STUDIES ON DOA ESTIMATION IN THE PRESENCE OF MULTIPATH IN A FREQUENCY HOPPING SYSTEM

BAO QIMING

NATIONAL UNIVERSITY OF SINGAPORE

2004

STUDIES ON DOA ESTIMATION IN THE PRESENCE OF MULTIPATH IN A FREQUENCY HOPPING SYSTEM

BAO QIMING

(*M.ENG.*)

(*SCHOOL OF ELECTRICAL ENGINEERING*)

(*XI'AN JIAOTONG UNIVERSITY, CHINA*)

A THESIS SUBMITTED FOR THE DEGREE OF M.ENG. ON

COMMUNICATIONS

DEPATMENT OF ELECTRICAL AND COMPUTER ENGINEERING

NATIONAL UNIVERSITY OF SINGAPORE

2004

Acknowledgements

I would like to express my sincere thanks to my supervisors, Professor Ko Chi Chung and Dr. Chin Francois. Their profound knowledge, invaluable guidance and encouragement are not only of great benefit to my research work, but to my life as well.

Special thanks to my parents and my wife. It is you who has given me everlasting care, encouragement and support.

Here I also want to give my thanks to Dr. Zhi Wanjun for her kind help in my research work.

Thanks to all the research students and research fellows in Communications Lab at ECE department of NUS for your friendship and helpful discussion.

Greatly appreciated for the research scholarship offered by the Institute of Infocomm Research (I^2R) of Singapore, by which I could carry out my research work.

Table of Contents

Summary

Direction Of Arrival (DOA) estimation of Frequency Hopping (FH) signal sources is important and useful in a number of applications. FCC requires the future mobile communication systems to have the ability to accurately locate emergency calls made from mobile phones. In military communications using FH technique, DOA estimation is also required for both non-cooperative signal interception and jammer localization.

In this thesis, we propose a new method for DOA estimation in the presence of multipath propagation and Mutual Coupling (MC) in a FH system. We take these two effects into account, and derive a Maximum Likelihood (ML) estimator for both MC matrix and DOA estimation. We then formulate an iterative Alternating Minimization (AM) algorithm for finding the MC and DOA parameters in an alternate manner. To illustrate the performance of the technique, we simulate the scenario where narrowband signals transmitted from one far-field source impinge on a Uniform Linear Array (ULA) of two half-wavelength spaced antenna elements via two paths. The simulation results presented illustrate the convergence of the algorithm and its statistical efficiency at high SNR.

In addition, we introduce a new method for joint time-delay and DOA estimation using two antennas and one receiver for a FH system in the presence of multipath. We derive a Least Square (LS) estimator for both time-delay and DOA estimation. We then formulate an iterative Alternating Maximization (AM) algorithm to jointly estimate the time-delays and DOA parameters. The simulation results demonstrate the performance of the algorithm.

List of Symbols

a	Scalar
$\boldsymbol{a}$	Column vector
$\boldsymbol{A}$	Matrix
$\boldsymbol{A}^{+}$	Pseudo-inverse of a matrix $\boldsymbol{A}$
$(\boldsymbol{A})_{ij}$	The $(i, j)^{th}$ entry of a matrix $\boldsymbol{A}$
$(\boldsymbol{A})^{T}$	Transpose of a matrix $\boldsymbol{A}$
$(\boldsymbol{A})^{H}$	Hermitian transpose of a matrix $\boldsymbol{A}$
$(\boldsymbol{A})^{-1}$	Inverse of a matrix $\boldsymbol{A}$
$tr\{\boldsymbol{A}\}$	Trace of a matrix $\boldsymbol{A}$
$\boldsymbol{I}$	Identity matrix
$\boldsymbol{0}$	Zero matrix or zero vector
$\boldsymbol{\Pi}_{A}$	Projection matrix of matrix $\boldsymbol{A}$ ($\boldsymbol{\Pi}_{A} = \boldsymbol{A}\boldsymbol{A}^{+}$)
$\boldsymbol{\Pi}_{A}^{\perp}$	Complementary projection matrix of matrix $\boldsymbol{A}$ ($\boldsymbol{\Pi}_{A}^{\perp} = \boldsymbol{I} - \boldsymbol{\Pi}_{A}$)
$diag(a_1, \cdots, a_m)$	$\begin{bmatrix} a_1 & & 0 \\ & \ddots & \\ 0 & & a_m \end{bmatrix}$
$\boldsymbol{A} = \begin{bmatrix} \boldsymbol{A}_1 \\ \vdots \\ \boldsymbol{A}_n \end{bmatrix}$	Matrix consisting of sub-matrices $\boldsymbol{A}_1, \cdots, \boldsymbol{A}_n$
$\boldsymbol{A} = \begin{bmatrix} \boldsymbol{A}_{11} & \cdots & \boldsymbol{A}_{1q} \\ \vdots & \ddots & \vdots \\ \boldsymbol{A}_{p1} & \cdots & \boldsymbol{A}_{pq} \end{bmatrix}$	Partitioned matrix consisting of sub-matrices $\boldsymbol{A}_{ij}$

$vec[\bullet]$	Vectorization operator
$\text{var}(\bullet)$	Variance of a random variable
$(\bullet)^*$	Conjugate
$\|\bullet\|$	Norm
$\|\bullet\|_F$	Frobenius norm
$\text{Re}\{\bullet\}$	Real part of a complex number
$\text{Im}\{\bullet\}$	Imaginary part of a complex number
$E[\bullet]$	Expectation
$\sum$	Summation
$\prod$	Product
max	Maximum
min	Minimum
log	Logarithm
$\log_{10}$	Denary Logarithm
$\frac{df}{dx}$	Derivative of the function $f(x)$ with respect to x
$\frac{\partial f}{\partial x}$	Partial derivative of the function $f(\bullet)$ with respect to x

List of Abbreviations

AM	Alternate Minimization (or Maximization)
AWGN	Additive White Gaussian Noise
BPSK	Binary Phase Shift keying
CDMA	Code Division Multiple Access
CRLB	Cramer Rao Lower Bound
DML	Deterministic Maximum Likelihood
DOA	Direct Of Arrival
DS-CDMA	Direct Sequence – Code Division Multiple Access
ESPRIT	Estimation of Signal Parameters via Rotational Invariance Technique
FCC	Federal Communications Commission
FFH	Fast Frequency Hopping
FH	Frequency Hopping
FIM	Fisher Information Matrix
LPI	Low Probability of Intercept
LS	Least Square
MC	Mutual Coupling
MFSK	M-ary Frequency Shift Keying
ML	Maximum Likelihood
MN	Minimum Norm
MOF	Minimum Norm formulation
MOM	Method Of Moment
MUSIC	MUltiple SIgnal Classification

MVU	Minimum Variance Unbiased
NSF	Noise Subspace Fitting
PDF	Probability Density Function
PN	Pseudo-random Noise
RF	Radio Frequency
RMSE	Root Mean Square Error
SFH	Slow Frequency Hopping
SML	Stochastic Maximum Likelihood
SNR	Signal Noise Ratio
SSF	Signal Subspace Fitting
ULA	Uniform Linear Array
WNSF	Weighted Noise Subspace Fitting
WSS	Wide Sense Stationary

List of Figures

List of Tables

Chapter 1

Introduction

Over the past three decades, sensor array signal processing emerged as an active area of research which centers on the ability to fuse data collected at several sensors in order to carry out a given estimation task. This framework takes advantage of prior information on the data acquisition system, for example, array geometry, sensor characteristics and so on. The corresponding signal processing methods are often used to solve several real-world problems such as DOA estimation in radar, sonar and wireless communications.

1.1 INTTRODUCTION TO A WIRELESS COMMUNICATION SYSTEM

The development of wireless communications stems from the works of Oersted, Faraday, Gauss, Maxwell and Hertz during the nineteenth century. In 1897, Marconi patented a radio telegraph system and established the Wireless Telegraph and Signal Company which was probably the first wireless communication system put into commercial use [1]. However, the ability to provide wireless communications to an entire population was not even conceived until the cellular concept developed by Bell Laboratories emerged in the 1960s. In 1970s, the highly reliable, miniature and solid-state RF hardware was developed. Then, the wireless communications era was born [2]. A typical functional block diagram of a wireless communication system is shown in Figure 1.1.

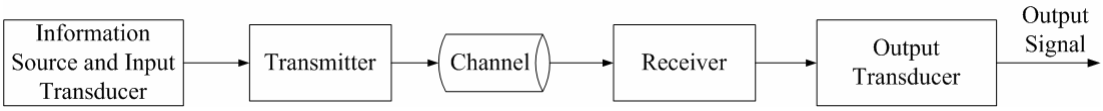


Figure 1.1 Functional block diagram of a wireless communication system [1].

1.2 WIRELESS COMMUNICATIONS THROUGH FADING MULTIPATH CHANNELS [3]

When electromagnetic wave propagates via sky wave in the high frequency range, a well-known phenomenon called signal multipath will occur. That is, the transmitted signal will arrive at the receiver via multiple propagation paths at different delays. Multipath propagation results in intersymbol interference in a digital communication system. Moreover, the signal components arriving via different propagation paths may add destructively, resulting in a phenomenon called signal fading.

1.2.1 Characterization of Fading Multipath Channels

In general, a narrowband band-pass signal takes the form

$$s(t) = \mathrm{Re}\left\{s_l(t)e^{j2\pi f_c t}\right\}, \tag{1.1}$$

where $s_l(t)$ is the equivalent low-pass signal, f_c is the carrier frequency, and $\mathrm{Re}\{\bullet\}$ denotes the real part of a complex-valued quantity. If we transmit this signal through the multipath propagation channel, a propagation delay and an attenuation factor will be associated with each path. The propagation delays and the attenuation factors are time-variant as a result of changes in the structure of the medium. Therefore the received band-pass signal can be expressed in the form

$$r(t) = \sum_n \alpha_n(t)s(t-\tau_n(t)), \tag{1.2}$$

where $\alpha_n(t)$ is the attenuation factor for the signal received on the *n*th path and $\tau_n(t)$ is the propagation delay for the *n*th path. And both of them are real-valued.

Substituting (1.1) back into (1.2) yields

$$r(t) = \mathrm{Re}\left\{\left[\sum_n \alpha_n(t) e^{-j2\pi f_c \tau_n(t)} s_l(t-\tau_n(t))\right] e^{j2\pi f_c t}\right\}. \tag{1.3}$$

Obviously, the equivalent low-pass received signal is

$$r_l(t) = \sum_n \alpha_n(t) e^{-j2\pi f_c \tau_n(t)} s_l(t-\tau_n(t)). \tag{1.4}$$

Note that if we define

$$\theta_n(t) = 2\pi f_c \tau_n(t), \tag{1.5}$$

then the received low-pass signal will be

$$r_l(t) = \sum_n \alpha_n(t) e^{-j\theta_n(t)} s_l(t-\tau_n(t)). \tag{1.6}$$

Further, by defining the complex-valued path attenuation factor

$$g_n(t) = \alpha_n(t) e^{-j\theta_n(t)}, \tag{1.7}$$

we finally get the received low-pass signal

$$r_l(t) = \sum_n g_n(t) s_l(t-\tau_n(t)). \tag{1.8}$$

Thus, it is the sum of a number of delayed versions of the transmitted baseband signal weighted by complex-valued path attenuation factors.

The multipath propagation model for the channel embodied in the received signal $r_l(t)$ given in (1.8), results in signal fading. The fading phenomenon is primarily due to the time variation in the phases $\{\theta_n(t)\}$. Note that large dynamic changes in the medium are required for $\alpha_n(t)$ to change sufficiently to cause a significant change in the complex-valued path attenuation factor $g_n(t)$ in equation

(1.7). However, on the other hand, $\theta_n(t)$ will change by 2π rad whenever $\tau_n(t)$ changes by $1/f_c$. Because $1/f_c$ is a very small number in the case of high frequency communications, $\theta_n(t)$ can change by 2π rad with relatively small motions of the medium.

The delays $\tau_n(t)$ associated with the different signal paths are usually expected to change at different rates and in an unpredictable manner. This implies that both the corresponding phase $\theta_n(t)$ and the received signal $r_l(t)$ can be modeled as a random processes. The fading phenomenon is primarily a result of the time variation in the phases set $\{\theta_n(t)\}$, or the stochastic disturbance of $\theta_n(t)$. That is, the randomly time variant phased $\{\theta_n(t)\}$ associated with the complex-valued path attenuation factors $\{g_n(t)\}$ result in the delayed versions of the transmitted baseband signal adding destructively. When that occurs, the resultant received signal $r_l(t)$ turns out to be very small. At other times, the delayed versions of the transmitted baseband signal may add constructively, so that $r_l(t)$ is large. Such amplitude variations in the received signal are termed signal fading. Moreover, it follows from (1.4) that the corresponding equivalent low-pass channel can be described by the linear time-variant impulse response

$$c(\tau;t)=\sum_n \alpha_n(t)e^{-j2\pi f_c\tau_n(t)}\delta(t-\tau_n(t)), \tag{1.9}$$

where $\delta(t)$ denotes the delta function.

Due to the assumption of random $\tau_n(t)$, the channel impulse response can also be modeled as a stochastic process. According to different statistical distribution, the linear time-variant channel can be divided into three main categories. If the channel

impulse response $c(\tau;t)$ is modeled as a zero-mean complex-valued Gaussian process, the channel is said to be Rayleigh fading channel for the envelope $|c(\tau;t)|$ being Rayleigh-distributed. In the presence of fixed scatterers or signal reflectors in the medium, in addition to randomly moving scatterers, the envelope $|c(\tau;t)|$ will no longer be Rayleigh-distributed but Ricean-distributed. Therefore, the corresponding channel is called Ricean fading channel. Another model is Nakagami fading channel, where the envelope $|c(\tau;t)|$ is Nakagami-*m* distributed.

Also, if we take the noise into consideration, the corresponding channel can be modeled as in Figure 1.2, where $n(t)$ is the Additive White Gaussian Noise (AWGN) which may come from atmospheric noise and thermal noise.

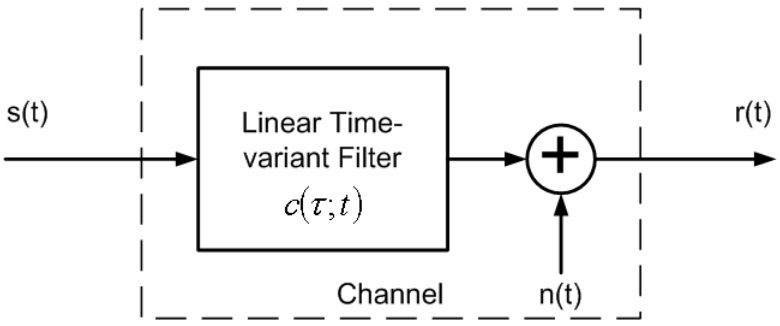


Figure 1.2 Linear time-variant model of the channel with additive noise [3].

1.2.2 Two Types of Distortions

Note that the equivalent low-pass received signal can be written as

$$r_l(t) = \int_{-\infty}^{+\infty} c(\tau;t)s_l(t-\tau)d\tau \,. \tag{1.10}$$

If we define $\boldsymbol{C}(f;t)$ and $\boldsymbol{S}_l(f)$ as the Fourier transformation of the channel impulse response $c(\tau;t)$ and transmitted baseband signal $s_l(t)$, respectively, then the received signal can also be expressed in the form

$$r_l(t) = \int_{-\infty}^{+\infty} \boldsymbol{C}(f;t)\boldsymbol{S}_l(f)e^{j2\pi ft}df \ . \tag{1.11}$$

Note that $\boldsymbol{C}(f;t)$ is called time-variant transfer function, where f is the frequency variable.

Suppose that we are transmitting digital information over the channel by modulating the basic pulse $s_l(t)$ at a rate T, where T is the signaling interval. We may suffer two types of distortion as follows.

A. *Frequency selectivity*

We assume that the channel impulse response $c(\tau;t)$ is Wide Sense Stationary (WSS). Then the corresponding autocorrelation function is defined as

$$\varphi_c(\tau_1,\tau_2;\Delta t) = \frac{1}{2}E\left[c^*(\tau_1;t)c(\tau_2;t+\Delta t)\right], \tag{1.12}$$

where $E[\bullet]$ denotes statistical expectation.

Assuming that the scattering at two different delays is uncorrelated, we get

$$\varphi_c(\tau_1,\tau_2;\Delta t) = \varphi_c(\tau_1;\Delta t)\delta(\tau_1-\tau_2). \tag{1.13}$$

If we let $\Delta t = 0$, the resulting autocorrelation function $\varphi_c(\tau;0)$ is simply the average power output of the channel as a function of the time delay τ. For this reason, $\varphi_c(\tau;0)$ is called the multipath intensity profile. The range of values of τ over which $\varphi_c(\tau;0)$ is essentially nonzero is termed multipath spread and is denoted by T_m.

Frequency selectivity is caused by multipath spread T_m or, equivalently, the coherence bandwidth $(\Delta f)_c$ which is the reciprocal of the multipath spread.

If the coherent bandwidth $(\Delta f)_c$ is small compared with the bandwidth of the transmitted signal, the channel is said to be frequency-selective. In this case, $\boldsymbol{S}_l(f)$ is subjected to different gains and phase shifts across the band, and the signal will be severely distorted by the channel. On the other hand, if $(\Delta f)_c$ is large in comparison with the bandwidth of the transmitted signal, the channel is said to be frequency-nonselective. That is, all the frequency components of the transmitted baseband signal $s_l(t)$ undergo the same attenuation and phase shift in transmission through the channel.

B. *Fading*

In the frequency domain, we define the autocorrelation function of the time-variant transfer function $\boldsymbol{C}(f;t)$ as

$$\Phi_C(\Delta f;\Delta t)=\frac{1}{2}E\left[C^*(f_1;t)C(f_2;t+\Delta t)\right], \tag{1.14}$$

where $\Delta f = f_1 - f_2$. Defining $S_C(\Delta f;\lambda)$ as the Fourier transform of $\Phi_C(\Delta f;\Delta t)$ with respect to the variable Δt, we get

$$S_C(\Delta f;\lambda)=\int_{-\infty}^{+\infty}\Phi_C(\Delta f;\Delta t)e^{-j2\pi\lambda\Delta t}d\Delta t, \tag{1.15}$$

where λ is the Doppler frequency. With Δf set to zero, the above equation becomes

$$S_C(0;\lambda)=\int_{-\infty}^{+\infty}\Phi_C(0;\Delta t)e^{-j2\pi\lambda\Delta t}d\Delta t. \tag{1.16}$$

The function $S_C(0;\lambda)$ is called the Doppler power spectrum of the channel, which is a power spectrum that gives the signal intensity as a function of the Doppler frequency λ. The range of values of λ over which $S_C(0;\lambda)$ is essentially nonzero is termed the Doppler spread B_d.

Fading is caused by the time variations of the channel, which are roughly characterized by the Doppler spread B_d or, equivalently, by the coherence time $(\Delta t)_c$ which is the reciprocal of the Doppler spread.

If the signaling interval T satisfies the condition $T << (\Delta t)_c$, then the channel attenuation and phase shift are essentially fixed for the duration of at least one signaling interval. When this condition holds, the corresponding channel is called a slow fading channel; Otherwise, it will be referred to as fast fading channel.

A channel model can be well described by the combination of the two types of distortions above. The frequency non-selective and slow fading channel model is by far the simplest channel model to analyze. It yields insight into the performance characteristics for digital signaling on a fading channel and serves to suggest the type of signal waveforms that are effective in overcoming the fading caused by the channel.

1.3 DOA ESTIMATION IN A WIRELESS COMMUNICATION SYSTEM

Receiving arrays and related estimation or detection techniques have long been used in high frequency wireless communications. And they are expected to play an important role in accommodating a multiuser personnel communication scenario in the presence of severe multipath.

For example, one of the most important problems in a multiuser asynchronous environment is the inter-user interference, which can degrade the performance quite severely. This is also the case in a practical Code Division Multiple Access (CDMA) system, because the varying delays of different users generate non-orthogonal codes. To combat fading due to the severe multipath, the base stations in mobile

communication systems have been using spatial diversity for a long time. However, using an antenna array gradually becomes preferable, because it can introduce additional degrees of freedom which can be used to obtain higher selectivity. An adaptive receiving array can be steered in the direction of one user at a time, while simultaneously nulling interference from other users. To do so, the DOA should be estimated first.

In addition, DOA estimation can make directive transmission possible in the downlink of Direct Sequence (DS)-CDMA base station. In detail, the base station first estimates the DOA using which the main part of the user signal is received. Then, based on the assumption of direction reciprocity, this estimated direction is used on downlink by choosing the weights of adaptive array so that the radiation pattern is a lobe (or lobes) directed toward the desired user.

As a typical spread spectrum communication method, Frequency Hopping (FH) shares the same basic advantage as direct sequence, namely, the Low Probability of Intercept (LPI). However, to some extent, an FH signal is preferred over a DS spread spectrum signal due to the fact that FH has a better performance of anti-jamming. Furthermore, FH does not need the stringent synchronization and power control, which is usually required in a DS spread spectrum system. Therefore, FH is the prevailing spread-spectrum technique in military communications. FH has also been adopted in two commercial standards, i.e., IEEE 802.11 and Bluetooth.

DOA estimation of frequency hopping signal sources is important in a number of applications. For instance, in military communications using FH technique, DOA estimation is required for both non-cooperative signal interception and jammer localization.

1.4 ORGANIZATION OF THIS THESIS

This thesis deals with the DOA estimation problem in a frequency hopping system. In Chapter 1, a wireless communication system is briefly introduced. And then the distortion caused by the wireless communication channel is stressed. Finally, the significance of DOA estimation in a wireless communication system is addressed. Chapter 1 forms the broad background of the research topic.

In Chapter 2, the main DOA estimation methods are summarized, and the frequency hopping system is introduced. Also, the DOA estimation methods for a frequency hopping system are reviewed. Chapter 2 is the cornerstone of the whole thesis, which also serves as a "reference" to the following chapters.

Chapter 3 concentrates on ML estimation and Cramer Rao Lower Bound (CRLB), which will be used in Chapter 4 and 5.

In Chapter 4, a new method for DOA estimation in the presence of Mutual Coupling (MC) and multipath propagation for a FH System is presented. We first present a signal model that takes MC and multipath propagation into account. Then, we derive a ML estimator for both MC and DOA estimation and introduce the associated Alternating Minimization algorithm. Furthermore, we present some simulation results to illustrate the performance of the technique, and finally we give some concluding remarks.

In Chapter 5, we propose a new method for joint time-delay and DOA estimation with a reduced number of receivers in the presence of multipath propagation for a frequency hopping system. First, we derive an LS estimator for both time delay and DOA estimation based on the corresponding signal model. Then, the

associated Alternating Maximization algorithm is introduced. Finally, we present some simulation results and concluding remarks.

Chapter 6 gives overall conclusion and some suggestions on future research work that can be done to extend the existing work.

