cramer-Rao bounds on the estimation error are developed in order to evaluate the performance of the estimators. Exact and approximate, more computationally efficient, Maximum-Likelihood estimation algorithms are developed for estimating the DOAs of all signals and the corresponding channel coefficients. We also derive compact expressions of the CRBs for multipath and single path cases.

Numerical examples are presented and compared with the theoretical results. The simulations give additional insight into the behavior of the problems addressed.

For a summary of the main results, refer to chapter 8 of this thesis.

The document is organized as follows: chapter 2 presents the basic theoretical background of the area of DOA estimation; chapter 3 presents a brief literature overview of relevant previous works; chapter 4 states the data model; chapter 5 presents the estimation algorithms; in chapter 6, the CRBs are derived; numerical examples are given in chapter 7 and the summary appears in chapter 8.

Chapter 2

Theoretical Background - Traditional Angle Estimation Algorithms

2.1 Introduction

Direction finding using an array of spatially distributed sensors has been a topic for extensive research in the past decades. Apart from theoretical aspects, this research effort has been motivated by a wide spectrum of real-world problems and applications, the classical application being the source localization in radar in sonar. In this section we will briefly overview the basic methods of direction finding for unknown signals.

2.1.1 Data model

Consider and array of sensors that collects samples of the impinging signals. Figure 2.1 illustrates an example of 2 sensors. Assuming that a signal with complex amplitude $s(t)$ is generated by a transmitter located in a far-field of the array, the

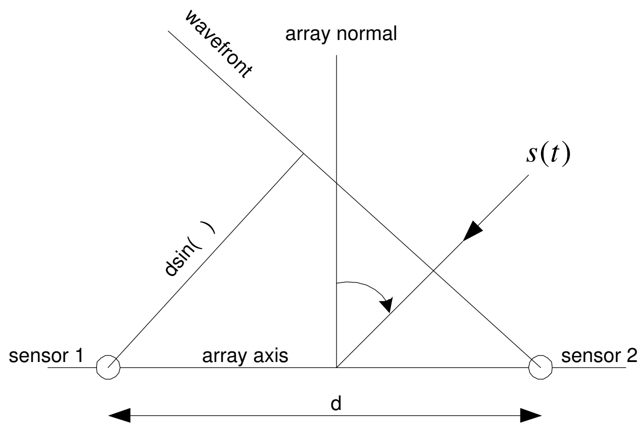

Figure 2.1: **A Uniform Linear Array with two sensor elements**

electromagnetic wave arriving at the sensor array is approximately plane. If the direction θ is different from zero, then sensor 1 experiences a time delay with respect to sensor 2 :

$$\tau = \frac{d\sin(\theta)}{v} \tag{2.1}$$

where d is the sensor spacing and v is the velocity of the plane wave. If $s(t)$ is a narrow band signal with carrier frequency f_0, then the time delay τ corresponds to a phase shift of

$$\phi = \frac{2\pi d}{\lambda_0}\sin\theta \tag{2.2}$$

where λ_0 is the wavelength corresponding to the carrier frequency, i.e.

$$\lambda_0 = \frac{v}{f_0} \tag{2.3}$$

For an array of M sensors and for K impinging signals, the measured vector $\mathbf{y}(t) \in C^{M\times 1}$ takes the following form

$$\mathbf{y}(t) = \mathbf{A}(\boldsymbol{\theta})\mathbf{s}(t) + \mathbf{n}(t) \tag{2.4}$$

where

$$\mathbf{A}(\boldsymbol{\theta}) = [\mathbf{a}(\theta_1) \dots \mathbf{a}(\theta_K)] \tag{2.5}$$

and

$$\mathbf{a}(\theta_k) = [1 \quad \phi_k \quad \dots \quad \phi_k^{M-1}]^T \tag{2.6}$$

$\mathbf{A}(\boldsymbol{\theta}) \in C^{M \times K}$ is called a *steering matrix*. And it's columns are the *steering vectors*, which contain the phase shifts of the signal as received at each of the sensors. For example, in case of Uniform Linear Array (ULA) with a sensor spacing of $\frac{\lambda}{2}$ the steering vector for a signal number k ($k = 1 \dots K$) would be

$$\phi_k = e^{-j\pi sin\theta_k} \tag{2.7}$$

$\mathbf{n}(t) \in C^{M \times 1}$ is the vector of additive noise composed of noises at each sensor.

For an in-depth analysis of antenna array structures, see [11].

2.2 Spectral Methods

2.2.1 Beamformer

The beamformer is the most basic direction finding algorithm. The idea is to "steer" the array in one direction at a time and measure the output power. The steering direction which results in maximum power yields the DOA estimate. The array response is steered by forming a linear combination of the sensor outputs

$$z(t) = \mathbf{w}^* \mathbf{y}(t) \tag{2.8}$$

where $\mathbf{w}$ is the weighting vector. Given the samples z(1),z(2),..,z(N), the output is measured by

$$P(\mathbf{w}) = \frac{1}{N} \sum_{t=1}^{N} \mid z(t) \mid^2 = \frac{1}{N} \sum_{t=1}^{N} \mathbf{w}^* \mathbf{y}(t) \mathbf{y}^*(t) \mathbf{w} = \mathbf{w}^* \hat{\mathbf{R}} \mathbf{w} \tag{2.9}$$

Different beamforming approaches correspond to different choices of weighting vector $\mathbf{w}$.

The advantage of this method is it's simplicity and robustness, while it's drawback is the resolution of closely spaced angles. A review of various beamforming algorithms appears in [6] and the references within.

The conventional (or Bartlett) beamformer is a natural extension to classical Fourier-based spectral analysis to sensor array data. This algorithm maximizes the power of the beamforming output for a given input signal, which is formulated as

$$\begin{aligned} \max_{\mathbf{w}} E\left\{\mathbf{w}^*\mathbf{y}(t)\mathbf{y}(t)^*\mathbf{w}\right\} &= \max_{\mathbf{w}} \mathbf{w}^* E\left\{\mathbf{y}(t)\mathbf{y}(t)^*\right\}\mathbf{w} \\ &= \max_{\mathbf{w}} \left\{E|\mathbf{s}(t)|^2|\mathbf{w}^*\mathbf{a}(\theta)|^2 + \sigma^2|\mathbf{w}|^2\right\} \end{aligned} \tag{2.10}$$

where the assumption of spatially white noise is used. To obtain non-trivial solution, the norm of $\mathbf{w}$ is constrained to $\mathbf{w} = 1$ when carrying out the above maximization. The resulting solution is then

$$\mathbf{w}_{BF} = \frac{\mathbf{a}(\theta)}{(\mathbf{a}^*(\theta)\mathbf{a}(\theta))^{1/2}} \tag{2.11}$$

The above vector can be interpreted as a spatial filter, which has been matched to the impinging signal. Intuitively, the array weighting equalizes the delays (and possibly attenuations) experienced by the signal on various sensors to maximally combine their respective contributions.

Inserting the weighting vector in (2.11) into (2.9), the classical spatial spectrum is obtained

$$P_{BF}(\theta) = \frac{\mathbf{a}^*(\theta)\hat{\mathbf{R}}_{yy}\mathbf{a}(\theta)}{\mathbf{a}^*(\theta)\mathbf{a}(\theta)} \tag{2.12}$$

2.2.2 Subspace Based Algorithms

Subspace-based methods are based on the eigenvector decomposition of the covariance matrix $\mathbf{R}_{yy}$

$$\mathbf{R}_{yy} = E\left\{\mathbf{y}(t)\mathbf{y}^*(t)\right\} \tag{2.13}$$

Using (2.4), we get

$$\mathbf{R}_{\mathbf{yy}} = \mathbf{A}\mathbf{R}_{ss}\mathbf{A}^* + \sigma^2\mathbf{I} = \mathbf{U}\mathbf{\Lambda}\mathbf{U}^* \tag{2.14}$$

where $\mathbf{\Lambda} = diag\{\lambda_1, \lambda_2, \ldots, \lambda_M\}$ is a diagonal matrix of real eigenvalues ordered such that $\lambda_1 \geq \lambda_2 \geq \ldots \lambda_M > 0$. Matrix $\mathbf{U}$ contains the eigenvectors or $\mathbf{R}_{yy}$ corresponding to the eigenvalues in $\mathbf{\Lambda}$. Any vector that is orthogonal to $\mathbf{A}$ is an eigenvector of $\mathbf{R}_{yy}$ with the eigenvalue σ^2. There are $M-K$ linearly independent such vectors. Since the remaining eigenvalues are all larger then σ^2, we can partition the eigenvalue/vector pairs into noise eigenvectors (corresponding to eigenvalues $\lambda_{K+1} = \ldots = \lambda_M$) and signal eigenvectors (corresponding to eigenvalues $\lambda_1, \ldots, \lambda_K$)

$$\mathbf{R}_{yy} = \mathbf{U}_{ss}\mathbf{\Delta}_s\mathbf{U}_{ss}^* + \mathbf{U}_{nn}\mathbf{\Delta}_n\mathbf{U}_{nn}^* \tag{2.15}$$

The simplest of algorithms that are based on the above stated decomposition is the MUSIC (MUltiple SIgnal Classification) algorithm.

Assume $K < M$ signal impinge on the array. Since the eigenvectors in $\mathbf{U}_{nn}$ (the noise eigenvectors) are orthogonal to $\mathbf{A}$, we have

$$\mathbf{U}_{nn}^*\mathbf{a}(\theta) = \mathbf{0}, \quad \theta \in \{\theta_1, \ldots, \theta_K\} \tag{2.16}$$

If the array is unambiguous (i.e. any collection of steering vectors corresponding to distinct DOAs forms a linearly independent set) and $\mathbf{R}_{ss}$ has full rank, then $\mathbf{A}\mathbf{R}_{ss}\mathbf{A}^*$ is also full rank. It then follows that $\{\theta_1, \ldots, \theta_K\}$ are the only possible solutions to the relation in (2.16), which could therefore be used to exactly locate the DOAs.

In practice, an estimate $\hat{\mathbf{R}}_{yy}$ of the covariance matrix is obtained,

$$\hat{\mathbf{R}}_{yy} = \frac{1}{N}\sum_{t=1}^{N}\mathbf{y}(t)\mathbf{y}^*(t) \tag{2.17}$$

and its eigenvectors are separated into the signal and noise eigenvectors as in (2.15). The orthogonal projector onto the noise subspace is estimated as

$$\hat{\mathbf{P}}^{\perp} = \mathbf{U}_{nn}\mathbf{U}_{nn}^{*} \tag{2.18}$$

The MUSIC "spatial spectrum" is then defined as

$$P_M(\theta) = \frac{\mathbf{a}^*(\theta)\mathbf{a}(\theta)}{\mathbf{a}^*(\theta)\hat{\mathbf{P}}^{\perp}\mathbf{a}(\theta)} \tag{2.19}$$

The peaks of this "spectrum" are the angle estimates. In contrast to the beamforming techniques, the MUSIC algorithm provides statistically consistent estimates.

The main drawback of MUSIC is that it breaks down for coherent signals. Though unlikely that one would deliberately transmit two coherent signals from distinct locations, such a phenomenon is not uncommon as either a natural result of a multipath propagation effect, or intentional unfriendly jamming. The end result is a rank deficiency in the source covariance matrix $\mathbf{R}_{ss}$. This, in turn, results in a divergence of a signal eigenvector into the noise subspace. Therefore, in general $\mathbf{U}_{nn}^{*}\mathbf{a}(\theta) \neq \mathbf{0}$ for any θ and the MUSIC "spectrum" may fail to produce peaks at the DOA locations.

2.3 Maximum Likelihood Methods

Perhaps the most well known and frequently used model-based approach in signal processing is the Maximum Likelihood (ML) technique. This methodology requires a statistical framework for the data generation process. While the background and receiver noise in the assumed data model can be thought of as emanating from a large number of independent sources, the same is usually not the case for the emitter signal. It is therefore natural to model the noise as a stationary gaussian white random process whereas the signal waveforms are deterministic. In this section we briefly discuss the conventional ML techniques that model the data as deterministic and unknown process. (This thesis treats a model where signals are assumed to be known, hence ML estimator for such case is extensively analyzed in chapters that follow).

For temporally and spatially white noise, the observation vector $\mathbf{y}(t)$ is a circularly symmetric and temporally white Gaussian random process, with mean $\mathbf{A}(\theta)\mathbf{s}(t)$ and covariance matrix $\sigma^2\mathbf{I}$. The *likelihood function* is the probability density function (PDF) of all observations given the unknown parameters. The PDF of one measurement vector $\mathbf{y}(t)$ is the complex M-variate Gaussian:

$$\frac{1}{(\pi\sigma^2)^M} exp\left\{-\|\mathbf{y}(t) - \mathbf{A}\mathbf{s}(t)\|^2/\sigma^2\right\} \tag{2.20}$$

Since the measurements are independent, the likelihood function is obtained as

$$L(\boldsymbol{\theta}, \mathbf{s}(t), \sigma^2) = \prod_{t=1}^{N} (\pi\sigma^2)^{-M} exp\left\{-\|\mathbf{y}(t) - \mathbf{A}\mathbf{s}(t)\|^2/\sigma^2\right\} \tag{2.21}$$

As indicated above, the unknown parameters in the likelihood function are the signal parameters $\boldsymbol{\theta}$, the signal waveforms $\mathbf{s(t)}$ and the noise variance σ^2. The ML estimates of these unknowns are calculated as the maximizing arguments of $L(\boldsymbol{\theta}, \mathbf{s}(t), \sigma^2)$, the rationale being that these values make the probability of the observation as large as possible. For convenience, the ML estimates are alternatively defined as the minimizing arguments of the negative log-likelihood function $-L(\boldsymbol{\theta}, \mathbf{s}(t), \sigma^2)$. Normalizing by N and ignoring the parameter-independent $Klog\pi$ term, we get:

$$L_{DML}(\boldsymbol{\theta}, \mathbf{s}(t), \sigma^2) = Klog\sigma^2 + \frac{1}{\sigma^2 N} \sum_{t=1}^{N} \|\mathbf{y}(t) - \mathbf{A}\mathbf{s}(t)\|^2 \tag{2.22}$$

whose minimizing arguments are the deterministic (referring to deterministic signals) maximum likelihood (DML) estimates.

It can be shown that explicit minima with respect to σ^2 and $\mathbf{s}(t)$ are given by

$$\hat{\sigma}^2 = \frac{1}{K} Tr\left\{\mathbf{P}_{\mathbf{A}}^{\perp} \hat{\mathbf{R}}_{yy}\right\} \tag{2.23}$$

and

$$\hat{\mathbf{s}}(t) = \mathbf{A}^{\dagger}\mathbf{y}(t) \tag{2.24}$$

where $\hat{\mathbf{R}}_{yy}$ is the sample covariance matrix, $\mathbf{A}^{\dagger}$ is the Moore-Penrose pseudo-inverse of $\mathbf{A}$ and $\mathbf{P}_{\mathbf{A}}^{\perp}$ is the orthogonal projector onto the nullspace of $\mathbf{A}^{*}$, i.e.

$$\hat{\mathbf{R}}_{yy} = \frac{1}{N}\sum_{t=1}^{N}\mathbf{y}(t)\mathbf{y}^{*}(t) \tag{2.25}$$

$$\mathbf{A}^{\dagger} = (\mathbf{A}^{*}\mathbf{A})^{-1}\mathbf{A}^{*} \tag{2.26}$$

$$\mathbf{P}_{\mathbf{A}} = \mathbf{A}\mathbf{A}^{\dagger} \tag{2.27}$$

$$\mathbf{P}_{\mathbf{A}}^{\perp} = \mathbf{I} - \mathbf{P}_{\mathbf{A}} \tag{2.28}$$

Substituting (2.23) and (2.24) into (2.21) shows that the DML signal parameter estimates are obtained by solving the following minimization problem

$$\hat{\boldsymbol{\theta}}_{DML} = arg\left\{\min_{\boldsymbol{\theta}} Tr\left\{\mathbf{P}_{\mathbf{A}}^{\perp}\hat{\mathbf{R}}_{yy}\right\}\right\} \tag{2.29}$$

The interpretation is that the measurements $\mathbf{y}(t)$ are projected onto a model subspace orthogonal to all anticipated signal components, and a power measurement $\frac{1}{N}\sum_{t=1}^{N}\|\mathbf{P}_{\mathbf{A}}^{\perp}\mathbf{y}(y)\|^2 = Tr\left\{\mathbf{P}_{\mathbf{A}}^{\perp}\hat{\mathbf{R}}_{yy}\right\}$ is evaluated. The energy should clearly be the smallest when the projector indeed removes all the true signal components, i.e., when $\theta = \theta_0$. Since only a finite number of noisy samples is available, the energy is not perfectly measured and $\hat{\theta}_{DML}$ will deviate from θ_0. However, if the scenario is stationary, the error will converge to zero as the number of samples is increased to infinity. This remains valid for correlated or even coherent signals, although the accuracy in finite samples is somewhat dependent upon signal correlations. Notice also that (2.29) reduces to Bartlett beamformer in the case of a single source (K=1).

To calculate DML estimates, the non-linear K-dimensional optimization problem must be solved numerically. Finding the signal waveform and the noise estimates is then straightforward, by inserting $\hat{\theta}_{DML}$ into (2.23) and (2.24)

Chapter 3

Literature Overview

The following is a brief overview of previous works on the subject of DOA estimation with known waveforms.

Angle estimation algorithms that incorporate knowledge of the transmitted waveforms were first introduced by Li and Compton in [7]. In this paper maximum likelihood algorithms were derived for multiple known waveforms with known or unknown gains for uniform linear arrays. The received signal was modelled as:

$$\mathbf{s}(t) = \mathbf{\Gamma}\mathbf{x}(t) \tag{3.1}$$

where $\mathbf{\Gamma} = diag\{\gamma_1, \ldots, \gamma_k\}$ is a diagonal matrix of complex gains of the received signals and $\mathbf{x}(t)$ is the known waveform.

They showed that incorporating knowledge of the signal waveform improves the angle estimates and proposed iterational methods that avoid the multidimensional search for a case of multiple signals. Later, Li et.al in [8] analyzed a case of multiple uncorrelated waveforms. They showed that given uncorrelated waveforms, the ML multidimensional search asymptotically decouples to a separate search for each signal which significantly reduces the computational complexity of the algorithm. The waveforms were assumed known up to a complex scaling. The algorithm presented in [8] was extended to handle coherent signals by Caderval and Moses in [4] .The matrix $\mathbf{\Gamma}$ is this case was modelled as :

$$\mathbf{\Gamma} = \begin{bmatrix} \gamma_{11} & \dots & \gamma_{1k_1} & 0 & & \dots & & & 0 \\ 0 & \dots & 0 & \gamma_{21} & \dots & \gamma_{2k_2} & 0 \dots & \dots & 0 \\ \vdots & & & & & & \ddots & \vdots & 0 \\ 0 & & \dots & & & 0 & \gamma_{C1} & \dots & \gamma_{Ck_C} \end{bmatrix} \tag{3.2}$$

Each index $\{k_i\}_{i=1}^{C}$ denotes the (known) number of incoming signals (multipaths) corresponding to the ith source signal $x_i(t)$. To estimate the angles of the coherent signals (which can result from reflections of one signal), an exact ML estimator was used. Although this algorithm was intended to provide a solution to the problem of estimating multipath reflections, no time delay of the multipath components was assumed. An algorithm for estimating the angle of reflections of a *single* known signal in conjunction with time delays was introduced in [18] . The data model used was

$$\mathbf{y}(t) = \sum_{l=1}^{c} \gamma_l \mathbf{a}(\theta_l) s(t - \tau_l) + \mathbf{n}(t) \tag{3.3}$$

They proposed an iterative algorithm for jointly estimating the angle and the delay for a case of a single source. [17] analyzed a similar model for a case of spatially correlated sensors.

None of the works, however, analyzed a case of *multiple* sources, possibly propagating through multiple paths (coherent), each characterized by multiple replicas of the signal impinging on the array with relative time delays (which cause ISI). This is the case analyzed in the following chapters.

Chapter 4

Data Model

This chapter states the data model for the estimation algorithms presented in this thesis. The goal is to describe mathematically the signals and noise received by the antenna array, accounting for the various scenarios of interest.

First, we present a data model for a case of multiple signals, each one propagating through a single path ISI channel. Then, this model is expanded to a case of multiple signals propagating through multiple paths (arriving from different angles), each propagating through a different ISI channel. The transmitted signals are assumed to be known. Channel coefficients and DOAs are assumed to be unknown and deterministic parameters.

4.1 Array Model

Consider an array of M sensors with arbitrary locations and arbitrary directional characteristics. Assume there are K sources (emitters) located at K distinct locations. Signals arriving from each one of the sources impinge on the array. The assembly of the received waveforms form the vector $\mathbf{s}(t)$:

$$\mathbf{s}(t) = [s_1(t) \quad s_2(t) \quad \dots \quad s_K(t)]^T \tag{4.1}$$

The angles of arrival of the impinging signals form the vector $\boldsymbol{\theta}$

$$\boldsymbol{\theta} \ = [\theta_1 \dots \theta_K] \tag{4.2}$$

The array response is described by the steering matrix $\mathbf{A}$ which is a transformation from the signal space of dimension K to the sensor space of dimension M.

$$\mathbf{A}(\boldsymbol{\theta}) = [\mathbf{a}(\theta_1) \dots \mathbf{a}(\theta_K)] \tag{4.3}$$

The columns of this matrix are called "steering vectors". The steering vectors contain the phase shifts of the signals in each of the sensors.

$$\mathbf{a}(\theta_k) = [1 \quad \phi_k \quad \dots \quad \phi_k^{M-1}]^T \tag{4.4}$$

For example, in a case of a uniform linear array (when the sensors are located in a straight line with equal spacings between them) and assuming sensor spacing of half a wavelength ($\lambda/2$), the phase shifts of the received signals are equal to :

$$\phi_k = e^{-j\pi \sin\theta_k} \tag{4.5}$$

The array manifold is assumed to be unambiguous (the steering vectors form a linearly independent set given all possible DOAs). This ensures that the estimates are unique. The signals are assumed to be narrow band, meaning that the bandwidth of the signals is much smaller than its center frequency. In this case the received signals on each sensor of the array differ only by a phase factor.

The array output is denoted by a vector $\mathbf{y}(t) \in C^{M\times 1}$, which in the presence of additive noise is given by the familiar equation

$$\mathbf{y}(t) = \mathbf{A}(\boldsymbol{\theta})\mathbf{s}(t) + \mathbf{n}(t) \tag{4.6}$$

This is the array model equation.

4.2 Transmitted Signal Model

In this section the assumed model of the transmitted signals is described.

As mentioned previously, the DEML algorithm in [8] treats a case of signals that are scaled and phase shifted. This assumption is suitable to a case of flat fading only.

A more general case is a selective fading, which can be modelled by parameterizing the channel as a linear frequency selective filter, applied on the transmitted symbols (which can be equivalently viewed as channel causing ISI). When this filter has only a single tap, the model becomes identical to the one in [8].

A situation where the signal arrives from only one angle but still suffers from ISI, occurs when there is a local scattering in the vicinity of the transmitter. The array receives few reflections of the transmitted signal, but since all reflections are local to the transmitter, the angle of their arrival is *the same* for all reflections. This happens due to the fact that the distance from the transmitter to the receiver is much larger than the distance to the scatterers. This scenario is suitable for a rural environment. For example, imagine that a cellular phone is located inside a house and is received by an antenna positioned on a high mast of a base station a few kilometers away. The only reflections that the signal may contain when it reaches the base station can origin from the walls of the house. When received at the base station, the relative angles of reflections are practically zero, but the relative time delays are not.

Also, there are cases where ISI is inherent in the transmitter due to a band limiting filter.

The unknown channel parameters must be estimated in order to compensate for the distortions caused by the channel, and in this way eliminate the difference between the known transmitted signal and the received signals. This will enable more accurate estimates of the DOA's.

The filter coefficients are assumed to be deterministic and unknown.

In the following sections we present the mathematical framework for the transmitted signal model for single-path and multi-path cases.

4.2.1 Single Path Case

Assume that we have K transmitting information sources $\{x_1(t), \quad x_2(t), \quad \dots \quad x_K(t)\}$. The waveform of each source is assumed to be completely known to the receiver, at each given t.