Chapter 2

DOA Estimation in A Frequency Hopping (FH) System

For decades, Direction Of Arrival (DOA) estimation has been a hot research topic of signal processing in radar, sonar or wireless communications. Its purpose is to find the direction of interested signal emitted by a source or reflected from an object.

DOA estimation is of importance. For example, FCC requires the future mobile communication systems to have the ability to accurately locate emergency calls made from mobile phones. In military communications using FH technique, DOA estimation is also required for both non-cooperative signal interception and jammer localization.

2.1 INTRODUCTION TO DOA ESTIMATION METHODS

For convenience, we first introduce the signal model commonly used in array signal processing. Assuming that M signals impinge on the antenna array consisting of K antennas with arbitrary geometrical distribution, after the signals are down-converted to baseband and sampled, the array output vector takes the form

$$\boldsymbol{x}(n) = \sum_{m=1}^{M} \boldsymbol{a}(\theta_m) s_m(n) + \boldsymbol{n}(n), \tag{2.1}$$

where $\boldsymbol{a}(\theta_m)$ is the steering vector, $\boldsymbol{n}(n)$ is the Additive White Gaussian Noise (AWGN) vector, and $s_m(n)$ is the m^{th} signal. For notational convenience, we also define

$$\boldsymbol{A}(\boldsymbol{\theta}) = [\boldsymbol{a}(\theta_1), \cdots, \boldsymbol{a}(\theta_M)] \tag{2.2}$$

and

$$\boldsymbol{s}(n) = [s_1(n), \cdots, s_M(n)], \tag{2.3}$$

where $\boldsymbol{A}(\boldsymbol{\theta})$ is the steering matrix, $\boldsymbol{s}(n)$ is the signal waveforms vector. Thus, (2.1) can be written as

$$\boldsymbol{x}(n) = \boldsymbol{A}(\boldsymbol{\theta})\boldsymbol{s}(n) + \boldsymbol{n}(n), \tag{2.4}$$

Considering the noise at all antenna elements are spatially white with a common variance σ^2, the corresponding spatial covariance matrix is then defined by

$$\boldsymbol{R} = \boldsymbol{A}\boldsymbol{P}\boldsymbol{A}^H + \sigma^2 \boldsymbol{I}, \tag{2.5}$$

where $\boldsymbol{P}$ is termed the source covariance matrix which takes the form

$$\boldsymbol{P} = E[\boldsymbol{s}(n)\boldsymbol{s}^H(n)], \tag{2.6}$$

and $\boldsymbol{I}$ is the identity matrix. Here, for simplicity, we have suppressed the parentheses portion of $\boldsymbol{A}(\boldsymbol{\theta})$.

2.1.1 DOA Estimation for Noncoherent Signals

Assuming matrix $\boldsymbol{P}$ is nonsingular, the corresponding spectral factorization can be implemented, which partitions $\boldsymbol{P}$ into two parts, namely,

$$\boldsymbol{R} = \boldsymbol{V}_s \boldsymbol{\Lambda}_s \boldsymbol{V}_s^H + \boldsymbol{V}_n \boldsymbol{\Lambda}_n \boldsymbol{V}_n^H, \tag{2.7}$$

where $\boldsymbol{V}_s$ and $\boldsymbol{V}_n$ are matrices containing the signal and noise eigenvectors, respectively, $\boldsymbol{\Lambda}_s$ and $\boldsymbol{\Lambda}_n$ are diagonal matrices consisting of signal and noise eigenvalues, respectively. The columns of $\boldsymbol{V}_s$ span the vector space called signal subspace, and the columns of $\boldsymbol{V}_n$ span the vector space called noise subspace. Note that in this case, $\boldsymbol{\Lambda}_n = \sigma^2 \boldsymbol{I}$. The projection matrices onto the signal and noise subspaces are defined, respectively, as

$$\boldsymbol{\Pi}_{\boldsymbol{A}} = \boldsymbol{V}_s \boldsymbol{V}_s^H = \boldsymbol{A}\boldsymbol{A}^+ = \boldsymbol{A}\left(\boldsymbol{A}^H \boldsymbol{A}\right)^{-1} \boldsymbol{A}^H, \tag{2.8}$$

$$\boldsymbol{\Pi}_{\boldsymbol{A}}^{\perp} = \boldsymbol{V}_n \boldsymbol{V}_n^H = \boldsymbol{I} - \boldsymbol{\Pi}_{\boldsymbol{A}}. \tag{2.9}$$

In retrospection, we can see that the DOA estimation methods of the early-stage are non-model based and are therefore referred to as conventional methods. Such methods are typically based on classical beamforming techniques, in which the beam is scanned over the angular space of interest and the corresponding output power is measured, resulting what is termed spatial spectrum. The DOA estimation is obtained by locating peaks in the spatial spectrum. For this reason, the conventional methods are also named as spectral-based methods, among which are Bartlett beamformer, Capon's beamformer and so on.

Although it is computationally attractive, such spectral-based methods have limited estimation accuracy and angular resolution. To overcome these shortages, alternatively, the underlying data model is exploited much more extensively, which gives rise to the model-based methods. Because of the remarkable improvement in resolution, the model-based methods are also named as super-resolution methods.

Roughly speaking, the model-based methods can be classified into three categories. The first one is Maximum Likelihood (ML) method, which will be discussed in detail in Chapter 3.

The second one is subspace-based method. Schmidt [4] was the first to fully exploit the input data model structure for the case of sensor array of arbitrary form. Since then, the subspace approach has held tremendous interest. In his famous paper published in 1986 [4], Schmidt gave a perfect geometric interpretation of the DOA estimation problem in the absence of noise, and a reasonable approximation was also

obtained by extending it to the case in the presence of noise. That is, the signal subspace is orthogonal to the noise subspace, which results in $\boldsymbol{V}_n^H \boldsymbol{a}(\theta_m) = 0$. Schmidt called his proposed technique the MUltiple SIgnal Classification (MUSIC) algorithm, which estimates the DOAs by locating the peaks of a so-called MUSIC spatial spectrum defined by

$$P_{MUSIC}(\boldsymbol{\theta}) = \frac{\boldsymbol{a}^H(\boldsymbol{\theta})\boldsymbol{a}(\boldsymbol{\theta})}{\boldsymbol{a}^H(\boldsymbol{\theta})\boldsymbol{V}_n\boldsymbol{V}_n^H\boldsymbol{a}(\boldsymbol{\theta})}. \tag{2.10}$$

Assuming both the steering matrix $\boldsymbol{A}$ and source covariance matrix $\boldsymbol{P}$ are full rank, i.e., the column vectors of $\boldsymbol{A}$ are linearly independent for any DOA set $\{\theta_m\}$ and the signals are of least correlation, then the array will be unambiguous. That is, the corresponding DOA estimated is unique.

Compared with beamforming techniques, MUSIC method can theoretically obtain the DOA estimation of an arbitrary accuracy as long as the data collection time is long enough or the SNR is sufficiently high, or the signal model is adequately accurate. That is, the MUSIC algorithm provides statistically consistent estimation.

However, in practice, it may be not the case. As we have stated before, for MUSIC method to work, the corresponding array steering matrix $\boldsymbol{A}(\boldsymbol{\theta})$ must be full rank, i.e., each pair of column vectors must be linearly independent. Moreover, the associated source covariance matrix $\boldsymbol{P}$ should be non-singular. Therefore, in small samples and at low SNR scenarios, MUSIC may fail to resolve two closely spaced signals. Further, if the array is not well-calibrated, or signals are highly correlated, the MUSIC algorithm also can not work.

Besides MUSIC, the Estimation of Signal Parameters via Rotational Invariance Technique (ESPRIT) algorithm proposed by Roy and Kailath [5] and the Minimum-Norm (MN) algorithm proposed by Kumaresan and Tufts [6] also make great contribution to the subspace-based approach. Like MUSIC method, ESPRIT also relies on properties of the eigen-decomposition of the array covariance matrix.

However, because ESPRIT produces the DOA estimates directly in terms of the eigenvalues, its computational and storage requirement is much less than that of MUSIC where an exhaustive search through all possible steering vectors to is needed. Also, the corresponding array calibration requirement is not as stringent as that of MUSIC. ESPRIT derives its advantage at the cost of a much more stringent array structure. That means the array should possess a translational invariance, and the sensors should occur in matched pairs with identical displacement.

The basic idea behind ESPRIT is to exploit the rotational invariance of the underlying signal subspace induced by the rotational invariance of the sensor array. The MN method is actually an extension of MUSIC applied to uniform linear array, and the corresponding spatial spectrum is defined by

$$\boldsymbol{P}_{MN}(\theta) = \frac{\boldsymbol{a}^H(\theta)\boldsymbol{a}(\theta)}{\boldsymbol{a}^H(\theta)\boldsymbol{\Pi}_A^{\perp}\boldsymbol{e}_1\boldsymbol{e}_1^H\boldsymbol{\Pi}_A^{\perp}\boldsymbol{a}(\theta)}, \tag{2.11}$$

where $\boldsymbol{e_1}$ denotes the first column of the $L \times L$ identity matrix.

Although subspace-based methods offer significant performance improvement compared with the conventional beamforming methods, they usually exhibit large bias in finite samples, leading to resolution problems. This problem is especially notable for high source correlations.

The third one is subspace fitting method. Essentially, it is the subspace-based ML approximation. In [7], Viberg and Ottersten proposed the method of Signal Subspace Fitting (SSF). The main idea is as follows.

Note that in the general case, the source covariance matrix $\boldsymbol{P}$ may not be guaranteed to be of full rank M, where M represents the number of signals. Consider the case where the rank of $\boldsymbol{P}$ is $\tilde{M}$ ($\tilde{M} \le M$), then the columns of $\boldsymbol{V}_s$ will span a $\tilde{M}$-dimensional subspace of $\boldsymbol{A}$ which is of M dimension. Using the fact that $\boldsymbol{V}_s\boldsymbol{V}_s^H + \boldsymbol{V}_n\boldsymbol{V}_n^H = \boldsymbol{I}$, the spectral factorization can further be written as

$$\boldsymbol{A}\boldsymbol{P}\boldsymbol{A}^H + \sigma^2\boldsymbol{V}_s\boldsymbol{V}_s^H = \boldsymbol{V}_s\boldsymbol{\Lambda}_s\boldsymbol{V}_s^H. \tag{2.12}$$

Post-multiplying on the right by $\boldsymbol{V}_s$ and noting that $\boldsymbol{V}_s^H\boldsymbol{V}_s = \boldsymbol{I}$ give

$$\boldsymbol{V}_s = \boldsymbol{A}\boldsymbol{T}, \tag{2.13}$$

where

$$\boldsymbol{T} = \boldsymbol{P}\boldsymbol{A}^H\boldsymbol{V}_s\left(\boldsymbol{\Lambda}_s - \sigma^2\boldsymbol{I}\right)^{-1}. \tag{2.14}$$

Equation (2.13) forms the basis for SSF approach. Using the least-square estimator, the SSF estimate is given by

$$\left\{\hat{\boldsymbol{\theta}}, \hat{\boldsymbol{T}}\right\} = \arg\left\{\min_{\boldsymbol{\theta},\boldsymbol{T}} \left\|\boldsymbol{V}_s - \boldsymbol{A}\boldsymbol{T}\right\|_F^2\right\}, \tag{2.15}$$

where $\|\bullet\|_F^2$ denotes the Frobenius norm. And the corresponding concentrated objective function is obtained as

$$\hat{\boldsymbol{\theta}}_{SSF} = \arg\min_{\boldsymbol{\theta}} tr\left\{\boldsymbol{\Pi}_A^{\perp}\boldsymbol{V}_s\boldsymbol{\Lambda}_s\boldsymbol{V}_s^H\right\}, \tag{2.16}$$

where $tr\{\bullet\}$ denotes the trace of a matrix. As an extension, the SSF approach can be generalized as

$$\hat{\boldsymbol{\theta}} = \arg\min_{\boldsymbol{\theta}} tr\left\{\boldsymbol{\Pi}_A^{\perp}\boldsymbol{V}_s\boldsymbol{W}\boldsymbol{V}_s^H\right\}, \tag{2.17}$$

where $\boldsymbol{W}$ is a weighting matrix. It has been shown that if $\boldsymbol{W}$ is chosen as

$$\boldsymbol{W}_{opt} = \left(\boldsymbol{\Lambda}_s - \sigma^2 \boldsymbol{I}\right)^2 \boldsymbol{\Lambda}_s^{-1}, \tag{2.18}$$

the corresponding estimation accuracy will be maximized. This results in the well-known method of Weighted Signal Subspace Fitting (WSSF).

Alternatively, if we replace signal subspace with noise subspace, it will lead to Weighted Noise Subspace Fitting (WNSF) approach which can be dated back to Stoica and Sharman's work in [8]. The associated criterion is given by

$$\hat{\boldsymbol{\theta}} = \arg\min_{\boldsymbol{\theta}} tr\left\{\boldsymbol{A}^H \boldsymbol{V}_n \boldsymbol{V}_n^H \boldsymbol{A} \tilde{\boldsymbol{W}}\right\}, \tag{2.19}$$

where $\tilde{\boldsymbol{W}}$ is some positive or positive semidefinite weighting matrix.

It has shown that the WSSF and WNSF will asymptotically coincide if the following relationship holds [8]

$$\tilde{\boldsymbol{W}} = \boldsymbol{A}^+(\theta_o)\boldsymbol{V}_s \boldsymbol{W} \boldsymbol{V}_s^H \left[\boldsymbol{A}^+(\theta_o)\right]^*, \tag{2.20}$$

where θ_o represents the optimal value of DOA. Note that the subspace fitting method also involves a multi-dimensional search.

2.1.2 DOA Estimation for Coherent Signals

The subspace-based methods work on the premise that the signals impinging on the array are independent or low correlated. If such premise is violated, the corresponding signal source covariance matrix will no longer be of full rank. The performances of subspace-based methods will no doubt greatly degrade in the case of highly correlated or coherent signals as encountered in multipath propagation scenario where multiple versions of the same signal arrive in one resolvable chip duration.

To make the subspace-based methods work in the presence of coherent signals, some modifications have been proposed. The well-known one is spatial smoothing [9], which first splits the array into many identical sub-arrays and then averages the covariance matrix of each. The spatial smoothing induces a random phase modulation which in turn tends to decorrelate the signals that caused the rank deficiency. Although spatial smoothing extends the subspace-based methods to the scenario where coherent signals exist while retaining the computational efficiency of the one-dimensional search, it reduces the effective aperture of array due to the fact that the subarrays are smaller than the original array.

The other is Multi-Dimensional MUSIC (MD-MUSIC) [10], where all of the coherent signals are grouped together as a single signal. Thus the full rank condition of the source covariance matrix still holds. However, now the steering matrix $\boldsymbol{A}$ will not consist of steering vectors corresponding to distinct DOA, but group of coherent signals instead. Therefore, MD-MUSIC has to search through all possible steering vectors to find peaks in the spectrum. For example, in the case where a single source arrives at an antenna array through M_{mp} paths, the corresponding MD-MUSIC has to perform a M_{mp} -dimensional search. Therefore, MD-MUSIC is computationally expensive.

Because of the drawbacks of spatial smoothing and MD-MUSIC, the ML method is frequently used in the scenario of coherent signals [11 - 15].

2.2 INTRODUCTION TO A FREQUENCY HOPPING SYSTEM

Spread spectrum signals are characterized by the fact that their bandwidth W is much bigger than the information rate R (bits per second). The two basic spread-spectrum techniques are Direct-Sequence (DS) and Frequency Hopping (FH).

In a FH spread spectrum communications system, the available bandwidth is divided into a large number of sub-bands referred to as frequency slots. The transmitted signal can occupy one or more of the frequency slots in any signaling duration. The selection of the corresponding frequency slot(s) in each signaling duration is made in a pseudorandom manner. Roughly speaking, taking a basic modulation and at the same time changing the carrier frequency by a pseudorandom mechanism is the FH approach to generate a spread-spectrum signal.

Based on the different modulation methods adopted, FH signals can be mainly classified into FH/BPSK and FH/MFSK. In some wireless communication applications, FH/MFSK is a little bit more popular than FH/BPSK. One reason for that is FH/MFSK system has a simpler receiver. That is to say, for FH/MFSK receiver, it only needs to detect the signal frequency. While for that of FH/BPSK, it has to detect both the signal amplitude and frequency. The other reason is that FH/MFSK is much more insensitive to rapid phase variations in the received signal that may occur in mobile communication channel. Figure 2.1 illustrates the typical FH system.

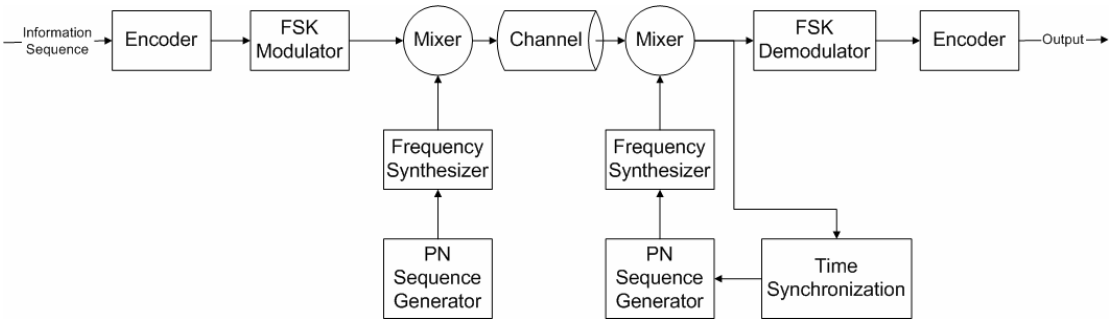


Figure 2.1 Block diagram of a FH spread spectrum system [3].

Also, according to the different hopping rate used, FH signal can be categorized into two classes, namely, Fast FH (FFH) and Slow FH (SFH). For a FFH signal, there are multiple hops per symbol. While for a SFH signal, there are one or multiple symbols per hop. SFH is often used in the scenario of flat (or frequency non-selective) fading and narrow-band signal, where the signal information bandwidth is centered at the sum of carrier frequency and instantaneous hopping frequency at any moment in time. And the instantaneous bandwidth is equal to the information bandwidth. Consequently, the total spread bandwidth is occupied only over longer time periods.

FFH is preferable for the scenario of frequency-selective fading and wideband signal, where the information signal is transmitted as a number of chips to fully exploit the frequency diversity. With each of them being transmitted at a different frequency, the chips will be faded independently, and thus only a portion of them will be greatly distorted. This is different form narrowband systems where the entire symbol can be faded.

As a typical spread spectrum communication method, frequency hopping shares the same basic advantage as direct sequence, namely, the Low Probability of Intercept (LPI). However, to some extent, an FH signal is preferred over a DS spread

spectrum signal due to the following fact. First, FH has a better performance of anti-jamming. For example, FFH can effectively prevent the follower jammer from creating interference. Second, FH does not need the stringent synchronization and power control, which is usually required in a DS spread spectrum system. Therefore, FH is the prevailing spread-spectrum technique in military communications. FH has also been adopted in two commercial standards, i.e., IEEE 802.11 and Bluetooth.

2.3 DOA ESTIMATION IN A FREQUENCY HOPPING SYSTEM

DOA estimation or localization of frequency hopping signal sources is important or useful in a number of applications. For instance, in military communications using FH technique, DOA estimation is required for both non-cooperative signal interception and jammer localization. Furthermore, the corresponding solution should be blind, i.e., unlike the mobile communication signals, both of the hopping sequence and the hopping instants are unknown and therefore have to be estimated blindly.

However, there are few papers on this topic [16 - 20]. In [16], Lemma et al. proposed a method to jointly estimate the DOA and frequency for SFH signals. But it will introduce some defects to the perfection of the algorithm because of the assumption that the hopping instants are known. Further, when generalizing from the scenario of one FH signal to that of multiple FH signals, the authors assume that only one signal hops at certain time instant, which may not the case in practice.

In [17], Wong presented a DOA estimation method for wideband-FFH signals using one electromagnetic vector-sensor. Unlike conventional spatially distributed arrays, the vector-sensor introduced by Wong eliminates the time-delay difference

caused by different DOAs and thus make the vector-sensor array-manifold independent of the impinging signal frequency. This characteristic will no doubt significantly simplify the received signal model. Although it is said that such novel electromagnetic vector-sensor is already commercially available, yet it has not been the mainstream of the relevant research work.

In [18 - 19], Liu et al. proposed a method to blindly localize and track multiple FH signals. The key point of their method is to model the snapshot as a 2-D harmonic mixture, and thus the DOA and frequency estimation is considered as a 2-D harmonic retrieval problem. However, their method is only applicable to the case of SFH/MFSK. And further, to get a high resolution solution, dynamic programming has to be used in their method, which will be of heavy computational load and the convergence of the algorithm can not be ensured as well.

In [20], sub-array was adopted to estimate the DOA for a HF system. The main purpose to do so is to reduce the number of receivers. The authors claimed that only two receivers were required. The one is fixed to a reference antenna, while the other is connected to the remaining elements of an antenna array by switching. Although such idea sounds like reasonable, yet there are some drawbacks. First, the parameter identifiability problem is not investigated, which is quite important in DOA estimation. Then, because the RF must switch at the same speed of hop, in FFH case, such RF switch may be unavailable. Finally, due to the basic assumption that the hopping sequence and hopping instants must be known, such a method is not applicable to military use.

What is more, while the methods discussed are efficient, they do not consider MC or multipath effect.

2.4 SUMMARY

In this chapter, we first give a thorough review on the research work done in the field of DOA estimation. Then, we also introduce the frequency hopping system. Finally, we review the DOA estimation method used in a FH system. This chapter is the cornerstone of the whole thesis, and it also serves as a “reference” to the following chapters.

Chapter 3

Maximum Likelihood (ML) DOA Estimation and Cramer-Rao Lower Bound (CRLB)

In this chapter, we briefly introduce the main mathematical tools which are used in the following chapters.

3.1 INTRODUCTION TO ESTIMATION IN SIGNAL PROCESSING

Modern signal processing systems such as radar, sonar, seismology and so on share a common problem of extracting parameters of interest based on continuous-time waveforms. Due to the use of digital computers to sample and store the continuous-time waveforms, this problem is equivalent to extract parameters of interest from a discrete-time waveform or data set.

Without loss of generality, mathematically, we have *N*-point data set $\{x(1), x(2), \cdots, x(N)\}$ which depends on a parameter of interest θ. For simplicity, we define the data vector

$$\boldsymbol{x} = [x(1), x(2), \cdots, x(N)]^T . \tag{3.1}$$

And we want to determine θ based on data or to define an estimator

$$\hat{\theta} = g(\boldsymbol{x}), \tag{3.2}$$

where g is some function. To determine a good estimator, the first step is to mathematically model the data. Because the data are inherently random, they are described by a Probability Density Function (PDF). The corresponding PDF will be different due to different assumptions of the parameter of interest θ.

3.1.1 Classical Estimation

If we assume that the parameter of interest θ is deterministic but unknown, the corresponding PDF will take the form $f_{\boldsymbol{X}(1),\cdots,\boldsymbol{X}(N)}(\boldsymbol{x};\theta)$. Estimation based on PDFs such as above is termed classical estimation. Note that the PDF is parameterized by the unknown parameter θ, i.e., we have a class of PDFs where each one is different due to a different value of θ. We will use a semicolon to denote this dependence.