The signal is modelled as following. Each signal $s_i(t), i = 1..K$ that arrives from angle θ_i is a known waveform $x(t)$ that passes through an unknown linear system, so that the received signal is a superposition of scaled past values of the transmitted signal.

In a digital case, where $t = \tau i, i = 1..N$ (i.e the signal is sampled at a symbol rate), the model can be viewed as a signal that passes through a linear FIR channel which causes Inter-Symbol Interference (ISI) . Thus, we can represent the vector of impinging signals by the following expression

$$\mathbf{s}(t) = \mathbf{H}\tilde{\mathbf{x}}(t) \tag{4.7}$$

H is the "channel" matrix, and has the following structure.

$$H = \begin{pmatrix} {h_1}^1 & \dots & {h_1}^d & 0 & \dots & 0 & & 0 & \dots & 0 \\ 0 & \dots & 0 & {h_2}^1 & \dots & {h_2}^d & & 0 & \dots & 0 \\ \vdots & \dots & \vdots & 0 & \dots & 0 & \ddots & \vdots & \dots & \vdots \\ 0 & \dots & 0 & 0 & \dots & 0 & & {h_K}^1 & \dots & {h_K}^d \end{pmatrix} \tag{4.8}$$

The matrix H is block diagonal, where each block is a row vector, which we define as the "channel" vector $\mathbf{h}_i(t) \in C^{1\times d}, i = 1\dots K$ as :

$$\mathbf{h}_i = \left[h_i^1 \dots h_i^d\right] \tag{4.9}$$

For the sake of convenience of the derivations that follow, the channel vector is defined as a row vector. The matrix H can be then rewritten as :

$$\mathbf{H} = \begin{pmatrix} \mathbf{h_1} & 0 & \dots & 0 \\ 0 & \mathbf{h_2} & & \vdots \\ \vdots & & \ddots & \\ 0 & \dots & & \mathbf{h_K} \end{pmatrix} \tag{4.10}$$

The vector of complex data values (digital symbols) at each t is given by,

$$\tilde{\mathbf{x}}(t) = vec \begin{bmatrix} x_1(t) & x_2(t) & \dots & x_K(t) \\ x_1(t-\tau) & x_2(t-\tau) & \dots & x_K(t-\tau) \\ x_1(t-2\tau) & x_2(t-2\tau) & \dots & x_K(t-2\tau) \\ \vdots & & & \vdots \\ x_1(t-(d-1)\tau) & x_2(t-(d-1)\tau) & \dots & x_K(t-(d-1)\tau) \end{bmatrix} \tag{4.11}$$

Defining a vector $\mathbf{x}_i(t) \in C^{d\times 1}, i = 1 \dots K$ as :

$$\mathbf{x}_i(t) = [x_i(t) \quad x_i(t-\tau) \dots x_i(t-(d-1)\tau)]^T \tag{4.12}$$

we can write :

$$\tilde{\mathbf{x}} = \left[\mathbf{x}_1^T \dots \mathbf{x}_K^T\right]^T \tag{4.13}$$

τ is the delay between the copies of the transmitted waveforms and is assumed to be known. For the digital signal transmission, this is the time spacing between symbols.

4.2.2 Multi-path Case

Next, we consider a situation of several known sources, arriving *each* from several angles.

This situation is more suitable for an urban environment, where in addition to local reflections, there will be reflections from surrounding buildings, multiplying the number of DOAs.

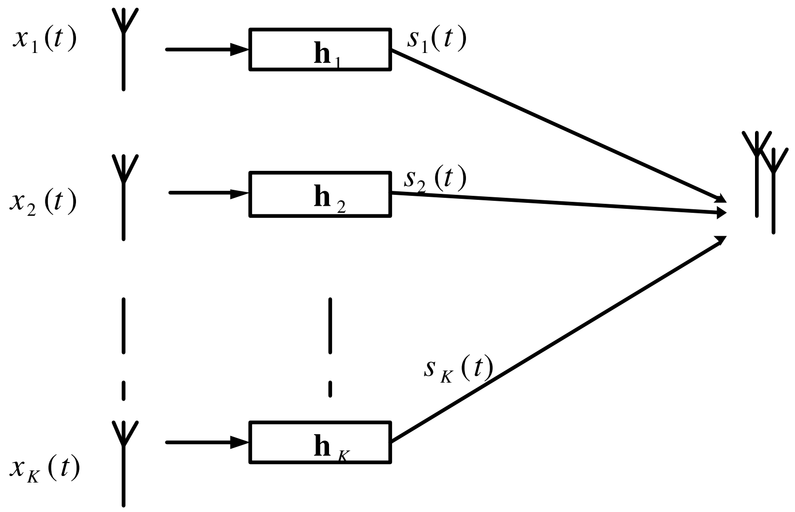

Figure 4.1: **Data model for a single-path case**

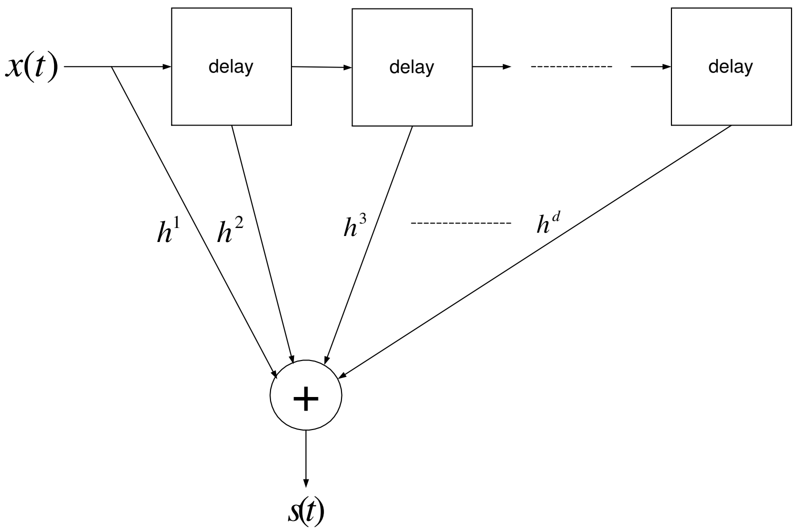

Figure 4.2: **Channel model**

For this case, we will redefine the semantics of the array manifold matrix and the vector of angles:

$$\boldsymbol{\theta} = [\theta_{11}\dots\theta_{1c}, \theta_{21}\dots\theta_{2c}, \dots, \theta_{K1}\dots\theta_{Kc}] \quad (4.14)$$

$$\mathbf{A}(\boldsymbol{\theta}) = [\mathbf{a}(\theta_{11})\dots\mathbf{a}(\theta_{1c}), \mathbf{a}(\theta_{21})\dots\mathbf{a}(\theta_{2c}), \dots, \mathbf{a}(\theta_{K1})\dots\mathbf{a}(\theta_{Kc})] \quad (4.15)$$

In the case of a multipath, the matrix H is still a block-diagonal matrix, but each block instead of a vector is now a '$c \times d$' matrix, where 'c' is the number of paths of the corresponding transmitting source and 'd' is the ISI channel length:

$$H = \begin{pmatrix} {h_{11}}^1 & \dots & {h_{11}}^d & 0 & \dots & 0 & & 0 & \dots & 0 \\ \vdots & \dots & \vdots & 0 & \dots & 0 & & 0 & \dots & 0 \\ {h_{1c}}^1 & \dots & {h_{1c}}^d & 0 & \dots & 0 & & 0 & \dots & 0 \\ 0 & \dots & 0 & {h_{21}}^1 & \dots & {h_{21}}^d & & 0 & \dots & 0 \\ 0 & \dots & 0 & \vdots & \dots & \vdots & & 0 & \dots & 0 \\ 0 & \dots & 0 & {h_{2c}}^1 & \dots & {h_{2c}}^d & & 0 & \dots & 0 \\ \vdots & \dots & \vdots & 0 & \dots & 0 & \ddots & \vdots & \dots & \vdots \\ 0 & \dots & 0 & 0 & \dots & 0 & & {h_{K1}}^1 & \dots & {h_{K1}}^d \\ 0 & \dots & 0 & 0 & \dots & 0 & & \vdots & \dots & \vdots \\ 0 & \dots & 0 & 0 & \dots & 0 & & {h_{Kc}}^1 & \dots & {h_{Kc}}^d \end{pmatrix} \quad (4.16)$$

$$\mathbf{H} = \begin{pmatrix} \mathbf{H_1} & 0 & \dots & 0 \\ 0 & \mathbf{H_2} & & \vdots \\ \vdots & & \ddots & \\ 0 & \dots & & \mathbf{H_K} \end{pmatrix} \quad (4.17)$$

$$\mathbf{H}_i = [\mathbf{h}_{i1}^T \dots \mathbf{h}_{ic}^T]^T \quad (4.18)$$

The number of paths is assumed to be known. For notational convenience it is also assumed to be equal for each source. In general, each signal may have different

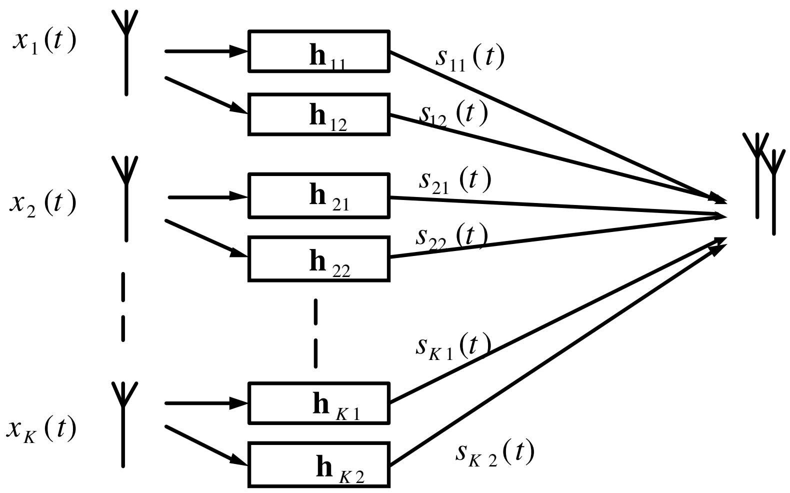

Figure 4.3: **Data model for a multi-path case. Two paths example.**

number of paths which needs to be either known or estimated in order to resolve the DOAs.

4.3 Noise Assumptions

Noise is the cause of the inaccuracy of the DOA estimates. Without noise we would have a set of equations which could be solved explicitly and yield exact estimates (given that the number of equations is not less then the number of parameters). We have two sources of noise. The first is the common and well known receiver noise (which is internal to the sensor array), and the second is an external interference. The receiver noise is temporally white and its distribution is complex normal. It is also "spatially" white, i.e. the noise at each sensor is uncorrelated with the noise on the other sensors of the array. The second source of noise is an interference that is received by the array and not generated internally, thus is not spatially white in general. The effect of such noise can be either modelled by covariance matrix of the sensor array, or in some cases incorporated into the received signal model (i.e interference treated as another signal rather than random additive noise). The latter assumption is used in this work. The additive noise is complex gaussian with mean **0** and covariance matrix **Q**

$$\mathbf{n}(t) \in \mathbf{C}^{Mx1} \sim \mathbf{N}(0, \mathbf{Q}) \tag{4.19}$$

The properties of the additive noise are the following:
property 1 - temporal whiteness

$$\mathbf{E}[\mathbf{n}(t_i)\mathbf{n}^*(t_j)] = \mathbf{Q}\delta_{i,j} \tag{4.20}$$

property 2 - circullar symmetrics

$$\mathbf{E}[\mathbf{n}(t_i)\mathbf{n}^T(t_j)] = 0 \tag{4.21}$$

Throughout this work we mostly deal with spatially white noise. i.e.,

$$\mathbf{Q} = \sigma^2\mathbf{I} \tag{4.22}$$

We address the case of spatially correlated noise in section 5.2.

4.4 Definition of correlation between known signals

Correlation between two random vector processes $\mathbf{x}(t)$ and $\mathbf{y}(t)$ is defined as the mean of their inner product

$$E\{\mathbf{x}(t)\mathbf{y}(t)^*\} \tag{4.23}$$

When the correlation is equal to zero, the random processes are said to be uncorrelated.

The following property holds for random uncorrelated processes:

$$\lim_{N\to\infty} \frac{1}{N} \sum_{t=1}^{N} \mathbf{x}(t)\mathbf{y}^*(t) = 0 \tag{4.24}$$

In our work we address a case of independent signal sources with known waveforms. Since the signals are known, they can not be treated as random, so we can not use the term ”correlation” in the same sense as in (4.23). But, it is clear that for a case of completely independent deterministic waveforms (4.24) is still true. Therefore, in this work we will use the term ”uncorrelated” for known waveforms in the sense of (4.24).

4.5 Notation of time dependent values

The only three time dependent values in this work are the source data sequence, the noise and the measured sensor response. The data that appears at the sensors is collected at N distinct times. Throughout this work we denote the time dependent values of some variable x as:

$$x(t), t = 1..N \tag{4.25}$$

which actually means

$$x(t_n), n = 1..N \tag{4.26}$$

where t_n in (4.26) are the actual times, and t in (4.25) is just an index of the samples. But notation (4.25) allows us a more general representation, which is not restricted to discrete time samples only. Since all the derivations in this work also hold for continuous time, we choose representation (4.25).

4.6 Definition of a ”white data sequence”

We refer to a data sequence as a ”white sequence”, if the following property holds:

$$E\{x(t)x^*(t+\tau)\} = 0 \qquad \tau \neq 0 \tag{4.27}$$

Which means that samples of the sequence taken at different time instances are uncorrelated. When dealing with known data, the samples cannot be treated as random, so the following assumption is used:

$$\lim_{N \longrightarrow \infty} \frac{1}{N} \sum_{t=1}^{N} x(t)x^*(t+\tau) = 0 \qquad \tau \neq 0 \tag{4.28}$$

A case of white sequences is very common in digital communications. Although digital transmission is usually "white" by its nature, often scramblers and interleavers are added to prevent correlation between the transmitted symbols, which allows better synchronization and error correction.

4.7 Estimation problem definition

To summarize, the estimation problem can be stated as follows.
Assuming the model with known source waveforms and unknown channel parameters,

$$\mathbf{y}(t) = \mathbf{A}(\boldsymbol{\theta})\mathbf{H}\tilde{\mathbf{x}}(t) + \mathbf{n}(t) \tag{4.29}$$

our goal is to jointly estimate the DOAs and the channel response based on N observations of the received array vector $\mathbf{y}(t_1), \ldots, \mathbf{y}(t_N)$

Chapter 5

ML channel and DOA estimation algorithm

In this chapter the estimation algorithm is presented. Exact Maximum Likelihood criterion is derived based on the data model and the noise distribution. Computationally efficient approximations to the exact ML are developed as well.

5.1 Spatially Uncorrelated Noise

As a consequence of the statistical assumptions on the noise, the observation vector $\mathbf{y}(t)$ is a circularly symmetric and temporally white Gaussian random process, with mean $\mathbf{A}(\boldsymbol{\theta})\mathbf{H}\tilde{\mathbf{x}}(t)$ and covariance matrix $\mathbf{Q} = \sigma^2\mathbf{I}$. The *likelihood function* is the probability density function (PDF) of all observations given the unknown parameters. The Probability function of a single measurement $\mathbf{y}(t)$ is then a complex M-variate Gaussian and is given by:

$$f(\mathbf{y}(t);\boldsymbol{\theta},\mathbf{H}) = \frac{1}{(2\pi)^M(\frac{\sigma^2}{2})^M} exp\left\{-\frac{1}{\sigma^2}\left[\mathbf{y}(t) - \mathbf{A}(\boldsymbol{\theta})\mathbf{H}\tilde{\mathbf{x}}(t)\right]^*\left[\mathbf{y}(t) - \mathbf{A}(\boldsymbol{\theta})\mathbf{H}\tilde{\mathbf{x}}(t)\right]\right\} \tag{5.1}$$

Since all measurement are independent, the likelihood function is obtained as:

$$L(\mathbf{y}(1)\ldots\mathbf{y}(N);\boldsymbol{\theta},,\mathbf{H})$$
$$=\prod_{t=1}^{N}\frac{1}{(2\pi)^m(\frac{\sigma^2}{2})^m}exp\left\{-\frac{1}{\sigma^2}\left[\mathbf{y}(t)-\mathbf{A}(\boldsymbol{\theta})\mathbf{H}\tilde{\mathbf{x}}(t)\right]^*\left[\mathbf{y}(t)-\mathbf{A}(\boldsymbol{\theta})\mathbf{H}\tilde{\mathbf{x}}(t)\right]\right\} \tag{5.2}$$

The unknown parameters in the likelihood function are the signal's DOAs $\boldsymbol{\theta}$ and the channel matrix $\mathbf{H}$. The Maximum Likelihood estimates of these unknowns are calculated as the maximizing arguments of $L(\mathbf{y}(1)\ldots\mathbf{y}(N);\boldsymbol{\theta},\mathbf{H})$, the rationale being that these values make the probability of the observation as large as possible. For convenience, the ML estimates are alternatively defined as the minimizing arguments of the *negative log-likelihood function*:

$$-\,logL(\mathbf{y}(1)\ldots\mathbf{y}(N);\boldsymbol{\theta},\mathbf{H})=$$
$$=log2\pi MN+log(\frac{\sigma^2}{2})MN+\sum_{t=1}^{N}\left\{-\frac{1}{\sigma^2}\left[\mathbf{y}(t)-\mathbf{A}(\boldsymbol{\theta})\mathbf{H}\tilde{\mathbf{x}}(t)\right]^*\left[\mathbf{y}(t)-\mathbf{A}(\boldsymbol{\theta})\mathbf{H}\tilde{\mathbf{x}}(t)\right]\right\} \tag{5.3}$$

Normalizing by N and ignoring the parameter independent terms, the ML estimates are obtained by minimizing the following cost function:

$$F=\min_{\boldsymbol{\theta},\mathbf{H}}\frac{1}{N}\sum_{t=1}^{N}\left\{\left[\mathbf{y}(t)-\mathbf{A}(\boldsymbol{\theta})\mathbf{H}\tilde{\mathbf{x}}(t)\right]^*\left[\mathbf{y}(t)-\mathbf{A}(\boldsymbol{\theta})\mathbf{H}\tilde{\mathbf{x}}(t)\right]\right\} \tag{5.4}$$

The definitions below will be used in later derivations:

$$\hat{\mathbf{R}}_{y\tilde{x}}=\frac{1}{N}\sum_{n=1}^{N}\mathbf{y}(t)\tilde{\mathbf{x}}^*(t) \tag{5.5}$$

$$\hat{\mathbf{R}}_{\tilde{x}y}=\frac{1}{N}\sum_{n=1}^{N}\tilde{\mathbf{x}}(t)\mathbf{y}^*(t) \tag{5.6}$$

$$\hat{\mathbf{R}}_{\tilde{x}\tilde{x}}=\frac{1}{N}\sum_{n=1}^{N}\tilde{\mathbf{x}}(t)\tilde{\mathbf{x}}^*(t) \tag{5.7}$$

$$\hat{\mathbf{R}}_{yy}=\frac{1}{N}\sum_{n=1}^{N}\mathbf{y}(t)\mathbf{y}^*(t) \tag{5.8}$$

The cost function (5.4) can be expressed in a following form:

$$\begin{aligned}
F &= \min_{\boldsymbol{\theta},\mathbf{H}} \frac{1}{N}\sum_{t=1}^{N}\left\{[\mathbf{y}(t)-\mathbf{A}(\boldsymbol{\theta})\mathbf{H}\tilde{\mathbf{x}}(t)]^{*}\,[\mathbf{y}(t)-\mathbf{A}(\boldsymbol{\theta})\mathbf{H}\tilde{\mathbf{x}}(t)]\right\} \\
&= \min_{\boldsymbol{\theta},\mathbf{H}} \frac{1}{N}\sum_{t=1}^{N} Tr\left\{[\mathbf{y}(t)-\mathbf{A}(\boldsymbol{\theta})\mathbf{H}\tilde{\mathbf{x}}(t)]\,[\mathbf{y}(t)-\mathbf{A}(\boldsymbol{\theta})\mathbf{H}\tilde{\mathbf{x}}(t)]^{*}\right\} \\
&= \min_{\boldsymbol{\theta},\mathbf{H}} \frac{1}{N}\sum_{t=1}^{N} Tr\left\{\hat{\mathbf{R}}_{yy}-\frac{1}{N}\sum_{t=1}^{N}\mathbf{A}(\boldsymbol{\theta})\mathbf{H}\tilde{\mathbf{x}}(t)\mathbf{y}^{*}-\frac{1}{N}\sum_{t=1}^{N}\mathbf{y}\,[\mathbf{A}(\boldsymbol{\theta})\mathbf{H}\tilde{\mathbf{x}}(t)]^{*}\right\} \\
&+ \min_{\boldsymbol{\theta},\mathbf{H}} \frac{1}{N}\sum_{t=1}^{N} Tr\left\{\mathbf{A}(\boldsymbol{\theta})\mathbf{H}\tilde{\mathbf{x}}(t)\,[\mathbf{A}(\boldsymbol{\theta})\mathbf{H}\tilde{\mathbf{x}}(t)]^{*}\right\} \\
&= \min_{\boldsymbol{\theta},\mathbf{H}} \frac{1}{N}\sum_{t=1}^{N} Tr\left\{\hat{\mathbf{R}}_{yy}-\mathbf{A}(\boldsymbol{\theta})\mathbf{H}\hat{\mathbf{R}}_{\tilde{x}y}-\hat{\mathbf{R}}_{y\tilde{x}}\mathbf{H}^{*}\mathbf{A}^{*}(\boldsymbol{\theta})+\mathbf{A}(\boldsymbol{\theta})\mathbf{H}\hat{\mathbf{R}}_{\tilde{x}\tilde{x}}\mathbf{H}^{*}\mathbf{A}^{*}(\boldsymbol{\theta})\right\}
\end{aligned} \tag{5.9}$$

using

$$\hat{\mathbf{R}}_{\tilde{x}y} = \hat{\mathbf{R}}^{*}_{y\tilde{x}} \tag{5.10}$$

and defining:

$$\mathbf{C} = \mathbf{A}(\boldsymbol{\theta})\mathbf{H} \tag{5.11}$$

we get

$$F = Tr\left\{\hat{\mathbf{R}}_{yy}-\mathbf{C}\hat{\mathbf{R}}^{*}_{y\tilde{x}}-\hat{\mathbf{R}}_{y\tilde{x}}\mathbf{C}^{*}+\mathbf{C}\hat{\mathbf{R}}_{\tilde{x}\tilde{x}}\mathbf{C}^{*}\right\} \tag{5.12}$$

which can be rewritten as:

$$F = Tr\left\{\hat{\mathbf{R}}_{yy}-\hat{\mathbf{R}}_{y\tilde{x}}\hat{\mathbf{R}}^{-1}_{\tilde{x}\tilde{x}}\hat{\mathbf{R}}^{*}_{y\tilde{x}}+(\mathbf{C}-\hat{\mathbf{R}}_{y\tilde{x}}\hat{\mathbf{R}}^{-1}_{\tilde{x}\tilde{x}})\hat{\mathbf{R}}_{\tilde{x}\tilde{x}}(\mathbf{C}-\hat{\mathbf{R}}_{y\tilde{x}}\hat{\mathbf{R}}^{-1}_{\tilde{x}\tilde{x}})^{*}\right\} \tag{5.13}$$

The first two terms in (5.13) do not depend on $\mathbf{C}$, and $\hat{\mathbf{R}}_{\tilde{x}\tilde{x}}$ is positive definite. Hence, F is minimized by the choice:

$$\hat{\mathbf{C}} = \hat{\mathbf{R}}_{y\tilde{x}}\hat{\mathbf{R}}^{-1}_{\tilde{x}\tilde{x}} \tag{5.14}$$

The negative log-likelihood cost function becomes:

$$F = Tr\left\{(\mathbf{A}\mathbf{H}-\hat{\mathbf{C}})\hat{\mathbf{R}}_{\tilde{x}\tilde{x}}(\mathbf{A}\mathbf{H}-\hat{\mathbf{C}})^{*}\right\} \tag{5.15}$$

Minimizing F in 5.15 over the space of angles in $\boldsymbol{\theta}$ gives us the exact ML estimator for $\boldsymbol{\theta}$

$$\hat{\boldsymbol{\theta}} = \arg\min_{\boldsymbol{\theta}} Tr\left\{(\mathbf{A}(\boldsymbol{\theta})\mathbf{H} - \hat{\mathbf{C}})\hat{\mathbf{R}}_{\tilde{x}\tilde{x}}(\mathbf{A}(\boldsymbol{\theta})\mathbf{H} - \hat{\mathbf{C}})^*\right\} \tag{5.16}$$

At this point we wish to analyze a case where the signals are uncorrelated with each other. The use of the term "correlation" for known waveforms has been addressed in section (4.4). When dealing with communication signals, it is very reasonable to assume that independent transmitters yeild uncorrelated sequences.

Uncorrelated sequences are defined as follows:

$$\mathbf{R}_{x_i x_j} = \lim_{N\to\infty} \frac{1}{N}\sum_{t=1}^{N} \mathbf{x}_i(t)\mathbf{x}_j^*(t) = 0 \qquad \text{for i} \neq \text{ j} \tag{5.17}$$

We'll use the following approximation:

$$\hat{\mathbf{R}}_{x_i x_j} = \frac{1}{N}\sum_{t=1}^{N} \mathbf{x}_i(t)\mathbf{x}_j^*(t) \cong 0 \qquad \text{for i} \neq \text{ j} \tag{5.18}$$

Although the correlation is zero only as N approaches infinity, we assume that it is negligible for a finite N.

In this case the covariance matrix of the transmitted signal vector takes block diagonal form.

Defining the autocorrelation function of the transmitted sequence as

$$\hat{r}(u) = \frac{1}{N}\sum_{t=1}^{N} \mathbf{x}^*(t)\mathbf{x}(t-u) \tag{5.19}$$

we get,

$$\hat{\mathbf{R}}_{\tilde{x}\tilde{x}} =$$

$$\begin{bmatrix} \hat{r}_{x_1}(0) & \hat{r}_{x_1}(\tau) & \dots & \hat{r}_{x_1}(\tau(d-1)) & & & & \dots & 0 \\ \hat{r}_{x_1}(\tau) & \hat{r}_{x_1}(0) & & \vdots & & & & & \\ \vdots & & \ddots & \vdots & & & & & \vdots \\ \hat{r}_{x_1}(\tau(d-1)) & \dots & \dots & \hat{r}_{x_1}(0) & & & & & \\ & & & & \ddots & & & & \\ & & & & & \hat{r}_{x_K}(0) & \hat{r}_{x_K}(\tau) & \dots & \hat{r}_{x_K}(\tau(d-1)) \\ \vdots & & & & & \hat{r}_{x_K}(\tau) & \hat{r}_{x_K}(0) & & \vdots \\ & & & & & \vdots & & \ddots & \vdots \\ 0 & & \dots & & & \hat{r}_{x_K}(\tau(d-1)) & \dots & \dots & \hat{r}_{x_K}(0) \end{bmatrix} \quad (5.20)$$

which can be written as

$$\hat{\mathbf{R}}_{\tilde{x}\tilde{x}} = \begin{bmatrix} \hat{\mathbf{R}}_{xx_1} & 0 & \dots & 0 \\ 0 & \hat{\mathbf{R}}_{xx_2} & & \vdots \\ \vdots & & \ddots & 0 \\ 0 & \dots & 0 & \hat{\mathbf{R}}_{xx_K} \end{bmatrix} \quad (5.21)$$

where

$$\hat{\mathbf{R}}_{xxi} = \frac{1}{N} \sum_{n=1}^{N} \mathbf{x}_i(t)\mathbf{x}_i^*(t) \quad (5.22)$$

By expanding the expression (5.15), one can see that given the block diagonal structure of $\hat{\mathbf{R}}_{\tilde{x}\tilde{x}}$, each ith element in the resulting K dimensional cost function consists only of factors involving the signal number i. Thus, the minimization of F can be decoupled into K independent minimization problems for each one of the signals, using the knowledge of the corresponding signal.

$$F = Tr\left\{(\mathbf{AH} - \hat{\mathbf{C}})\hat{\mathbf{R}}_{\tilde{x}\tilde{x}}(\mathbf{AH} - \hat{\mathbf{C}})^*\right\} = Tr\{f_1\} + Tr\{f_2\} + \dots + Tr\{f_K\} \quad (5.23)$$

Recall that the angle vector is defined as follows:

$$\mathbf{A}(\boldsymbol{\theta}_i) = [\mathbf{a}(\theta_{i1}), \ldots, \mathbf{a}(\theta_{ic})] \tag{5.24}$$

and

$$\mathbf{A}(\boldsymbol{\theta}) = [\mathbf{A}(\boldsymbol{\theta}_1) \ldots \mathbf{A}(\boldsymbol{\theta}_K)] \tag{5.25}$$

So, given that the sources are uncorrelated, we get separate minimization problems of dimension equal to the number of paths. We have K such minimization problems:

$$\min_{\theta_i, \mathbf{H}_i} f_i = Tr\left\{(\mathbf{A}(\theta_i)\mathbf{H}_i - \hat{\mathbf{C}}_i)\hat{\mathbf{R}}_{xx_i}(\mathbf{A}(\theta_i)\mathbf{H}_i - \hat{\mathbf{C}}_i)^*\right\} \qquad i = 1 \ldots K \tag{5.26}$$

Where $\hat{\mathbf{C}}_i \in C^{M \times d}$ is defined as:

$$\hat{\mathbf{C}}_i = \hat{\mathbf{R}}_{yx_i}\hat{\mathbf{R}}_{xx_i}^{-1} \tag{5.27}$$

and

$$\hat{\mathbf{R}}_{yx_i} = \frac{1}{N}\sum_{t=1}^{N} \mathbf{y}(t)\mathbf{x}_i^*(t) \tag{5.28}$$

and is part of the matrix $\hat{\mathbf{C}}$ as indicated below

$$\hat{\mathbf{C}} = \left[\hat{\mathbf{C}}_1\hat{\mathbf{C}}_2 \ldots \hat{\mathbf{C}}_K\right] \tag{5.29}$$

A matrix $\hat{\mathbf{C}}_i$ is a matrix whose columns are the columns 'i' to '$i+c-1$' of $\hat{\mathbf{C}}$.

We've seen that the exact ML decouples to separate estimation problems for each source. Hence, we continue the analysis referring to a single source case (and dropping the signal index 'i').

Claim:

The channel matrix $\mathbf{H}$ which minimizes f in (5.26) is given by:

$$\hat{\mathbf{H}} = \mathbf{A}^{\dagger}(\boldsymbol{\theta})\hat{\mathbf{C}} \tag{5.30}$$

where $\mathbf{A}^{\dagger} = [\mathbf{A}^*\mathbf{A}]^{-1}\mathbf{A}^*$ is the pseudo-inverse of $\mathbf{A}$.

Proof:

Observe the following property of the trace operator. For any 3 matrices $\mathbf{U}, \mathbf{V}, \mathbf{W}$,

$$Tr\{\mathbf{UVW}\} = Tr\{\mathbf{VWU}\} \tag{5.31}$$

given that the products exist.(proof is straightforward)

Hence, we can write

$$\begin{aligned} &Tr\left\{(\mathbf{A}(\theta)\mathbf{H} - \hat{\mathbf{C}})\hat{\mathbf{R}}_{xx}(\mathbf{A}(\theta)\mathbf{H} - \hat{\mathbf{C}})^*\right\} = \\ &= Tr\left\{\hat{\mathbf{R}}_{xx}(\mathbf{A}(\theta)\mathbf{H} - \hat{\mathbf{C}})^*(\mathbf{A}(\theta)\mathbf{H} - \hat{\mathbf{C}})\right\} \end{aligned} \tag{5.32}$$

and the cost function f can then be expressed in the following form:

$$\begin{aligned} &Tr\left\{\hat{\mathbf{R}}_{xx}\left(\mathbf{AH} - \hat{\mathbf{C}}\right)^*\left(\mathbf{AH} - \hat{\mathbf{C}}\right)\right\} = \\ =&Tr\left\{\hat{\mathbf{R}}_{xx}\mathbf{H}^*\mathbf{A}^*\mathbf{AH} - \hat{\mathbf{R}}_{xx}\hat{\mathbf{C}}^*\mathbf{AH} - \mathbf{R}_{xx}\mathbf{H}^*\mathbf{A}^*\hat{\mathbf{C}} + \hat{\mathbf{R}}_{xx}\hat{\mathbf{C}}^*\hat{\mathbf{C}}\right\} = \\ =&Tr\{\hat{\mathbf{R}}_{xx}\left[\mathbf{H} - (\mathbf{A}^*\mathbf{A})^{-1}\mathbf{A}^*\hat{\mathbf{C}}\right]^*\mathbf{A}^*\mathbf{A}\left[\mathbf{H} - (\mathbf{A}^*\mathbf{A})^{-1}\mathbf{A}^*\hat{\mathbf{C}}\right] + \\ &+ \hat{\mathbf{R}}_{xx}\hat{\mathbf{C}}^*\hat{\mathbf{C}} - \hat{\mathbf{R}}_{xx}\hat{\mathbf{C}}^*\mathbf{A}(\mathbf{A}^*\mathbf{A})^{-1}\mathbf{A}^*\hat{\mathbf{C}}\} \end{aligned} \tag{5.33}$$

It can be shown that $\mathbf{A}^*\mathbf{A}$ and $\hat{\mathbf{R}}_{xx}$ are positive definite matrices. For any given vector $\mathbf{u}$ and $\mathbf{v}$

$$\mathbf{u}^T\mathbf{A}^*\mathbf{Au} = \mathbf{bb}^* > 0 \tag{5.34}$$

$$\sum \mathbf{v}^T\mathbf{xx}^*\mathbf{v} = \sum \mathbf{dd}^* > 0 \tag{5.35}$$

where $\mathbf{b} = \mathbf{u}^T\mathbf{A}^*$ and $\mathbf{d} = \mathbf{v}^T\mathbf{x}$

Since $\mathbf{A}^*\mathbf{A}$ is positive definite and the second and third terms in (5.33) do not depend on $\mathbf{H}$, it's clear that $\mathbf{H}$ is minimized by the choice in (5.30) ▲

Finally, inserting (5.30) into (5.26) we get the following cost function:

$$f(\boldsymbol{\theta}) = tr\left\{\hat{\mathbf{R}}_{xx}[\mathbf{A}(\boldsymbol{\theta})\mathbf{A}(\boldsymbol{\theta})^{\dagger}\hat{\mathbf{C}} - \hat{\mathbf{C}}]^*[\mathbf{A}(\boldsymbol{\theta})\mathbf{A}(\boldsymbol{\theta})^{\dagger}\hat{\mathbf{C}} - \hat{\mathbf{C}}]\right\} \tag{5.36}$$

Which is a minimization problem of dimension equal to c (number of paths).

In summary, the exact Maximum Likelihood estimator of a single source DOA and the Decoupled Maximum Likelihood for a case of multiple sources is obtained by solving the following minimization problem:

$$\begin{aligned}\hat{\boldsymbol{\theta}}_i &= \arg\min_{\boldsymbol{\theta}_i}(Tr\left\{\hat{\mathbf{R}}_{xxi}[\mathbf{P}^{\perp}_{\mathbf{A}_i}\hat{\mathbf{C}}_i]^*[\mathbf{P}^{\perp}_{\mathbf{A}_i}\hat{\mathbf{C}}_i]\right\}) \\ &= \arg\min_{\boldsymbol{\theta}_i}(Tr\left\{[\mathbf{P}^{\perp}_{\mathbf{A}_i}\hat{\mathbf{C}}_i]\hat{\mathbf{R}}_{xxi}[\mathbf{P}^{\perp}_{\mathbf{A}_i}\hat{\mathbf{C}}_i]^*\right\}) \quad i = 1\ldots K\end{aligned} \tag{5.37}$$

where

$$\mathbf{P}^{\perp}_{\mathbf{A}_i} = \mathbf{A}(\theta_i)\mathbf{A}^{\dagger}(\theta_i) - \mathbf{I} = \mathbf{P}_{\mathbf{A}_i} - \mathbf{I} \tag{5.38}$$

In order to get the estimate for the channel, the estimated angle vector $\hat{\boldsymbol{\theta}}_i$ is inserted into (5.30).

The decoupled estimator is a "large sample" approximation to the exact ML. When the number of samples goes to infinity, the correlation, as defined in (4.4), approaches zero and the decoupled ML approaches the exact ML. The decoupling property allows significant improvement in the computational complexity of the algorithm. It lowers the dimension of the search from '$K \times c$' to 'c', for each of the 'K' sources. Unlike the traditional DOA estimation algorithms for unknown waveforms, the decoupled ML estimator does not suffer any degradation when the estimated angles are closely spaced. There is also no demand on the number of array sensors to be greater than the number of signals, unlike some other computationally efficient estimators, such as MUSIC ([13]) for example.

This algorithm can be especially useful in a busy and dense signal environment, where there are many interfering signals. The interferers are uncorrelated with the desired signal, so the DOAs of all its reflections and corresponding channel coefficients can be estimated using (5.37) ignoring the interferers.

When dealing with data communications, the data transmitted is usually independently distributed. We refer to such data sequence as "white" (as explained in greater detail in section 4.6).

For an i.i.d source sequence, $\hat{\mathbf{R}}_{xx_i}$ becomes diagonal for large enough N,

$$\lim_{N\to\infty} \hat{\mathbf{R}}_{xx_i} = \rho\mathbf{I} \tag{5.39}$$

hence we can omit it from the minimization expression and get an approximate, more computationally light expression:

$$\hat{\boldsymbol{\theta}}_i = \arg\min_{\boldsymbol{\theta}_i}(tr\{[\mathbf{P}^{\perp}_{\mathbf{A}_i}\hat{\mathbf{R}}_{yx_i}][\mathbf{P}^{\perp}_{\mathbf{A}_i}\hat{\mathbf{R}}_{yx_i}]^*\}) \tag{5.40}$$

$$\boxed{\hat{\boldsymbol{\theta}}_i = \arg\min_{\boldsymbol{\theta}_i}(\| \mathbf{P}^{\perp}_{\mathbf{A}_i}\hat{\mathbf{R}}_{yx_i} \|^2_F)} \tag{5.41}$$

where $\| \cdot \|_F$ is a "Frobinius norm of a matrix", and is defined as

$$\| \mathbf{F} \|^2_F = tr\{\mathbf{F}\mathbf{F}^*\} \tag{5.42}$$

The above approximation is equal to the exact 1 source ML asymptotically. It is an approximation of the decoupled "large sample" maximum likelihood and is itself "large sample" maximum likelihood, provided that the signal waveform is "white". And as can be seen in the numerical examples, in practice, the approximation is equal to the exact ML for a relatively small N given i.i.d source sequence.

The computational advantage of this approximation is the avoidance of computation and inversion of the signal covariance matrix $\hat{\mathbf{R}}_{xx}$

Single-Path Case

In a single-path case there is only one signal for each source impinging on the array. For this case, the decoupled minimization problems are one dimensional and are given by:

$$\min_{\theta_i,\mathbf{h}_i} Tr\left\{(\mathbf{a}(\theta_i)\mathbf{h}_i - \hat{\mathbf{C}}_i)\hat{\mathbf{R}}_{xx_i}(\mathbf{a}(\theta_i)\mathbf{h}_i - \hat{\mathbf{C}}_i)^*\right\} \qquad i = 1\ldots K \tag{5.43}$$

with

$$\hat{\mathbf{h}}_i = \mathbf{a}^{\dagger}(\theta_i)\hat{\mathbf{R}}_{yx_i}\hat{\mathbf{R}}^{-1}_{xx_i} \quad i = 1,\ldots,K \tag{5.44}$$

where $\mathbf{a}^{\dagger} = [\mathbf{a}^{*}\mathbf{a}]^{-1}\mathbf{a}^{*}$

The exact one source (decoupled) ML DOA estimator is given by:

$$\hat{\theta}_i = \arg\min_{\theta_i}(Tr\{[\mathbf{P}_{\mathbf{a}_i}^{\perp}\hat{\mathbf{C}}_i]\hat{\mathbf{R}}_{xxi}[\mathbf{P}_{\mathbf{a}_i}^{\perp}\hat{\mathbf{C}}_i]^{*}\}) \tag{5.45}$$

where

$$\mathbf{P}_{\mathbf{a}_i} = \mathbf{a}(\theta_i)\mathbf{a}^{\dagger}(\theta_i) \tag{5.46}$$

and

$$\mathbf{P}_{\mathbf{a}_i}^{\perp} = \mathbf{a}(\theta_i)\mathbf{a}^{\dagger}(\theta_i) - \mathbf{I} = \mathbf{P}_{\mathbf{a}_i} - \mathbf{I} \tag{5.47}$$

which is a one-dimensional minimization problem.

The approximate ML using the white sequence assumption is given by:

$$\hat{\theta}_i = \arg\min_{\theta_i}(Tr\{[\mathbf{P}_{\mathbf{a}_i}^{\perp}\hat{\mathbf{C}}_i][\mathbf{P}_{\mathbf{a}_i}^{\perp}\hat{\mathbf{C}}_i]^{*}\}) = \arg\min_{\theta_i} \parallel \mathbf{P}_{\mathbf{a}_i}^{\perp}\hat{\mathbf{R}}_{yxi} \parallel_F^2 \tag{5.48}$$

5.2 Spatially Correlated Noise

When spatial correlation of noise exists, there is a need to estimate the spatial cross correlation matrix $\mathbf{Q}$ in addition to the DOAs and the channel parameters. It can be done in a way similar to the one described in [8] and [17].
In this work we address a case of spatially uncorrelated sensors. And the reason is as follows.

The main cause of spatial correlation is an interfering signal. In our model we show that when the desired signal (the one whose DOA needs to be estimated) is known and is uncorrelated with the interfering signal (which is a reasonable assumption when dealing with digital communication signals), the estimation accuracy degradation due to interference is small. On the other hand, assuming no spatial correlation simplifies the estimation algorithm and makes it computationally lighter than the one with sensor cross-correlation estimated.

5.3 Minimization of the cost function

In order to obtain the DOAs there is a need to find the minimum of the cost functions described above. This process in general is computationally burdensome. For a case of single path, there are only one dimensional searches involved, one for each signal. For a case of multi-path the search has dimension equal to the number of paths. There are many algorithms for finding a minimum of cost function, some are iterational. In general, the algorithms described in [7], [8],[4] and [18] can be applied to find the DOA estimates. The c-dimensional search that is required to estimate the multipath reflections can be solved iteratively using the IQML algorithm that is described in [7] and [3].

Chapter 6

Cramer-Rao Bound

6.1 Introduction

Cramer-Rao Bound is a lower bound on the variance of the estimation error of any unbiased estimator (an estimator is unbiased when the mean of the estimation error is zero). It is important to calculate this bound since it can serve as a theoretical limit for the achievable accuracy of the estimates. The proximity of the estimates to the bound can serve as a benchmark for the estimator performance. An estimation process that produces an error who's variance achieves the bound is called an "efficient estimator".

For a vector of parameters $\boldsymbol{\omega}$, the lower bound on covariance matrix of the estimation error is given by

$$\left[E\{(\hat{\boldsymbol{\omega}}-\boldsymbol{\omega})(\hat{\boldsymbol{\omega}}-\boldsymbol{\omega})^T\}\right]^{-1} \geqslant E\left[\left(\frac{\partial LogL}{\partial \boldsymbol{\omega}}\right)\left(\frac{\partial LogL}{\partial \boldsymbol{\omega}}\right)^T\right] \tag{6.1}$$

The expression on the right hand side of (6.1) is called Fisher Information Matrix (FIM). Bound on the variance of the estimation error for each parameter ω_i is given by the elements on the diagonal of the inverse of FIM.

$$Var(\omega_i-\omega_i) = E\{(\hat{\omega}_i-\omega_i)(\hat{\omega}_i-\omega_i)^T\} \geqslant \left[E\left[\left(\frac{\partial LogL}{\partial \boldsymbol{\omega}}\right)\left(\frac{\partial LogL}{\partial \boldsymbol{\omega}}\right)^T\right]\right]^{-1}_{ii} \tag{6.2}$$

In this chapter we present the computation of the bound for estimating the channel and the DOA parameters. Starting from a case of single path, and expanding it to multipath.
We will show that the exact CRB decouples to separate bounds for each of the sources, when the sources are uncorrelated. We will also develop compact expressions for single source (decoupled) bounds.

All the derivatives of vectors in the following calculations are defined as column vectors.

6.2 Single path case

Define the vector of parameters,

$$\omega = [\bar{\mathbf{h}}_{\mathbf{1}}, \breve{\mathbf{h}}_{\mathbf{1}}, \bar{\mathbf{h}}_{\mathbf{2}}, \breve{\mathbf{h}}_{\mathbf{2}}, \ldots, \bar{\mathbf{h}}_{\mathbf{K}}, \breve{\mathbf{h}}_{\mathbf{K}}, \theta_1, \ldots, \theta_K] \tag{6.3}$$

which consists of the channel vectors (real part of each vector followed by the complex part), followed by the vector of angles.

Recall the signal model and the negative log-likelihood function:

$$\mathbf{y}(t) = \mathbf{A}(\boldsymbol{\theta})\mathbf{H}\tilde{\mathbf{x}}(t) + \mathbf{n}(t) \tag{6.4}$$

$$\begin{aligned} &- logL(\mathbf{y}(1) \ldots \mathbf{y}(N); \boldsymbol{\theta}, \mathbf{H}) \\ &= log2\pi MN + log(\frac{\sigma^2}{2})MN + \sum_{t=1}^{N} \left\{ -\frac{1}{\sigma^2} \left[\mathbf{y}(t) - \mathbf{A}(\boldsymbol{\theta})\mathbf{H}\tilde{\mathbf{x}}(t)\right]^* \left[\mathbf{y}(t) - \mathbf{A}(\boldsymbol{\theta})\mathbf{H}\tilde{\mathbf{x}}(t)\right] \right\} \end{aligned} \tag{6.5}$$

and

$$\mathbf{n}(t) = \mathbf{y}(t) - \mathbf{A}\mathbf{H}\tilde{\mathbf{x}}(t) \tag{6.6}$$

$\mathbf{n}(t) \in C^{M\times 1}$ is the noise array vector. It is a white gaussian process with spacial covariance:

$$E\mathbf{n}(t)\mathbf{n}^*(t) = \eta\mathbf{I} \tag{6.7}$$

and

$$E\mathbf{n}(t)\mathbf{n}^T(s) = 0 \qquad \text{for all t and s} \tag{6.8}$$

Calculation of derivatives with respect to parameters

Derivative of the log-likelihood function with respect to the DOAs

For each angle number i, out of K angles, the derivative of the log-likelihood function is given by

$$\begin{aligned}\frac{\partial LogL}{\partial\theta_i} &= \frac{2}{\eta}\sum_{t=1}^{N} Re\left[(\mathbf{H}\tilde{\mathbf{x}}(t))^*\frac{\partial\mathbf{A}^*}{\partial\theta}_i\mathbf{n}(t)\right] \\ &= \frac{2}{\eta}\sum_{t=1}^{N} Re\left[\mathbf{x}_i(t)^*\mathbf{h}_i^*\mathbf{d}^*(\theta_i)\mathbf{n}(t)\right]\end{aligned} \tag{6.9}$$

where $\mathbf{d}(\theta)$ is the derivative of steering vector by the angle

$$\mathbf{d}(\theta_i) = \frac{\partial\mathbf{a}(\theta_i)}{\partial\theta_i} \tag{6.10}$$

stacking the derivatives of steering vectors into a matrix we get

$$\mathbf{D}(\boldsymbol{\theta}) = [\mathbf{d}(\theta_i)\ldots\mathbf{d}(\theta_K)] \tag{6.11}$$

The derivative by the vector of all angles is equal to

$$\frac{\partial LogL}{\partial\boldsymbol{\theta}} = \frac{2}{\eta}\sum_{t=1}^{N} Re\left[\mathbf{S}^*(t)\mathbf{D}^*(\boldsymbol{\theta})\mathbf{n}(t)\right] \tag{6.12}$$

$$\mathbf{S}(t) = \begin{bmatrix} \mathbf{h_1}\mathbf{x}_1(t) & 0 & \ldots & 0 \\ 0 & \mathbf{h_2}\mathbf{x}_2(t) & & \vdots \\ \vdots & & \ddots & \\ 0 & \ldots & & \mathbf{h_K}\mathbf{x}_K(t) \end{bmatrix} \tag{6.13}$$

where $\mathbf{S}(t)$ is a diagonal matrix consisting of the received signal components.