3.1.2 Bayesian Estimation

If we assume that the parameter of interest θ is no longer deterministic but a random variable and assign it a PDF, the parameter we are attempting to estimate is then viewed as a realization of the random variable θ. And the corresponding PDF will be a joint PDF described by

$$f_{\boldsymbol{X}(1),\cdots,\boldsymbol{X}(N),\theta}(\boldsymbol{x},\theta) = f_{\boldsymbol{X}(1),\cdots,\boldsymbol{X}(N)|\theta}(\boldsymbol{x}\mid\theta) f_{\theta}(\theta), \tag{3.3}$$

where $f_{\theta}(\theta)$ is the prior PDF that summarizes our knowledge about θ before any data are observed, and $f_{\boldsymbol{X}(1),\cdots,\boldsymbol{X}(N)|\theta}(\boldsymbol{x}\mid\theta)$ is a conditional PDF summarizing our knowledge provided by the data $\boldsymbol{x}$ conditioned on knowing θ. Such an approach above is termed Bayesian estimation. Note that the results above can be extended to the case where we wish to estimate a vector parameter $\boldsymbol{\theta} = [\theta_1, \theta_2, \cdots, \theta_M]^T$ by replacing θ with $\boldsymbol{\theta}$.

When the PDF is viewed as a function of the unknown parameter (with $\boldsymbol{x}$ fixed), it is termed the likelihood function.

3.2 ML DOA ESTIMATION

ML estimation is achieved by finding out the value of the unknown parameter that maximizes the likelihood function. According to the different statistical assumption of signal generation process, ML methods for DOA estimation can be classified into two categories, namely, Deterministic ML (DML) method and Stochastic ML (SML) method.

3.2.1 DML DOA Estimation

In DML method, the noise is modeled as a stationary Additive White Gaussian Noise (AWGN) random process, whereas the signal waveforms are deterministic (arbitrary) and unknown. And the carrier frequencies are also assumed to be known. Therefore, DML DOA estimation belongs to the category of classical estimation.

Considering the signal model introduced in Chapter 2, after the signals are down-converted to baseband and sampled, the PDF of one snapshot is a *K*-variate complex Gaussian process

$$f_{X_1(n),\cdots,X_K(n)}\left(\boldsymbol{x}(n)\right)=\frac{1}{\left(\pi\sigma^2\right)^K}e^{-\|\boldsymbol{x}(n)-\boldsymbol{A}(\boldsymbol{\theta})\boldsymbol{s}(n)\|^2/\sigma^2}. \tag{3.4}$$

where $\boldsymbol{x}(n)$ is defined by

$$\boldsymbol{x}(n)=\left[x_1(n),x_2(n),\cdots,x_K(n)\right]^T. \tag{3.5}$$

Note that a snapshot is defined as the set of array outputs collected at a particular instant of time. Since the measurements are independent, the normalized negative log-likelihood function for *N* snapshots takes the form [41]

$$L_{Nor.\_Neg.}(\boldsymbol{\theta}, \boldsymbol{s}(n), \sigma^2) = -\frac{1}{N}\log L\left(\boldsymbol{\theta}, \boldsymbol{s}(n), \sigma^2\right) = K\log\sigma^2 + \frac{1}{N\sigma^2}\sum_{n=1}^{N}\left\|\boldsymbol{x}(n) - \boldsymbol{A}(\boldsymbol{\theta})\boldsymbol{s}(n)\right\|^2 . \tag{3.6}$$

And the corresponding DOA estimates can be obtained as

$$\hat{\boldsymbol{\theta}}_{DML} = \arg\left\{\min_{\boldsymbol{\theta}} tr\left\{\boldsymbol{\Pi}_{\boldsymbol{A}}^{\perp}\hat{\boldsymbol{R}}\right\}\right\}, \tag{3.7}$$

where

$$\hat{\boldsymbol{R}} = \frac{1}{N}\sum_{n=1}^{N}\boldsymbol{x}(n)\boldsymbol{x}(n)^H . \tag{3.8}$$

3.2.2 SML DOA Estimation

In SML method, however, the signal waveforms are modeled as AWGN random process. Therefore, SML DOA estimation belongs to the category of Bayesian estimation. The corresponding negative log-likelihood function is proportional to

$$\frac{1}{N}\sum_{n=1}^{N}\left\|\boldsymbol{\Pi}_{\boldsymbol{A}}^{\perp}\boldsymbol{x}(n)\right\|^2 = tr\left\{\boldsymbol{\Pi}_{\boldsymbol{A}}^{\perp}\hat{\boldsymbol{R}}\right\}. \tag{3.9}$$

And the DOA estimates are given by [11]

$$\hat{\boldsymbol{\theta}}_{SML} = \arg\left\{\min_{\boldsymbol{\theta}}\log\left\|\boldsymbol{A}\hat{\boldsymbol{P}}_{SML}(\boldsymbol{\theta})\boldsymbol{A}^H + \hat{\sigma}^2_{SML}(\boldsymbol{\theta})\boldsymbol{I}\right\|^2\right\}, \tag{3.10}$$

where

$$\hat{\boldsymbol{P}}_{SML}(\boldsymbol{\theta}) = \boldsymbol{A}^{+}(\hat{\boldsymbol{R}} - \hat{\sigma}^2_{SML}(\boldsymbol{\theta})\boldsymbol{I})(\boldsymbol{A}^{+})^H , \tag{3.11}$$

$$\hat{\sigma}^2_{SML}(\boldsymbol{\theta}) = \frac{1}{L-M}tr\left\{\boldsymbol{\Pi}_{\boldsymbol{A}}^{\perp}\right\}. \tag{3.12}$$

Note that the cost functions of ML DOA estimation are highly non-linear, and therefore non-linear programming techniques must be used. Although the ML method costs somewhat high computation, yet it achieves the optimal DOA estimation.

3.3 CRAMER-RAO LOWER BOUND (CRLB) [21]

In practice, being able to set a lower bound on the variance of any unbiased estimator proves to be extremely useful. At best, it allows us to assert that an estimator is the Minimum Variance Unbiased (MVU) estimator. This will be the case if the estimator attains the bound for all values of the unknown parameter. At worst, it provides us with a benchmark against which we can compare the performance of any unbiased estimator. Furthermore, it alerts us to the physical impossibility of finding an unbiased estimator whose variance is less than the bound. The latter is often useful in signal processing feasibility studies.

Although there are many such variance bounds, the Cramer-Rao Lower Bound (CRLB) is by far the easiest to determine. Also, the theory allows us to immediately determine if an estimator exists that attains the bound. If no such estimator exists, then all is not lost since estimators can be found that attain the bound in an approximate sense.

3.3.1 CRLB of a Scalar Parameter

Assume that the *PDF* $f_{\boldsymbol{X}(1),\cdots,\boldsymbol{X}(N)}(\boldsymbol{x};\theta)$ satisfies the "regularity" condition

$$E\left[\frac{\partial \log f_{\boldsymbol{X}(1),\cdots,\boldsymbol{X}(N)}(\boldsymbol{x};\theta)}{\partial\theta}\right]=0 \quad \text{for all } \theta, \tag{3.13}$$

where the expectation is taken with respect to $f_{\boldsymbol{X}(1),\cdots,\boldsymbol{X}(N)}(\boldsymbol{x};\theta)$. Then the variance of any unbiased estimator $\hat{\theta}$ satisfies

$$\mathrm{var}(\hat{\theta}) \geq \frac{1}{-E\left[\dfrac{\partial^2 \log f_{\boldsymbol{X}(1),\cdots,\boldsymbol{X}(N)}(\boldsymbol{x};\theta)}{\partial\theta^2}\right]}, \tag{3.14}$$

where the derivative is evaluated at the true value of θ and the expectation is taken with respect to $f_{\boldsymbol{X}(1),\cdots,\boldsymbol{X}(N)}(\boldsymbol{x};\theta)$. Furthermore, an unbiased estimator may be found that attains the bound for all θ if and only if

$$\frac{\partial \log f_{\boldsymbol{X}(1),\cdots,\boldsymbol{X}(N)}(\boldsymbol{x};\theta)}{\partial\theta} = Z(\theta)[g(\boldsymbol{x}) - \theta], \tag{3.15}$$

for some functions *g* and *Z*. That estimator, which is the MVU estimator, is $\hat{\theta} = g(\boldsymbol{x})$. And the minimum variance is $1/Z(\theta)$.

3.3.2 CRLB of a Vector Parameter

Assume that the *PDF* $f_{\boldsymbol{X}(1),\cdots,\boldsymbol{X}(N)}(\boldsymbol{x};\boldsymbol{\theta})$ satisfies the "regularity" conditions

$$E\left[\frac{\partial \log f_{\boldsymbol{X}(1),\cdots,\boldsymbol{X}(N)}(\boldsymbol{x};\boldsymbol{\theta})}{\partial\boldsymbol{\theta}}\right] = \boldsymbol{0} \quad \text{for all } \boldsymbol{\theta}, \tag{3.16}$$

where the expectation is taken with respect to $f_{\boldsymbol{X}(1),\cdots,\boldsymbol{X}(N)}(\boldsymbol{x};\boldsymbol{\theta})$ and $\boldsymbol{\theta} = [\theta_1, \theta_2, \cdots, \theta_M]^T$. Then the covariance matrix of any unbiased estimator $\hat{\boldsymbol{\theta}}$ satisfies

$$\boldsymbol{C}_{\hat{\theta}} - \boldsymbol{Z}^{-1}(\boldsymbol{\theta}) \geq \boldsymbol{0}, \tag{3.17}$$

where greater or equal to zero is interpreted as meaning that the matrix is positive semidefinite.

The $p \times p$ Fisher Information Matrix (FIM) is defined by

$$[\boldsymbol{Z}(\boldsymbol{\theta})]_{ij} = -E\left[\frac{\partial^2 \log f_{\boldsymbol{X}(1),\cdots,\boldsymbol{X}(N)}(\boldsymbol{x};\boldsymbol{\theta})}{\partial\theta_i \partial\theta_j}\right]. \tag{3.18}$$

where the derivatives are evaluated at the true value of $\boldsymbol{\theta}$ and the expectation is taken with respect to $f_{\boldsymbol{X}(1),\cdots,\boldsymbol{X}(N)}(\boldsymbol{x};\boldsymbol{\theta})$. An unbiased estimator may be found that attains the bound in that $\boldsymbol{C}_{\hat{\theta}} = \boldsymbol{Z}^{-1}(\boldsymbol{\theta})$ if and only if

$$\frac{\partial \log f_{\boldsymbol{X}(1),\cdots,\boldsymbol{X}(N)}(\boldsymbol{x};\boldsymbol{\theta})}{\partial\boldsymbol{\theta}} = \boldsymbol{Z}(\boldsymbol{\theta})[g(\boldsymbol{x}) - \boldsymbol{\theta}], \tag{3.19}$$

for some *p*-dimensional function $\boldsymbol{g}$ and some $p \times p$ matrix $\boldsymbol{Z}$. That estimator, which is the MVU estimator, is $\hat{\boldsymbol{\theta}} = g(\boldsymbol{x})$, and its covariance matrix is $\boldsymbol{Z}^{-1}(\boldsymbol{\theta})$.

3.4 SUMMARY

In this chapter, we briefly introduce the main mathematical tools which will be used in the following chapters. Due to the fact that it can be implemented for complicated estimation problems, ML estimator is a quite popular approach. In addition, for most cases of practical interest, the ML performance is optimal for large enough data records.

The Cramer-Rao Lower Bound (CRLB) is by far the easiest to determine. Also, it allows us to immediately determine if an estimator exists that attains the bound. Therefore, CRLB is extensively used in performance comparison.

Chapter 4

DOA Estimation in the Presence of Unknown Mutual Coupling and Multipath Propagation in a FH System

In early DOA estimation works, Mutual Coupling (MC) between the antenna elements was ignored and the antenna elements were assumed to be isotropic point sensors which sample but do not re-radiate the incident fields. However, in practice, MC is unavoidable.

Actually, when an antenna element receives the incident electromagnetic fields, it will at the same time reradiate part of them to other elements. The re-radiated EM fields interact with the other elements causing the sensors to be mutually coupled. This results in MC which will in turn affect the array manifold.

4.1 INTRODUCTION TO DOA ESTIMATION IN THE PRESENCE OF UNKNOWN MUTUAL COUPLING

As a matter of fact, with the emergence of high-resolution DOA estimation algorithms, MC becomes an important factor to be taken into consideration in the model. As we know, the high-resolution DOA estimation algorithms need an accurate knowledge of the array manifold (i.e. array response for any DOA). Otherwise, the corresponding performance will be significantly deteriorated.

To a great extent, the uncertainty in the array manifold (or model mismatch) stems from the electromagnetic perturbations of the antenna array, and an important characteristic of which is none other than MC.

Gupta and Ksienski may be the first who systematically studied the effect of MC on the performance of adaptive arrays [22]. By regarding an antenna array as a bilateral network responding to an outside source, they proposed that the MC effect can be characterized by a mutual impedance matrix (or MC matrix). The "true" vector of element output voltages should be the normalized by the MC matrix. In their work, they found that MC affects the performance of adaptive arrays even for large interelement spacing. For small interelement spacing case, such effect is much more remarkable and the corresponding array output SNR is significantly lower than that obtained in the scenario where MC is ignored.

In fact, a number of algorithms and techniques have been proposed for DOA estimation in the presence of MC in the past two decades. The "true" array steering vector is then obtained as the product of MC matrix and the "ideal" one, based on which succeeding DOA estimation can be performed with ordinary methods. So, how to get the MC matrix and do the corresponding compensation is of most importance. Generally, there are two conventional methods for MC compensation.

4.1.1 Array Calibration

Array calibration involves accurate measurement of the RF channel associated with each element of the array, a basic form of which is to collect the corresponding array manifold information. This is usually done by moving a pilot source over a fine grid of directions covered by the array and measuring the corresponding array response vectors for these directions. And the array response vector between the grid steps is usually estimated by interpolation. Also, it can be performed by rotating the array relative to a fixed source at closely spaced angles, say, ϕ_i.

Throughout the whole calibration process, the array response at each angle, $\boldsymbol{u}(\theta_i)$, is recorded and normalized to estimate the steering vectors $\boldsymbol{a}(\theta_i)$ which will further be mapped to corresponding DOA. However, the process of array calibration can be time consuming and costly. That is, array calibration has to be done from time to time, because array manifold may change over time due to many factors such as the behavior of sensor itself, environment around sensor array, location of the sensors and so on.

Again, storing the array manifold once it has been measured may take a great number of system memories. In practice, array calibration may fail to achieve the precision of array manifold that super-resolution DOA estimation algorithms require. Thus the associated system performance will be remarkably degraded, sometimes making it worse than that of the conventional methods. What is more, in the case of unknown multipath propagation, it is hardly possible to implement array calibration.

To overcome this problem, a method referred to as self-calibration came out, which make use of the received signal to perfect the array calibration. In [23 - 24] Rockah and Schultheiss studied in detail the self-calibration problem for the scenario of passive sensor arrays with sensor locations being imprecisely known.

In [25] Paulraj and Kailath developed a method to estimate DOA and meanwhile calibrate gain and phase perturbation. However, the method is significantly limited in that there exists a rotation uncertainty in the corresponding DOA estimation.

In [26], Brown et al. proposed a method to calibrate gain, phase and MC uncertainty using a single source of unknown position. This method works based on

the fundamental assumption that the initial DOA estimation is roughly near the true value.

In [27] Pierre and Kaveh developed a calibration procedure based on least square fitting. In [28] Jaffer proposed an array calibration method which estimates the sparse distortion matrix subject to linear constraints that force zeros in known position in the estimated distortion matrix. Note that because at least a single source has to be known or partially known, the method introduced in [27 - 28] is not a self-calibration in strict sense.

4.1.2 Fundamental Electromagnetic Viewpoint

This method deals with the electromagnetic model of the sensor array. The MC parameters are estimated from the actual antenna voltages. Generally speaking, this method can be subdivided into two classes according the different antenna model used.

The infinite array model fits the central elements of a large but finite array quite well and implicitly includes the MC effect. The corresponding MC compensation method is discussed in [29]. The finite phased array model is especially critical in adaptive array applications. This model gives much more practical information than the former, but much more difficult to analyze as well. For this model, the Method of Moment (MOM) is quite often proposed to model and eliminate the MC for signals impinging on the array from a given DOA, which converts close-form integrodifferential equations to a numerical matrix equation [30 -34].

However, due to the complex numerical computation, the MOM method is almost unacceptable for the case where the number of DOAs is greater than 1. Instead,

two viewpoints for MC compensation based on the finite phased array model have been proposed. The first one is MC compensation using open-circuit voltages which was proposed by Gupta and Ksienski [22]. They argue that the open-circuit voltages are free of MC because no terminal current exists and therefore the reradiated fields are reduced. And they derive the open-circuits voltages from the measured voltages. However, their theory is valid only for half-wave dipoles with half-wavelength spacing.

The second one is MC compensation using the Minimum Norm Formulation (MOF). In [34], Adve and Sarkar presented an MC compensation method for direct data domain adaptive algorithms, in which the MOM admittance matrix is used to estimate the incident fields that would generate the received voltages. However, their method is only applicable for linear arrays of linearly polarized dipole elements. This method is further generalized to general antenna arrays with arbitrary-shaped elements, called MOF [35]. Although MOF does not require the solution to the entire MOM problem, yet its computation load is still heavy, which may prevent it from being popular.

Alternatively, we can jointly estimate both the MC matrix and DOA directly by carrying out some statistical signal processing on the received signals [36 - 40]. In [36], Friedlander and Weiss proposed a modified MUSIC algorithm to estimate DOA and array calibration parameters including mutual coupling for a narrowband circular array. This method is iterative in nature. However, this algorithm was proven to have nonunique solutions for some special cases [37].

In some recent studies, Svantesson [38] introduced an algorithm based on a modified version of Noise Subspace Fitting (NSF), where the MC matrix is simplified

to a banded symmetric Toeplitz matrix. The main idea is as follows. First, the conventional NSF algorithm is extended to the case where MC is present. Then concentrate the NSF criterion with respect to MC parameters, and the DOAs are accordingly obtained by a numerical search. However, the global convergence of the algorithm can not be ensured because the performance surface on which the multi-dimensional search is done usually has locally optimum solutions.

In [39], Mao et al. proposed a separable dimension subspace method to estimate signal frequencies, direction of arrivals and MC jointly. The frequency is estimated first by subspace method in temporal dimension, and after that, the estimates of DOA and MC are obtained by subspace method in spatial dimension. Because it is not iterative, the accuracy of DOA and MC estimation greatly depends on that of the frequency estimation in the first step. Therefore, any small error generated in the first step may cause considerable errors in DOA or MC estimations.

In [40], Jaffer proposed a method to iteratively estimate the DOA and distortion matrix which accounts for the gain and phase errors and MC effect. The author argued that the MC effect can be negligible for elements separated by a few wavelengths and therefore the distortion matrix should be a sparse one. Based on this argument, a constrained MC estimation method is developed.

In summary, all these algorithms are only effective in scenarios where the transmitted signals are noncoherent and do not exhibit multipath propagation. In typical wireless communication systems where multipath propagation is unavoidable, these algorithms may not be effective.

4.2 SIGNAL MODEL

Consider a multipath propagation scenario where a far-field narrowband signal impinges on a Uniform Linear Array (ULA) with K antenna elements via M distinct paths (as shown in Figure 4.1). In the presence of mutual coupling, the MC matrix of the antenna array will be symmetric and Toeplitz [38].

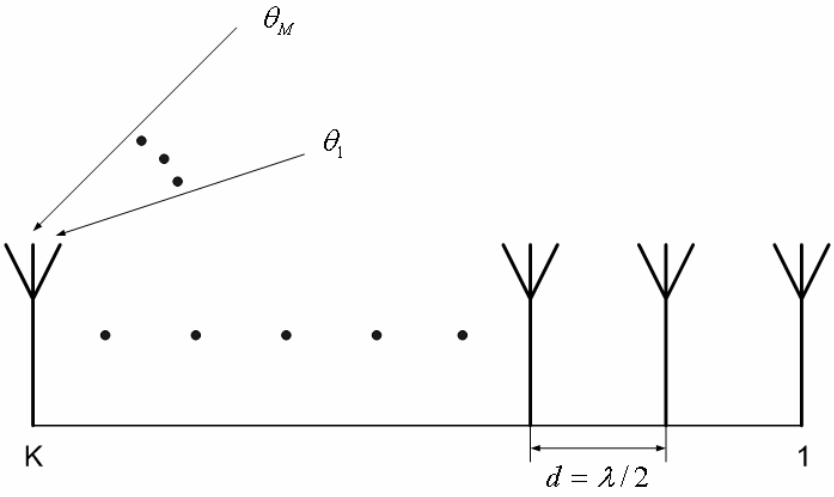


Figure 4.1 Receiving antenna array.

Assume also that frequency hopping is used and the transmitted signal in the h^{th} hop is

$$\varphi_h(t) = \mathrm{Re}\left\{ s_h(t) e^{j2\pi(f_c + f_h)t} \right\}. \tag{4.1}$$

where f_c is the center frequency, while f_h denotes the hop frequency. After down conversion, the array output vector becomes

$$\boldsymbol{x}_h(t) = \left[\boldsymbol{x}_{1h}, \cdots, \boldsymbol{x}_{Kh} \right]^T = \boldsymbol{C}\boldsymbol{A}_h(\boldsymbol{\theta})\boldsymbol{\Phi}_h \boldsymbol{G}\boldsymbol{s}_h(t) e^{j2\pi f_h t} + \boldsymbol{n}_h(t), \tag{4.2}$$

where $\boldsymbol{C}$ is the $K \times K$ MC matrix, $\boldsymbol{A}_h(\boldsymbol{\theta})$ is the steering matrix, $\boldsymbol{\Phi}_h$ is the phase shift matrix, $\boldsymbol{G}$ is the multipath attenuation matrix and $\boldsymbol{s}_h(t)$ is a M-dimensional baseband signal vector given by

$$\boldsymbol{C} = \begin{bmatrix} c_1 & c_2 & \cdots & c_K \\ c_2 & c_1 & \ddots & \vdots \\ \vdots & \ddots & \ddots & c_2 \\ c_K & \cdots & c_2 & c_1 \end{bmatrix} = \begin{bmatrix} 1 & c_2 & \cdots & c_K \\ c_2 & 1 & \ddots & \vdots \\ \vdots & \ddots & \ddots & c_2 \\ c_K & \cdots & c_2 & 1 \end{bmatrix}, \tag{4.3}$$

$$\boldsymbol{A}_h(\boldsymbol{\theta}) = \left[\boldsymbol{a}_h(\theta_1), \cdots, \boldsymbol{a}_h(\theta_M)\right], \tag{4.4}$$

$$\boldsymbol{a}_h(\theta_i) = \left[1, e^{-j2\pi(f_c+f_h)\beta_i}, \cdots, e^{-j2\pi(f_c+f_h)(K-1)\beta_i}\right]^T,\ i = 1,...,M, \tag{4.5}$$

$$\boldsymbol{\Phi}_h = diag\left(e^{-j2\pi(f_c+f_h)\tau_1}, \cdots, e^{-j2\pi(f_c+f_h)\tau_M}\right), \tag{4.6}$$

$$\boldsymbol{G} = diag\left(g_1, \cdots, g_M\right) \tag{4.7}$$

and

$$\boldsymbol{s}_h(t) = \left[s_h(t-\tau_1), \cdots, s_h(t-\tau_M)\right]^T. \tag{4.8}$$

Note that in (4.5), β_i is given by $\beta_i = \dfrac{d\cos\theta_i}{v_{lt}}$, where d is the spacing between two neighboring antennas, and v_{lt} is the speed of light.

Without loss of generality, suppose that $s_h(t) = 1$ for H hops as when pilots are sent during initial training. The output vector of the k^{th} antenna in all H hops therefore becomes

$$\boldsymbol{x}^{(k)}(t) = \left[x_{k1}, \cdots, x_{kH}\right]^T = \sum_{m=1}^{M} g_m \boldsymbol{\Gamma}\boldsymbol{\zeta}_m + \boldsymbol{n}^{(k)}(t),\ k = 1, \cdots, K, \tag{4.9}$$

where

$$\boldsymbol{\Gamma} = diag\left(e^{-j2\pi f_1 t}, \cdots, e^{-j2\pi f_H t}\right), \tag{4.10}$$

$$\boldsymbol{\zeta}_m = \left[e^{-j2\pi(f_c+f_1)\tau_m}\boldsymbol{c}_k^T\boldsymbol{a}_1(\theta_m), \cdots, e^{-j2\pi(f_c+f_H)\tau_m}\boldsymbol{c}_k^T\boldsymbol{a}_H(\theta_m)\right]^T, \tag{4.11}$$

and $\boldsymbol{c}_k$ is the k^{th} column of the MC matrix $\boldsymbol{C}$. Pre-multiplying (4.9) by the matrix $\boldsymbol{\Gamma}^{-1}$ yields the normalized vector

$$\boldsymbol{x}_{Nor.}^{(k)}(t)=\sum_{m=1}^{M} g_m \boldsymbol{\zeta}_m + \boldsymbol{n}_{Nor.}^{(k)}(t). \tag{4.12}$$

The received signals at the K antennas in all H hops can then be characterized by the KH-dimensional signal vector

$$\boldsymbol{x}(t)=vec\left[\boldsymbol{x}_{Nor.}^{(1)}(t),\cdots,\boldsymbol{x}_{Nor.}^{(K)}(t)\right]=\boldsymbol{F}(\boldsymbol{\gamma})\boldsymbol{D}(\boldsymbol{\eta})\boldsymbol{g}+\boldsymbol{n}(t), \tag{4.13}$$

where $vec[\bullet]$ denotes the column vectorizing operator which stacks the column of a matrix in a column vector, $\boldsymbol{n}(t)$ is the noise vector, $\boldsymbol{\gamma}$ is a $(K-1)$-dimensional MC parameter vector, $\boldsymbol{\eta}$ is a $2M$-dimensional delay-DOA joint parameter vector, $\boldsymbol{g}$ is a M-dimensional multipath attenuation parameter vector, $\boldsymbol{F}(\boldsymbol{\gamma})$ is a $KH\times KH$ symmetric Toeplitz sparse matrix, $\boldsymbol{d}_m$ is a LH -dimensional vector and $\boldsymbol{D}(\boldsymbol{\eta})$ is a $KH\times M$ matrix given by

$$\boldsymbol{\gamma}=[c_2,...,c_K]^T, \tag{4.14}$$

$$\boldsymbol{\eta}=\left[\tau_{1,}\tau_{2,}\cdots,\tau_{M,}\beta_1,\beta_{2,}\cdots,\beta_M\right]^T, \tag{4.15}$$

$$\boldsymbol{g}=\left[g_1,g_2,\cdots,g_M\right], \tag{4.16}$$

$$\boldsymbol{F}(\boldsymbol{\gamma})=\begin{bmatrix} \overbrace{\begin{matrix} c_1 & 0 & \cdots & 0 \\ 0 & c_1 & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & c_1 \end{matrix}}^{H} & \cdots & \overbrace{\begin{matrix} c_K & 0 & \cdots & 0 \\ 0 & c_K & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & c_K \end{matrix}}^{H} \\ \vdots & \ddots & \vdots \\ \begin{matrix} c_K & 0 & \cdots & 0 \\ 0 & c_K & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & c_K \end{matrix} & \cdots & \begin{matrix} c_1 & 0 & \cdots & 0 \\ 0 & c_1 & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & c_1 \end{matrix} \end{bmatrix}_{KH}, \tag{4.17}$$

$$\boldsymbol{d}_m=\left[e^{-j\varpi_1\tau_m},\cdots,e^{-j\varpi_H\tau_m},\cdots,e^{-j\varpi_1[\tau_m+(K-1)\beta_m]},\cdots,e^{-j\varpi_H[\tau_m+(K-1)\beta_m]}\right]^T, \tag{4.18}$$

$$\boldsymbol{D}(\boldsymbol{\eta})=\left[\boldsymbol{d}_1,\cdots,\boldsymbol{d}_M\right] \tag{4.19}$$

Notice that hereafter we define $\omega_h = 2\pi\left(f_c + f_h\right),\ h = 1, \cdots, H$.

4.3 ML ESTIMATOR

Figure 4.2 illustrates the receiver structure of the *l* th element of the antenna array.

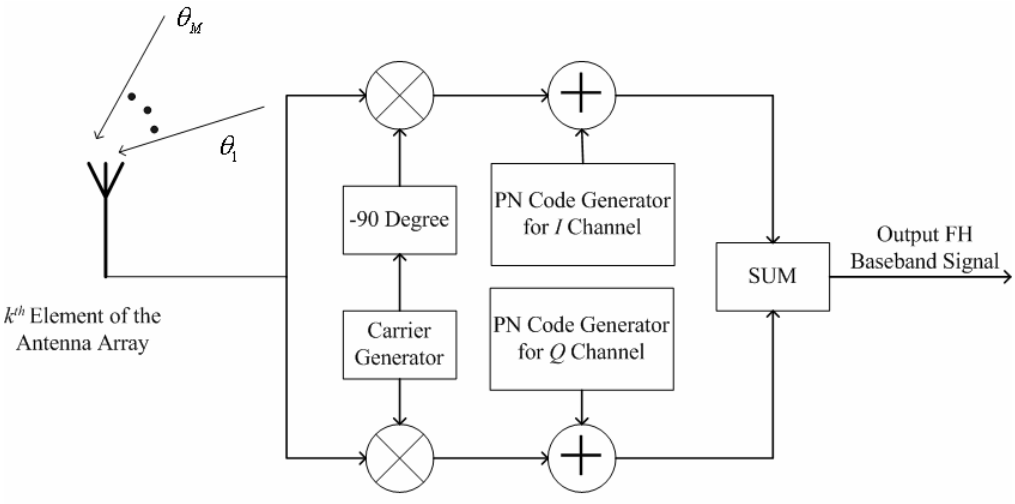


Figure 4.2 Block diagram of the receiver structure of the k^{th} element of the antenna array.

After sampling, the likelihood function of *N* snapshots takes the form

$$l\left(\sigma^2, \boldsymbol{\eta}, \boldsymbol{\gamma}, \boldsymbol{g}\right) = \prod_{n=1}^{N} \left(\pi\sigma^2\right)^{-K\times H} e^{-\frac{\left\|\boldsymbol{x}(n) - \boldsymbol{F}(\boldsymbol{\gamma})\boldsymbol{D}(\boldsymbol{\eta})\boldsymbol{g}\right\|^2}{\sigma^2}}, \tag{4.20}$$

where σ^2 is the variance of the sampled noise. The log-likelihood function is

$$L\left(\sigma^2, \boldsymbol{\eta}, \boldsymbol{\gamma}, \boldsymbol{g}\right) = -(K \times H \times N)\log\sigma^2 - \frac{1}{\sigma^2}\sum_{n=1}^{N}\left\|\boldsymbol{x}(n) - \boldsymbol{F}(\boldsymbol{\gamma})\boldsymbol{D}(\boldsymbol{\eta})\boldsymbol{g}\right\|^2 . \tag{4.21}$$

And the normalized negative log-likelihood function is [41]

$$L_{Nor.\_Neg.}\left(\sigma^2, \boldsymbol{\eta}, \boldsymbol{\gamma}, \boldsymbol{g}\right) = (K \times H)\log\sigma^2 + \frac{1}{N\sigma^2}\sum_{n=1}^{N}\left\|\boldsymbol{x}(n) - \boldsymbol{F}(\boldsymbol{\gamma})\boldsymbol{D}(\boldsymbol{\eta})\boldsymbol{g}\right\|^2 . \tag{4.22}$$

Therefore, the ML estimator can be obtained from solving

$$\left\{\hat{\sigma}^2, \hat{\boldsymbol{\eta}}, \hat{\boldsymbol{\gamma}}, \hat{\boldsymbol{g}}\right\} = \arg\min_{\sigma^2, \boldsymbol{\eta}, \boldsymbol{\gamma}, \boldsymbol{G}} \left\{(K \times H)\log\sigma^2 + \frac{1}{N\sigma^2}\sum_{n=1}^{N}\left\|\boldsymbol{x}(n) - \boldsymbol{F}(\boldsymbol{\gamma})\boldsymbol{D}(\boldsymbol{\eta})\boldsymbol{g}\right\|^2\right\} . \tag{4.23}$$

Minimizing with respect to σ^2 yields

$$\hat{\sigma}^2 = \frac{1}{(K \times H)N} \sum_{n=1}^{N} \left\| \boldsymbol{x}(n) - \boldsymbol{F}(\boldsymbol{\gamma}) \boldsymbol{D}(\boldsymbol{\eta}) \boldsymbol{g} \right\|^2 . \tag{4.24}$$

Substituting (4.24) back into (4.23) gives

$$\{\hat{\boldsymbol{\eta}}, \hat{\boldsymbol{\gamma}}, \hat{\boldsymbol{g}}\} = \arg \min_{\boldsymbol{\eta}, \boldsymbol{\gamma}, \boldsymbol{G}} J(\boldsymbol{\eta}, \boldsymbol{\gamma}, \boldsymbol{g}), \tag{4.25}$$

where the objective function $J(\boldsymbol{\eta}, \boldsymbol{\gamma}, \boldsymbol{g})$ is

$$J(\boldsymbol{\eta}, \boldsymbol{\gamma}, \boldsymbol{g}) = \frac{1}{N} \sum_{n=1}^{N} \left\| \boldsymbol{x}(n) - \boldsymbol{F}(\boldsymbol{\gamma}) \boldsymbol{D}(\boldsymbol{\eta}) \boldsymbol{g} \right\|^2 . \tag{4.26}$$

Note that equation (4.25) is a nonlinear $[2(K-1)+4M]$ -dimensional optimization problem. That is, there are $(K-1)$ complex-valued mutual coupling parameters, *M* real-valued time-delay parameters, *M* real-valued DOA parameters and *M* complex-valued multipath attenuation parameters. Furthermore, the unknown parameter vectors $\boldsymbol{\eta}$, $\boldsymbol{\gamma}$ and $\boldsymbol{g}$ cannot be completely separated. Therefore, straightforward minimization of (4.26) with respect to $\boldsymbol{\eta}$, $\boldsymbol{\gamma}$ and $\boldsymbol{g}$ is rarely feasible. To tackle this problem, we will now derive an Alternating Minimization (AM) algorithm for finding the unknown parameters in an iterative manner. Specifically, similar to that in [42], instead of minimizing (4.26) directly, the AM algorithm will iteratively perform minimization to obtain the two ML estimators in succession as follows.

A. *ML estimator for delay-DOA joint parameter vector*

With the MC parameter vector $\boldsymbol{\gamma}$ equal to $\hat{\boldsymbol{\gamma}}$, the ML estimation of $\boldsymbol{\eta}$ and $\boldsymbol{g}$ can be derived from (4.25) by minimizing

$$J_1(\boldsymbol{\eta}, \hat{\boldsymbol{\gamma}}, \boldsymbol{g}) = \frac{1}{N} \sum_{n=1}^{N} \left\| \boldsymbol{x}(n) - \boldsymbol{F}(\hat{\boldsymbol{\gamma}}) \boldsymbol{D}(\boldsymbol{\eta}) \boldsymbol{g} \right\|^2 . \tag{4.27}$$

Defining the matrix

$$\boldsymbol{B} = \boldsymbol{F}(\hat{\gamma})\boldsymbol{D}(\boldsymbol{\eta}), \tag{4.28}$$

we have, from (4.27),

$$\hat{\boldsymbol{g}} = \boldsymbol{B}^{+}\boldsymbol{m}_{x}, \tag{4.29}$$

where

$$\boldsymbol{B}^{+} = \left(\boldsymbol{B}^{H}\boldsymbol{B}\right)^{-1}\boldsymbol{B}^{H}, \tag{4.30}$$

$$\boldsymbol{m}_{x} = \frac{1}{N}\sum_{n=1}^{N}\boldsymbol{x}(n). \tag{4.31}$$

Substituting (4.29) back into (4.27) yields the concentrated objective function

$$J_{2}(\boldsymbol{\eta},\hat{\gamma}) = \frac{1}{N}\sum_{n=1}^{N}\left\|\boldsymbol{x}(n) - \boldsymbol{B}\boldsymbol{B}^{+}\boldsymbol{m}_{x}\right\|^{2} = tr\left\{\hat{\boldsymbol{R}} - \boldsymbol{\Pi}_{\boldsymbol{B}}\boldsymbol{M}_{x}\right\}, \tag{4.32}$$

where

$$\hat{\boldsymbol{R}} = \frac{1}{N}\sum_{n=1}^{N}\boldsymbol{x}(n)\boldsymbol{x}(n)^{H}, \tag{4.33}$$

$$\boldsymbol{\Pi}_{\boldsymbol{B}} = \boldsymbol{B}\boldsymbol{B}^{+} = \boldsymbol{B}\left(\boldsymbol{B}^{H}\boldsymbol{B}\right)^{-1}\boldsymbol{B}^{H}, \tag{4.34}$$

$$\boldsymbol{M}_{x} = \boldsymbol{m}_{x}\boldsymbol{m}_{x}^{H}. \tag{4.35}$$

Therefore, minimizing (4.27) is equivalent to maximizing

$$J_{3}(\boldsymbol{\eta},\hat{\gamma}) = tr\{\boldsymbol{\Pi}_{\boldsymbol{B}}\boldsymbol{M}_{x}\}, \tag{4.36}$$

or

$$\hat{\boldsymbol{\eta}} = \arg\max_{\boldsymbol{\eta}} tr\{\boldsymbol{\Pi}_{\boldsymbol{B}}\boldsymbol{M}_{x}\}. \tag{4.37}$$

Note that equation (4.37) is a nonlinear $2M$ -dimensional optimization problem. Thus the problem is not strictly well posed unless $K \times H \geq 2M$. If the number of antennas K is fixed, the number of hopping frequencies used should satisfy $H \geq \frac{2M}{K}$.

B. *ML estimator for MC parameter vector*

Likewise, with $\boldsymbol{\eta}$ and $\boldsymbol{g}$ given by $\hat{\boldsymbol{\eta}}$ and $\hat{\boldsymbol{g}}$, respectively, the ML estimation of the MC parameter vector $\boldsymbol{\gamma}$ can also be derived from (4.26) by minimizing

$$J_4(\hat{\boldsymbol{\eta}},\boldsymbol{\gamma},\hat{\boldsymbol{g}}) = \frac{1}{N}\sum_{n=1}^{N}\left\|\boldsymbol{x}(n) - \boldsymbol{F}(\boldsymbol{\gamma})\boldsymbol{D}(\hat{\boldsymbol{\eta}})\hat{\boldsymbol{g}}\right\|^2 . \tag{4.38}$$

Defining the KH-dimensional vector

$$\boldsymbol{u} = \boldsymbol{D}(\hat{\boldsymbol{\eta}})\hat{\boldsymbol{g}}, \tag{4.39}$$

(4.38) can be rewritten as

$$J_4(\hat{\boldsymbol{\eta}},\boldsymbol{\gamma},\hat{\boldsymbol{g}}) = \frac{1}{N}\sum_{n=1}^{N}\left\|\boldsymbol{x}(n) - \boldsymbol{Q}(\boldsymbol{u})\boldsymbol{\gamma}\right\|^2 , \tag{4.40}$$

where $\boldsymbol{Q}(\boldsymbol{u})$ is a $KH \times K$ matrix given by the sum of two $KH \times K$ matrices

$$\left[\boldsymbol{\Psi}_1\right]_{pq} = \begin{cases} \left[\boldsymbol{u}\right]_{p+(q-1)\times H}, & p+(q-1)\times H \le (K\times H) \\ 0, & otherwise \end{cases} \tag{4.41}$$

and

$$\left[\boldsymbol{\Psi}_2\right]_{pq} = \begin{cases} \left[\boldsymbol{u}\right]_{p-(q-1)\times H}, & p \ge H, q \ge 2, p-(q-1)\times H \ge 1 \\ 0, & otherwise \end{cases} . \tag{4.42}$$

Note that $c_1 = 1$ implies that minimizing equation (4.40) with respect to $\boldsymbol{\gamma}$ is a constrained optimization problem, with the constraint given by

$$\boldsymbol{\gamma}^H \boldsymbol{w} = 1, \tag{4.43}$$

and $\boldsymbol{w}$ being the K-dimensional vector

$$\boldsymbol{w} = \left[1,0,\cdots,0\right]^T . \tag{4.44}$$

Using the method of Lagrange Multipliers, we finally obtain

$$\hat{\boldsymbol{\gamma}} = \boldsymbol{Q}^{+}\boldsymbol{m}_X - \frac{\boldsymbol{w}^H\boldsymbol{Q}^{+}\boldsymbol{m}_X - 1}{\boldsymbol{w}^H\left[\boldsymbol{Q}^H\boldsymbol{Q}\right]^{-1}\boldsymbol{w}}\left[\boldsymbol{Q}^H\boldsymbol{Q}\right]^{-1}\boldsymbol{w}, \tag{4.45}$$

where $\boldsymbol{Q}^{+}$ denotes the pseudo-inverse of the matrix $\boldsymbol{Q}(\boldsymbol{u})$

$$\boldsymbol{Q}^{+} = \left(\boldsymbol{Q}^{H}\boldsymbol{Q}\right)^{-1}\boldsymbol{Q}^{H}. \tag{4.46}$$

Here, for simplicity, we have suppressed the parentheses portion of $\boldsymbol{Q}(\boldsymbol{u})$.

In the light of the results in (4.37) and (4.45), the corresponding AM algorithm can be formulated to find the MC and DOA parameters individually in succession. The algorithm is described in detail below.

A. *Global Initialization*

A.1. Set the iteration counter number k to 0.

A.2. Set the MC parameter vector $\boldsymbol{\gamma} = [c_1, c_2, ..., c_K]^T$ to $\hat{\boldsymbol{\gamma}}^{(0)} = [1,0,...,0]^T$. This is identical to the case where no MC exists.

A.3. Select the global initial values for the delay-DOA joint parameter vector $\boldsymbol{\eta}$ as $\hat{\boldsymbol{\eta}}^{(0)}$ which may come from some previous knowledge such as coarse estimation and so on.

B. *Iteration*

B.1. Estimating the delay-DOA joint parameter vector $\boldsymbol{\eta}$ and the path attenuation parameter vector **g**.

i. For the k$^{\text{th}}$ ($k \geq 1$) iteration, hold $\boldsymbol{\gamma}$ fixed as $\hat{\boldsymbol{\gamma}}^{(k-1)}$.

ii. Set the initial value of $\boldsymbol{\eta}$ as $\hat{\boldsymbol{\eta}}^{(k-1)}$. Obtain $\hat{\boldsymbol{\eta}}^{(k)}$ using the SD algorithm [43]: $\boldsymbol{\eta}^{(k)}(m+1) = \boldsymbol{\eta}^{(k)}(m) - \mu \nabla_{\boldsymbol{\eta}} J_3\left(\boldsymbol{\eta}, \hat{\boldsymbol{\gamma}}^{(k-1)}\right)\Big|_{\boldsymbol{\eta}=\boldsymbol{\eta}^{(k)}(m)}$ (the gradient of $J_3(\boldsymbol{\eta}, \hat{\boldsymbol{\gamma}})$ with respect to $\boldsymbol{\eta}$ is given in Section 4.3.1). Of course, we can also resort to the Gauss-Newton algorithm [43], if necessary (the Hessian matrix is given in Section 4.3.2).

iii. $\hat{\boldsymbol{g}}^{(k)} = \boldsymbol{B}^{+}(\hat{\boldsymbol{\eta}}^{(k-1)}, \hat{\boldsymbol{\gamma}}^{(k-1)})\boldsymbol{m}_{x}$.

B.2. Estimating the MC parameter vector $\boldsymbol{\gamma}$.

i. Fix $\boldsymbol{\eta}$ as $\hat{\boldsymbol{\eta}}^{(k)}$ and $\mathbf{g}$ as $\hat{\boldsymbol{g}}^{(k)}$.

ii. $$\hat{\boldsymbol{\gamma}}^{(k)} = \boldsymbol{Q}^{+}\boldsymbol{m}_{x} - \frac{\boldsymbol{w}^{H}\boldsymbol{Q}^{+}\boldsymbol{m}_{X} - 1}{\boldsymbol{w}^{H}\left[\boldsymbol{Q}^{H}\boldsymbol{Q}\right]^{-1}\boldsymbol{w}}\left[\boldsymbol{Q}^{H}\boldsymbol{Q}\right]^{-1}\boldsymbol{w}.$$

Once $\hat{\boldsymbol{\gamma}}^{(k)}$ is obtained, let $k = k+1$ and go back to step B.1 until convergence.