Derivatives with respect to channel coefficients

Derivative of the log-likelihood function with respect to channel coefficient h_i^j

$$\frac{\partial LogL}{\partial \bar{h}_i^j} = \frac{2}{\eta}\sum_{t=1}^{N} Re\left[\tilde{\mathbf{x}}^*(t)\frac{\partial \mathbf{H}^*}{\partial \bar{h}_i^j}\mathbf{A}(\boldsymbol{\theta})^*\mathbf{n}(t)\right] \tag{6.14}$$

$$= \frac{2}{\eta}\sum_{t=1}^{N} Re\left[x_i^+(t-(j-1)\tau)\mathbf{a}^*(\boldsymbol{\theta})\mathbf{n}(t)\right] \quad j = 1..d \tag{6.15}$$

Assembling the values in (6.15) for each channel coefficient into a vector, we get the derivative by the channel vector:

$$\frac{\partial LogL}{\partial \bar{\mathbf{h}}_i} = \frac{2}{\eta}\sum_{t=1}^{N} Re\left[\mathbf{x}_i^+(t)\mathbf{a}^*(\theta_i)\mathbf{n}(t)\right] \tag{6.16}$$

Similarly, the derivative by the complex part of the channel vector is given by

$$\frac{\partial LogL}{\partial \breve{\mathbf{h}}_i} = \frac{2}{\eta}\sum_{t=1}^{N} Im\left[\mathbf{x}_i^+(t)\mathbf{a}^*(\theta_i)\mathbf{n}(t)\right] \tag{6.17}$$

Derivation of FIM elements

The following properties of complex vectors are essential for later derivations:

$$Re(\mathbf{x})Re(\mathbf{y}^T) = \frac{1}{2}\left[Re(\mathbf{x}\mathbf{y}^T) + Re(\mathbf{x}\mathbf{y}^*)\right] \tag{6.18}$$

$$Im(\mathbf{x})Im(\mathbf{y}^T) = -\frac{1}{2}\left[Re(\mathbf{x}\mathbf{y}^T) - Re(\mathbf{x}\mathbf{y}^*)\right] \tag{6.19}$$

$$Im(\mathbf{x})Re(\mathbf{y}^T) = \frac{1}{2}\left[Im(\mathbf{x}\mathbf{y}^T) + Im(\mathbf{x}\mathbf{y}^*)\right] \tag{6.20}$$

$$Re(\mathbf{x})Im(\mathbf{y}^T) = \frac{1}{2}\left[Im(\mathbf{x}\mathbf{y}^T) - Im(\mathbf{x}\mathbf{y}^*)\right] \tag{6.21}$$

<u>FIM element of the vector of DOAs</u>

$$
\begin{aligned}
\mathbf{J}_{\boldsymbol{\theta}\boldsymbol{\theta}} =& E\left[\frac{\partial LogL}{\partial \boldsymbol{\theta}}\right]\left[\frac{\partial LogL}{\partial \boldsymbol{\theta}}\right]^T = \\
=& E\left(\frac{2}{\eta}\sum_{t=1}^{N} Re\left[\mathbf{S}^*(t)\mathbf{D}^*(\boldsymbol{\theta})\mathbf{n}(t)\right]\frac{2}{\eta}\sum_{t=1}^{N} Re\left[\mathbf{S}^*(t)\mathbf{D}^*(\boldsymbol{\theta})\mathbf{n}(t)\right]^T\right) \\
=& \frac{4}{\eta^2}\frac{1}{2}\sum_{t=1}^{N} Re\left[\mathbf{S}^*(t)\mathbf{D}^*(\boldsymbol{\theta})\mathbf{n}(t)\mathbf{n}^*(t)\mathbf{D}(\boldsymbol{\theta})\mathbf{S}(t)\right] \\
=& \frac{2}{\eta}\sum_{t=1}^{N} Re\left[\mathbf{S}^*(t)\mathbf{D}^*(\boldsymbol{\theta})\mathbf{D}(\boldsymbol{\theta})\mathbf{S}(t)\right] \triangleq \boldsymbol{\Gamma}
\end{aligned}
\tag{6.22}
$$

To go from $2nd$ line in (6.22) to the $3rd$ line, the equalities (6.8) and (6.18) were used. To go from $3rd$ to $4th$ line, (6.7) was used.

In the derivations that follow the same technique is utilized to calculate the remaining FIM elements.

<u>FIM element for the channel vector</u>

$$
\begin{aligned}
\mathbf{J}_{\bar{\mathbf{h}}_i\bar{\mathbf{h}}_j} =& E\left[\frac{\partial LogL}{\partial \bar{\mathbf{h}}_i}\right]\left[\frac{\partial LogL}{\partial \bar{\mathbf{h}}_j}\right]^T = \\
=& E\left(\frac{2}{\eta}\sum_{t=1}^{N} Re\left[\mathbf{x}_i^+(t)\mathbf{a}^*(\theta_i)\mathbf{n}(t)\right]\frac{2}{\eta}\sum_{t=1}^{N} Re\left[\mathbf{x}_j^+(t)\mathbf{a}^*(\theta_j)\mathbf{n}(t)\right]^T\right) \\
=& \frac{4}{\eta^2}\frac{1}{2}\sum_{t=1}^{N}\sum_{s=1}^{N} Re\left[E\left\{\mathbf{x}_i^+(t)\mathbf{a}^*(\theta_i)\mathbf{n}(t)\mathbf{n}^*(s)\mathbf{a}(\theta_j)\mathbf{x}_j^T(s)\right\}\right] \\
=& \frac{2}{\eta}\sum_{t=1}^{N} Re\left[\mathbf{x}_i^+(t)\mathbf{a}^*(\theta_i)\mathbf{a}(\theta_j)\mathbf{x}_j^T(t)\right] \triangleq \bar{\boldsymbol{\Phi}}_{ij} \quad i,j = 1\ldots K
\end{aligned}
\tag{6.23}
$$

$$\begin{aligned}
\mathbf{J}_{\breve{\mathbf{h}}_i\breve{\mathbf{h}}_j} =& E\left[\frac{\partial LogL}{\partial \breve{\mathbf{h}}_i}\right]\left[\frac{\partial LogL}{\partial \breve{\mathbf{h}}_j}\right]^T = \\
=& E\left(\frac{2}{\eta}\sum_{t=1}^{N} Im\left[\mathbf{x}_i^{+}(t)\mathbf{a}^*(\theta_i)\mathbf{n}(t)\right]\frac{2}{\eta}\sum_{t=1}^{N} Im\left[\mathbf{x}_j^{+}(t)\mathbf{a}^*(\theta_j)\mathbf{n}(t)\right]^T\right) \\
=& \frac{4}{\eta^2}\frac{1}{2}\sum_{t=1}^{N}\sum_{s=1}^{N} Re\left[E\left\{\mathbf{x}_i^{+}(t)\mathbf{a}^*(\theta_i)\mathbf{n}(t)\mathbf{n}^*(s)\mathbf{a}(\theta_j)\mathbf{x}_j^T(s)\right\}\right] \\
=& \frac{2}{\eta}\sum_{t=1}^{N} Re\left[\mathbf{x}_i^{+}(t)\mathbf{a}^*(\theta_i)\mathbf{a}(\theta_j)\mathbf{x}_j^T(t)\right] \triangleq \bar{\mathbf{\Phi}}_{ij} \quad i,j = 1\ldots K
\end{aligned} \tag{6.24}$$

$$\begin{aligned}
\mathbf{J}_{\bar{\mathbf{h}}_i\breve{\mathbf{h}}_j} =& E\left[\frac{\partial LogL}{\partial \bar{\mathbf{h}}_i}\right]\left[\frac{\partial LogL}{\partial \breve{\mathbf{h}}_j}\right]^T = \\
=& E\left(\frac{2}{\eta}\sum_{t=1}^{N} Re\left[\mathbf{x}_i^{+}(t)\mathbf{a}^*(\theta_i)\mathbf{n}(t)\right]\frac{2}{\eta}\sum_{t=1}^{N} Im\left[\mathbf{x}_j^{+}(t)\mathbf{a}^*(\theta_j)\mathbf{n}(t)\right]^T\right) \\
=& -\frac{4}{\eta^2}\frac{1}{2}\sum_{t=1}^{N}\sum_{s=1}^{N} Im\left[E\left\{\mathbf{x}_i^{+}(t)\mathbf{a}^*(\theta_i)\mathbf{n}(t)\mathbf{n}^*(s)\mathbf{a}(\theta_j)\mathbf{x}_j^T(s)\right\}\right] \\
=& -\frac{2}{\eta}\sum_{t=1}^{N} Im\left[\mathbf{x}_i^{+}(t)\mathbf{a}^*(\theta_i)\mathbf{a}(\theta_j)\mathbf{x}_j^T(t)\right] \triangleq -\breve{\mathbf{\Phi}}_{ij} \quad i,j = 1\ldots K
\end{aligned} \tag{6.25}$$

$$\begin{aligned}
\mathbf{J}_{\breve{\mathbf{h}}_i\bar{\mathbf{h}}_j} =& E\left[\frac{\partial LogL}{\partial \breve{\mathbf{h}}_i}\right]\left[\frac{\partial LogL}{\partial \bar{\mathbf{h}}_j}\right]^T \\
=& E\left(\frac{2}{\eta}\sum_{t=1}^{N} Im\left[\mathbf{x}_i^{+}(t)\mathbf{a}^*(\theta_i)\mathbf{n}(t)\right]\frac{2}{\eta}\sum_{t=1}^{N} Re\left[\mathbf{x}_j^{+}(t)\mathbf{a}^*(\theta_j)\mathbf{n}(t)\right]^T\right) \\
=& \frac{4}{\eta^2}\frac{1}{2}\sum_{t=1}^{N}\sum_{s=1}^{N} Im\left[E\left\{\mathbf{x}_i^{+}(t)\mathbf{a}^*(\theta_i)\mathbf{n}(t)\mathbf{n}^*(s)\mathbf{a}(\theta_j)\mathbf{x}_j^T(s)\right\}\right] \\
=& \frac{2}{\eta}\sum_{t=1}^{N} Im\left[\mathbf{x}_i^{+}(t)\mathbf{a}^*(\theta_i)\mathbf{a}(\theta_j)\mathbf{x}_j^T(t)\right] \triangleq \breve{\mathbf{\Phi}}_{ij} \quad i,j = 1\ldots K
\end{aligned} \tag{6.26}$$

$$\begin{aligned}
\mathbf{J}_{\bar{\mathbf{h}}_i\boldsymbol{\theta}} =& E\left[\frac{\partial LogL}{\partial \bar{\mathbf{h}}_i}\right]\left[\frac{\partial LogL}{\partial \boldsymbol{\theta}}\right]^T \\
=& E\left(\frac{2}{\eta}\sum_{t=1}^{N} Re\left[\mathbf{x}_i^{+}(t)\mathbf{a}^*(\theta_i)\mathbf{n}(t)\right]\frac{2}{\eta}\sum_{t=1}^{N} Re\left[\mathbf{S}^*(t)\mathbf{D}^*(\boldsymbol{\theta})\mathbf{n}(t)\right]^T\right) = \\
=& \frac{4}{\eta^2}\frac{1}{2}\sum_{t=1}^{N}\sum_{s=1}^{N} Re\left[E\left\{\mathbf{x}_i^{+}(t)\mathbf{a}^*(\theta_i)\mathbf{n}(t)\mathbf{n}^*(s)\mathbf{D}(\boldsymbol{\theta})\mathbf{S}(s)\right\}\right] \\
=& \frac{2}{\eta}\sum_{t=1}^{N} Re\left[\mathbf{x}_i^{+}(t)\mathbf{a}^*(\theta_i)\mathbf{D}(\boldsymbol{\theta})\mathbf{S}(t)\right] \triangleq \bar{\boldsymbol{\Delta}}_i \quad i = 1\ldots K
\end{aligned} \tag{6.27}$$

$$\begin{aligned}
\mathbf{J}_{\breve{\mathbf{h}}_i\boldsymbol{\theta}} =& E\left[\frac{\partial LogL}{\partial \breve{\mathbf{h}}_i}\right]\left[\frac{\partial LogL}{\partial \boldsymbol{\theta}}\right]^T = \\
=& E\left(\frac{2}{\eta}\sum_{t=1}^{N} Im\left[\mathbf{x}_i^{+}(t)\mathbf{a}^*(\theta_i)\mathbf{n}(t)\right]\frac{2}{\eta}\sum_{t=1}^{N} Re\left[\mathbf{S}^*(t)\mathbf{D}^*(\boldsymbol{\theta})\mathbf{n}(t)\right]^T\right) = \\
=& \frac{4}{\eta^2}\frac{1}{2}\sum_{t=1}^{N}\sum_{s=1}^{N} Im\left[E\left\{\mathbf{x}_i^{+}(t)\mathbf{a}^*(\theta_i)\mathbf{n}(t)\mathbf{n}^*(s)\mathbf{D}(\boldsymbol{\theta})\mathbf{S}(s)\right\}\right] \\
=& \frac{2}{\eta}\sum_{t=1}^{N} Im\left[\mathbf{x}_i^{+}(t)\mathbf{a}^*(\theta_i)\mathbf{D}(\boldsymbol{\theta})\mathbf{S}(t)\right] \triangleq \breve{\boldsymbol{\Delta}}_i \quad i = 1\ldots K
\end{aligned} \tag{6.28}$$

Similarly

$$\begin{aligned}
\mathbf{J}_{\boldsymbol{\theta}\bar{\mathbf{h}}_i} =& E\left[\frac{\partial LogL}{\partial \boldsymbol{\theta}}\right]\left[\frac{\partial LogL}{\partial \bar{\mathbf{h}}_i}\right]^T \\
=& E\left(\frac{2}{\eta}\sum_{t=1}^{N} Re\left[\mathbf{S}^*(t)\mathbf{D}^*(\boldsymbol{\theta})\mathbf{n}(t)\right]\frac{2}{\eta}\sum_{t=1}^{N} Re\left[\mathbf{x}_i^{+}(t)\mathbf{a}^*(\theta_i)\mathbf{n}(t)\right]^T\right) = \\
=& \frac{4}{\eta^2}\frac{1}{2}\sum_{t=1}^{N}\sum_{s=1}^{N} Re\left[E\left\{\mathbf{S}^*(\boldsymbol{\theta})\mathbf{D}^*\mathbf{n}(t)\mathbf{n}^*(s)\mathbf{a}(\theta_i)\mathbf{x}_i^T(t)\right\}\right] \\
=& \frac{2}{\eta}\sum_{t=1}^{N} Re\left[\mathbf{S}^*(\boldsymbol{\theta})\mathbf{D}^*\mathbf{a}(\theta_i)\mathbf{x}_i^T(t)\right] \triangleq \bar{\boldsymbol{\Delta}}_i^T \quad i = 1\ldots K
\end{aligned} \tag{6.29}$$

$$\begin{aligned}\mathbf{J}_{\boldsymbol{\theta}\breve{\mathbf{h}}_i} =& E\left[\frac{\partial LogL}{\partial \boldsymbol{\theta}}\right]\left[\frac{\partial LogL}{\partial \breve{\mathbf{h}}_i}\right]^T \\ =& E\left(\frac{2}{\eta}\sum_{t=1}^{N} Re\left[\mathbf{S}^*(t)\mathbf{D}^*(\boldsymbol{\theta})\mathbf{n}(t)\right]\frac{2}{\eta}\sum_{t=1}^{N} Im\left[\mathbf{x}_i^+(t)\mathbf{a}^*(\theta_i)\mathbf{n}(t)\right]^T\right) = \\ =& -\frac{4}{\eta^2}\frac{1}{2}\sum_{t=1}^{N}\sum_{s=1}^{N} Im\left[E\left\{\mathbf{S}^*(\boldsymbol{\theta})\mathbf{D}^*\mathbf{n}(t)\mathbf{n}^*(s)\mathbf{a}(\theta_i)\mathbf{x}_i^T(t)\right\}\right] \\ =& -\frac{2}{\eta}\sum_{t=1}^{N} Im\left[\mathbf{S}^*(\boldsymbol{\theta})\mathbf{D}^*\mathbf{a}(\theta_i)\mathbf{x}_i^T(t)\right] \triangleq \breve{\boldsymbol{\Delta}}_i^T \quad i = 1\dots K\end{aligned} \tag{6.30}$$

Using the definitions in (6.23) - (6.30) we can present the Fisher Information Matrix for the vectors of parameters $\omega = [\bar{\mathbf{h_1}}, \breve{\mathbf{h_1}}, \bar{\mathbf{h_2}}, \breve{\mathbf{h_2}}\dots, \bar{\mathbf{h_K}}, \breve{\mathbf{h_K}}\theta_1, \dots, \theta_K]$ in the following form

$$FIM = \begin{bmatrix} \bar{\boldsymbol{\Phi}}_{11} & -\breve{\boldsymbol{\Phi}}_{11} & \dots & \bar{\boldsymbol{\Phi}}_{K1} & -\breve{\boldsymbol{\Phi}}_{K1} & \bar{\boldsymbol{\Delta}}_1 \\ \breve{\boldsymbol{\Phi}}_{11} & \bar{\boldsymbol{\Phi}}_{11} & \dots & \breve{\boldsymbol{\Phi}}_{K1} & \bar{\boldsymbol{\Phi}}_{K1} & \breve{\boldsymbol{\Delta}}_1 \\ \vdots & \vdots & \ddots & \vdots & \vdots & \vdots \\ \bar{\boldsymbol{\Phi}}_{1K} & -\breve{\boldsymbol{\Phi}}_{1K} & \dots & \bar{\boldsymbol{\Phi}}_{KK} & -\breve{\boldsymbol{\Phi}}_{KK} & \bar{\boldsymbol{\Delta}}_K \\ \breve{\boldsymbol{\Phi}}_{1K} & \bar{\boldsymbol{\Phi}}_{1K} & \dots & \breve{\boldsymbol{\Phi}}_{KK} & \bar{\boldsymbol{\Phi}}_{KK} & \breve{\boldsymbol{\Delta}}_K \\ \bar{\boldsymbol{\Delta}}_1^T & \breve{\boldsymbol{\Delta}}_1^T & \dots & \bar{\boldsymbol{\Delta}}_K^T & \breve{\boldsymbol{\Delta}}_K^T & \boldsymbol{\Gamma} \end{bmatrix} \tag{6.31}$$

partitioning to blocks we have:

$$FIM = \begin{bmatrix} \boldsymbol{\Phi} & \boldsymbol{\Delta} \\ \boldsymbol{\Delta}^T & \boldsymbol{\Gamma} \end{bmatrix} \tag{6.32}$$

Where

$$\boldsymbol{\Phi} = \begin{bmatrix} \bar{\boldsymbol{\Phi}}_{11} & -\breve{\boldsymbol{\Phi}}_{11} & \dots & \bar{\boldsymbol{\Phi}}_{K1} & -\breve{\boldsymbol{\Phi}}_{K1} \\ \breve{\boldsymbol{\Phi}}_{11} & \bar{\boldsymbol{\Phi}}_{11} & \dots & \breve{\boldsymbol{\Phi}}_{K1} & \bar{\boldsymbol{\Phi}}_{K1} \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ \bar{\boldsymbol{\Phi}}_{1K} & -\breve{\boldsymbol{\Phi}}_{1K} & \dots & \bar{\boldsymbol{\Phi}}_{KK} & -\breve{\boldsymbol{\Phi}}_{KK} \\ \breve{\boldsymbol{\Phi}}_{1K} & \bar{\boldsymbol{\Phi}}_{1K} & \dots & \breve{\boldsymbol{\Phi}}_{KK} & \bar{\boldsymbol{\Phi}}_{KK} \end{bmatrix} \tag{6.33}$$

and

$$\boldsymbol{\Delta} = \begin{bmatrix} \bar{\boldsymbol{\Delta}}_1 \\ \breve{\boldsymbol{\Delta}}_1 \\ \vdots \\ \bar{\boldsymbol{\Delta}}_K \\ \breve{\boldsymbol{\Delta}}_K \end{bmatrix} \tag{6.34}$$

Then, using the inverse of a partitioned matrix lemma, we get the CRB for the vector of DOAs:

$$CRB(\boldsymbol{\theta}) = \left(\boldsymbol{\Gamma} - \boldsymbol{\Delta}^T \boldsymbol{\Phi}^{-1} \boldsymbol{\Delta}\right)^{-1} \tag{6.35}$$

This is the exact Cramer-Rao bound for estimating the angles of arrival of K signals with known waveforms propagating through unknown channels.

Uncorrelated signals

Next we treat a case of uncorrelated signals (as defined in section 4.4), which is a common case in digital communication. Using the following approximation:

$$\hat{\mathbf{R}}_{x_i x_j} = \frac{1}{N} \sum_{t=1}^{N} \mathbf{x}(t)\mathbf{x}^*(t) \cong 0 \qquad \text{for i} \neq \text{ j} \tag{6.36}$$

we get:

$$\boldsymbol{\Phi}_{ij} \cong 0 \qquad for \quad i \neq j \tag{6.37}$$

$$\boldsymbol{\Gamma} = \frac{2}{\eta} \sum_{t=1}^{N} Re\left[\mathbf{S}^*(t)\mathbf{D}^*(\boldsymbol{\theta})\mathbf{D}(\boldsymbol{\theta})\mathbf{S}(t)\right] \cong diag([\gamma_1 \dots \gamma_K]) \tag{6.38}$$

where

$$\gamma_i = \frac{2}{\eta} \sum_{t=1}^{N} Re\left[\mathbf{x}_i^*(t)\mathbf{h}_i^*\mathbf{d}^*(\theta_i)\mathbf{d}(\theta_i)\mathbf{h}_i\mathbf{x}_i(t)\right] \quad i = 1 \dots K \tag{6.39}$$

Also

$$\bar{\boldsymbol{\Delta}}_i \cong [0 \dots \bar{\boldsymbol{\delta}}_i \dots 0] \tag{6.40}$$

$$\breve{\boldsymbol{\Delta}}_i \cong [0 \dots \breve{\boldsymbol{\delta}}_i \dots 0] \tag{6.41}$$

$\bar{\mathbf{\Delta}}_i$ and $\breve{\mathbf{\Delta}}_i$ are matrices with all columns equal to zero except a column number i, which is equal to the vectors $\breve{\boldsymbol{\delta}}_i$ and $\bar{\boldsymbol{\delta}}_i$. Where :

$$\bar{\boldsymbol{\delta}}_i = \frac{2}{\eta} \sum_{t=1}^{N} Re\left[\mathbf{x}_i^{+}(t)\mathbf{a}^{*}(\theta_i)\mathbf{d}(\theta_i)\mathbf{h}_i\mathbf{x}_i(t)\right] \tag{6.42}$$

$$\breve{\boldsymbol{\delta}}_i = \frac{2}{\eta} \sum_{t=1}^{N} Im\left[\mathbf{x}_i^{+}(t)\mathbf{a}^{*}(\theta_i)\mathbf{d}(\theta_i)\mathbf{h}_i\mathbf{x}_i(t)\right] \tag{6.43}$$

The FIM matrix of the parameter vector $\omega = [\bar{\mathbf{h_1}}, \breve{\mathbf{h_1}}, \bar{\mathbf{h_2}}, \breve{\mathbf{h_2}} \ldots, \bar{\mathbf{h_K}}, \breve{\mathbf{h_K}}\theta_1, \ldots, \theta_K]$ for uncorrelated signals takes the following form:

$$FIM \cong \begin{bmatrix} \bar{\mathbf{\Phi}}_{11} & -\breve{\mathbf{\Phi}}_{11} & \ldots & 0 & 0 & \bar{\boldsymbol{\delta}}_1 & \ldots & 0 \\ \breve{\mathbf{\Phi}}_{11} & \bar{\mathbf{\Phi}}_{11} & \ldots & 0 & 0 & \breve{\boldsymbol{\delta}}_1 & \ldots & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \ldots & \bar{\mathbf{\Phi}}_{KK} & -\breve{\mathbf{\Phi}}_{KK} & 0 & \ldots & \bar{\boldsymbol{\delta}}_K \\ 0 & 0 & \ldots & \breve{\mathbf{\Phi}}_{KK} & \bar{\mathbf{\Phi}}_{KK} & 0 & \ldots & \breve{\boldsymbol{\delta}}_K \\ \bar{\boldsymbol{\delta}}_1^T & \breve{\boldsymbol{\delta}}_1^T & \ldots & 0 & 0 & \gamma_1 & \ldots & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \ldots & \bar{\boldsymbol{\delta}}_K^T & \breve{\boldsymbol{\delta}}_K^T & 0 & \ldots & \gamma_K \end{bmatrix} \tag{6.44}$$

The Cramer Rao bound on the variance of an error for estimating the ith angle θ_i $(i = 1 \ldots K)$, is equal to the element in FIM^{-1} with the same coordinates as γ_i in FIM.

One can see from the structure of FIM, that the CRBs of θ_i are decoupled one from another. Note that all the matrices in 6.35 are block diagonal, and that the element in FIM^{-1} with coordinates of γ_i depends only on $\mathbf{h}_i$ and θ_i.

Hence, for uncorrelated signals, the CRB for each source asymptotically decouples to a one source CRB, which is given by

$$CRB(\theta_i) = \left(\gamma_i - [\bar{\boldsymbol{\delta}}_i^T \quad \breve{\boldsymbol{\delta}}_i^T]\tilde{\boldsymbol{\Phi}}_i^{-1}[\bar{\boldsymbol{\delta}}_i^T \quad \breve{\boldsymbol{\delta}}_i^T]^T\right)^{-1} \tag{6.45}$$

where

$$\tilde{\boldsymbol{\Phi}}_i = \begin{bmatrix} \bar{\boldsymbol{\Phi}}_i & -\breve{\boldsymbol{\Phi}}_i \\ \breve{\boldsymbol{\Phi}}_i & \bar{\boldsymbol{\Phi}}_i \end{bmatrix} \tag{6.46}$$

Claim

$$\tilde{\boldsymbol{\Phi}}_i^{-1} = \begin{bmatrix} \bar{\boldsymbol{\Phi}}_i & -\breve{\boldsymbol{\Phi}}_i \\ \breve{\boldsymbol{\Phi}}_i & \bar{\boldsymbol{\Phi}}_i \end{bmatrix}^{-1} = \begin{bmatrix} \bar{\boldsymbol{\Phi}}_i^{-1} & -\breve{\boldsymbol{\Phi}}_i^{-1} \\ \breve{\boldsymbol{\Phi}}_i^{-1} & \bar{\boldsymbol{\Phi}}_i^{-1} \end{bmatrix} \tag{6.47}$$

proof

The equality above can be equivalently rewritten as

$$\bar{\boldsymbol{\Phi}}_i\bar{\boldsymbol{\Phi}}_i^{-1} - \breve{\boldsymbol{\Phi}}_i\breve{\boldsymbol{\Phi}}_i^{-1} = \mathbf{I} \tag{6.48}$$

$$\bar{\boldsymbol{\Phi}}_i\breve{\boldsymbol{\Phi}}_i^{-1} - \bar{\boldsymbol{\Phi}}_i\breve{\boldsymbol{\Phi}}_i = \mathbf{0} \tag{6.49}$$

which certainly must hold since

$$\begin{aligned} \mathbf{I} = \tilde{\boldsymbol{\Phi}}_i\tilde{\boldsymbol{\Phi}}_i^{-1} &= (\bar{\boldsymbol{\Phi}}_i + j\breve{\boldsymbol{\Phi}}_i)(\bar{\boldsymbol{\Phi}}_i^{-1} + j\breve{\boldsymbol{\Phi}}_i^{-1}) \\ =&(\bar{\boldsymbol{\Phi}}_i\bar{\boldsymbol{\Phi}}_i^{-1} - \breve{\boldsymbol{\Phi}}_i\breve{\boldsymbol{\Phi}}_i^{-1}) + j(\bar{\boldsymbol{\Phi}}_i\breve{\boldsymbol{\Phi}}_i^{-1} - \bar{\boldsymbol{\Phi}}_i\breve{\boldsymbol{\Phi}}_i) \end{aligned} \tag{6.50}$$

▲

Also,

$$\begin{aligned} &\begin{bmatrix} \bar{\boldsymbol{\Phi}}_i^{-1} & -\breve{\boldsymbol{\Phi}}_i^{-1} \\ \breve{\boldsymbol{\Phi}}_i^{-1} & \bar{\boldsymbol{\Phi}}_i^{-1} \end{bmatrix} \begin{bmatrix} \bar{\boldsymbol{\delta}}_i \\ \breve{\boldsymbol{\delta}}_i \end{bmatrix} = \\ &= \begin{bmatrix} \bar{\boldsymbol{\Phi}}_i^{-1}\bar{\boldsymbol{\delta}}_i - \breve{\boldsymbol{\Phi}}_i^{-1}\breve{\boldsymbol{\delta}}_i & \breve{\boldsymbol{\Phi}}_i^{-1}\bar{\boldsymbol{\delta}}_i + \bar{\boldsymbol{\Phi}}_i^{-1}\breve{\boldsymbol{\delta}}_i \end{bmatrix} = \begin{bmatrix} Re\{\boldsymbol{\Phi}_i^{-1}\boldsymbol{\delta}_i\} & Im\{\boldsymbol{\Phi}_i^{-1}\boldsymbol{\delta}_i\} \end{bmatrix} \end{aligned} \tag{6.51}$$

and

$$\begin{bmatrix} \bar{\boldsymbol{\delta}}_i^T & \breve{\boldsymbol{\delta}}_i^T \end{bmatrix} \begin{bmatrix} Re\{\boldsymbol{\Phi}_i^{-1}\boldsymbol{\delta}_i\} & Im\{\boldsymbol{\Phi}_i^{-1}\boldsymbol{\delta}_i\} \end{bmatrix} = Re\{\boldsymbol{\delta}_i^*\boldsymbol{\Phi}_i^{-1}\boldsymbol{\delta}_i\} \tag{6.52}$$

Hence,

$$\begin{bmatrix} \bar{\boldsymbol{\delta}}_i^T & \breve{\boldsymbol{\delta}}_i^T \end{bmatrix} \tilde{\boldsymbol{\Phi}}_i^{-1} \begin{bmatrix} \bar{\boldsymbol{\delta}}_i^T & \breve{\boldsymbol{\delta}}_i^T \end{bmatrix}^T = Re\{\boldsymbol{\delta}_i^*\boldsymbol{\Phi}_i^{-1}\boldsymbol{\delta}_i\} \tag{6.53}$$

Finally, the Cramer Rao bound for estimating the DOA of signal number i, given that it's uncorrelated with the other signals is asymptotically equal to:

$$CRB(\theta_i) = \left(\gamma_i - Re\{\boldsymbol{\delta}_i^* \boldsymbol{\Phi}_i^{-1} \boldsymbol{\delta}_i\}\right)^{-1} \tag{6.54}$$

Which is a CRB for a single signal. As the number of samples goes to infinity, the decoupled bound approaches the exact bound. We've seen in chapter 5 that Maximum Likelihood criterion for multiple uncorrelated signals also decouples for uncorrelated signals.

We can further simplify the expressions for $\boldsymbol{\delta}$, γ and $\boldsymbol{\Phi}$:

$$\begin{aligned}\boldsymbol{\delta} =& \frac{2}{\eta}\sum_{t=1}^{N}\left[\mathbf{x}(t)\mathbf{a}^*(\theta)\mathbf{d}(\theta_i)\mathbf{h}\mathbf{x}(t)\right]\\ =& \frac{2}{\eta}\mathbf{a}^*(\theta)\mathbf{d}(\theta)\sum_{t=1}^{N}\mathbf{x}^+(t)\mathbf{h}\mathbf{x}(t)\\ =& \frac{2N}{\eta}\mathbf{a}^*(\theta)\mathbf{d}(\theta)\hat{\mathbf{R}}_{xx}^T\mathbf{h}^T\end{aligned} \tag{6.55}$$

$$\begin{aligned}\gamma =& \frac{2}{\eta}\sum_{t=1}^{N} Re\left[\mathbf{x}^*(t)\mathbf{h}^*\mathbf{d}^*(\theta)\mathbf{d}(\theta)\mathbf{h}\mathbf{x}(t)\right]\\ =& \frac{2}{\eta}\mathbf{d}^*(\theta)\mathbf{d}(\theta)\sum_{t=1}^{N} Re\left[\mathbf{x}^*(t)\mathbf{h}^*\mathbf{h}\mathbf{x}(t)\right] =\\ =& \frac{2}{\eta}\mathbf{d}^*(\theta)\mathbf{d}(\theta)\sum_{t=1}^{N} Re\left[\mathbf{h}\mathbf{x}(t)\mathbf{x}^*(t)\mathbf{h}^*\right]\\ =& \frac{2N}{\eta}\|\mathbf{d}(\theta)\|^2\mathbf{h}\hat{\mathbf{R}}_{xx}\mathbf{h}^*\end{aligned} \tag{6.56}$$

$$\begin{aligned}\boldsymbol{\Phi} =& \frac{2}{\eta}\sum_{t=1}^{N}\left[\mathbf{x}^+(t)\mathbf{a}^*(\theta)\mathbf{a}(\theta)\mathbf{x}^T(t)\right] =\\ =& \frac{2}{\eta}\mathbf{a}^*(\theta)\mathbf{a}(\theta)\sum_{t=1}^{N}\left[\mathbf{x}^+(t)\mathbf{x}^T(t)\right] = \frac{2N}{\eta}\|\mathbf{a}(\theta)\|^2\hat{\mathbf{R}}_{xx}^T\end{aligned} \tag{6.57}$$

Then

$$
\begin{aligned}
&Re\left[\boldsymbol{\delta}\boldsymbol{\Phi}^{-1}\boldsymbol{\delta}^*\right] = \\
=&Re\left[\left[\frac{2N}{\eta}\mathbf{a}^*(\theta)\mathbf{d}(\theta)\hat{\mathbf{R}}_{xx_i}^T\mathbf{h}_i^T\right]^*\left[\frac{2}{\eta}\|\mathbf{a}(\theta)\|^2\hat{\mathbf{R}}_{xx}^T\right]^{-1}\left[\frac{2N}{\eta}\mathbf{a}^*(\theta)\mathbf{d}(\theta)\hat{\mathbf{R}}_{xx_i}^T\mathbf{h}^T\right]\right] \\
=&Re\left[\frac{2N}{\eta}\frac{|\mathbf{a}^*(\theta)\mathbf{d}(\theta)|^2}{\|\mathbf{a}(\theta)\|^2}\left[\hat{\mathbf{R}}_{xx}^T\mathbf{h}^T\right]^*\left[\hat{\mathbf{R}}_{xx}^T\right]^{-1}\left[\hat{\mathbf{R}}_{xx_i}^T\mathbf{h}^T\right]\right]
\end{aligned}
\tag{6.58}
$$

since $\hat{\mathbf{R}}_{xx}^* = \hat{\mathbf{R}}_{xx}$, we can write

$$
\begin{aligned}
&\left[\hat{\mathbf{R}}_{xx}^T\mathbf{h}^T\right]^*\left[\hat{\mathbf{R}}_{xx}^T\right]^{-1}\left[\hat{\mathbf{R}}_{xx_i}^T\mathbf{h}^T\right] \\
=&\left[\hat{\mathbf{R}}_{xx}^T\mathbf{h}^T\right]^*\mathbf{h}^T = \mathbf{h}\left[\left[\hat{\mathbf{R}}_{xx}^T\mathbf{h}^T\right]^*\right]^T \\
=&\mathbf{h}\left[\mathbf{h}\hat{\mathbf{R}}_{xx}\right]^* = \mathbf{h}\hat{\mathbf{R}}_{xx}\mathbf{h}^*
\end{aligned}
\tag{6.59}
$$

we have

$$
Re\left[\boldsymbol{\delta}\boldsymbol{\Phi}^{-1}\boldsymbol{\delta}^*\right] = \frac{2N}{\eta}\frac{|\mathbf{a}^*(\theta)\mathbf{d}(\theta)|^2}{\|\mathbf{a}(\theta)\|^2}\mathbf{h}\hat{\mathbf{R}}_{xx}\mathbf{h}^* \tag{6.60}
$$

Hence,

$$
\begin{aligned}
CRB^{-1}(\theta) =&\gamma - Re\left[\boldsymbol{\delta}\boldsymbol{\Phi}^{-1}\boldsymbol{\delta}^*\right] = \\
=&\frac{2N}{\eta}\|\mathbf{d}(\theta_i)\|^2\mathbf{h}\hat{\mathbf{R}}_{xx}\mathbf{h}^* - \frac{2N\rho}{\eta}\frac{|\mathbf{a}^*(\theta)\mathbf{d}(\theta)|^2}{\|\mathbf{a}(\theta)\|^2}\mathbf{h}\hat{\mathbf{R}}_{xx}\mathbf{h}^* \\
=&\frac{2N}{\eta}\mathbf{h}\hat{\mathbf{R}}_{xx}\mathbf{h}^*\left[\frac{\|\mathbf{a}(\theta)\|^2\|\mathbf{d}(\theta)\|^2 - |\mathbf{a}^*(\theta)\mathbf{d}(\theta)|^2}{\|\mathbf{a}(\theta)\|^2}\right]
\end{aligned}
\tag{6.61}
$$

The CRB for one-source case and the asymptotic CRB for a case of multiple uncorrelated sources with no multipath, is given by

$$
\boxed{CRB(\theta) = \frac{\eta}{2N}\left[\frac{\|\mathbf{a}(\theta)\|^2}{\mathbf{h}\hat{\mathbf{R}}_{xx}\mathbf{h}^*\left[\|\mathbf{a}(\theta)\|^2\|\mathbf{d}(\theta)\|^2 - |\mathbf{a}^*(\theta)\mathbf{d}(\theta)|^2\right]}\right]} \tag{6.62}
$$

6.2.1 "White sequences"

For a case of white data sequences (as defined in section 4.6) where

$$\lim_{N\to\infty} \hat{\mathbf{R}}_{xx_i} = \rho\mathbf{I} \tag{6.63}$$

we get the following expression:

$$\boxed{CRB(\theta) = \frac{1}{2\cdot SNR\cdot N}\left[\frac{\|\mathbf{a}(\theta)\|^2}{\|\mathbf{h}\|^2\left[\|\mathbf{a}(\theta)\|^2\|\mathbf{d}(\theta)\|^2 - |\mathbf{a}^*(\theta)\mathbf{d}(\theta)|^2\right]}\right]} \tag{6.64}$$

Where the SNR is defined as:

$$SNR = \frac{\rho}{\eta} \tag{6.65}$$

This bound is equal asymptotically to a single source CRB.

6.3 Multipath case

In this section, the CRB for a case of multiple paths of the same source is derived.

In the derivations that follow we use the same guidelines as for single path case, and extend it to multiple paths.

In contrast to a single path case, we have several AOAs and channels for each transmitting source. Hence, the parameter vector for this case is redefined as:

$$\omega = [\bar{\mathbf{h}}_{\mathbf{11}}, \breve{\mathbf{h}}_{\mathbf{11}}, \ldots, \bar{\mathbf{h}}_{\mathbf{1c}}, \breve{\mathbf{h}}_{\mathbf{1c}}, \bar{\mathbf{h}}_{\mathbf{21}}, \breve{\mathbf{h}}_{\mathbf{21}}, \ldots, \bar{\mathbf{h}}_{\mathbf{2c}}, \breve{\mathbf{h}}_{\mathbf{2c}}, \ldots, \bar{\mathbf{h}}_{\mathbf{K1}}, \breve{\mathbf{h}}_{\mathbf{K1}}, \ldots, \bar{\mathbf{h}}_{\mathbf{Kc}}, \breve{\mathbf{h}}_{\mathbf{Kc}}, \boldsymbol{\theta}_1^T, \ldots, \boldsymbol{\theta}_K^T] \tag{6.66}$$

where c is the number of paths for each source. The angle vector is redefined as follows:

$$\mathbf{A}(\boldsymbol{\theta}_i) = [\mathbf{a}(\theta_{i1}), \ldots, \mathbf{a}(\theta_{ic})] \tag{6.67}$$

with

$$\mathbf{A}(\boldsymbol{\theta}) = [\mathbf{A}(\boldsymbol{\theta}_1)\ldots\mathbf{A}(\boldsymbol{\theta}_K)] \tag{6.68}$$

and

$$\mathbf{D}(\boldsymbol{\theta}_i) = [\mathbf{d}(\theta_{i1}), \ldots, \mathbf{d}(\theta_{ic})] \tag{6.69}$$

where:

$$\mathbf{D}(\boldsymbol{\theta}) = [\mathbf{D}(\boldsymbol{\theta}_1)\ldots\mathbf{D}(\boldsymbol{\theta}_K)] \tag{6.70}$$

and

$$\mathbf{d}(\theta_i) = \frac{\partial \mathbf{a}(\theta_i)}{\partial \theta_i} \tag{6.71}$$

Calculation of derivatives with respect to parameters

Derivative of the log-likelihood function with respect to the DOAs

$$\begin{aligned} \frac{\partial LogL}{\partial \theta_{ij}} &== \frac{2}{\eta}\sum_{t=1}^{N} Re\left[(\mathbf{H}\tilde{\mathbf{x}}(t))^{*}\frac{\partial \mathbf{A}^{*}}{\partial \theta\ _i}\mathbf{n}(t)\right] \\ &= \frac{2}{\eta}\sum_{t=1}^{N} Re\left[\mathbf{x}_i(t)^{*}\mathbf{h}_i^{*}\mathbf{d}^{*}(\theta_{ij})\mathbf{n}(t)\right] \end{aligned} \tag{6.72}$$

where $i = 1\ldots K$ is the index for sources and $j = 1\ldots c$ is the index for reflections (paths).

Define a diagonal matrix consisting of the received signal components

$$\mathbf{S}_i(t) = \begin{bmatrix} \mathbf{h_i^1}\mathbf{x}_i(t) & 0 & \ldots & 0 \\ 0 & \mathbf{h_i^2}\mathbf{x}_i(t) & & \vdots \\ \vdots & & \ddots & \\ 0 & \ldots & & \mathbf{h_i^c}\mathbf{x}_i(t) \end{bmatrix} \tag{6.73}$$

Then,

$$\frac{\partial LogL}{\partial \boldsymbol{\theta}_i} = \frac{2}{\eta}\sum_{t=1}^{N} Re\left[\mathbf{S}_i^{*}(t)\mathbf{D}^{*}(\boldsymbol{\theta}_i)\mathbf{n}(t)\right] \tag{6.74}$$

Derivative of the log-likelihood function with respect channel coefficients

$$\frac{\partial LogL}{\partial \bar{h_ij}^{k}} = \frac{2}{\eta}\sum_{t=1}^{N} Re\left[\tilde{\mathbf{x}}^{*}(t)\frac{\partial \mathbf{H}^{*}}{\partial \bar{h}_i^{k}}\mathbf{A}(\boldsymbol{\theta})^{*}\mathbf{n}(t)\right] \tag{6.75}$$

$$= \frac{2}{\eta}\sum_{t=1}^{N} Re\left[x_i^{+}(t-(k-1)\tau)\mathbf{a}^*(\boldsymbol{\theta}_{ij})\mathbf{n}(t)\right] \quad k = 1..d \tag{6.76}$$

Assembling to get the derivative by the channel vector, we get:

$$\frac{\partial LogL}{\partial \bar{\mathbf{h}}_{ij}} = \frac{2}{\eta}\sum_{t=1}^{N} Re\left[\mathbf{x}_i^{+}(t)\mathbf{a}^*(\theta_{ij})\mathbf{n}(t)\right] \tag{6.77}$$

Similarly

$$\frac{\partial LogL}{\partial \bar{\mathbf{h}}_{ij}} = \frac{2}{\eta}\sum_{t=1}^{N} Im\left[\mathbf{x}_i^{+}(t)\mathbf{a}^*(\theta_{ij})\mathbf{n}(t)\right] \tag{6.78}$$

Calculation of FIM elements

The FIM elements are calculated using the same technique as in previous section (single path case).

We redefine $\mathbf{S}(t)$ as block diagonal matrix, where$\mathbf{S}_i(t)$ is defined in (6.79)

$$\mathbf{S}(t) = \begin{bmatrix} \mathbf{S}_1(t) & 0 & \dots & 0 \\ 0 & \mathbf{S}_2(t) & & \vdots \\ \vdots & & \ddots & \\ 0 & \dots & & \mathbf{S}_K(t) \end{bmatrix} \tag{6.79}$$

Then

$$\begin{aligned} \mathbf{J}_{\boldsymbol{\theta}\boldsymbol{\theta}} =& E\left[\frac{\partial LogL}{\partial \boldsymbol{\theta}}\right]\left[\frac{\partial LogL}{\partial \boldsymbol{\theta}}\right]^T = \\ =& \frac{2}{\eta}\sum_{t=1}^{N} Re\left[\mathbf{S}^*(t)\mathbf{D}^*(\boldsymbol{\theta})\mathbf{D}(\boldsymbol{\theta})\mathbf{S}(t)\right] \triangleq \boldsymbol{\Gamma} \end{aligned} \tag{6.80}$$

Where $\boldsymbol{\theta} = [\boldsymbol{\theta}_1 \dots \boldsymbol{\theta}_c]$, and

$$\begin{aligned} \mathbf{J}_{\boldsymbol{\theta}_i\boldsymbol{\theta}_j} =& E\left[\frac{\partial LogL}{\partial \boldsymbol{\theta}_i}\right]\left[\frac{\partial LogL}{\partial \boldsymbol{\theta}_j}\right]^T = \\ =& \frac{2}{\eta}\sum_{t=1}^{N} Re\left[\mathbf{S}_i^*(t)\mathbf{D}^*(\boldsymbol{\theta}_i)\mathbf{D}(\boldsymbol{\theta}_j)\mathbf{S}_j(t)\right] \triangleq \boldsymbol{\Gamma}_{ij} \qquad i,j = 1\dots K \end{aligned} \tag{6.81}$$