4.3.1 Gradient of $J_3(\boldsymbol{\eta}, \hat{\boldsymbol{\gamma}})$ with respect to $\boldsymbol{\eta}$

The gradient of $J_3(\boldsymbol{\eta}, \hat{\boldsymbol{\gamma}})$ with respect to $\boldsymbol{\eta}$ takes the form

$$\nabla_{\boldsymbol{\eta}} J_3(\boldsymbol{\eta}, \hat{\boldsymbol{\gamma}}) = \left[\frac{\partial J_3}{\partial \eta_1}, \cdots, \frac{\partial J_3}{\partial \eta_{2M}}\right]^{T}, \tag{4.47}$$

where

$$\frac{\partial J_3}{\partial \eta_i} = tr\left\{\frac{\partial \boldsymbol{\Pi}_{\boldsymbol{B}}}{\partial \eta_i}\boldsymbol{M}_{x}\right\}, \quad i = 1, \cdots, 2M, \tag{4.48}$$

$$\frac{\partial \boldsymbol{\Pi}_{\boldsymbol{B}}}{\partial \eta_i} = \boldsymbol{\Pi}_{\boldsymbol{B}}^{\perp}\dot{\boldsymbol{B}}_{\eta_i}\boldsymbol{B}^{+} + \left(\boldsymbol{\Pi}_{\boldsymbol{B}}^{\perp}\dot{\boldsymbol{B}}_{\eta_i}\boldsymbol{B}^{+}\right)^{H}, \tag{4.49}$$

$$\boldsymbol{\Pi}_{\boldsymbol{B}}^{\perp} = \boldsymbol{I} - \boldsymbol{\Pi}_{\boldsymbol{B}}, \tag{4.50}$$

$$\dot{\boldsymbol{B}}_{\eta_i} = \frac{\partial \boldsymbol{B}}{\partial \eta_i} = \boldsymbol{F}(\hat{\boldsymbol{\gamma}})\frac{\partial \boldsymbol{D}(\boldsymbol{\eta})}{\partial \eta_i} = \boldsymbol{F}(\hat{\boldsymbol{\gamma}})\dot{\boldsymbol{d}}_{\eta_i}, \tag{4.51}$$

$$\dot{\boldsymbol{d}}_{\eta_i} = \frac{\partial \boldsymbol{D}(\boldsymbol{\eta})}{\partial \eta_i} = \left[(\dot{\boldsymbol{d}}_1)_{\eta_i}, \cdots, (\dot{\boldsymbol{d}}_m)_{\eta_i}, \cdots, (\dot{\boldsymbol{d}}_M)_{\eta_i}\right]^{T}, \tag{4.52}$$

$$(\dot{\boldsymbol{d}}_m)_{\eta_i} = \frac{\partial \boldsymbol{d}_m}{\partial \eta_i} = \begin{cases} -j\delta_{m,j}\boldsymbol{\Omega}_1\boldsymbol{d}_m, & 1 \le i \le M; \\ -j\delta_{m,(i-M)}\boldsymbol{\Omega}_2\boldsymbol{d}_m, & M+1 \le i \le 2M \end{cases}, \tag{4.53}$$

$$\delta_{m,j} = \begin{cases} 1, & m = j \\ 0, & m \ne j \end{cases}, \tag{4.54}$$

where both of $\boldsymbol{\Omega}_1$ and $\boldsymbol{\Omega}_2$ are $KH \times KH$ diagonal matrices, and they can be expressed as

$$\boldsymbol{\Omega}_1 = diag\left(\omega_1, \cdots, \omega_H, \omega_1, \cdots, \omega_H, \cdots\cdots, \omega_1, \cdots, \omega_H\right), \tag{4.55}$$

$$\boldsymbol{\Omega}_2 = diag\left(0, \cdots, 0, \omega_1, \cdots, \omega_H, \cdots\cdots, (K-1)\omega_1, \cdots, (K-1)\omega_H\right). \tag{4.56}$$

4.3.2 Derivation of Hessian Matrix

Define $\left(\boldsymbol{H}\right)_{ij}$ as the $(i, j)^{th}$ entry of the Hessian matrix, we have

$$\left(\boldsymbol{H}\right)_{ij} = \frac{\partial^2 J_3}{\partial \eta_i \partial \eta_j} = 2\,\mathrm{Re}\left\{tr\left\{\left(\boldsymbol{\Pi}_B^{\perp}\,\dot{\boldsymbol{B}}_{\eta_i}\,\frac{\partial \boldsymbol{B}^{+}}{\partial \eta_j} + \boldsymbol{\Pi}_B^{\perp}\,\ddot{\boldsymbol{B}}_{\eta_i\eta_j}\,\boldsymbol{B}^{+} - \frac{\partial \boldsymbol{\Pi}_B}{\partial \eta_j}\dot{\boldsymbol{B}}_{\eta_i}\,\boldsymbol{B}^{+}\right)\boldsymbol{M}_x\right\}\right\}, \tag{4.57}$$

where

$$\frac{\partial \boldsymbol{B}^{+}}{\partial \eta_j} = \left(\boldsymbol{B}^H\boldsymbol{B}\right)^{-1}\left(\dot{\boldsymbol{B}}_{\eta_j}\right)^H \boldsymbol{\Pi}_B^{\perp} - \boldsymbol{B}^{+}\,\dot{\boldsymbol{B}}_{\eta_j}\,\boldsymbol{B}^{+}, \tag{4.58}$$

$$\ddot{\boldsymbol{B}}_{\eta_i\eta_j} = \boldsymbol{F}(\hat{\boldsymbol{\gamma}})\,\ddot{\boldsymbol{D}}_{\eta_i\eta_j} = \boldsymbol{F}(\hat{\boldsymbol{\gamma}})\left[(\ddot{\boldsymbol{d}}_1)_{\eta_i\eta_j}, \cdots, (\ddot{\boldsymbol{d}}_m)_{\eta_i\eta_j}, \cdots, (\ddot{\boldsymbol{d}}_M)_{\eta_i\eta_j}\right], \tag{4.59}$$

$$(\ddot{\boldsymbol{d}}_m)_{\eta_i\eta_j} = \begin{cases} -\delta_{m,i}\delta_{m,j}\boldsymbol{\Omega}_1^2\boldsymbol{d}_m, & 1 \le i, j \le M \\ -\delta_{m,i}\delta_{m,(j-M)}\boldsymbol{\Omega}_1\boldsymbol{\Omega}_2\boldsymbol{d}_m, & 1 \le i \le M, M+1 \le j \le 2M \\ -\delta_{m,(i-M)}\delta_{m,j}\boldsymbol{\Omega}_2\boldsymbol{\Omega}_1\boldsymbol{d}_m, & M+1 \le i \le 2M, 1 \le j \le M \\ -\delta_{m,(i-M)}\delta_{m,(j-M)}\boldsymbol{\Omega}_2^2\boldsymbol{d}_m, & otherwise \end{cases}. \tag{4.60}$$

4.3.3 Derivation of CRLB

For simplicity, we now break the delay-DOA joint parameter vector $\boldsymbol{\eta}$ into two parts $\boldsymbol{\eta} = \left[\boldsymbol{\tau}^T, \boldsymbol{\theta}^T\right]^T$, where $\boldsymbol{\tau}$ denotes the delay vector, and $\boldsymbol{\theta}$ denotes the DOA vector

$$\boldsymbol{\tau} = \left[\tau_1, \cdots, \tau_M\right]^T, \tag{4.61}$$

$$\boldsymbol{\theta} = \left[\theta_1, \cdots, \theta_M\right]^T . \tag{4.62}$$

For arbitrary complex variable $x = \mathrm{Re}\{x\} + j\,\mathrm{Im}\{x\}$, we define its partial derivative as $\frac{\partial}{\partial x} = \frac{1}{2}\left[\frac{\partial}{\mathrm{Re}\{x\}} + j\frac{\partial}{\mathrm{Im}\{x\}}\right]$. Hereafter, for notational simplicity, we will omit the parentheses portions of $\boldsymbol{F}(\boldsymbol{\gamma})$ and $\boldsymbol{D}(\boldsymbol{\eta})$.

A. *First order partial derivatives*

A.1. *First order partial derivatives with respect to multipath delays*

$$\frac{\partial L}{\partial \tau_i} = -2\frac{N}{\sigma^2}\mathrm{Re}\left\{\frac{1}{N}\sum_{n=1}^{N} tr\left\{\left[\boldsymbol{FDg} - \boldsymbol{x}(n)\right]\boldsymbol{g}^H\left(\dot{\boldsymbol{D}}_{\tau_i}\right)^H \boldsymbol{F}^H\right\}\right\}, \tag{4.63}$$

where

$$\dot{\boldsymbol{D}}_{\tau_i} = \left[\frac{\partial \boldsymbol{d}_1}{\partial \tau_i}, \cdots, \frac{\partial \boldsymbol{d}_m}{\partial \tau_i}, \cdots, \frac{\partial \boldsymbol{d}_M}{\partial \tau_i}\right]^T, \tag{4.64}$$

$$\frac{\partial \boldsymbol{d}_m}{\partial \tau_i} = \left(\dot{\boldsymbol{d}}_m\right)_{\tau_i} = -j\delta_{m,i}\boldsymbol{\Omega}_1\boldsymbol{d}_m, \quad 1 \le i, m \le M. \tag{4.65}$$

A.2. *First order partial derivatives with respect to multipath DOA*

$$\frac{\partial L}{\partial \theta_i} = -2\frac{N}{\sigma^2}\mathrm{Re}\left\{\frac{1}{N}\sum_{n=1}^{N} tr\left\{\left[\boldsymbol{FDg} - \boldsymbol{x}(n)\right]\boldsymbol{g}^H\left(\dot{\boldsymbol{D}}_{\theta_i}\right)^H \boldsymbol{F}^H\right\}\right\}, \tag{4.66}$$

where

$$\dot{\boldsymbol{D}}_{\theta_i} = \left[\frac{\partial \boldsymbol{d}_1}{\partial \theta_i}, \cdots, \frac{\partial \boldsymbol{d}_m}{\partial \theta_i}, \cdots, \frac{\partial \boldsymbol{d}_M}{\partial \theta_i}\right]^T, \tag{4.67}$$

$$\frac{\partial \boldsymbol{d}_m}{\partial \theta_i} = \left(\dot{\boldsymbol{d}}_m\right)_{\theta_i} = j\delta_{m,i}\boldsymbol{\Omega}_2\boldsymbol{d}_m\frac{d}{v_{lt}}\sin\theta_i, \quad 1 \le i, m \le M. \tag{4.68}$$

A.3. *First order partial derivatives with respect to multipath attenuations*

$$\frac{\partial L}{\partial \mathrm{Re}\{g_i\}} = -2\frac{N}{\sigma^2}\mathrm{Re}\left\{\frac{1}{N}\sum_{n=1}^{N} tr\left\{\left[\boldsymbol{FDg} - \boldsymbol{x}(n)\right]\dot{\boldsymbol{g}}_i \boldsymbol{D}^H \boldsymbol{F}^H\right\}\right\}, \tag{4.69}$$

$$\frac{\partial L}{\partial \mathrm{Im}\{g_i\}} = -2\frac{N}{\sigma^2}\mathrm{Im}\left\{\frac{1}{N}\sum_{n=1}^{N} tr\left\{\left[\boldsymbol{FDg} - \boldsymbol{x}(n)\right]\dot{\boldsymbol{g}}_i \boldsymbol{D}^H \boldsymbol{F}^H\right\}\right\}, \tag{4.70}$$

where

$$\dot{\boldsymbol{g}}_i = \frac{\partial \boldsymbol{g}^H}{\partial g_i}. \tag{4.71}$$

Note that $\dot{\boldsymbol{g}}_i$ is an *M*-dimension row vector with 1 in the i^{th} entry and zeros elsewhere.

A.4. *First order partial derivatives with respect to MC parameters*

$$\frac{\partial L}{\partial \mathrm{Re}\{c_i\}} = -2\frac{N}{\sigma^2}\mathrm{Re}\left\{\frac{1}{N}\sum_{n=1}^{N} tr\left\{\left[\boldsymbol{FDg} - \boldsymbol{x}(n)\right]^H \boldsymbol{gD}^H \dot{\boldsymbol{F}}_i\right\}\right\}, \tag{4.72}$$

$$\frac{\partial L}{\partial \mathrm{Im}\{c_i\}} = -2\frac{N}{\sigma^2}\mathrm{Im}\left\{\frac{1}{N}\sum_{n=1}^{N} tr\left\{\left[\boldsymbol{FDg} - \boldsymbol{x}(n)\right]\boldsymbol{g}^H \boldsymbol{D}^H \dot{\boldsymbol{F}}_i\right\}\right\}, \tag{4.73}$$

$$\dot{\boldsymbol{F}}_i = \frac{\partial \boldsymbol{F}^H}{\partial c_i}. \tag{4.74}$$

B. *Expectation of Second order partial derivatives*

B.1. *Expectation of Second order partial derivatives with respect to multipath delays*

$$E\left\{\frac{\partial^2 L}{\partial \tau_i \partial \tau_j}\right\} = -2\frac{N}{\sigma^2}\mathrm{Re}\left\{tr\left\{\boldsymbol{F}\,\dot{\boldsymbol{D}}_{\tau_j}\,\boldsymbol{gg}^H\left(\dot{\boldsymbol{D}}_{\tau_i}\right)^H \boldsymbol{F}^H\right\}\right\}. \tag{4.75}$$

B.2. *Expectation of Second order partial derivatives with respect to multipath DOA*

$$E\left\{\frac{\partial^2 L}{\partial \theta_i \partial \theta_j}\right\} = -2\frac{N}{\sigma^2}\mathrm{Re}\left\{tr\left\{\boldsymbol{F}\,\dot{\boldsymbol{D}}_{\theta_j}\,\boldsymbol{gg}^H\left(\dot{\boldsymbol{D}}_{\theta_i}\right)^H \boldsymbol{F}^H\right\}\right\}. \tag{4.76}$$

B.3. *Expectation of Second order partial derivatives with respect to multipath attenuations*

$$E\left[\frac{\partial^2 L}{\partial \mathrm{Re}\{g_i\}\partial \mathrm{Re}\{g_j\}}\right] = -2\frac{N}{\sigma^2}\mathrm{Re}\left\{tr\left\{\boldsymbol{F}\boldsymbol{D}\left(\dot{\boldsymbol{g}}_i\right)^H \dot{\boldsymbol{g}}_j \boldsymbol{D}^H \boldsymbol{F}^H\right\}\right\}, \tag{4.77}$$

$$E\left[\frac{\partial^2 L}{\partial \mathrm{Re}\{g_i\}\partial \mathrm{Im}\{g_j\}}\right] = -2\frac{N}{\sigma^2}\mathrm{Im}\left\{tr\left\{\boldsymbol{F}\boldsymbol{D}\left(\dot{\boldsymbol{g}}_i\right)^H \dot{\boldsymbol{g}}_j \boldsymbol{D}^H \boldsymbol{F}^H\right\}\right\}, \tag{4.78}$$

$$E\left[\frac{\partial^2 L}{\partial \mathrm{Im}\{g_i\}\partial \mathrm{Im}\{g_j\}}\right] = -2\frac{N}{\sigma^2}\mathrm{Re}\left\{tr\left\{\boldsymbol{F}\boldsymbol{D}\left(\dot{\boldsymbol{g}}_i\right)^H \dot{\boldsymbol{g}}_j \boldsymbol{D}^H \boldsymbol{F}^H\right\}\right\}. \tag{4.79}$$

B.4. *Expectation of Second order partial derivatives with respect to MC parameters*

$$E\left[\frac{\partial^2 L}{\partial \mathrm{Re}\{c_i\}\partial \mathrm{Re}\{c_j\}}\right] = -2\frac{N}{\sigma^2}\mathrm{Re}\left\{tr\left\{\dot{\boldsymbol{F}}_i \boldsymbol{D}\boldsymbol{g}\boldsymbol{g}^H \boldsymbol{D}^H \dot{\boldsymbol{F}}_j\right\}\right\}, \tag{4.80}$$

$$E\left[\frac{\partial^2 L}{\partial \mathrm{Re}\{c_i\}\partial \mathrm{Im}\{c_j\}}\right] = -2\frac{N}{\sigma^2}\mathrm{Im}\left\{tr\left\{\dot{\boldsymbol{F}}_i \boldsymbol{D}\boldsymbol{g}\boldsymbol{g}^H \boldsymbol{D}^H \dot{\boldsymbol{F}}_j\right\}\right\}, \tag{4.81}$$

$$E\left[\frac{\partial^2 L}{\partial \mathrm{Im}\{c_i\}\partial \mathrm{Im}\{c_j\}}\right] = -2\frac{N}{\sigma^2}\mathrm{Re}\left\{tr\left\{\dot{\boldsymbol{F}}_i \boldsymbol{D}\boldsymbol{g}\boldsymbol{g}^H \boldsymbol{D}^H \dot{\boldsymbol{F}}_j\right\}\right\}. \tag{4.82}$$

B.5. *Expectation of Second order partial derivatives with respect to multipath delay-DOA cross terms*

$$E\left[\frac{\partial^2 L}{\partial \tau_i \partial \theta_j}\right] = -2\frac{N}{\sigma^2}\mathrm{Re}\left\{tr\left\{\boldsymbol{F}\,\dot{\boldsymbol{D}}_{\theta_j}\,\boldsymbol{g}\boldsymbol{g}^H\left(\dot{\boldsymbol{D}}_{\tau_i}\right)^H \boldsymbol{F}^H\right\}\right\}. \tag{4.83}$$

B.6. *Expectation of Second order partial derivatives with respect to multipath delay-attenuation cross terms*

$$E\left[\frac{\partial^2 L}{\partial \tau_i \partial \mathrm{Re}\{g_j\}}\right] = -2\frac{N}{\sigma^2}\mathrm{Re}\left\{tr\left\{\boldsymbol{F}\,\dot{\boldsymbol{D}}_{\tau_i}\,\boldsymbol{g}\,\dot{\boldsymbol{g}}_j \boldsymbol{D}^H \boldsymbol{F}^H\right\}\right\}, \tag{4.84}$$

$$E\left[\frac{\partial^2 L}{\partial \tau_i \partial \operatorname{Im}\{g_j\}}\right] = -2\frac{N}{\sigma^2}\operatorname{Im}\left\{tr\left\{\boldsymbol{F}\,\dot{\boldsymbol{D}}_{\tau_i}\,\boldsymbol{g}\dot{\boldsymbol{g}}_j\boldsymbol{D}^H\boldsymbol{F}^H\right\}\right\}. \tag{4.85}$$

B.7. *Expectation of Second order partial derivatives with respect to multipath delay-MC parameter cross terms*

$$E\left[\frac{\partial^2 L}{\partial \tau_i \partial \operatorname{Re}\{c_j\}}\right] = -2\frac{N}{\sigma^2}\operatorname{Re}\left\{tr\left\{\boldsymbol{F}\,\dot{\boldsymbol{D}}_{\tau_i}\,\boldsymbol{g}\boldsymbol{g}^H\boldsymbol{D}^H\,\dot{\boldsymbol{F}}_j\right\}\right\}, \tag{4.86}$$

$$E\left[\frac{\partial^2 L}{\partial \tau_i \partial \operatorname{Im}\{c_j\}}\right] = -2\frac{N}{\sigma^2}\operatorname{Im}\left\{tr\left\{\boldsymbol{F}\,\dot{\boldsymbol{D}}_{\tau_i}\,\boldsymbol{g}\boldsymbol{g}^H\boldsymbol{D}^H\,\dot{\boldsymbol{F}}_j\right\}\right\}. \tag{4.87}$$

B.8. *Expectation of Second order partial derivatives with respect to multipath DOA-attenuation cross terms*

$$E\left[\frac{\partial^2 L}{\partial \theta_i \partial \operatorname{Re}\{g_j\}}\right] = -2\frac{N}{\sigma^2}\operatorname{Re}\left\{tr\left\{\boldsymbol{F}\,\dot{\boldsymbol{D}}_{\theta_i}\,\boldsymbol{g}\dot{\boldsymbol{g}}_j\boldsymbol{D}^H\boldsymbol{F}^H\right\}\right\}, \tag{4.88}$$

$$E\left[\frac{\partial^2 L}{\partial \theta_i \partial \operatorname{Im}\{g_j\}}\right] = -2\frac{N}{\sigma^2}\operatorname{Im}\left\{tr\left\{\boldsymbol{F}\,\dot{\boldsymbol{D}}_{\theta_i}\,\boldsymbol{g}\dot{\boldsymbol{g}}_j\boldsymbol{D}^H\boldsymbol{F}^H\right\}\right\}. \tag{4.89}$$

B.9. *Expectation of Second order partial derivatives with respect to multipath DOA-MC parameter cross terms*

$$E\left[\frac{\partial^2 L}{\partial \theta_i \partial \operatorname{Re}\{c_j\}}\right] = -2\frac{N}{\sigma^2}\operatorname{Re}\left\{tr\left\{\boldsymbol{F}\,\dot{\boldsymbol{D}}_{\theta_i}\,\boldsymbol{g}\boldsymbol{g}^H\boldsymbol{D}^H\,\dot{\boldsymbol{F}}_j\right\}\right\}, \tag{4.90}$$

$$E\left[\frac{\partial^2 L}{\partial \theta_i \partial \operatorname{Im}\{c_j\}}\right] = -2\frac{N}{\sigma^2}\operatorname{Im}\left\{tr\left\{\boldsymbol{F}\,\dot{\boldsymbol{D}}_{\theta_i}\,\boldsymbol{g}\boldsymbol{g}^H\boldsymbol{D}^H\,\dot{\boldsymbol{F}}_j\right\}\right\}. \tag{4.91}$$

B.10. *Expectation of Second order partial derivatives with respect to multipath attenuation-MC parameter cross terms*

$$E\left[\frac{\partial^2 L}{\partial \mathrm{Re}\{g_i\}\partial \mathrm{Re}\{c_j\}}\right] = -2\frac{N}{\sigma^2}\mathrm{Re}\left\{tr\left\{\boldsymbol{FD}\left(\dot{\boldsymbol{g}}_i\right)^H \boldsymbol{g}^H \boldsymbol{D}^H \dot{\boldsymbol{F}}_j\right\}\right\}, \tag{4.92}$$

$$E\left[\frac{\partial^2 L}{\partial \mathrm{Re}\{g_i\}\partial \mathrm{Im}\{c_j\}}\right] = -2\frac{N}{\sigma^2}\mathrm{Im}\left\{tr\left\{\boldsymbol{FD}\left(\dot{\boldsymbol{g}}_i\right)^H \boldsymbol{g}^H \boldsymbol{D}^H \dot{\boldsymbol{F}}_j\right\}\right\}, \tag{4.93}$$

$$E\left[\frac{\partial^2 L}{\partial \mathrm{Im}\{g_i\}\partial \mathrm{Im}\{c_j\}}\right] = -2\frac{N}{\sigma^2}\mathrm{Re}\left\{tr\left\{\boldsymbol{FD}\left(\dot{\boldsymbol{g}}_i\right)^H \boldsymbol{g}^H \boldsymbol{D}^H \dot{\boldsymbol{F}}_j\right\}\right\}, \tag{4.94}$$

$$E\left[\frac{\partial^2 L}{\partial \mathrm{Im}\{g_i\}\partial \mathrm{Re}\{c_j\}}\right] = -2\frac{N}{\sigma^2}\mathrm{Im}\left\{tr\left\{\boldsymbol{FD}\left(\dot{\boldsymbol{g}}_i\right)^H \boldsymbol{g}^H \boldsymbol{D}^H \dot{\boldsymbol{F}}_j\right\}\right\}. \tag{4.95}$$