FIM element for the channel vector:

$$\mathbf{J}_{\bar{\mathbf{h}}_{iu}\bar{\mathbf{h}}_{jv}} = E\left[\frac{\partial LogL}{\partial \bar{\mathbf{h}}_{iu}}\right]\left[\frac{\partial LogL}{\partial \bar{\mathbf{h}}_{jv}}\right]^T = \\ = \frac{2}{\eta}\sum_{t=1}^{N} Re\left[\mathbf{x}_i^+(t)\mathbf{a}^*(\theta_{iu})\mathbf{a}(\theta_{iv})\mathbf{x}_j^T(t)\right] \triangleq \bar{\mathbf{\Phi}}_{u,v}^{ij} \quad i,j=1\dots K \quad u,v=1\dots c \tag{6.82}$$

$$\mathbf{J}_{\breve{\mathbf{h}}_{iu}\breve{\mathbf{h}}_{jv}} = E\left[\frac{\partial LogL}{\partial \breve{\mathbf{h}}_{iu}}\right]\left[\frac{\partial LogL}{\partial \breve{\mathbf{h}}_{jv}}\right]^T = \\ = \frac{2}{\eta}\sum_{t=1}^{N} Re\left[\mathbf{x}_i^+(t)\mathbf{a}^*(\theta_{iu})\mathbf{a}(\theta_{iv})\mathbf{x}_j^T(t)\right] \triangleq \bar{\mathbf{\Phi}}_{u,v}^{ij} \quad i,j=1\dots K \quad u,v=1\dots c \tag{6.83}$$

$$\mathbf{J}_{\bar{\mathbf{h}}_{iu}\breve{\mathbf{h}}_{jv}} = E\left[\frac{\partial LogL}{\partial \bar{\mathbf{h}}_{iu}}\right]\left[\frac{\partial LogL}{\partial \breve{\mathbf{h}}_{jv}}\right]^T = \\ = -\frac{2}{\eta}\sum_{t=1}^{N} Im\left[\mathbf{x}_i^+(t)\mathbf{a}^*(\theta_{iu})\mathbf{a}(\theta_{iv})\mathbf{x}_j^T(t)\right] \triangleq -\breve{\mathbf{\Phi}}_{u,v}^{ij} \quad i,j=1\dots K \quad u,v=1\dots c \tag{6.84}$$

$$\mathbf{J}_{\breve{\mathbf{h}}_{iu}\bar{\mathbf{h}}_{jv}} = E\left[\frac{\partial LogL}{\partial \breve{\mathbf{h}}_{iu}}\right]\left[\frac{\partial LogL}{\partial \bar{\mathbf{h}}_{jv}}\right]^T = \\ = \frac{2}{\eta}\sum_{t=1}^{N} Im\left[\mathbf{x}_i^+(t)\mathbf{a}^*(\theta_{iu})\mathbf{a}(\theta_{iv})\mathbf{x}_j^T(t)\right] \triangleq \breve{\mathbf{\Phi}}_{u,v}^{ij} \quad i,j=1\dots K \quad u,v=1\dots c \tag{6.85}$$

$$\mathbf{J}_{\bar{\mathbf{h}}_{iu}\boldsymbol{\theta}_j} = E\left[\frac{\partial LogL}{\partial \bar{\mathbf{h}}_{i}}\right]\left[\frac{\partial LogL}{\partial \boldsymbol{\theta}_j}\right]^T \\ = \frac{2}{\eta}\sum_{t=1}^{N} Re\left[\mathbf{x}_i^+(t)\mathbf{a}^*(\theta_{iu})\mathbf{D}(\boldsymbol{\theta}_j)\mathbf{S}_j(t)\right] \triangleq \bar{\mathbf{\Delta}}_{u}^{ij} \quad i,j=1\dots K, \quad u=1\dots c \tag{6.86}$$

$$\mathbf{J}_{\breve{\mathbf{h}}_{iu}\boldsymbol{\theta}_j} = E\left[\frac{\partial LogL}{\partial \breve{\mathbf{h}}_{i}}\right]\left[\frac{\partial LogL}{\partial \boldsymbol{\theta}_j}\right]^T \\ = \frac{2}{\eta}\sum_{t=1}^{N} Im\left[\mathbf{x}_i^+(t)\mathbf{a}^*(\theta_{iu})\mathbf{D}(\boldsymbol{\theta}_j)\mathbf{S}_j(t)\right] \triangleq \breve{\mathbf{\Delta}}_{u}^{ij} \quad i,j=1\dots K, \quad u=1\dots c \tag{6.87}$$

Similarly

$$\mathbf{J}_{\boldsymbol{\theta}_j \bar{\mathbf{h}}_{iu}} = E\left[\frac{\partial LogL}{\partial \boldsymbol{\theta}_j}\right]\left[\frac{\partial LogL}{\partial \bar{\mathbf{h}}_{iu}}\right]^T$$
$$= \frac{2}{\eta}\sum_{t=1}^{N} Re\left[\mathbf{S}_j^*(t)\mathbf{D}^*(\boldsymbol{\theta}_j)\mathbf{a}(\theta_{iu})\mathbf{x}_i^T(t)\right] \triangleq \{\bar{\boldsymbol{\Delta}}_u^{ij}\}^T \quad i,j = 1\ldots K, \quad u = 1\ldots c \tag{6.88}$$

$$\mathbf{J}_{\boldsymbol{\theta}_j \breve{\mathbf{h}}_{iu}} = E\left[\frac{\partial LogL}{\partial \boldsymbol{\theta}_j}\right]\left[\frac{\partial LogL}{\partial \breve{\mathbf{h}}_{iu}}\right]^T$$
$$= -\frac{2}{\eta}\sum_{t=1}^{N} Im\left[\mathbf{S}_j^*(t)\mathbf{D}^*(\boldsymbol{\theta}_j)\mathbf{a}(\theta_{iu})\mathbf{x}_i^T(t)\right] \triangleq \{\breve{\boldsymbol{\Delta}}_u^{ij}\}^T \quad i,j = 1\ldots K, \quad u = 1\ldots c \tag{6.89}$$

The FIM for the vector of parameters $[\bar{\mathbf{h_1}}, \breve{\mathbf{h_1}}, \bar{\mathbf{h_2}}, \breve{\mathbf{h_2}} \ldots, \bar{\mathbf{h_K}}, \breve{\mathbf{h_K}}, \boldsymbol{\theta}_1^T, \ldots, \boldsymbol{\theta}_K^T]$ takes the following form:

$$FIM = \begin{bmatrix} \bar{\boldsymbol{\Phi}}_{11} & -\breve{\boldsymbol{\Phi}}_{11} & \ldots & \bar{\boldsymbol{\Phi}}_{K1} & -\breve{\boldsymbol{\Phi}}_{K1} & \bar{\boldsymbol{\Delta}}_{11} & \ldots & \bar{\boldsymbol{\Delta}}_{1K} \\ \breve{\boldsymbol{\Phi}}_{11} & \bar{\boldsymbol{\Phi}}_{11} & \ldots & \breve{\boldsymbol{\Phi}}_{K1} & \bar{\boldsymbol{\Phi}}_{K1} & \breve{\boldsymbol{\Delta}}_{11} & \ldots & \breve{\boldsymbol{\Delta}}_{1K} \\ \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & & \\ \bar{\boldsymbol{\Phi}}_{1K} & -\breve{\boldsymbol{\Phi}}_{1K} & \ldots & \bar{\boldsymbol{\Phi}}_{KK} & -\breve{\boldsymbol{\Phi}}_{KK} & \bar{\boldsymbol{\Delta}}_{K1} & \ldots & \bar{\boldsymbol{\Delta}}_{KK} \\ \breve{\boldsymbol{\Phi}}_{1K} & \bar{\boldsymbol{\Phi}}_{1K} & \ldots & \breve{\boldsymbol{\Phi}}_{KK} & \bar{\boldsymbol{\Phi}}_{KK} & \breve{\boldsymbol{\Delta}}_{K1} & \ldots & \breve{\boldsymbol{\Delta}}_{KK} \\ \bar{\boldsymbol{\Delta}}_{11}^T & \breve{\boldsymbol{\Delta}}_{11}^T & \ldots & \bar{\boldsymbol{\Delta}}_{K1}^T & \breve{\boldsymbol{\Delta}}_{K1}^T & \boldsymbol{\Gamma}_{11} & \ldots & \boldsymbol{\Gamma}_{1K} \\ \vdots & \vdots & & \vdots & \vdots & \vdots & & \vdots \\ \bar{\boldsymbol{\Delta}}_{1K}^T & \breve{\boldsymbol{\Delta}}_{1K}^T & \ldots & \bar{\boldsymbol{\Delta}}_{KK}^T & \breve{\boldsymbol{\Delta}}_{KK}^T & \boldsymbol{\Gamma}_{K1} & \ldots & \boldsymbol{\Gamma}_{KK} \end{bmatrix} \tag{6.90}$$

Let's partition the FIM in (6.90)to

$$FIM = \begin{bmatrix} \boldsymbol{\Phi} & \boldsymbol{\Delta}^T \\ \boldsymbol{\Delta} & \boldsymbol{\Gamma} \end{bmatrix} \tag{6.91}$$

Where $\boldsymbol{\Phi}$ is given by

$$\mathbf{\Phi} = \begin{bmatrix} \bar{\mathbf{\Phi}}_{11} & -\breve{\mathbf{\Phi}}_{11} & \dots & \bar{\mathbf{\Phi}}_{K1} & -\breve{\mathbf{\Phi}}_{K1} \\ \breve{\mathbf{\Phi}}_{11} & \bar{\mathbf{\Phi}}_{11} & \dots & \breve{\mathbf{\Phi}}_{K1} & \bar{\mathbf{\Phi}}_{K1} \\ \vdots & \vdots & \ddots & & \\ \bar{\mathbf{\Phi}}_{1K} & -\breve{\mathbf{\Phi}}_{1K} & \dots & \bar{\mathbf{\Phi}}_{KK} & -\breve{\mathbf{\Phi}}_{KK} \\ \breve{\mathbf{\Phi}}_{1K} & \bar{\mathbf{\Phi}}_{1K} & \dots & \breve{\mathbf{\Phi}}_{KK} & \bar{\mathbf{\Phi}}_{KK} \end{bmatrix} \tag{6.92}$$

Each matrix $\mathbf{\Phi}_{ij}$ in 6.92 is given by

$$\mathbf{\Phi}_{ij} = \begin{bmatrix} \mathbf{\Phi}^{ij}_{1,1} & \dots & \mathbf{\Phi}^{ij}_{1,c} \\ & \dots & \\ \mathbf{\Phi}^{ij}_{c,1} & \dots & \mathbf{\Phi}^{ij}_{c,c} \end{bmatrix} \tag{6.93}$$

where each element in $\mathbf{\Phi}_{ij}$ is given by

$$\mathbf{\Phi}^{ij}_{u,v} = \frac{2}{\eta} \sum_{t=1}^{N} \left[\mathbf{x}_i^{+}(t) \mathbf{a}^*(\theta_{iu}) \mathbf{a}(\theta_{iv}) \mathbf{x}_j^T(t) \right] \quad i,j = 1 \dots K \quad u,v = 1 \dots c \tag{6.94}$$

it follows that $\mathbf{\Phi}_{ij}$ can be written as

$$\mathbf{\Phi}_{ij} = \frac{2N}{\eta} \left[\mathbf{A}^*(\boldsymbol{\theta}_i) \mathbf{A}(\boldsymbol{\theta}_j) \right] \otimes \hat{\mathbf{R}}^T_{x_{ij}} \tag{6.95}$$

Where $\otimes$ is the "kronecker product" of matrices.

$\mathbf{\Delta}$ in (6.91) is given by

$$\mathbf{\Delta} = \begin{bmatrix} \bar{\mathbf{\Delta}}_{11} & \dots & \bar{\mathbf{\Delta}}_{1K} \\ \breve{\mathbf{\Delta}}_{11} & \dots & \breve{\mathbf{\Delta}}_{1K} \\ \vdots & \vdots & \vdots \\ \bar{\mathbf{\Delta}}_{K1} & \dots & \bar{\mathbf{\Delta}}_{KK} \\ \breve{\mathbf{\Delta}}_{K1} & \dots & \breve{\mathbf{\Delta}}_{KK} \end{bmatrix} \tag{6.96}$$

where

$$\mathbf{\Delta}_{ij} = \begin{bmatrix} \mathbf{\Delta}^{ij}_1 \\ \vdots \\ \mathbf{\Delta}^{ij}_c \end{bmatrix} \tag{6.97}$$

and

$$\mathbf{\Gamma} = \begin{bmatrix} \mathbf{\Gamma}_{11} & \dots & \mathbf{\Gamma}_{1K} \\ \vdots & \ddots & \vdots \\ \mathbf{\Gamma}_{K1} & \dots & \mathbf{\Gamma}_{KK} \end{bmatrix} \tag{6.98}$$

where

$$\mathbf{\Gamma}_{ij} = \frac{2}{\eta} \sum_{t=1}^{N} Re \left[\mathbf{S}_i^*(t) \mathbf{D}^*(\boldsymbol{\theta}_i) \mathbf{D}(\boldsymbol{\theta}_j) \mathbf{S}_j(t) \right] \tag{6.99}$$

We can further simplify the expressions for $\mathbf{\Gamma}$. Observe that for any vector $\mathbf{v}$ and any square matrix $\mathbf{P}$, the following property holds

$$diag\{\mathbf{v}\}\mathbf{P} \left[diag\{\mathbf{v}\}\right]^* = P \odot \left[\mathbf{v}\mathbf{v}^*\right] \tag{6.100}$$

Then,

$$\begin{aligned} \mathbf{\Gamma}_{ij} &= \frac{2}{\eta} \sum_{t=1}^{N} Re \left[\mathbf{S}_i^*(t) \mathbf{D}^*(\boldsymbol{\theta}_i) \mathbf{D}(\boldsymbol{\theta}_j) \mathbf{S}_j(t) \right] \\ &= \frac{2}{\eta} \sum_{t=1}^{N} Re \left[diag\{\mathbf{H}\mathbf{x}_i(t)\} \mathbf{D}^*(\boldsymbol{\theta}_i) \mathbf{D}(\boldsymbol{\theta}_j) diag\{\mathbf{H}\mathbf{x}_j(t)\} \right] \\ &= \frac{2}{\eta} \sum_{t=1}^{N} Re \left[\mathbf{H}\mathbf{x}_i(t) \left[\mathbf{H}\mathbf{x}_j(t)\right]^* \odot \left[\mathbf{D}^*(\boldsymbol{\theta}_i) \mathbf{D}(\boldsymbol{\theta}_j)\right] \right] \\ &= \frac{2N}{\eta} Re \left[\left[\mathbf{H}\hat{\mathbf{R}}_{x_i x_j} \mathbf{H}^* \right] \odot \left[\mathbf{D}^*(\boldsymbol{\theta}_i) \mathbf{D}(\boldsymbol{\theta}_j)\right] \right] \end{aligned} \tag{6.101}$$

Using the inverse of a partitioned matrix lemma, we get the CRB for the vector of DOAs:

$$CRB(\boldsymbol{\theta}) = \left(\mathbf{\Gamma} - \mathbf{\Delta} \mathbf{\Phi}^{-1} \mathbf{\Delta}^T \right)^{-1} \tag{6.102}$$

This is the exact Cramer Rao bound on the estimation of vector of angles of K sources with c multipath reflections each.

Uncorrelated signals

Proceeding with the analysis for uncorrelated signals (as defined in section 4.4), we use the following approximation:

$$\hat{\mathbf{R}}_{x_i x_j} = \frac{1}{N}\sum_{t=1}^{N}\mathbf{x}(t)\mathbf{x}^*(t) \cong 0 \qquad \text{for i} \neq \text{ j} \tag{6.103}$$

Then we get:

$$\mathbf{\Phi}_{ij}, \mathbf{\Delta}_{ij}, \mathbf{\Gamma}_{ij} \cong 0 \qquad for \quad i \neq j \tag{6.104}$$

$$\mathbf{\Gamma} = \frac{2}{\eta}\sum_{t=1}^{N} Re\left[\mathbf{S}^*(t)\mathbf{D}^*(\boldsymbol{\theta})\mathbf{D}(\boldsymbol{\theta})\mathbf{S}(t)\right] \cong \begin{bmatrix} \mathbf{\Gamma}_1 & \dots & 0 \\ 0 & \ddots & 0 \\ 0 & \dots & \mathbf{\Gamma}_K \end{bmatrix} \tag{6.105}$$

The FIM for uncorrelated signals takes the following form:

$$FIM \cong \begin{bmatrix} \bar{\mathbf{\Phi}}_{11} & -\breve{\mathbf{\Phi}}_{11} & \dots & 0 & 0 & \bar{\mathbf{\Delta}}_{11} & \dots & 0 \\ \breve{\mathbf{\Phi}}_{11} & \bar{\mathbf{\Phi}}_{11} & \dots & 0 & 0 & \breve{\mathbf{\Delta}}_{11} & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & \bar{\mathbf{\Phi}}_{KK} & -\breve{\mathbf{\Phi}}_{KK} & 0 & \dots & \bar{\bar{\mathbf{\Delta}}}_{KK} \\ 0 & 0 & \dots & \breve{\mathbf{\Phi}}_{KK} & \bar{\mathbf{\Phi}}_{KK} & 0 & \dots & \breve{\breve{\mathbf{\Delta}}}_{KK} \\ \bar{\mathbf{\Delta}}_{11}^T & \breve{\mathbf{\Delta}}_{11}^T & \dots & 0 & 0 & \mathbf{\Gamma}_1 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & \bar{\mathbf{\Delta}}_{KK}^T & \breve{\mathbf{\Delta}}_{KK}^T & 0 & \dots & \mathbf{\Gamma}_K \end{bmatrix} \tag{6.106}$$

The Cramer Rao bound on the variance of an error for estimating the vector of angles of the *ith* source, $\boldsymbol{\theta}_i$ $(i = 1 \dots K)$, is equal to the elements in FIM^{-1} with the same coordinates as $\mathbf{\Gamma}_i$ in FIM.

It can be seen from the structure of FIM, that the CRBs of $\boldsymbol{\theta}_i$ are decoupled one from another. Note that all the matrices in 6.102 are block diagonal, and that the elements in FIM^{-1} with coordinates of $\mathbf{\Gamma}_i$ depend only on $\mathbf{H}_i$ and $\boldsymbol{\theta}_i$.

Hence, similar to the single path case, the Cramer Rao bound for estimating the angles of ith source, given that it's uncorrelated with the other signals is asymptotically equal to:

$$CRB(\boldsymbol{\theta}_i) = \left(\boldsymbol{\Gamma}_i - \begin{bmatrix} \bar{\boldsymbol{\Delta}}_{ii}^T & \breve{\boldsymbol{\Delta}}_{ii}^T \end{bmatrix} \boldsymbol{\Phi}_{ii}^{-1} \begin{bmatrix} \bar{\boldsymbol{\Delta}}_{ii} \\ \breve{\boldsymbol{\Delta}}_{ii} \end{bmatrix}\right)^{-1} \tag{6.107}$$

Which is the CRB for estimating the vector of angles of reflections of one source.

Then, following the same guidelines as in previous section (single path case), using (6.47) we get

$$CRB(\boldsymbol{\theta}_i) = \left(\boldsymbol{\Gamma}_i - Re\left\{\boldsymbol{\Delta}_{ii}^* \boldsymbol{\Phi}_{ii}^{-1} \boldsymbol{\Delta}_{ii}\right\}\right)^{-1} \tag{6.108}$$

From now on we drop the index i, and refer to a singe-source CRB.

For a case of one source we have

$$\Gamma = \frac{2}{\eta} Re\left[\mathbf{D}^*(\boldsymbol{\theta})\mathbf{D}(\boldsymbol{\theta}) \odot \left[\mathbf{H}\hat{\mathbf{R}}_{xx}\mathbf{H}^*\right]\right] \tag{6.109}$$

and

$$\boldsymbol{\Delta} = [\boldsymbol{\Delta}_1 \dots \boldsymbol{\Delta}_c] \tag{6.110}$$

Claim:

$$\begin{aligned} &[\boldsymbol{\Delta}_1^* \dots \boldsymbol{\Delta}_c^*]\, \boldsymbol{\Phi}^{-1} \begin{bmatrix} \boldsymbol{\Delta}_1 \\ \vdots \\ \boldsymbol{\Delta}_c \end{bmatrix} = \\ =&\frac{2N}{\eta} \mathbf{D}^*(\boldsymbol{\theta})\mathbf{A}(\boldsymbol{\theta}) \left[\mathbf{A}^*(\boldsymbol{\theta})\mathbf{A}(\boldsymbol{\theta})\right]^{-1} \mathbf{A}^*(\boldsymbol{\theta})\mathbf{D}(\boldsymbol{\theta}) \odot \left[\mathbf{H}\hat{\mathbf{R}}_{xx}\mathbf{H}^*\right] \end{aligned} \tag{6.111}$$

proof

The inverse of $\mathbf{\Phi}$ is given by:

$$\mathbf{\Phi}^{-1} = \frac{\eta}{2N} \left[\mathbf{A}^*\mathbf{A}\right]^{-1} \otimes \left[\hat{\mathbf{R}}_{xx}^T\right]^{-1} \tag{6.112}$$

Let's partition $\mathbf{\Phi}^{-1}$, which is a '$c \cdot d \times c \cdot d$' matrix to '$d \cdot d$' matrices in the following way

$$\mathbf{\Phi}^{-1} \triangleq \mathbf{\Omega} = \begin{bmatrix} \tilde{\mathbf{\Omega}}_{11} & \dots & \tilde{\mathbf{\Omega}}_{1c} \\ \vdots & \ddots & \vdots \\ \tilde{\mathbf{\Omega}}_{c1} & \dots & \tilde{\mathbf{\Omega}}_{cc} \end{bmatrix} \tag{6.113}$$

then

$$\tilde{\mathbf{\Omega}}_{ij} = \frac{\eta}{2N} \left[\mathbf{A}^*\mathbf{A}\right]_{ij}^{-1} \hat{\mathbf{R}}_{xx}^T \tag{6.114}$$

also

$$\begin{aligned} &\left[\mathbf{\Delta}_1^* \dots \mathbf{\Delta}_c^*\right] \mathbf{\Omega} \begin{bmatrix} \mathbf{\Delta}_1 \\ \vdots \\ \mathbf{\Delta}_c \end{bmatrix} \\ &= \sum_{i=1}^{c} \sum_{j=1}^{c} \mathbf{\Delta}_i \tilde{\mathbf{\Omega}}_{ij} \mathbf{\Delta}_j \end{aligned} \tag{6.115}$$

Let's rewrite $\mathbf{\Delta}_i, i = 1 \dots c$ in the following form

$$\begin{aligned} \mathbf{\Delta}_i =& \frac{2}{\eta} \sum_{t=1}^{N} \left[\mathbf{x}^+(t)\mathbf{a}^*(\theta_i)\mathbf{D}(\boldsymbol{\theta})\mathbf{S}(t)\right] = \\ &= \begin{bmatrix} \frac{2}{\eta} \sum_{t=1}^{N} \left[\mathbf{x}^+(t)\mathbf{a}^*(\theta_i)\mathbf{d}(\boldsymbol{\theta}_1)\mathbf{h}_1\mathbf{x}(t)\right] & \dots & \frac{2}{\eta} \sum_{t=1}^{N} \left[\mathbf{x}^+(t)\mathbf{a}^*(\theta_i)\mathbf{d}(\boldsymbol{\theta}_c)\mathbf{h}_c\mathbf{x}(t)\right] \end{bmatrix} \\ &= \begin{bmatrix} \frac{2N}{\eta}\mathbf{a}^*(\theta_i)\mathbf{d}(\theta_1)\hat{\mathbf{R}}_{xx}^T\mathbf{h}_1^T & \dots & \frac{2N}{\eta}\mathbf{a}^*(\theta_i)\mathbf{d}(\theta_c)\hat{\mathbf{R}}_{xx}^T\mathbf{h}_c^T \end{bmatrix} \end{aligned} \tag{6.116}$$

then,

$$\begin{aligned} \sum_{i=1}^{c} \sum_{j=1}^{c} \mathbf{\Delta}_i \tilde{\mathbf{\Omega}}_{ij} \mathbf{\Delta}_j = \sum_{i=1}^{c} \sum_{j=1}^{c} \frac{2N}{\eta} &\begin{bmatrix} \mathbf{a}^*(\theta_i)\mathbf{d}(\theta_1)\hat{\mathbf{R}}_{xx}^T\mathbf{h}_1^T & \dots & \mathbf{a}^*(\theta_i)\mathbf{d}(\theta_c)\hat{\mathbf{R}}_{xx}^T\mathbf{h}_c^T \end{bmatrix}^* \tilde{\mathbf{\Omega}}_{ij} \cdot \\ &\cdot \frac{2N}{\eta} \begin{bmatrix} \mathbf{a}^*(\theta_j)\mathbf{d}(\theta_1)\hat{\mathbf{R}}_{xx}^T\mathbf{h}_1^T & \dots & \mathbf{a}^*(\theta_j)\mathbf{d}(\theta_c)\hat{\mathbf{R}}_{xx}^T\mathbf{h}_c^T \end{bmatrix} \end{aligned} \tag{6.117}$$

$$\sum_{i=1}^{c}\sum_{j=1}^{c}\mathbf{\Delta}_i\tilde{\mathbf{\Omega}}_{ij}\mathbf{\Delta}_j = \sum_{i=1}^{c}\sum_{j=1}^{c}\frac{2N}{\eta}\begin{bmatrix}\mathbf{d}^*(\theta_1)\mathbf{a}(\theta_i)\left[\hat{\mathbf{R}}_{xx}^T\mathbf{h}_1^T\right]^* \\ \vdots \\ \mathbf{d}^*(\theta_c)\mathbf{a}(\theta_i)\left[\hat{\mathbf{R}}_{xx}^T\mathbf{h}_c^T\right]^*\end{bmatrix}\frac{\eta}{2N}\left[\mathbf{A}^*\mathbf{A}\right]_{ij}^{-1}\left(\hat{\mathbf{R}}_{xx}^T\right)^{-1}\cdot$$
$$\cdot\frac{2N}{\eta}\begin{bmatrix}\mathbf{a}^*(\theta_j)\mathbf{d}_1(\theta)\hat{\mathbf{R}}_{xx}^T\mathbf{h}_1^T & \dots & \mathbf{a}^*(\theta_j)\mathbf{d}_c(\theta)\hat{\mathbf{R}}_{xx}^T\mathbf{h}_c^T\end{bmatrix} \tag{6.118}$$

using (6.59) we have

$$= \sum_{i=1}^{c}\sum_{j=1}^{c}\frac{2N}{\eta}\left[\mathbf{A}^*\mathbf{A}\right]_{ij}^{-1}\cdot$$
$$\cdot\begin{bmatrix}\mathbf{d}^*(\theta_1)\mathbf{a}(\theta_i)\mathbf{a}^*(\theta_j)\mathbf{d}(\theta_1)\mathbf{h}_1\hat{\mathbf{R}}_{xx}\mathbf{h}_1^* & \dots & \mathbf{d}^*(\theta_1)\mathbf{a}(\theta_i)\mathbf{a}^*(\theta_j)\mathbf{d}(\theta_c)\mathbf{h}_1\hat{\mathbf{R}}_{xx}\mathbf{h}_c^* \\ \vdots & \ddots & \vdots \\ \mathbf{d}^*(\theta_c)\mathbf{a}(\theta_i)\mathbf{a}^*(\theta_j)\mathbf{d}(\theta_1)\mathbf{h}_c\hat{\mathbf{R}}_{xx}\mathbf{h}_1^* & \dots & \mathbf{d}^*(\theta_c)\mathbf{a}(\theta_i)\mathbf{a}^*(\theta_j)\mathbf{d}(\theta_c)\mathbf{h}_c\hat{\mathbf{R}}_{xx}\mathbf{h}_c^*\end{bmatrix}$$
$$= \sum_{i=1}^{c}\sum_{j=1}^{c}\frac{2N}{\eta}\mathbf{D}^*\frac{\mathbf{a}(\theta_i)\mathbf{a}^*(\theta_j)}{\left[\mathbf{A}^*\mathbf{A}\right]_{ij}^{-1}}\mathbf{D}\odot\left[\mathbf{H}\hat{\mathbf{R}}_{xx}\mathbf{H}^*\right] \tag{6.119}$$