The FIM is given by the following partitioned matrix

$$\boldsymbol{Z} = \begin{bmatrix} \boldsymbol{Z}_{\tau\tau} & \boldsymbol{Z}_{\tau\theta} & \boldsymbol{Z}_{\tau\,\mathrm{Re}\{g\}} & \boldsymbol{Z}_{\tau\,\mathrm{Im}\{g\}} & \boldsymbol{Z}_{\tau\,\mathrm{Re}\{\gamma\}} & \boldsymbol{Z}_{\tau\,\mathrm{Im}\{\gamma\}} \\ \boldsymbol{Z}_{\theta\tau} & \boldsymbol{Z}_{\theta\theta} & \boldsymbol{Z}_{\theta\,\mathrm{Re}\{g\}} & \boldsymbol{Z}_{\theta\,\mathrm{Im}\{g\}} & \boldsymbol{Z}_{\theta\,\mathrm{Re}\{\gamma\}} & \boldsymbol{Z}_{\theta\,\mathrm{Im}\{\gamma\}} \\ \boldsymbol{Z}_{\mathrm{Re}\{g\}\tau} & \boldsymbol{Z}_{\mathrm{Re}\{g\}\theta} & \boldsymbol{Z}_{\mathrm{Re}\{g\}\mathrm{Re}\{g\}} & \boldsymbol{Z}_{\mathrm{Re}\{g\}\mathrm{Im}\{g\}} & \boldsymbol{Z}_{\mathrm{Re}\{g\}\mathrm{Re}\{\gamma\}} & \boldsymbol{Z}_{\mathrm{Re}\{g\}\mathrm{Im}\{\gamma\}} \\ \boldsymbol{Z}_{\mathrm{Im}\{g\}\tau} & \boldsymbol{Z}_{\mathrm{Im}\{g\}\theta} & \boldsymbol{Z}_{\mathrm{Im}\{g\}\mathrm{Re}\{g\}} & \boldsymbol{Z}_{\mathrm{Im}\{g\}\mathrm{Im}\{g\}} & \boldsymbol{Z}_{\mathrm{Im}\{g\}\mathrm{Re}\{\gamma\}} & \boldsymbol{Z}_{\mathrm{Im}\{g\}\mathrm{Im}\{\gamma\}} \\ \boldsymbol{Z}_{\mathrm{Re}\{\gamma\}\tau} & \boldsymbol{Z}_{\mathrm{Re}\{\gamma\}\theta} & \boldsymbol{Z}_{\mathrm{Re}\{\gamma\}\mathrm{Re}\{g\}} & \boldsymbol{Z}_{\mathrm{Re}\{\gamma\}\mathrm{Im}\{g\}} & \boldsymbol{Z}_{\mathrm{Re}\{\gamma\}\mathrm{Re}\{\gamma\}} & \boldsymbol{Z}_{\mathrm{Re}\{\gamma\}\mathrm{Im}\{\gamma\}} \\ \boldsymbol{Z}_{\mathrm{Im}\{\gamma\}\tau} & \boldsymbol{Z}_{\mathrm{Im}\{\gamma\}\theta} & \boldsymbol{Z}_{\mathrm{Im}\{\gamma\}\mathrm{Re}\{g\}} & \boldsymbol{Z}_{\mathrm{Im}\{\gamma\}\mathrm{Im}\{g\}} & \boldsymbol{Z}_{\mathrm{Im}\{\gamma\}\mathrm{Re}\{\gamma\}} & \boldsymbol{Z}_{\mathrm{Im}\{\gamma\}\mathrm{Im}\{\gamma\}} \end{bmatrix}, \tag{4.96}$$

where each sub-matrix has the same structure as $\boldsymbol{Z}_{\rho\xi}$, with both $\boldsymbol{\rho}$ and $\boldsymbol{\xi}$ being vectors. And the corresponding $(i,j)^{th}$ entry takes the form

$$\left(\boldsymbol{Z}_{\rho\xi}\right)_{ij} = -E\left[\frac{\partial^2 L}{\partial \rho_i \partial \boldsymbol{\xi}_j}\right], \tag{4.97}$$

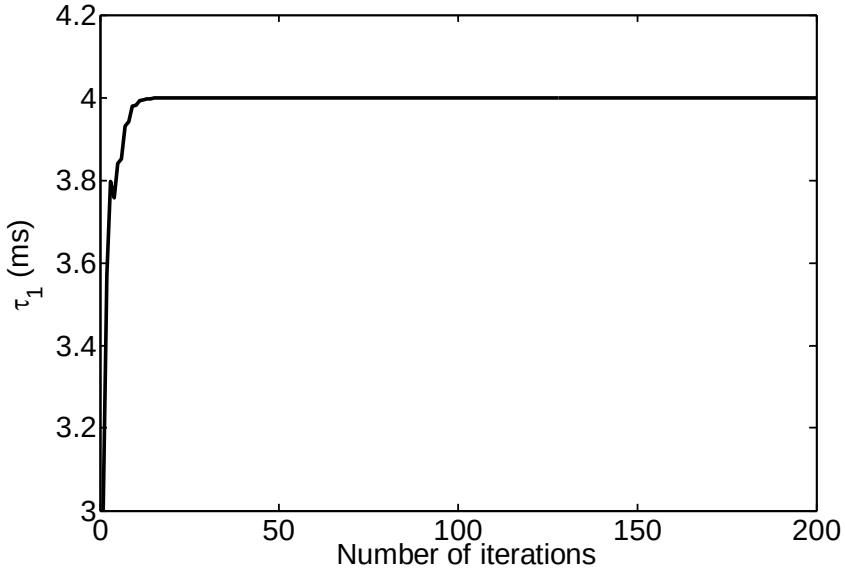
where ρ_i denotes the i^{th} entry of the vector $\boldsymbol{\rho}$, and $\boldsymbol{\xi}_j$ represents the j^{th} entry of the vector $\boldsymbol{\xi}$.

4.4 SIMULATION RESULTS

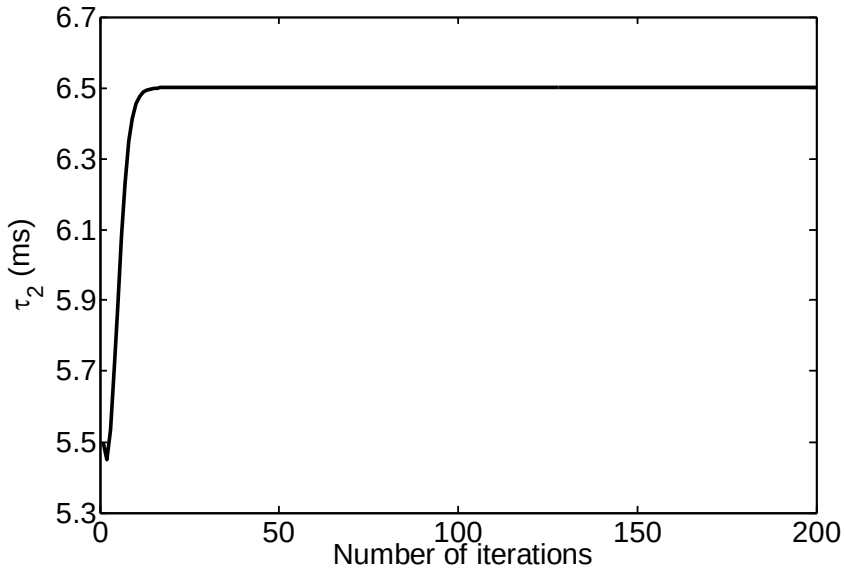
In this section, we will present some simulation results to illustrate the performance of the proposed algorithm.

In the following experiments, we simulated the scenario where narrowband signals transmitted from one far-field source impinge on an ULA of two half-wavelength spaced antenna elements via $M = 2$ paths. The DOAs are given by $\theta_1 = 30^0$ and $\theta_2 = 65^0$. The time delays are $\tau_1 = 4ms$ and $\tau_2 = 6.5ms$, while the path attenuations are $g_1 = 0.85 - 0.3j$ and $g_2 = 0.5 + 0.2j$. The complex MC parameter vector is $\boldsymbol{\gamma} = [1, c]^T$, where $c = 0.06 - 0.03j$. The transmitted signal adopts FH operating at the center frequency $f_c = 1GHz$. A set of 32 hopping carriers with 2.5MHz spacing are allocated in the hopping frequency band with a bandwidth of about 80 MHz. While the parameters are not exactly the same, this is similar to the IEEE 802.11 FHSS Standard [18]. During the initial training period, pilots are sent using $H = 4$ hops, where the corresponding instantaneous frequencies are $f_1 = 2.5MHz$, $f_2 = 12.5MHz$, $f_3 = 65MHz$ and $f_4 = 75MHz$.

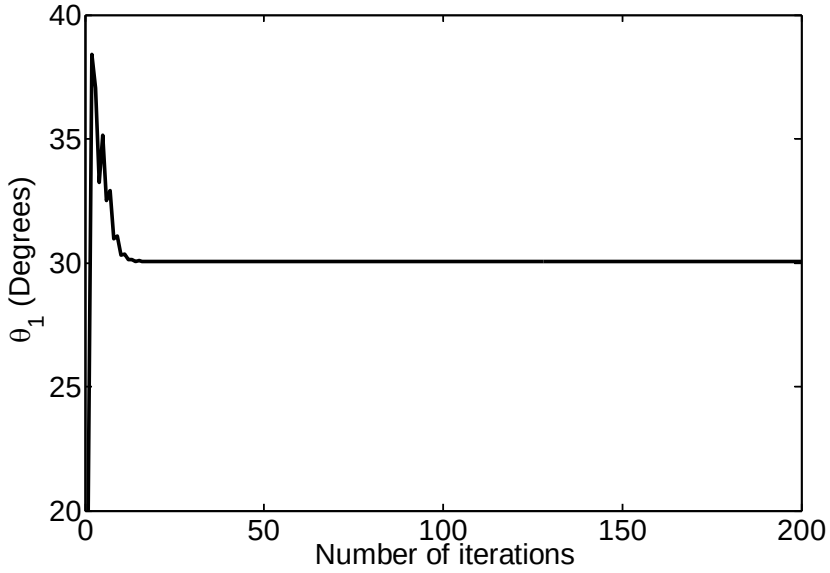
In the first experiment, we demonstrated the convergence of the algorithm using 250 snapshots at 30dB SNR. The SNR is defined as $SNR = 10\log_{10}\left(\sigma_s^2 / \sigma_n^2\right)$, where σ_s^2 and σ_n^2 are the average power of the source and the receiver noise, respectively. The initial parameter values are given by $\hat{\tau}_1^{(0)} = 3ms$, $\hat{\tau}_2^{(0)} = 5.5ms$, $\hat{\theta}_1^{(0)} = 20^0$, and $\hat{\theta}_2^{(0)} = 55^0$. Figure 4.3 shows the convergence of all the parameters being estimated.



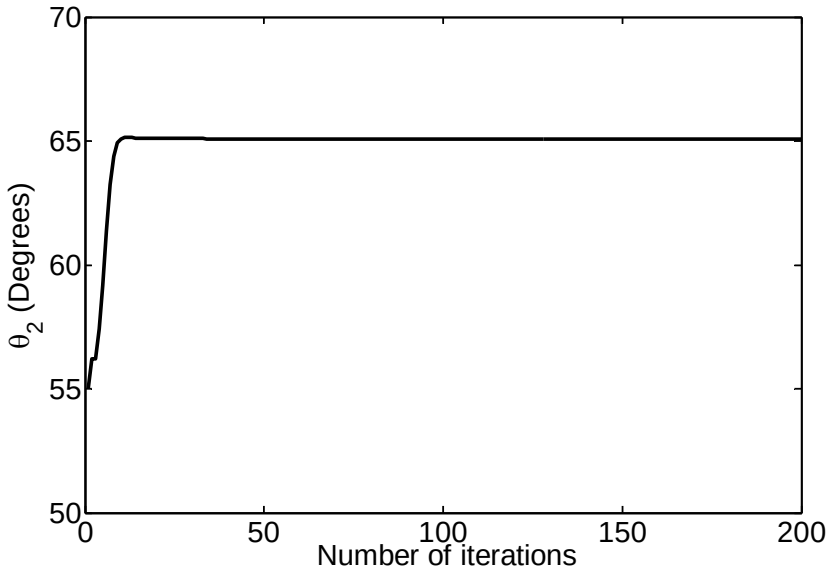
(a) τ_1 versus number of iterations.



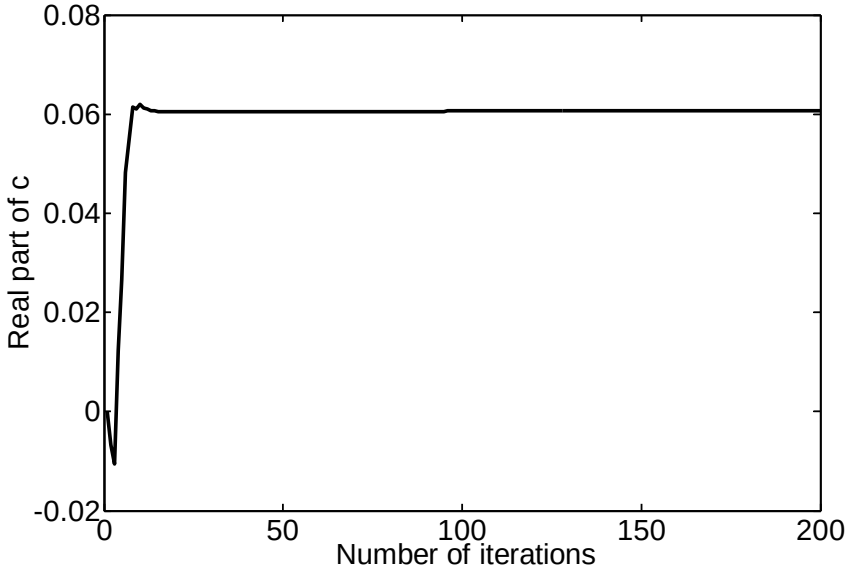
(b) τ_2 versus number of iterations.



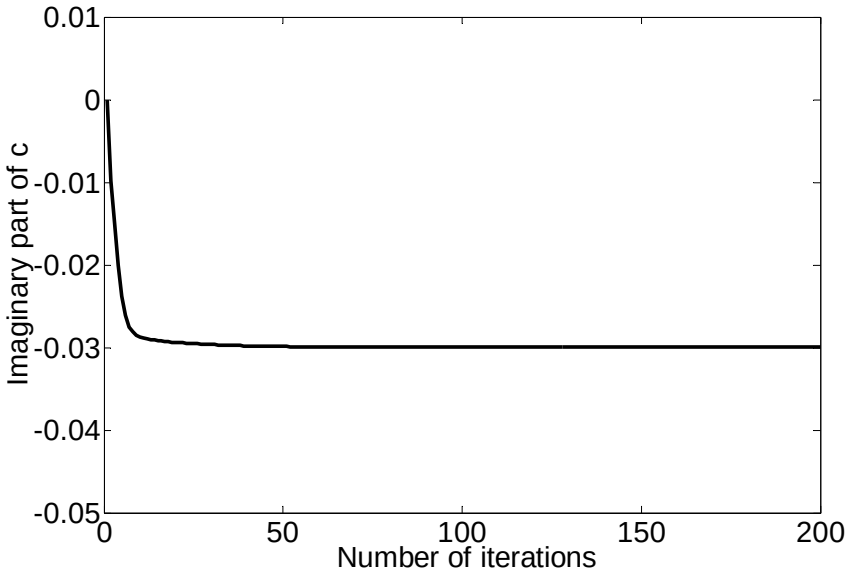
(c) θ_1 versus number of iterations.



(d) θ_2 versus number of iterations.



(e) $\mathrm{Re}\{c\}$ versus number of iterations.

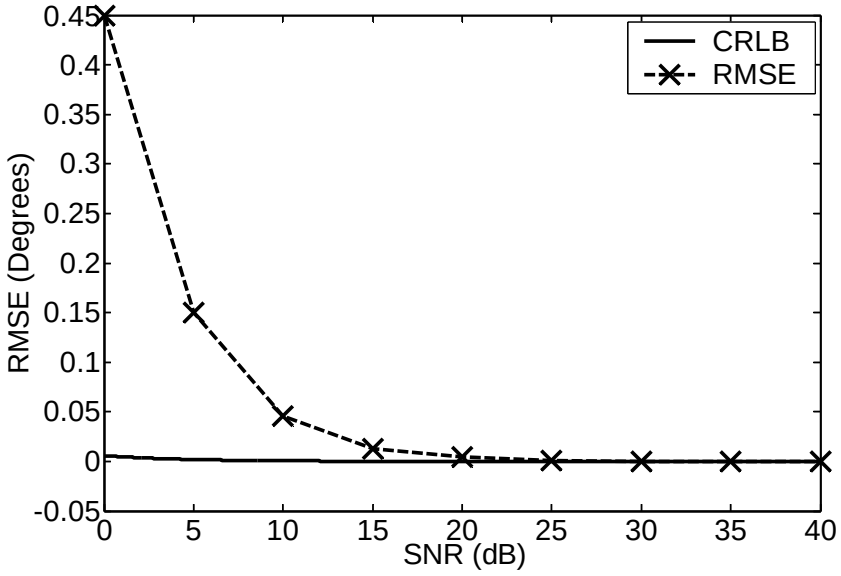


(f) $\mathrm{Im}\{c\}$ versus number of iterations.

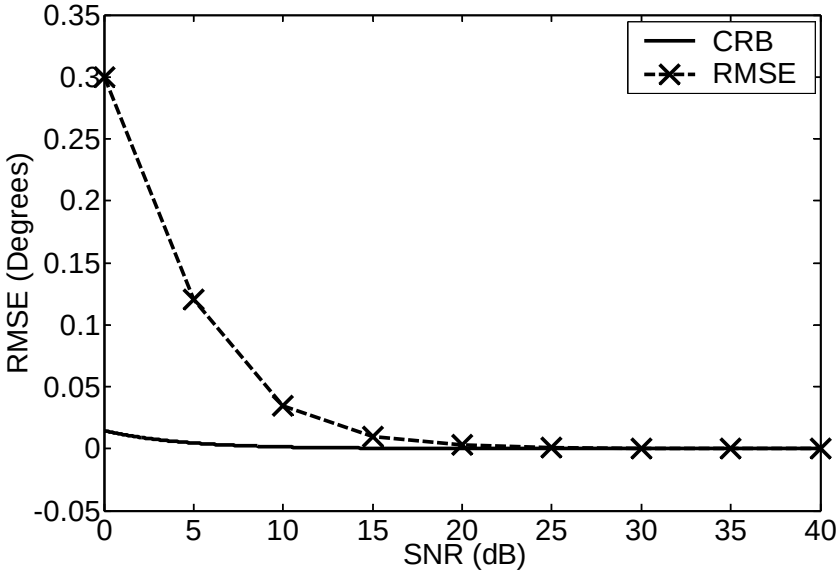
Figure 4.3 Convergence curves of all the parameters estimated.

Obviously, the parameters converge quickly in about 50 iterations.

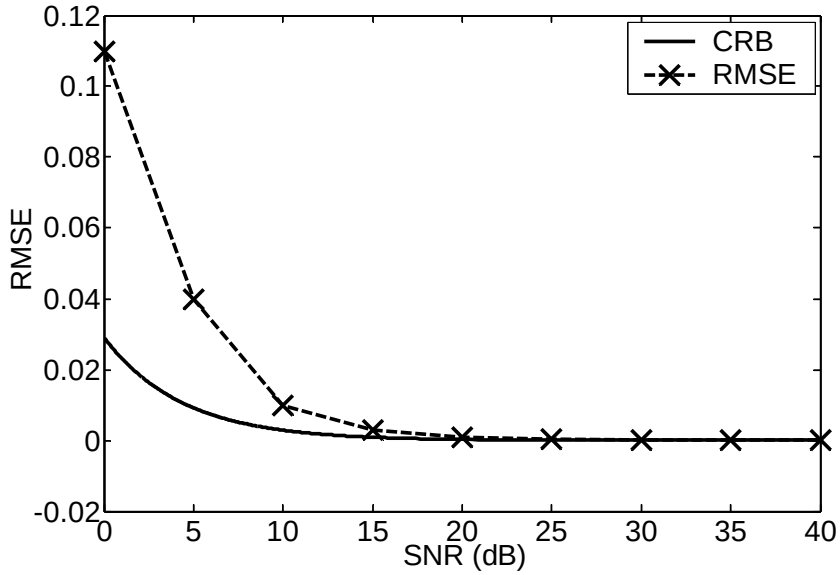
In the second experiment, we investigated the statistical efficiency of the AM algorithm at different SNR. For each SNR, we performed 100 Monte Carlo runs of 250 snapshots and computed the associated Root Mean Square Error (RMSE). We also derived the Cramer-Rao Lower Bound (CRLB) to compare with the RMSE obtained. Figure 4.4 shows the RMSE of $\theta_1, \theta_2,$ $\mathrm{Re}\{c\}$ and $\mathrm{Im}\{c\}$, respectively, versus SNR. The corresponding CRLB curves are also plotted for comparison. The results show that the AM algorithm is statistically efficient, and for $SNR \geq 25dB$, it achieves the CRLB which is plotted using solid lines in the figure. Also, If we follow the standard used by [35] and [36], which tend to set the SNR value corresponding to the case where DOA RMSE equals to 0.1^0 as a threshold, then our algorithm works even in a fairly low SNR case.



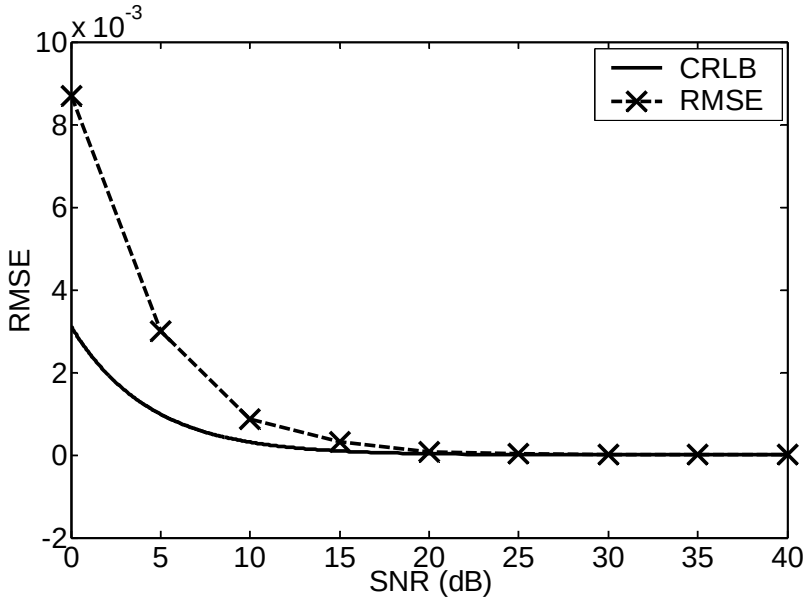
(a) RMSE of θ_1 versus SNR.



(b) RMSE of θ_2 versus SNR.



(c) RMSE of $\mathrm{Re}\{c\}$ versus SNR.



(d) RMSE of $\mathrm{Im}\{c\}$ versus SNR.

Figure 4.4 RMSE versus SNR. The solid lines give the CRLB theoretical value. The asterisk points denote the RMSE at a corresponding SNR, obtained by 100 Monte Carlo runs of 250 snapshots.

4.5 SUMMARY

In this chapter, we propose a new method for DOA estimation in the presence of multipath propagation and unknown MC for a FH system. We present a signal model that takes these effects into account and derived the ML estimators for both DOA and MC parameters. We then formulate an iterative Alternating Minimization (AM) algorithm for finding the MC and DOA parameters in an alternate manner. The simulation results presented illustrate the convergence of the algorithm and its statistical efficiency at high SNR.

Chapter 5

DOA Estimation with A Reduced Number of Receivers in the Presence of Multipath Propagation in A FH System

Detection and localization of multiple narrowband sources by a passive sensor array is a fundamental problem in radar, communication, sonar, seismology and radio-astronomy. The existing super-resolution solutions to this problem require that the whole array be sampled simultaneously and therefore demand that the number of receivers be equal to the number of sensors. This requirement may prevent such solutions from practical implementation especially in cases where the number of sensors is large, due to the high cost of receivers and difficulty in array calibration.

To facilitate the implementation of super-resolution techniques in practice, a reduction in receiver channels is preferable, which is often referred to as “sequential scanning” of the angular space. And some propositions were presented in recent years.

5.1 INTRODUCTION TO DOA ESTIMATION USING FEWER RECEIVERS THAN ANTENNAS

In a loose sense, such propositions can be classified into two categories. The first one is the introduction of subarray. The primitive idea of DOA estimation using subarray may come from [44], where a switching technique is used in conventional interferometers to sequentially sample different “baselines” of the array.

However, there are certain difficulties in performing DOA estimation of multiple sources using subarray data. First, no single subarray can implement DOA estimation on its own. Moreover, when only part of the different pairs of the sensors

are sampled, the existing high-resolution techniques are not applicable since no estimate of the whole array covariance matrix can be constructed.

In [45], Sheinvald and Wax proposed a subarray-based method to localize multiple narrow band signals. In their method, the SML estimator is used, and it is further approximated by a generalized LS estimator to reduce the computational complexity. However, because the corresponding premise is covariance matrix of subarray, consequently the minimum antenna number of a subarray should be no less than two.

In [46 - 47], the problem of DOA estimation using a time-varying array was studied, where the corresponding elements move during the observation interval in an arbitrary but known way. Such scenario is comparable to the DOA estimation using subarray. However, the main drawback is that only one source case is considered therein.

The second one is using a preprocessing beamforming network. In [48], a linear, time-varying network was used as preprocessing stage which performed the dimension reduction. In [49], the author considered the case of multiple source DOA estimation using a preprocessing network with two output ports, i.e., only two receivers were needed. However, the method can not be applicable to the multipath propagation scenario in that it is eigen-decomposition-based. In [50 - 51], the corresponding preprocessing transformation matrix is assumed to be known. In [52], a method to derive the optimal transformation matrix was presented, which was based on CRLB of DOA estimation error variance.

In this chapter, we briefly introduce a new method for joint time delay and DOA estimation using two antennas and one receiver for a FH system in the presence of multipath.

5.2 SIGNAL MODEL

Consider a multipath propagation scenario where a far-field narrowband signal impinges on a Uniform Linear Array (ULA) with two antenna elements via *M* distinct paths (as shown in Figure 5.1). The two antennas share the same receiver by sequential switching.

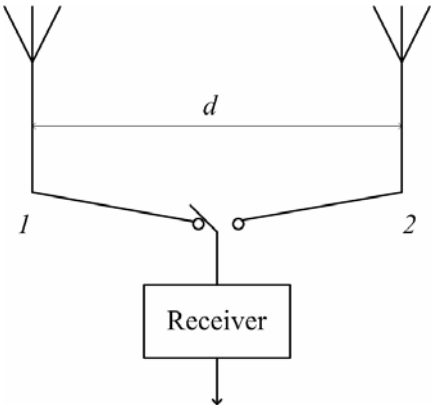


Figure 5.1 The structure with two antennas and one receiver.

Assume also that slow frequency hopping is used and the transmitted signal in the h^{th} hop is

$$\varphi_h(t) = \mathrm{Re}\left\{ s_h(t) e^{j2\pi(f_c + f_h)t} \right\}, \tag{5.1}$$

where f_c is the center frequency, while f_h denotes the hop frequency. Of course, we can also express the whole baseband signal as

$$s(t) = \sum_{h=1}^{H} e^{j2\pi f_h [t-(h-1)T_c]} p[t-(h-1)T_c], \tag{5.2}$$

where T_c denotes the FH chip period and $p(t)$ represents a rectangular function with the form

$$p(t)=\begin{cases}1, & 0\le t\le T_c\\ 0, & otherwise\end{cases}. \tag{5.3}$$

To solve the problem, we assume that the RF switch time Δ is exactly the same as the FH chip period T_c, and we sequentially switch receiver from antenna 1 to antenna 2, with the starting time point being t_0. Also, t_0 and the time-delay of every path are assumed to be less than the FH chip period T_c. Such assumptions are reasonable for a suitable T_c.

Assume that *2N* samples are collected after down-conversion and sampling, and they are labeled with the numbers from 1 to *2N*. Then we can see that the samples with odd numbers are obtained from antenna 1. The corresponding values are

$$\begin{aligned}y_{2n-1} &= x_1(t_0+2(n-1)T_c)\\ &=\sum_{m=1}^{M} g_m s(t_0+2(n-1)T_c-\tau_m)+n_1(t_0+2(n-1)T_c),\\ &=\sum_{m=1}^{M} g_m e^{j2\pi f_{(2n-1)}[t_0+2(n-1)T_c-\tau_m]}+n_1(t_0+2(n-1)T_c)\end{aligned}$$

$$n=1,2,\cdots,N. \tag{5.4}$$

Further we define a sample sequence (or time series) as $\{u(1),\cdots,u(N)\}$, where $u(n)=y_{2n-1}$. This time series can be broken into *P* neighboring time windows of length *L* (without loss of generality, we assume $N=P\times L$). Thus we get

$$\boldsymbol{u}_p=[u((p-1)L+1),...,u(pL)]^T=[y_{(2(p-1)L+1)},...,y_{(2pL-1)}]^T=\boldsymbol{\Phi}_{1p}(\boldsymbol{\tau})\boldsymbol{g}+\boldsymbol{n}_{1p},$$

$$p=1,2,\cdots,P. \tag{5.5}$$

where

$$\boldsymbol{g}=[g_1,g_2,\cdots,g_M]^T, \tag{5.6}$$

$$\boldsymbol{\Phi}_{1p}(\boldsymbol{\tau})=\begin{bmatrix} e^{j2\pi f_{(2(P-1)L+1)}[t_0+2(p-1)LT_c-\tau_1]} & \cdots & e^{j2\pi f_{(2(P-1)L+1)}[t_0+2(p-1)LT_c-\tau_M]} \\ \vdots & \ddots & \vdots \\ e^{j2\pi f_{(2pL-1)}[t_0+2(pL-1)T_c-\tau_1]} & \cdots & e^{j2\pi f_{(2pL-1)}[t_0+2(pL-1)T_c-\tau_M]} \end{bmatrix}^{L\times M}, \tag{5.7}$$

$$\boldsymbol{\tau}=\left[\tau_1,\tau_2,\cdots,\tau_M\right]^T, \tag{5.8}$$

$$\boldsymbol{n}_{1p}=\left[n_1(t_0+2(p-1)LT_c),\cdots,n_1(t_0+2(pL-1)T_c)\right]. \tag{5.9}$$

Hence, the corresponding data matrix can be written as

$$\boldsymbol{U}=\begin{bmatrix}\boldsymbol{u}_1^T\\ \vdots \\ \boldsymbol{u}_P^T\end{bmatrix}=\begin{bmatrix}\boldsymbol{g}^T\boldsymbol{\Phi}_{11}^T(\boldsymbol{\tau})\\ \vdots \\ \boldsymbol{g}^T\boldsymbol{\Phi}_{1P}^T(\boldsymbol{\tau})\end{bmatrix}+\begin{bmatrix}\boldsymbol{n}_{11}^T\\ \vdots \\ \boldsymbol{n}_{1P}^T\end{bmatrix}. \tag{5.10}$$

Likewise, the samples with even numbers are obtained from antenna 2, which take the form

$$\begin{aligned} y_{2n} &= x_2(t_0+(2n-1)T_c) \\ &= \sum_{m=1}^{M} g_m e^{j\alpha d\cos\theta_m} e^{j2\pi f_{2n}[t_0+(2n-1)T_c-\tau_m]} + n_2(t_0+(2n-1)T_c)' \end{aligned} \tag{5.11}$$

where $\alpha=\dfrac{2\pi}{\lambda}$ is the phase propagation factor.

As we have mentioned before, we only consider the scenario of narrow band frequency hopping signal herein, i.e., $f_h << f_c$, so α is a constant. Further, let $d=\dfrac{\lambda}{2}$. We can rewrite (5.11) as

$$y_{2n}=\sum_{m=1}^{M} g_m e^{j\pi\cos\theta_m} e^{j2\pi f_{2n}[t_0+(2n-1)T_c-\tau_m]} + n_2(t_0+(2n-1)T_c). \tag{5.12}$$

Similarly, we define a sample sequence as $\{v(1),\cdots,v(N)\}$, where $v(n)=y_{2n}$. This time series can be broken into *P* neighboring time windows of length *L* (without loss of generality, we assume $N=P\times L$). Thus we get

$$\begin{aligned}\boldsymbol{v}_p &= \left[v\left((p-1)L+1\right),\ldots,v\left(pL\right)\right]^T \\ &= \left[y_{(2(p-1)L+2)},\ldots,y_{(2pL)}\right]^T = \boldsymbol{\Phi}_{2p}(\boldsymbol{\tau})\boldsymbol{G}(\boldsymbol{g})\boldsymbol{q}(\boldsymbol{\theta})+\boldsymbol{n}_{2p}\end{aligned}, \tag{5.13}$$

where

$$\boldsymbol{\Phi}_{2p}(\boldsymbol{\tau}) = \begin{bmatrix} e^{j2\pi f_{(2(P-1)L+2)}\left[t_0+(2(p-1)L+1)T_c-\tau_1\right]} & \cdots & e^{j2\pi f_{(2(P-1)L+2)}\left[t_0+(2(p-1)L+1)T_c-\tau_M\right]} \\ \vdots & \ddots & \vdots \\ e^{j2\pi f_{(2pL)}\left[t_0+(2pL-1)T_c-\tau_1\right]} & \cdots & e^{j2\pi f_{(2pL)}\left[t_0+(2pL-1)T_c-\tau_M\right]} \end{bmatrix}^{L\times M}, \tag{5.14}$$

$$\boldsymbol{G}(\boldsymbol{g}) = diag\left(g_1, g_2, \cdots, g_M\right), \tag{5.15}$$

$$\boldsymbol{q}(\boldsymbol{\theta}) = \left[e^{-j\pi\cos\theta_1}, \cdots, e^{-j\pi\cos\theta_M}\right]^T, \tag{5.16}$$

$$\boldsymbol{n}_2 = \left[n_2\left(t_0+T_c-\tau_m\right), \cdots, n_2\left(t_0+\left(2N-1\right)T_c-\tau_m\right)\right]. \tag{5.17}$$

Therefore, the associated data matrix can be expressed as

$$\boldsymbol{V} = \begin{bmatrix} \boldsymbol{v}_1^T \\ \vdots \\ \boldsymbol{v}_P^T \end{bmatrix} = \begin{bmatrix} \boldsymbol{q}^T(\boldsymbol{\theta})\boldsymbol{G}(\boldsymbol{g})\boldsymbol{\Phi}_{21}^T(\boldsymbol{\tau}) \\ \vdots \\ \boldsymbol{q}^T(\boldsymbol{\theta})\boldsymbol{G}(\boldsymbol{g})\boldsymbol{\Phi}_{2P}^T(\boldsymbol{\tau}) \end{bmatrix} + \begin{bmatrix} \boldsymbol{n}_{21}^T \\ \vdots \\ \boldsymbol{n}_{2P}^T \end{bmatrix}. \tag{5.18}$$

5.3 LEAST SQUARE ESTIMATOR

The main idea of our joint estimation method is first to obtain $\hat{\boldsymbol{\tau}}$ from equation (5.5), based on which the DOA estimate $\hat{\boldsymbol{\theta}}$ is then achieved from equation (5.11). The first step, time-delay estimation, is crucial for our joint estimation algorithm.

Note that the time windows in (5.5) are non-overlapping, therefore the corresponding noise vector are pair-wise uncorrelated. Based on (5.10), the associated cost function is given by

$$J_1(\boldsymbol{\tau},\boldsymbol{g}) = \left\| \boldsymbol{U} - \begin{bmatrix} \boldsymbol{g}^T \boldsymbol{\Phi}_{11}^T(\boldsymbol{\tau}) \\ \vdots \\ \boldsymbol{g}^T \boldsymbol{\Phi}_{1P}^T(\boldsymbol{\tau}) \end{bmatrix} \right\|_F^2 = \sum_{p=1}^{P} \left\| \boldsymbol{u}_p^T - \boldsymbol{g}^T \boldsymbol{\Phi}_{1p}^T(\boldsymbol{\tau}) \right\|^2 = \sum_{p=1}^{P} \left\| \boldsymbol{u}_p - \boldsymbol{\Phi}_{1p}(\boldsymbol{\tau}) \boldsymbol{g} \right\|^2 , \tag{5.19}$$

where $\|\bullet\|_F^2$ denotes Frobenius norm. Equation (5.19) can be rewritten as

$$J_1(\boldsymbol{\tau},\boldsymbol{g}) = \left\| \boldsymbol{u} - \boldsymbol{\Phi}_1(\boldsymbol{\tau}) \boldsymbol{g} \right\|^2 , \tag{5.20}$$

where

$$\boldsymbol{u} = \left[\boldsymbol{u}_1^T, \cdots \boldsymbol{u}_P^T \right]^T , \tag{5.21}$$

$$\boldsymbol{\Phi}_1(\boldsymbol{\tau}) = \begin{bmatrix} \boldsymbol{\Phi}_{11}(\boldsymbol{\tau}) \\ \vdots \\ \boldsymbol{\Phi}_{1P}(\boldsymbol{\tau}) \end{bmatrix}^{N \times M} . \tag{5.22}$$

Minimizing (5.20) with respect to $\boldsymbol{g}$ yields

$$\hat{\boldsymbol{g}} = \boldsymbol{\Phi}_1^{+}(\boldsymbol{\tau}) \boldsymbol{u} , \tag{5.23}$$

where

$$\boldsymbol{\Phi}_1^{+}(\boldsymbol{\tau}) = \left[\boldsymbol{\Phi}_1^{H}(\boldsymbol{\tau}) \boldsymbol{\Phi}_1(\boldsymbol{\tau}) \right]^{-1} \boldsymbol{\Phi}_1(\boldsymbol{\tau})^{H} . \tag{5.24}$$

Substituting (5.23) back into (5.20) yields the concentrated cost function

$$\tilde{J}_1(\boldsymbol{\tau}) = \left\| \boldsymbol{u} - \boldsymbol{\Phi}_1(\boldsymbol{\tau}) \boldsymbol{\Phi}_1^{+}(\boldsymbol{\tau}) \boldsymbol{u} \right\|^2 = \left\| \boldsymbol{\Pi}_{\boldsymbol{\Phi}_1(\boldsymbol{\tau})}^{\perp} \boldsymbol{u} \right\|^2 , \tag{5.25}$$

where

$$\boldsymbol{\Pi}_{\boldsymbol{\Phi}_1(\boldsymbol{\tau})} = \boldsymbol{\Phi}_1(\boldsymbol{\tau}) \boldsymbol{\Phi}_1^{+}(\boldsymbol{\tau}), \tag{5.26}$$

$$\boldsymbol{\Pi}_{\boldsymbol{\Phi}_1(\boldsymbol{\tau})}^{\perp} = \boldsymbol{I} - \boldsymbol{\Pi}_{\boldsymbol{\Phi}_1(\boldsymbol{\tau})} . \tag{5.27}$$

Therefore, the Least Square (LS) estimation of time-delay becomes

$$\hat{\boldsymbol{\tau}} = \arg\min_{\boldsymbol{\tau}} \left\{ \tilde{J}_1(\boldsymbol{\tau}) \right\} = \arg\min_{\boldsymbol{\tau}} \left\| \boldsymbol{\Pi}_{\boldsymbol{\Phi}_1(\boldsymbol{\tau})}^{\perp} \boldsymbol{u} \right\|^2 = \arg\max_{\boldsymbol{\tau}} \left\| \boldsymbol{\Pi}_{\boldsymbol{\Phi}_1(\boldsymbol{\tau})} \boldsymbol{u} \right\|^2 . \tag{5.28}$$

Further, from equation (5.18), we get the corresponding cost function

$$J_2(\boldsymbol{\tau},\boldsymbol{g},\boldsymbol{\theta}) = \left\| \boldsymbol{V} - \begin{bmatrix} \boldsymbol{q}^T(\boldsymbol{\theta})\boldsymbol{G}(\boldsymbol{g})\boldsymbol{\Phi}_{21}^T(\boldsymbol{\tau}) \\ \vdots \\ \boldsymbol{q}^T(\boldsymbol{\theta})\boldsymbol{G}(\boldsymbol{g})\boldsymbol{\Phi}_{2P}^T(\boldsymbol{\tau}) \end{bmatrix} \right\|^2 = \sum_{p=1}^{P} \left\| \boldsymbol{v}_p^T - \boldsymbol{q}^T(\boldsymbol{\theta})\boldsymbol{G}(\boldsymbol{g})\boldsymbol{\Phi}_{2p}^T(\boldsymbol{\tau}) \right\|^2 . \quad (5.29)$$

Equation (5.29) can be rewritten as

$$J_2(\boldsymbol{\tau},\boldsymbol{g},\boldsymbol{\theta}) = \left\| \boldsymbol{v} - \boldsymbol{\Phi}_2(\boldsymbol{\tau})\boldsymbol{G}(\boldsymbol{g})\boldsymbol{q}(\boldsymbol{\theta}) \right\|^2 . \quad (5.30)$$

where

$$\boldsymbol{v} = \left[\boldsymbol{v}_1^T, \cdots \boldsymbol{v}_P^T \right]^T , \quad (5.31)$$

$$\boldsymbol{\Phi}_2(\boldsymbol{\tau}) = \begin{bmatrix} \boldsymbol{\Phi}_{21}(\boldsymbol{\tau}) \\ \vdots \\ \boldsymbol{\Phi}_{2P}(\boldsymbol{\tau}) \end{bmatrix} . \quad (5.32)$$

Minimizing (5.30) with respect to $\boldsymbol{q}$ yields

$$\hat{\boldsymbol{q}}(\boldsymbol{\theta}) = \left[\boldsymbol{\Phi}_2(\boldsymbol{\tau})\boldsymbol{G}(\boldsymbol{g}) \right]^+ \boldsymbol{v} . \quad (5.33)$$

Replacing $\boldsymbol{\tau}$ and $\boldsymbol{g}$ with $\hat{\boldsymbol{\tau}}$ and $\hat{\boldsymbol{g}}$ coming from (5.28) and (5.23) respectively, we get

$$\hat{\boldsymbol{q}}(\boldsymbol{\theta}) = \left[\boldsymbol{\Phi}_2(\hat{\boldsymbol{\tau}})\boldsymbol{G}(\hat{\boldsymbol{g}}) \right]^+ \boldsymbol{v} . \quad (5.34)$$

Comparing (5.16) with (5.34), we get the corresponding DOA estimates as

$$\hat{\theta}_m = \cos^{-1}\left[-\frac{\arg(\hat{q}_m)}{\pi} \right], \; m = 1, 2, \cdots, M , \quad (5.35)$$

where $\hat{q}_m$ represents the m^{th} entry of the estimated vector $\hat{\boldsymbol{q}}(\boldsymbol{\theta})$ in (5.34).

Once the estimators have been derived, the associated algorithm is quite self-explanatory. In detail, it consists of two main steps as follows.

A. *Time-delay estimation*

To solve the non-linear *M*-dimensional optimization problem of (5.28), we resort to the Alternating Maximization (AM) algorithm, which performs *M* one-dimensional searches iteratively instead of a direct *M*-dimensional search.