Since,

$$\sum_{i=1}^{c}\sum_{j=1}^{c}\frac{\mathbf{a}(\theta_i)\mathbf{a}^*(\theta_j)}{\left[\mathbf{A}^*\mathbf{A}\right]_{ij}^{-1}} = \mathbf{A}\left[\mathbf{A}\mathbf{A}\right]^{-1}\mathbf{A}^* \tag{6.120}$$

the claim is proved ▲

Inserting (6.111) and (6.109) into (6.107) we get:

$$CRB(\boldsymbol{\theta})^{-1} =$$
$$\frac{2N}{\eta}Re\left[\mathbf{D}^*(\boldsymbol{\theta})\mathbf{D}(\boldsymbol{\theta})\odot\left[\mathbf{H}\hat{\mathbf{R}}_{xx}\mathbf{H}^*\right]\right]$$
$$-\frac{2N}{\eta}Re\left[\mathbf{D}^*(\boldsymbol{\theta})\mathbf{A}(\boldsymbol{\theta})\left[\mathbf{A}^*(\boldsymbol{\theta})\mathbf{A}(\boldsymbol{\theta})\right]^{-1}\mathbf{A}^*(\boldsymbol{\theta})\mathbf{D}(\boldsymbol{\theta})\odot\left[\mathbf{H}\hat{\mathbf{R}}_{xx}\mathbf{H}^*\right]\right]$$
$$=\frac{2N}{\eta}Re\left[\left[\mathbf{H}\hat{\mathbf{R}}_{xx}\mathbf{H}^*\right]\odot\left[\mathbf{D}^*(\boldsymbol{\theta})\mathbf{D}(\boldsymbol{\theta})-\mathbf{D}^*(\boldsymbol{\theta})\mathbf{A}(\boldsymbol{\theta})\left[\mathbf{A}^*(\boldsymbol{\theta})\mathbf{A}(\boldsymbol{\theta})\right]^{-1}\mathbf{A}^*(\boldsymbol{\theta})\mathbf{D}(\boldsymbol{\theta})\right]\right]$$
$$=\frac{2N}{\eta}Re\left[\left[\mathbf{H}\hat{\mathbf{R}}_{xx}\mathbf{H}^*\right]\odot\left[\mathbf{D}^*(\boldsymbol{\theta})\left[I-\mathbf{A}(\boldsymbol{\theta})\left[\mathbf{A}^*(\boldsymbol{\theta})\mathbf{A}(\boldsymbol{\theta})\right]^{-1}\mathbf{A}^*(\boldsymbol{\theta})\right]\mathbf{D}(\boldsymbol{\theta})\right]\right] \tag{6.121}$$

Finally, the decoupled CRB for estimating DOAs of multiple reflections of multiple uncorrelated sources (which equals to the exact CRB for a known single source) with unknown channel parameters is given by

$$\boxed{CRB(\boldsymbol{\theta}) = \frac{\eta}{2N}\left[Re\left\{\left[\mathbf{H}\hat{\mathbf{R}}_{xx}\mathbf{H}^*\right] \odot \left[\mathbf{D}^*(\boldsymbol{\theta})\left[I - \mathbf{A}(\boldsymbol{\theta})\left[\mathbf{A}^*(\boldsymbol{\theta})\mathbf{A}(\boldsymbol{\theta})\right]^{-1}\mathbf{A}^*(\boldsymbol{\theta})\right]\mathbf{D}(\boldsymbol{\theta})\right]\right\}\right]^{-1}} \tag{6.122}$$

6.3.1 "White sequences"

For a case of white data sequences (as defined in section 4.6) where

$$\lim_{N\to\infty} \hat{\mathbf{R}}_{xx_i} = \rho\mathbf{I} \tag{6.123}$$

we get the following expression:

$$\boxed{CRB(\boldsymbol{\theta}) = \frac{1}{2\cdot SNR\cdot N}\left[Re\left\{\left[\mathbf{H}\mathbf{H}^*\right] \odot \left[\mathbf{D}^*(\boldsymbol{\theta})\left[I - \mathbf{A}(\boldsymbol{\theta})\left[\mathbf{A}^*(\boldsymbol{\theta})\mathbf{A}(\boldsymbol{\theta})\right]^{-1}\mathbf{A}^*(\boldsymbol{\theta})\right]\mathbf{D}(\boldsymbol{\theta})\right]\right\}\right]^{-1}} \tag{6.124}$$

$$SNR = \frac{\rho}{\eta} \tag{6.125}$$

Chapter 7

Simulation Study

This chapter presents numerical examples of the problems addressed in this thesis. The estimation algorithms and the corresponding CRBs were simulated for a case of linear uniform array of M sensors.The data waveform used is a digital sequence of symbols uniformly and independently distributed between $1+j, 1-j, -1+j, -1-j$ (QPSK Modulation).

$$x(t) = I(t) + jQ(t) \tag{7.1}$$

where

$$I(t), Q(t) \propto U(\{-1; 1\}) \tag{7.2}$$

A typical known data waveform is shown in figure (7.1).

The simulation was performed using Monte-Carlo method with P trials. The ISI channel was simulated using Rayleigh distribution, and its power was normalized to 1.

$$\mathbf{h} = \frac{1}{\sqrt{2d}}(\bar{\mathbf{h}} + j\breve{\mathbf{h}}) \tag{7.3}$$

with

$$\bar{\mathbf{h}}, \breve{\mathbf{h}} \propto N(0, I) \tag{7.4}$$

where I is a unitary matrix of size d.

The charts that follow present the angle estimation accuracy, which is evaluated by its Root Mean Square Error (RMSE)

$$RMSE = \sqrt{\frac{1}{P}\sum_{i=1}^{P} e_i^2} \tag{7.5}$$

where e_i is the angle estimation error at each trial. Subsequently, the Cramer-Rao bound presented in the charts is the bound on the standard deviation of the error, and is equal to square root of the CRB derived in chapter 6.

In some of the charts, the results of ML estimates for different random waveforms were averaged using A averages.

At the top of each chart, various simulation parameters are depicted, such as the signal to noise (SNR), number of data samples (N), channel length (d), number of array sensors (M), etc.

Simulations for single path case were performed assuming quarter wavelength spacings between antenna elements ($\lambda/4$), simulations for multipath case were performed assuming half wavelength spacings ($\lambda/2$).

7.1 Single Path Case

First, we present numerical examples for a case of one signal that impinges on the array from a single angle with unknown ISI channel coefficients. The exact and the approximate (white sequence) ML estimates for this case are obtained using (5.45) and (5.48) respectively. Figure 7.2 depicts the angle estimation accuracy as a function of the number of samples. As expected, the ML estimator is efficient asymptotically, i.e the estimates coincide with the CRB after enough measurements had been accumulated. Since the waveform used is independently distributed, the estimates using the white sequence assumption are very close to the exact ML estimates and are also asymptotically efficient. Hence, for large enough N, it's beneficial to use (5.48) rather than (5.45), since it involves less computational operations.

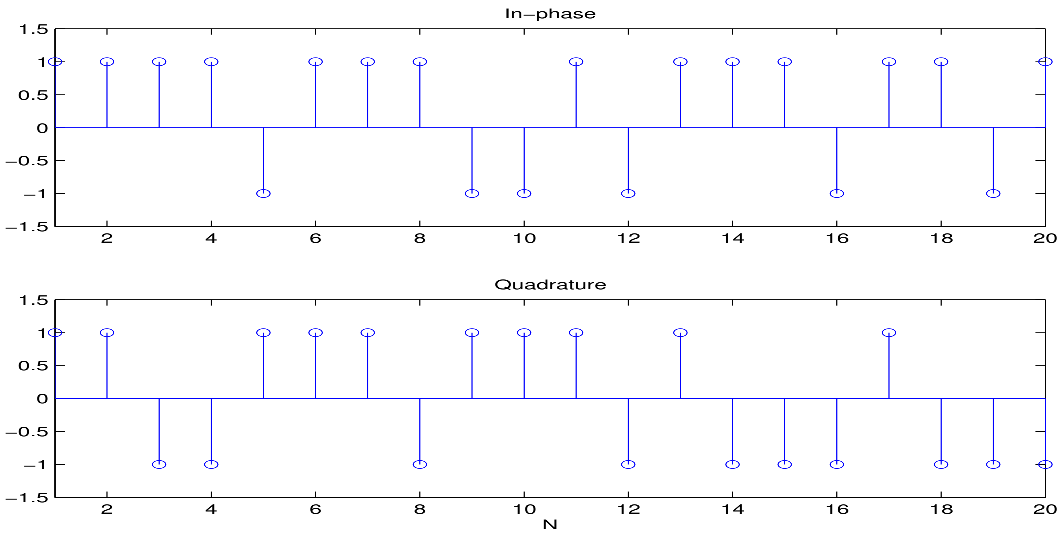

Figure 7.1: **Typical known waveform**

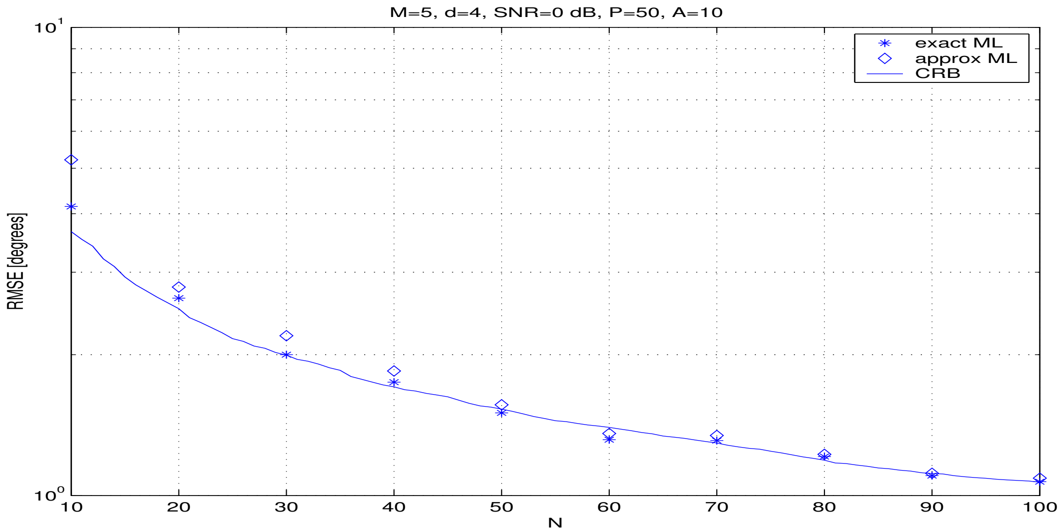

Figure 7.2: **Angle estimation accuracy as a function of the number of symbols, for a single source. Solid line is for exact CRB. The symbols "⋄" and "*" are for the exact and for the approximate ML estimators, respectively. In this simulation, one signal impinges on the array from an angle of 0 degrees**

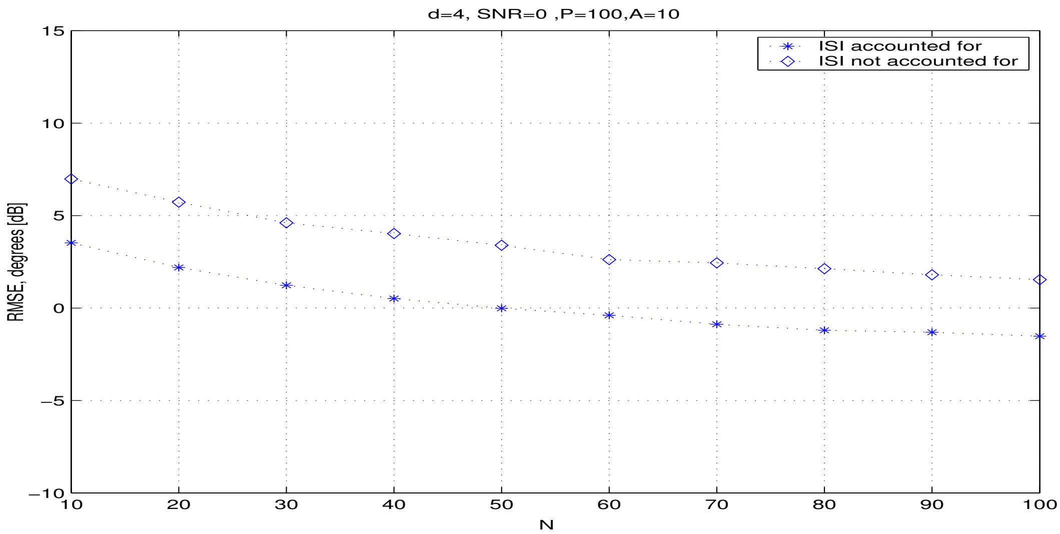

Figure 7.3: **Comparison of estimation accuracy between an algorithm that estimates ISI components and one that does not.**
In this simulation a single signal with 4 ISI components impinges on the array. The top line (denoted by "$\diamond$")is the angle estimation ignoring the channel parameters and assuming only complex scaling. The bottom line (denoted by "$*$") is for estimation accounting for and jointly estimating ISI channel components. In this particular example, there is a 3db improvement when ISI is estimated.

Figure (7.3) portrays the gain in the accuracy of estimation when accounting for the ISI components. In this simulation a single signal with 4 ISI components impinges on the array. Estimates using the algorithm presented in chapter 5 are compared to an estimation assuming only complex scaling (the DEML algorithm in [8]). In this example we have a 3dB improvement in RMSE when the channel coefficient are estimated.

The decoupling property of the algorithm for known waveforms is demonstrated in Figure 7.4, which shows the simulation results of the case of two sources with known waveforms. ML angle estimates of one of the sources are presented and compared to the one-source CRB. Although there are two signals that impinge on the array, the ML estimates asymptotically yield one-source performance. This is due to the fact that the data waveforms are weakly correlated (in this example

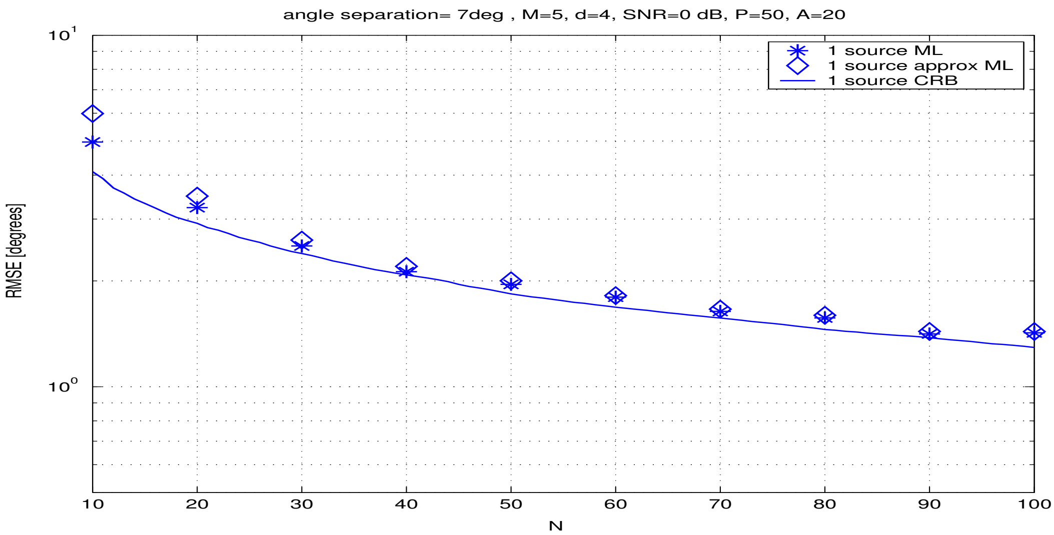

Figure 7.4: **Decoupled ML estimator: accuracy as a function of the number of symbols, for two known, randomly chosen data sequences.**
In this example, two signals with known waveforms impinge on the array from angles 0 and 7 degrees. The chart shows the ML estimates of the first angle, using only the knowledge of the first waveform. The estimation accuracy is compared to the exact CRB for a single source (solid line). Since the waveforms are uncorrelated, the estimates reach one source accuracy for large enough N. The decoupled ML estimates are denoted by "*", and the approximate (white data assumption) estimates are denoted by "$\diamond$".

both waveforms are randomly selected). Since the minimization problems for each signal are decoupled from each other, only the knowledge of the desired waveform is required for the estimation of its angle.

Figure 7.5 demonstrates the behavior of the decoupled estimator as a function of the SNR. At low SNR, where the AGWN is the main cause of estimation error, the estimates coincide with the one-source CRB. At high SNR, the achievable error due to white noise is small, and the interference of the other signal becomes the main contributor to the error. At some point the estimates diverge from the CRB and converge to a certain RMSE limit. Beyond this limit the performance of the algorithm is no longer influenced by the AWGN. Instead, it is the correlation between the data waveforms that limits the achievable error. The lower the correlation, the lower the threshold point. Another factor that determines this threshold is the ratio between the signals amplitudes, which is equal to 1 in our example. Further analysis of this error limit appears in appendix A.

We can also see from figure 7.6 that the decoupled ML has no sensitivity to angle separation (as opposed to exact ML). This is also a result of the fact that the minimization problems are separate for uncorrelated data waveforms.

7.2 Multipath Case

Next we turn to the case of multiple paths of the same source. The corresponding estimators are given by (5.37) and (5.41). Simulations for multipath case were performed assuming half wavelength spacings between antenna elements.

Figure 7.8 presents the Cramer-Rao Bounds for a case of two paths of a single source as a function of the array size. As we've seen, for large enough N, the angle estimates coincide with the corresponding CRB. Hence, we can view figure 7.8 as a prediction of the estimator performance. For an angle separation of 5 degrees, the one path and two paths CRBs coincide when the number of sensor equals 27,

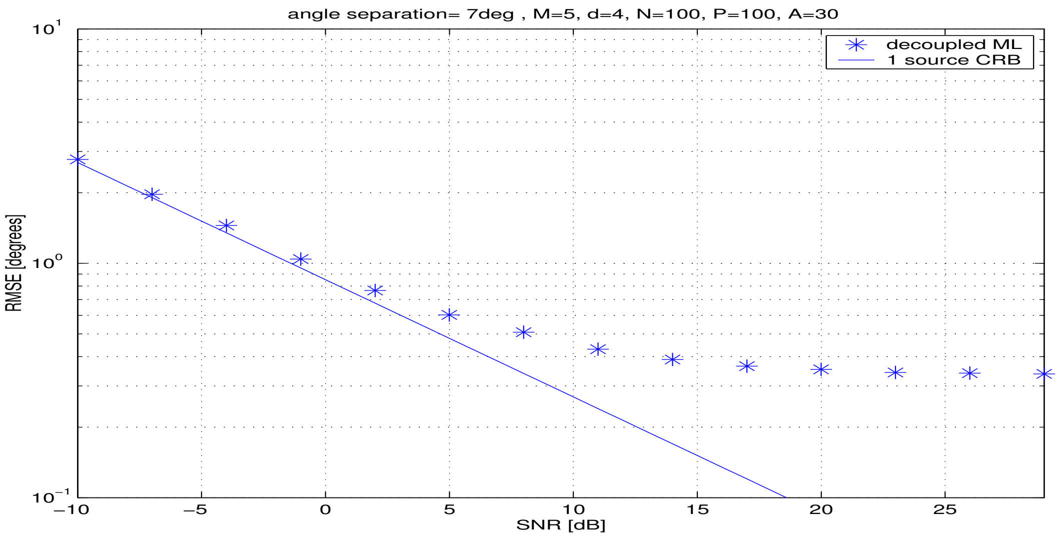

Figure 7.5: **Estimation accuracy as a function of the SNR, for two sources. In this example two signals with known waveforms impinge on the array from angles 0 and 7 degrees. The chart shows the ML estimates of the first angle (denoted by "*"), using only the knowledge of the first waveform. The estimation accuracy is compared with the exact CRB for a single source (solid line). One can see that for a low SNR, where the AWGN noise is the main contributor to the error, the decoupled ML coincides with the one-source CRB. However at high SNR, where the second (interfering) waveform is the main contributor to the error, the estimates diverge from the CRB.**

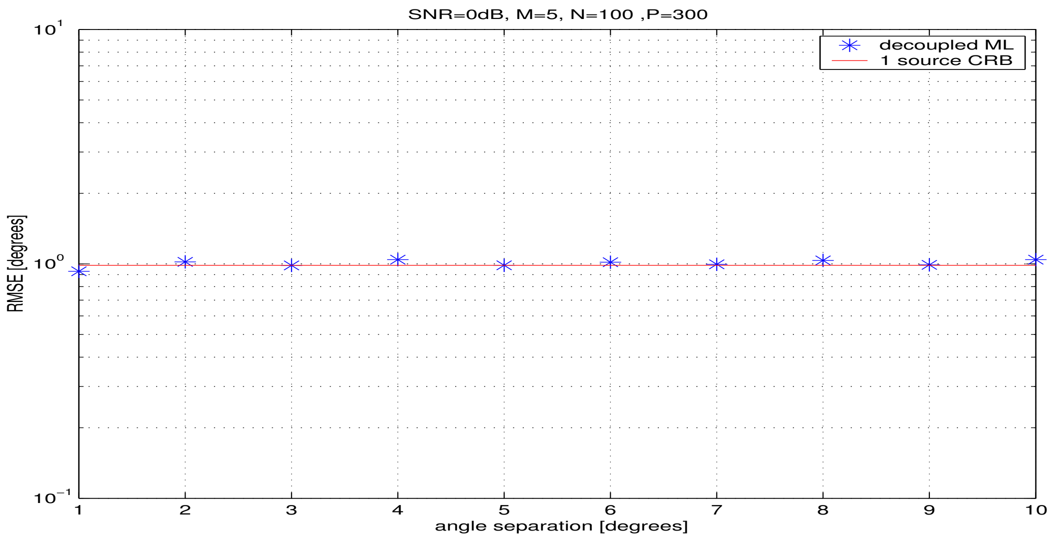

Figure 7.6: **Estimation accuracy of the decoupled ML estimator as a function of the angle separation, for two sources.**

In this example two signals with known waveforms impinge on the array. The chart shows the ML estimates of the first signal with DOA of 0 degrees (denoted by "*"), while the other signal arrives from different angles which form the 'x' axis of this chart. The estimation accuracy is compared to the exact CRB for a single source (solid line). As one can see, the decoupled ML is absolutely independent of angle separation. And in this particular example, since N is sufficiently large, it also coincides with the exact one-source CRB.

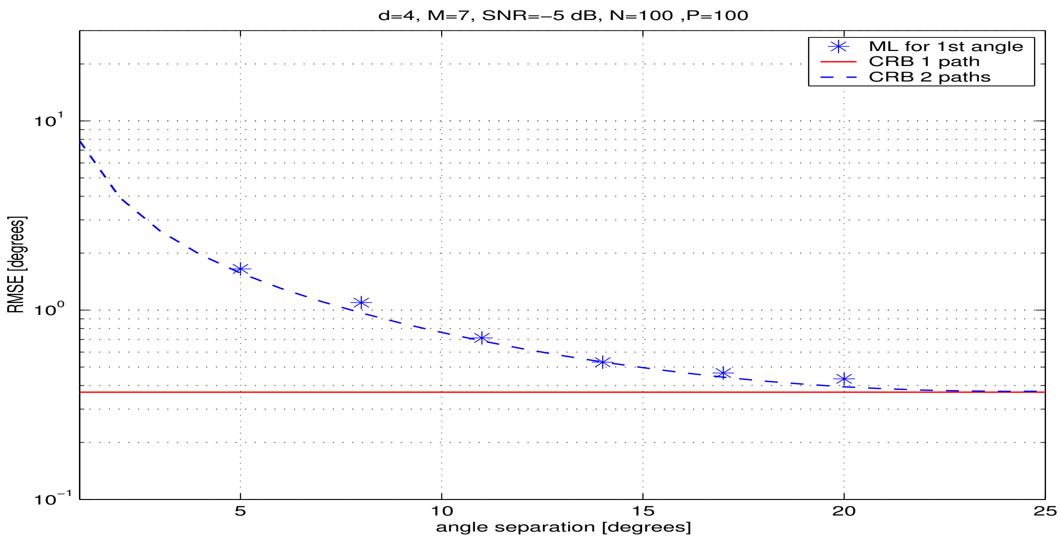

Figure 7.7: **Estimation accuracy of ML estimator as a function of the angle separation, for two reflections of a single source.**
In this example two reflections (”multipaths”) of the same known waveform impinge on the array. The chart shows the ML estimates of the angle of the first path that impinges from 0 degrees (denoted by ”*”), while the other signal arrives from different angles which form the ’x’ axis of this chart. The estimation accuracy is compared to the exact CRBs (dashed and solid lines). Since N is large enough in this example, the estimates coincide with the CRB. As one can see, the estimation accuracy for two paths is worse than for one-path when array aperture (3d bandwidth of 14.5 degrees for M=7) is larger than the angle separation, but as the angle separation increases the lines coincide.

which corresponds to a 3 db beam width of approx 3.8 degrees. Which means that when the estimated angles are 1.3 beam widths apart, both estimators have equal accuracy.

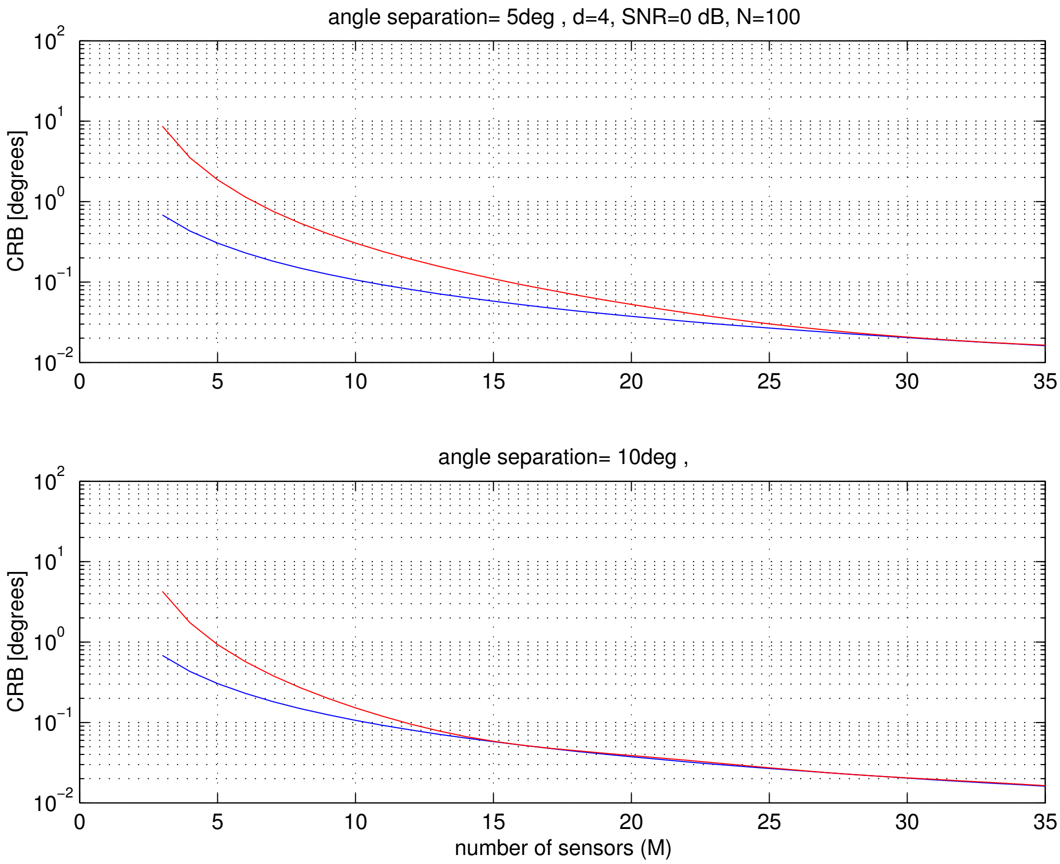

Figure 7.8: **CRB vs number of sensors, for a case of multiple and single path. The top line of each figure is the CRB for estimating the first angle out of two reflections of the same source separated by 5 degrees and 10 degrees. Bottom line is the CRB for estimating the angle of a single path. As the array aperture narrows (number of sensors grows) the lines coincide.**

Chapter 8

Summary and Conclusions

This thesis deals with estimating the angles of arrival of multiple signals in a channel with memory (ISI channel), assuming known waveforms. The case of known waveforms is common in digital communication when training sequence is used or when the transmitted symbols are demodulated. The case of known waveforms was recently addressed in several works ([7],[8],[4],[18]), however none treated a case of ISI channel.

The main contribution of this thesis is the development of ML estimators and corresponding performance bounds (Cramer-Rao bounds) for a case of multiple sources, possibly propagating through multiple paths (reflections), each passing through a selective fading (ISI) channel. Exact and approximate, computationally efficient, Maximum Likelihood estimators for jointly estimating the DOAs and the channel coefficients were derived for single-path and multi-path propagation.

Estimation algorithms presented in this thesis are listed below:

- Exact ML estimator for a case of multiple signals with known waveforms propagating through ISI channel (Eq. 5.16).
- Decoupled ML channel and DOA estimator for a case of multiple signals with uncorrelated waveforms (Eq. 5.45).

- Decoupled ML for a case of multiple uncorrelated waveforms, assuming i.i.d data sequences (Eq. 5.48).
- Decoupled ML channel and DOA estimator for a case of multiple uncorrelated waveforms, propagating through multiple paths each (Eq. 5.37).
- Decoupled ML channel and DOA estimator for a case of multiple uncorrelated waveforms, propagating through multiple paths each, assuming i.i.d data sequences (Eq. 5.41).

The decoupled ML estimators are especially useful since they avoid the multidimensional search required for exact ML estimator. For K uncorrelated signals with c reflections each, the dimension of optimization problem reduces from $K \times c$ to K c - dimensional searches, which significantly reduces the computational complexity.

The decoupled estimators converge to the exact ML asymptotically when the signal sources are uncorrelated with one another.

When dealing with communication signals, the sequentially transmitted symbols are also uncorrelated with each other (i.i.d sequence), and for this case further simplification of the estimator was derived. Although the search dimension remains the same, the need to compute the inverse of the source covariance matrix is avoided. Simulations show that this approximate estimator quickly coincides with the exact ML.

The Cramer-Rao bounds (CRB) for all of the scenarios mentioned in the list above were derived in this thesis. It was shown that given uncorrelated sources, the exact CRB (similarly to the estimator) asymptotically decouples to a one source CRB for each source.

Compact expressions for the decoupled CRBs (one source case) were derived for multipath and single path propagation (Eq. 6.122 and 6.62), as well as approximate CRBs for a case of white data sequences (Eq. 6.124 and 6.64).

Numerical examples were presented for all estimators and compared with the corresponding bounds. All simulations approve the theoretical results.

An interesting observation from the numerical examples is the threshold point in figure 7.5. It demonstrates the achievable accuracy limit of the decoupled estimator as a function of SNR, when two uncorrelated signals impinge on the array. At low SNR, where the AWGN is the main cause of estimation error, the estimates coincide with the one-source CRB. At high SNR, the achievable error due to white noise is small, and the interference of the other signal becomes the main contributor to the error. At some point the estimates diverge from the CRB and converge to a limit, beyond which the performance of the algorithm is no more influenced by the AWGN. Instead, it is the correlation between the data waveforms that limits the achievable error. Factors that determine this threshold point are the correlation between the sequences and the relative amplitude ratio (which is assumed 1 in our examples).

The achieved accuracy gain of the proposed estimator over the estimator in [8] for ISI case was also presented.

8.1 Further Research

An interesting topic for further research is a continuation of the analysis that appears in the appendix of the thesis, which is the development of an explicit expression of the threshold point in figure 7.5. An ability to predict the limit in the achievable error can serve as a benchmark for the decoupled estimator.

In this thesis, the transmitted waveforms were assumed completely known. Another topic for further work could be a development of algorithms for partially known digital waveforms. For example when the alpha-bet of the signal is known, but not the specific waveforms transmitted. Algorithms using some knowledge of the structure of the waveform previously appeared in the literature. One example is [10], where a case of constant modulus signals is analyzed. This algorithm can be extended to signals with a known digital modulation pattern, like high order QAM for example.

Appendix A

Small Error Analysis

In this chapter we will investigate the behavior of the algorithm when the estimation error is small. When white noise is the only factor causing the inaccuracy of the estimates, the error becomes small when the number of samples is large (thus the small error analysis is often called asymptotic analysis). In our model, in addition to the white noise, signals coming from other angles also contribute to the error. As the signal to noise ratio grows, the white noise affect on the estimation error becomes smaller and the interference of other signals becomes the major contributor to the error. We start first by analyzing a case of two signals. We want to estimate the DOA of one signal while the other is treated as an interference. Both signals obey the same model as shown in chapter 4.

$$\mathbf{y}(t) = \mathbf{a}(\theta_1)\mathbf{h}_1\mathbf{x}_1(t) + \mathbf{a}(\theta_2)\mathbf{h}_2\mathbf{x}_2(t) + \mathbf{n}(t) \tag{A.1}$$

Recall that the angle of arrival for each signal is found by minimizing the following cost function

$$\hat{\theta}_1 = \arg\min_{\theta_1}(tr\{[\mathbf{P}^{\perp}_{\mathbf{a}_1}\hat{\mathbf{C}}_1]\hat{\mathbf{R}}_{xx1}[\mathbf{P}^{\perp}_{\mathbf{a}_1}\hat{\mathbf{C}}_1]^*\}) \tag{A.2}$$

Denote the cost function as F:

$$F(\theta_1, \mathbf{y}(1)\ldots\mathbf{y}(N)) = (tr\{[\mathbf{P}^{\perp}_{\mathbf{a}_1}\hat{\mathbf{C}}_1]\hat{\mathbf{R}}_{xx1}[\mathbf{P}^{\perp}_{\mathbf{a}_1}\hat{\mathbf{C}}_1]^*\}) \tag{A.3}$$

Define a derivative of F by the angle as Q:

$$Q(\theta_1, \mathbf{y}(1)\ldots\mathbf{y}(N)) = \frac{\partial}{\partial\theta_1}F(\theta_1, \mathbf{y}(1)\ldots\mathbf{y}(N)) \tag{A.4}$$

Let $\overline{\theta_1}$ be the true angle of arrival. And let $\overline{\mathbf{y}}(1)\ldots\overline{\mathbf{y}}(N)$ be the corresponding measurements of the array when there is no noise and interference. Using Taylor series let's expand Q around the point $\theta_1 = \overline{\theta_1}$ and $\mathbf{y}(1)\ldots\mathbf{y}(N) = \overline{\mathbf{y}}(1)\ldots\overline{\mathbf{y}}(N)$ The second order term approximations yields:

$$\begin{aligned}
&Q(\theta_1, \mathbf{y}(1)\ldots\mathbf{y}(N)) = \\
&+ Q(\overline{\theta_1}, \overline{\mathbf{y}}(1)\ldots\overline{\mathbf{y}}(N)) \\
&+ \frac{\partial}{\partial\theta_1}\{Q(\overline{\theta_1}, \overline{\mathbf{y}}(1)\ldots\overline{\mathbf{y}}(N))\}(\theta_1 - \overline{\theta_1}) \\
&+ \sum_{t=1}^{N}\frac{\partial}{\partial\mathbf{y}(t)}\{Q(\overline{\theta_1}, \overline{\mathbf{y}}(1)\ldots\overline{\mathbf{y}}(N))\}(\mathbf{y}(t) - \overline{\mathbf{y}}(t)) \\
&+ \frac{1}{2}\frac{\partial^2}{\partial\theta_1\partial\theta_1}\{Q(\overline{\theta_1}, \overline{\mathbf{y}}(1)\ldots\overline{\mathbf{y}}(N))\}(\theta_1 - \overline{\theta_1})^2 \\
&+ \frac{1}{2}\sum_{t=1}^{N}\frac{\partial^2}{\partial\mathbf{y}(t)\partial\theta_1}\{Q(\overline{\theta_1}, \overline{\mathbf{y}}(1)\ldots\overline{\mathbf{y}}(N))\}(\mathbf{y}(t) - \overline{\mathbf{y}}(t))(\theta_1 - \overline{\theta_1}) \\
&+ \frac{1}{2}\sum_{t=1}^{N}\sum_{s=1}^{N}\frac{\partial^2}{\partial\mathbf{y}(t)\mathbf{y}(s)}\{Q(\overline{\theta_1}, \overline{\mathbf{y}}(1)\ldots\overline{\mathbf{y}}(N))\}(\mathbf{y}(t) - \overline{\mathbf{y}}(t))(\mathbf{y}(s) - \overline{\mathbf{y}}(s))
\end{aligned} \tag{A.5}$$

Higher orders of the Taylor series can be neglected due to the following:
(a) The derivative of the cost function Q is a square function of the measurements, thus the third order derivatives are equal to zero.
(b) Square terms of the angle error can be neglected assuming that the the angle error is small.
Let's evaluate Q at the point of the estimated angle $\theta = \hat{\theta}$ and the corresponding measurements. F reaches it's maximum at the estimated angle and the corresponding measurements and at true angle and the corresponding measurements . Hence Q is zero at these two points.

$$Q(\hat{\theta_1}, \mathbf{y}(1)\ldots\mathbf{y}(N)) = 0 \tag{A.6}$$

$$Q(\overline{\theta_1}, \overline{\mathbf{y}}(1)\ldots\overline{\mathbf{y}}(N)) = 0 \tag{A.7}$$

When the angle error is small, the term $\frac{1}{2}\frac{\partial^2}{\partial\theta_1\partial\theta_1}\{Q(\overline{\theta_1}, \overline{\mathbf{y}}(1)\ldots\overline{\mathbf{y}}(N))\}(\theta_1 - \overline{\theta_1})^2$ can be neglected and we get:

$$\begin{aligned}
&-\frac{\partial}{\partial\theta_1}\{Q(\overline{\theta_1}, \overline{\mathbf{y}}(1)\ldots\overline{\mathbf{y}}(N))\}(\hat{\theta_1} - \overline{\theta_1}) \\
&-\frac{1}{2}\sum_{t=1}^{N}\frac{\partial^2}{\partial\mathbf{y}(t)\partial\theta_1}\{Q(\overline{\theta_1}, \overline{\mathbf{y}}(1)\ldots\overline{\mathbf{y}}(N))\}(\mathbf{y}(t) - \overline{\mathbf{y}}(t))(\hat{\theta_1} - \overline{\theta_1}) = \\
&\sum_{t=1}^{N}\frac{\partial}{\partial\mathbf{y}(t)}\{Q(\overline{\theta_1}, \overline{\mathbf{y}}(1)\ldots\overline{\mathbf{y}}(N))\}(\mathbf{y}(t) - \overline{\mathbf{y}}(t)) \\
&+\frac{1}{2}\sum_{t=1}^{N}\sum_{s=1}^{N}\frac{\partial^2}{\partial\mathbf{y}(t)\mathbf{y}(s)}\{Q(\overline{\theta_1}, \overline{\mathbf{y}}(1)\ldots\overline{\mathbf{y}}(N))\}(\mathbf{y}(t) - \overline{\mathbf{y}}(t))(\mathbf{y}(s) - \overline{\mathbf{y}}(s))
\end{aligned} \tag{A.8}$$

Denoting $(\hat{\theta_1} - \overline{\theta_1}) = \Delta\theta$ and $(\mathbf{y}(t) - \overline{\mathbf{y}}(t)) = \Delta\mathbf{y}(t)$ we can rewrite the above equation:

$$\begin{aligned}
&-\Delta\theta\left[\frac{\partial Q}{\partial\theta_1} + \frac{1}{2}\sum_{t=1}^{N}\frac{\partial^2 Q}{\partial\mathbf{y}(t)\partial\theta_1}\Delta\mathbf{y}(t)\right] = \\
&\sum_{t=1}^{N}\frac{\partial Q}{\partial\mathbf{y}(t)}\Delta\mathbf{y}(t) + \frac{1}{2}\sum_{t=1}^{N}\sum_{s=1}^{N}\frac{\partial^2 Q}{\partial\mathbf{y}(t)\mathbf{y}(s)}\Delta\mathbf{y}(t)\Delta\mathbf{y}(s)
\end{aligned} \tag{A.9}$$

$$\boxed{\Delta\theta = -\frac{\sum_{t=1}^{N}\frac{\partial Q}{\partial\mathbf{y}(t)}\Delta\mathbf{y}(t) + \frac{1}{2}\sum_{t=1}^{N}\sum_{s=1}^{N}\frac{\partial^2 Q}{\partial\mathbf{y}(t)\mathbf{y}(s)}\Delta\mathbf{y}(t)\Delta\mathbf{y}(s)}{\left[\frac{\partial Q}{\partial\theta_1} + \frac{1}{2}\sum_{t=1}^{N}\frac{\partial^2 Q}{\partial\mathbf{y}(t)\partial\theta_1}\Delta\mathbf{y}(t)\right]}} \tag{A.10}$$

where

$$\Delta\mathbf{y}(t) = \mathbf{a}(\theta_2)\mathbf{h}_2\mathbf{x}_2(t) + \mathbf{n}(t) \tag{A.11}$$

At high SNR, assuming negligible noise ($\mathbf{n}(t) \cong 0$), $(\Delta\theta)^2$ gives the achievable square error due to interference only. When noise is present, we need to evaluate $E\{\Delta\theta\}^2$ to get an asymptotic MSE due to both noise and interference.

Bibliography

[1] David Asztely, "On antenna arrays in mobile communication systems - fast fading and GSM base station receiver algorithms", *report no. IR-S3-SB-9611, Upsalla University,* march 1996

[2] O. Besson and P. Stoica, "Decoupled estimation of DOA and angular spread for a spatially distributed source", *IEEE Trans.Signal Processing,* vol. 48,No. 7,pp. 1872-1882, July. 2000

[3] Y. Bresler and A.Macovsky, "Exact maximum likelihood parameter estimation of superimposed esponential signals in noise", *IEEE Trans. Acoust., Speech, Signal Processing* vol AssP-34, pp. 1081-1089, Mar. 1986

[4] M. Cedervall and R.L. Moses, "Efficient maximum likelihood DOA Estimation for signals with known waveforms in the presense of multipath", *IEEE Trans.Signal Processing,* vol. 45,pp. 808-810, March. 1997

[5] A. Doganzic, and A. Nehorai, "Space-time fading channel estimation and symbol detection in unknown spatially correlated noise", *IEEE Trans.Signal Processing,* vol. 50,No. 3,pp. 457-473, March. 2002

[6] H.Krim and M.Viberg, "Two decades of array processing research", *IEEE Signal Processing Magazine,* pp. 67-93, July. 1996

[7] J.Li and R.T.Compton, "Maximum Likelihood angle estimation for signals with known waveforms", *IEEE Trans.Signal Processing,* vol. 41,pp. 2850-2862, Sept. 1993

[8] J.Li, B. Halder, P. Stoica, and M. Viberg, "Computationaly Efficient Angle Estimation for signals with known waveforms", *IEEE Trans.Signal Processing,* vol. 43,pp. 2154-2163, Sept. 1995

[9] Y. U. Lee,J. Choi, I. Song, S. R. Lee, "Distributed Source modeling and direction-of-arrival estimation techniques", *IEEE Trans.Signal Processing,* vol. 45,No. 4,pp. 960-968, April. 1997

[10] A. Leshem, and A. van der Veen, "Direction-of-arrival estimation for constant modulus signals", *IEEE Trans.Signal Processing,* vol. 47,pp. 3125-3129, Nov. 1999

[11] M.T. Ma, *Theory and aplication of antenna arrays*, Wiley-Interscience,1974

[12] J.G. Proakis, *Digital Communications*, 3rd ed. New York: McGraw-Hill,1995

[13] H.O. Schmidt, "Multiple emitter location and signal parameter estimation", *IEEE Trans. Antennas Propagat.,* vol AP-34, no. 3,pp. 276-280, Mar. 1986

[14] P. Stoica, E.G. Larsson and A.B. Gershman, "The stochastic CRB for array processing: A textbook derivation", *IEEE Sig. Proc. Letters,* vol. 8,No. 5, pp. 148-150, May. 2001

[15] P. Stoica, and A. Nehorai, "MUSIC, Maximum Likelihood and Cramer-Rao bound", *IEEE Acoust., Speech,Signal Processing,* vol. 37,pp. 720-741, May. 1989

[16] H.L. Van Trees, "Detection, estimation and modulation theory", 2001, Wiley & Sons.

[17] M. Viberg, P. Stoica, and B. Ottersten, "Maximum likelihood array processing in spatially correlated noise fields using parameterized signals", *IEEE Trans.Signal Processing,* vol. 45,pp. 996-1004, April. 1995

[18] M. Wax and A. Leshem, "Joint estimation of time delays and directions of arrival of multiple reflections of a known signal ", *IEEE Trans.Signal Processing,* vol. 45,pp. 2477-2484, Oct. 1997

[19] M.Wax and I.Ziskind, "On unique localization of multiple sources by passive sensor arrays", *IEEE Trans.Acoust.,Speech,Signal Processing,* vol. 37,pp. 996-1000, July. 1989

[20] Anthony J.Weiss, "On the accuracy of a cellular location system based on RSS measurements", *IEEE Trans. Vehicular Technology* vol. 52,pp. 1508- 1518, Nov. 2003

[21] A. Zeira and B. Friedlander, "Direction of arrival estimation using parametric signal models", *IEEE Trans.Signal Processing,* vol. 44,pp. 339-349, Febr. 1996

[22] I.Ziskind and M.Wax, "Maximum Likelihood localization of multiple sources by alternating projection", *IEEE Trans.Acoust.,Speech,Signal Processing,* vol. 36,pp. 1553-1560, Oct. 1988