A.1. *Initialization*

The initialization procedure consists of M steps. For the m^{th} step, we solve the maximization problem assuming that there exists m path(s), and the associated time-delay estimation will be [42]

$$\hat{\tau}_m^{(0)} = \arg\max_{\tau_m} \left\{ \boldsymbol{b}^H\left(\tau_m, \hat{\boldsymbol{\tau}}_m^{(0)}\right) \boldsymbol{R}_{\boldsymbol{uu}} \boldsymbol{b}\left(\tau_m, \hat{\boldsymbol{\tau}}_m^{(0)}\right) \right\}, \tag{5.36}$$

where

$$\hat{\boldsymbol{\tau}}_m^{(0)} = \left[\hat{\tau}_1^{(0)}, \cdots, \hat{\tau}_{m-1}^{(0)}\right], \tag{5.37}$$

$$\boldsymbol{b}^H\left(\tau_m, \hat{\boldsymbol{\tau}}_m^{(0)}\right) = \frac{\boldsymbol{\varphi}\left(\tau_m\right)_{\boldsymbol{\Phi}_1\left(\hat{\boldsymbol{\tau}}_m^{(0)}\right)}}{\left\| \boldsymbol{\varphi}\left(\tau_m\right)_{\boldsymbol{\Phi}_1\left(\hat{\boldsymbol{\tau}}_m^{(0)}\right)} \right\|}, \tag{5.38}$$

$$\boldsymbol{\varphi}\left(\tau_m\right)_{\boldsymbol{\Phi}_1\left(\hat{\boldsymbol{\tau}}_m^{(0)}\right)} = \boldsymbol{\Pi}^{\perp}_{\boldsymbol{\Phi}_1\left(\hat{\boldsymbol{\tau}}_m^{(0)}\right)} \boldsymbol{\varphi}\left(\tau_m\right), \tag{5.39}$$

$$\boldsymbol{R}_{\boldsymbol{uu}} = \boldsymbol{u}\boldsymbol{u}^H. \tag{5.40}$$

Note that $\boldsymbol{\varphi}\left(\tau_m\right)$ denotes the m^{th} column vector of the $N \times M$ matrix $\boldsymbol{\Phi}_1\left(\boldsymbol{\tau}\right)$, and $\boldsymbol{\Phi}_1\left(\hat{\boldsymbol{\tau}}_m^{(0)}\right)$ is a $N \times \left(m-1\right)$ matrix given by

$$\boldsymbol{\Phi}_1\left(\hat{\boldsymbol{\tau}}_m^{(0)}\right) = \begin{bmatrix} \boldsymbol{\Phi}_{11}\left(\hat{\boldsymbol{\tau}}_m^{(0)}\right) \\ \vdots \\ \boldsymbol{\Phi}_{1P}\left(\hat{\boldsymbol{\tau}}_m^{(0)}\right) \end{bmatrix}^{N \times (m-1)}, \tag{5.41}$$

$$\boldsymbol{\Phi}_{1p}\left(\hat{\boldsymbol{\tau}}_m^{(0)}\right) = \begin{bmatrix} e^{j2\pi f_{(2(P-1)L+1)}\left[t_0+2(p-1)LT_c-\hat{\tau}_1^{(0)}\right]} & \cdots & e^{j2\pi f_{(2(P-1)L+1)}\left[t_0+2(p-1)LT_c-\hat{\tau}_{m-1}^{(0)}\right]} \\ \vdots & \ddots & \vdots \\ e^{j2\pi f_{(2pL-1)}\left[t_0+2(pL-1)T_c-\hat{\tau}_1^{(0)}\right]} & \cdots & e^{j2\pi f_{(2pL-1)}\left[t_0+2(pL-1)T_c-\hat{\tau}_{m-1}^{(0)}\right]} \end{bmatrix}^{L \times (m-1)}. \tag{5.42}$$

A.2. *Iteration*

For $m = 1, 2, \cdots, M$, do the following maximization until $\left|\hat{\tau}_m^{(n+1)} - \hat{\tau}_m^{(n)}\right| < \varepsilon$.

$$\hat{\tau}_m^{(n+1)} = \arg\max_{\tau_m} \left\{ \boldsymbol{b}^H\left(\tau_m, \hat{\boldsymbol{\tau}}_m^{(n)}\right) \boldsymbol{R}_{\boldsymbol{uu}} \boldsymbol{b}\left(\tau_m, \hat{\boldsymbol{\tau}}_m^{(n)}\right) \right\}, \tag{5.43}$$

where

$$\hat{\boldsymbol{\tau}}_m^{(n)} = \left[\hat{\tau}_1^{(n)}, \hat{\tau}_2^{(n)}, \cdots, \hat{\tau}_{m-1}^{(n)}, \hat{\tau}_{m+1}^{(n)}, \cdots, \hat{\tau}_M^{(n)}\right], \tag{5.44}$$

$$\boldsymbol{b}^H\left(\tau_m, \hat{\boldsymbol{\tau}}_m^{(n)}\right) = \frac{\boldsymbol{\varphi}\left(\tau_m\right)_{\boldsymbol{\Phi}_1\left(\hat{\boldsymbol{\tau}}_m^{(n)}\right)}}{\left\|\boldsymbol{\varphi}\left(\tau_m\right)_{\boldsymbol{\Phi}_1\left(\hat{\boldsymbol{\tau}}_m^{(n)}\right)}\right\|}, \tag{5.45}$$

$$\boldsymbol{\varphi}\left(\tau_m\right)_{\boldsymbol{\Phi}_1\left(\hat{\boldsymbol{\tau}}_m^{(n)}\right)} = \boldsymbol{\Pi}^{\perp}_{\boldsymbol{\Phi}_1\left(\hat{\boldsymbol{\tau}}_m^{(n)}\right)} \boldsymbol{\varphi}\left(\tau_m\right). \tag{5.46}$$

Note that $\boldsymbol{\Phi}_1\left(\hat{\boldsymbol{\tau}}_m^{(n)}\right)$ is a $N \times (M-1)$ matrix given by

$$\boldsymbol{\Phi}_1\left(\hat{\boldsymbol{\tau}}_m^{(n)}\right) = \begin{bmatrix} \boldsymbol{\Phi}_{11}\left(\hat{\boldsymbol{\tau}}_m^{(n)}\right) \\ \vdots \\ \boldsymbol{\Phi}_{1P}\left(\hat{\boldsymbol{\tau}}_m^{(n)}\right) \end{bmatrix}^{N\times(M-1)}, \tag{5.47}$$

$$\boldsymbol{\Phi}_{1p}\left(\hat{\boldsymbol{\tau}}_m^{(n)}\right) = \begin{bmatrix} e^{j2\pi f_{(2(P-1)L+1)}\left[t_0+2(p-1)LT_c-\hat{\tau}_1^{(n)}\right]} & \cdots & e^{j2\pi f_{(2(P-1)L+1)}\left[t_0+2(p-1)LT_c-\hat{\tau}_{m-1}^{(n)}\right]} \\ \vdots & \ddots & \vdots \\ e^{j2\pi f_{(2pL-1)}\left[t_0+2(pL-1)T_c-\hat{\tau}_1^{(n)}\right]} & \cdots & e^{j2\pi f_{(2pL-1)}\left[t_0+2(pL-1)T_c-\hat{\tau}_{m-1}^{(n)}\right]} \\ e^{j2\pi f_{(2(P-1)L+1)}\left[t_0+2(p-1)LT_c-\hat{\tau}_{m+1}^{(n)}\right]} & \cdots & e^{j2\pi f_{(2(P-1)L+1)}\left[t_0+2(p-1)LT_c-\hat{\tau}_M^{(n)}\right]} \\ \vdots & \ddots & \vdots \\ e^{j2\pi f_{(2pL-1)}\left[t_0+2(pL-1)T_c-\hat{\tau}_{m+1}^{(n)}\right]} & \cdots & e^{j2\pi f_{(2pL-1)}\left[t_0+2(pL-1)T_c-\hat{\tau}_M^{(n)}\right]} \end{bmatrix}^{L\times(M-1)}. \tag{5.48}$$

B. *DOA estimation*

Once the time-delay estimation $\hat{\boldsymbol{\tau}}$ is achieved, the corresponding DOA estimation $\hat{\boldsymbol{\theta}}$ will be readily obtained from equation (5.35).

5.4 SIMULATION RESULTS

In this section, some simulation results will be presented to illustrate the performance of the proposed algorithm.

We simulated the case where narrowband signals transmitted from one far-field source impinge on a ULA of two half-wavelength spaced antenna elements via $M = 2$ paths. The corresponding DOAs are given by $\theta_1 = 30^0$ and $\theta_2 = 45^0$. The

time-delays are $\tau_1 = \frac{1}{8}T_C$ and $\tau_2 = \frac{1}{4}T_C$, where T_c represents the chip duration equal to $1\mu s$. And the associated complex path attenuations are $g_1 = 0.8 - j0.5$ and $g_2 = 0.6 + j0.35$. The transmitted signal adopts FH operating at the center frequency $f_c = 1GHz$. Also, there are twenty hopping frequencies used, i.e., $f_h = 0.3h\ MHz$ and $h = 1, \cdots, 20$. The initial sampling time is $t_0 = T_C/2$.

For the stage of time-delay estimation, the method of unit step search [43] was used, with step size being $\mu = T_C/200$. Figure 5.2 shows the initialization of the time-delays τ_1 and τ_2 in the absence of noise. By locating the peak values of the corresponding cost functions, the initial estimation values of τ_1 and τ_2 are given by $0.160\mu s$ and $0.235\mu s$, respectively.

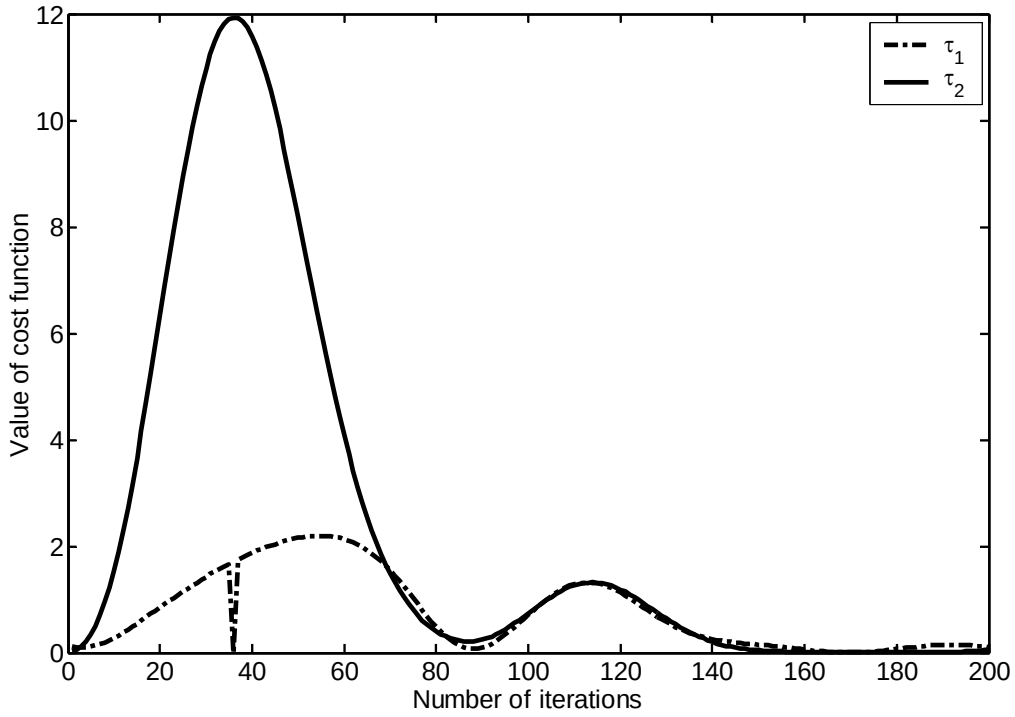


Figure 5.2 Initialization of the time-delays τ_1 and τ_2 in the absence of noise.

Based on the initial estimation, the following iteration is performed to find the optimal estimation of τ_1 and τ_2. Table 5.1 shows this process for the scenarios of no noise and $SNR = 5dB$. The SNR is defined as $SNR = 10\log_{10}\left(\sigma_s^2/\sigma_n^2\right)$, where σ_s^2 and σ_n^2 are the average power of the source and the receiver noise, respectively. We can see that the iteration algorithm converges well after 2 iterations.

Table 5.1 Time-delay estimation

Iteration No. / *SNR* (dB)		0*	1	2	3
∞ (No Noise)	$\tau_1(\mu s)$	0.160	0.125	0.125	0.125
	$\tau_2(\mu s)$	0.235	0.250	0.250	0.250
5	$\tau_1(\mu s)$	0.155	0.125	0.130	0.130
	$\tau_2(\mu s)$	0.240	0.245	0.245	0.245

* Iteration number 0 corresponds to initial value estimation of τ_1 and τ_2.

Also, we computed the corresponding Root Mean Square Error (RMSE) of DOA estimation with 50 Monte Carlo runs. The associated curve is shown as Figure 5.3.

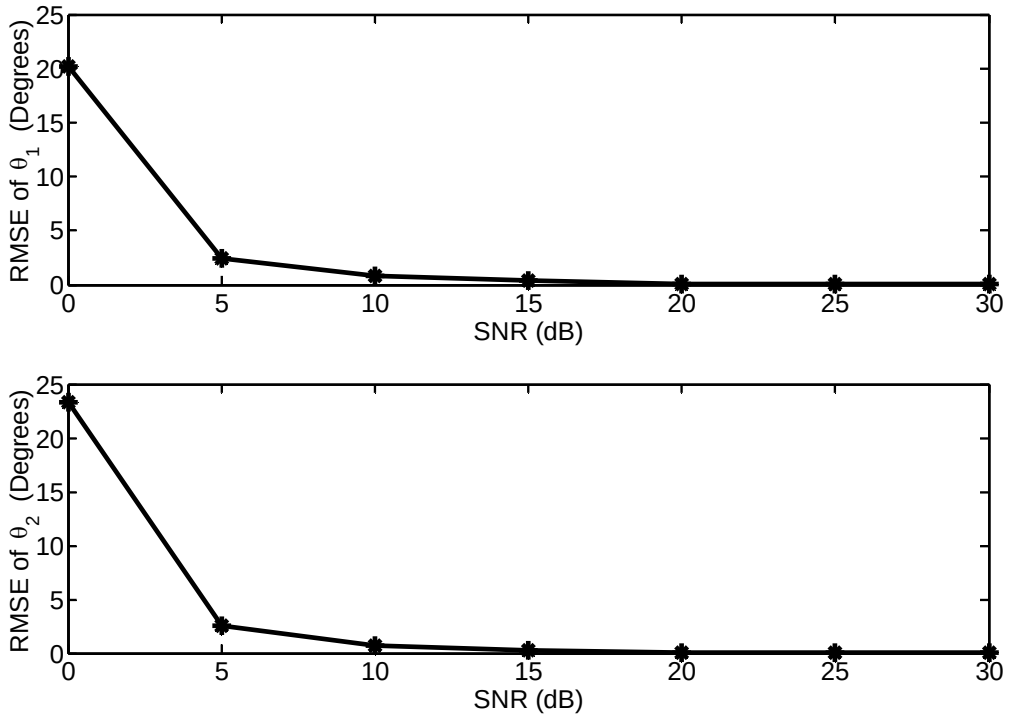


Figure 5.3 Root Mean Square Error (RMSE) of DOA estimation.

From the above curve, we can see that, for $SNR \geq 15dB$, the RMSE is smaller than 0.1^0 . If we follow the standard used by [35] and [36], which tend to set the SNR value corresponding to the case where DOA RMSE equals to 0.1^0 as a threshold, then our method operates properly for $SNR \geq 15dB$.

5.5 SUMMARY

In this chapter, we briefly introduce a new method for joint time delay and DOA estimation using two antennas and one receiver for a FH system in the presence of multipath. We derive a Least Square (LS) estimator for both time delay and DOA estimation. We then formulate an iterative Alternating Maximization (AM) algorithm to jointly estimate the time delays and DOA parameters. The simulation results demonstrate the performance of the algorithm.

Chapter 6

Conclusions and Future Work

6.1 CONCLUSIONS

DOA estimation or localization of frequency hopping signal sources is important or useful in a number of applications. FCC requires the future mobile communication systems to have the ability to accurately locate emergency calls made from mobile phones. In military communications using FH technique, DOA estimation is also required for both non-cooperative signal interception and jammer localization.

Although some work has been done in this area, they did not consider the effect of unknown mutual coupling. As a matter of fact, super-resolution DOA estimation algorithms require an accurate knowledge of the array manifold, especially the unknown mutual coupling between the antenna elements. Otherwise, the corresponding performance will be dramatically deteriorated.

Furthermore, all the algorithms of the work mentioned above are only effective in the scenario where the transmitted signals are noncoherent and do not exhibit multipath propagation. However, in typical wireless communication systems where multipath propagation is unavoidable, these algorithms may not be effective.

In this thesis, we propose a new method for DOA estimation in the presence of multipath propagation and mutual coupling for a FH system [52 - 53]. We take these two effects into account and derive a Maximum Likelihood (ML) estimator for both

MC matrix and DOA estimation. We then formulate an iterative Alternating Minimization (AM) algorithm for finding the MC and DOA parameters in an alternate manner. To illustrate the performance of the technique, we simulate the scenario where narrowband signals transmitted from one far-field source impinge on an ULA of two half-wavelength spaced antenna elements via two paths. The simulation results presented illustrate the convergence of the algorithm and its statistical efficiency at high SNR.

In addition, we introduce a new method for joint time delay and DOA estimation using two antennas and one receiver for a FH system in the presence of multipath. We derive a Least Square (LS) estimator for both time delay and DOA estimation. We then formulate an iterative Alternating Maximization (AM) algorithm to jointly estimate the time delays and DOA parameters. The simulation results demonstrate the performance of the algorithm.

6.2 FUTURE WORK

Note that in this thesis, we only consider the scenario of cooperative communication. That is, both the hopping sequence and the hopping instants are assumed to be known.

However, if we want to extend our work to the scenario of non-cooperative communication where both the hopping sequence and the hopping instants are unknown, we have to estimate these parameters first. Perhaps we can get some inspiration from the work done in [16 - 18].

Appendix A

Proof of Equation (4.49)

Note that

$$\frac{\partial \boldsymbol{\Pi}_{\boldsymbol{B}}}{\partial \eta_i} = \frac{\partial\left(\boldsymbol{B}\boldsymbol{B}^{+}\right)}{\partial \eta_i} = \dot{\boldsymbol{B}}\,\boldsymbol{B}^{+} + \boldsymbol{B}\frac{\partial \boldsymbol{B}^{+}}{\partial \eta_i}. \tag{A.1}$$

To prove Equation (4.49), we have to evaluate $\frac{\partial \boldsymbol{B}^{+}}{\partial \eta_i}$ first, which is given by

$$\frac{\partial \boldsymbol{B}^{+}}{\partial \eta_i} = \frac{\partial\left[\left(\boldsymbol{B}^{H}\boldsymbol{B}\right)^{-1}\boldsymbol{B}^{H}\right]}{\partial \eta_i} = \left(\boldsymbol{B}^{H}\boldsymbol{B}\right)^{-1}\frac{\partial \boldsymbol{B}^{H}}{\partial \eta_i} + \frac{\partial\left(\boldsymbol{B}^{H}\boldsymbol{B}\right)^{-1}}{\partial \eta_i}\boldsymbol{B}^{H}. \tag{A.2}$$

Let $\boldsymbol{A}(x)$ be a non-singular square matrix whose elements are differentiable functions with respect to real-valued variable *x*, we get [54]

$$\frac{d}{dx}\boldsymbol{A}^{-1}(x) = -\boldsymbol{A}^{-1}(x)\frac{d\boldsymbol{A}(x)}{dx}\boldsymbol{A}^{-1}(x), \tag{A.3}$$

$$\frac{d}{dx}\boldsymbol{A}^{H}(x) = \left[\frac{d\boldsymbol{A}(x)}{dx}\right]^{H}. \tag{A.4}$$

Therefore, Equation (A.2) can be rewritten as

$$\begin{aligned}
\frac{\partial \boldsymbol{B}^{+}}{\partial \eta_i} &= \left(\boldsymbol{B}^{H}\boldsymbol{B}\right)^{-1}\left(\dot{\boldsymbol{B}}\right)^{H} - \left[\left(\boldsymbol{B}^{H}\boldsymbol{B}\right)^{-1}\frac{\partial\left(\boldsymbol{B}^{H}\boldsymbol{B}\right)}{\partial \eta_i}\left(\boldsymbol{B}^{H}\boldsymbol{B}\right)^{-1}\right]\boldsymbol{B}^{H} \\
&= \left(\boldsymbol{B}^{H}\boldsymbol{B}\right)^{-1}\left(\dot{\boldsymbol{B}}\right)^{H} - \left\{\left(\boldsymbol{B}^{H}\boldsymbol{B}\right)^{-1}\left[\left(\dot{\boldsymbol{B}}\right)^{H}\boldsymbol{B} + \boldsymbol{B}^{H}\dot{\boldsymbol{B}}\right]\left(\boldsymbol{B}^{H}\boldsymbol{B}\right)^{-1}\right\}\boldsymbol{B}^{H} \\
&= \left(\boldsymbol{B}^{H}\boldsymbol{B}\right)^{-1}\left(\dot{\boldsymbol{B}}\right)^{H} - \left(\boldsymbol{B}^{H}\boldsymbol{B}\right)^{-1}\left(\dot{\boldsymbol{B}}\right)^{H}\boldsymbol{B}\left(\boldsymbol{B}^{H}\boldsymbol{B}\right)^{-1}\boldsymbol{B}^{H} - \left(\boldsymbol{B}^{H}\boldsymbol{B}\right)^{-1}\boldsymbol{B}^{H}\dot{\boldsymbol{B}}\left(\boldsymbol{B}^{H}\boldsymbol{B}\right)^{-1}\boldsymbol{B}^{H}
\end{aligned} \tag{A.5}$$

Note that the pseudo-inverse and projection matrix for matrix $\boldsymbol{B}$ are defined by

$$\boldsymbol{B}^{+} = \left(\boldsymbol{B}^{H}\boldsymbol{B}\right)^{-1}\boldsymbol{B}^{H}, \tag{A.6}$$

$$\boldsymbol{\Pi}_{\boldsymbol{B}} = \boldsymbol{B}\boldsymbol{B}^{+}, \tag{A.7}$$

respectively. We can rewrite Equation (A.5) as

$$\begin{aligned}\frac{\partial \boldsymbol{B}^{+}}{\partial \eta_i} &= \left(\boldsymbol{B}^{H}\boldsymbol{B}\right)^{-1}\left(\dot{\boldsymbol{B}}\right)^{H} - \left(\boldsymbol{B}^{H}\boldsymbol{B}\right)^{-1}\left(\dot{\boldsymbol{B}}\right)^{H}\boldsymbol{B}^{+}\boldsymbol{B} - \boldsymbol{B}^{+}\dot{\boldsymbol{B}}\boldsymbol{B}^{+} \\ &= \left(\boldsymbol{B}^{H}\boldsymbol{B}\right)^{-1}\left(\dot{\boldsymbol{B}}\right)^{H} - \left(\boldsymbol{B}^{H}\boldsymbol{B}\right)^{-1}\left(\dot{\boldsymbol{B}}\right)^{H}\boldsymbol{\Pi}_{\boldsymbol{B}} - \boldsymbol{B}^{+}\dot{\boldsymbol{B}}\boldsymbol{B}^{+} \\ &= \left(\boldsymbol{B}^{H}\boldsymbol{B}\right)^{-1}\left(\dot{\boldsymbol{B}}\right)^{H}\boldsymbol{\Pi}_{B}^{\perp} - \boldsymbol{B}^{+}\dot{\boldsymbol{B}}\boldsymbol{B}^{+}\end{aligned}, \tag{A.8}$$

where $\boldsymbol{\Pi}_{B}^{\perp}$ denotes the complementary projection matrix for matrix $\boldsymbol{B}$, that is,

$$\boldsymbol{\Pi}_{B}^{\perp} = \boldsymbol{I} - \boldsymbol{\Pi}_{\boldsymbol{B}}, \tag{A.9}$$

with $\boldsymbol{I}$ being the identity matrix. Substituting (A.8) back into (A.1), we get

$$\begin{aligned}\frac{\partial \boldsymbol{\Pi}_{\boldsymbol{B}}}{\partial \eta_i} &= \dot{\boldsymbol{B}}\boldsymbol{B}^{+} + \boldsymbol{B}\left[\left(\boldsymbol{B}^{H}\boldsymbol{B}\right)^{-1}\left(\dot{\boldsymbol{B}}\right)^{H}\boldsymbol{\Pi}_{B}^{\perp} - \boldsymbol{B}^{+}\dot{\boldsymbol{B}}\boldsymbol{B}^{+}\right] \\ &= \dot{\boldsymbol{B}}\boldsymbol{B}^{+} + \boldsymbol{B}\left(\boldsymbol{B}^{H}\boldsymbol{B}\right)^{-1}\left(\dot{\boldsymbol{B}}\right)^{H}\boldsymbol{\Pi}_{B}^{\perp} - \boldsymbol{B}\boldsymbol{B}^{+}\dot{\boldsymbol{B}}\boldsymbol{B}^{+} \\ &= \dot{\boldsymbol{B}}\boldsymbol{B}^{+} + \boldsymbol{B}\left(\boldsymbol{B}^{H}\boldsymbol{B}\right)^{-1}\left(\dot{\boldsymbol{B}}\right)^{H}\boldsymbol{\Pi}_{B}^{\perp} - \boldsymbol{\Pi}_{\boldsymbol{B}}\dot{\boldsymbol{B}}\boldsymbol{B}^{+} \\ &= \boldsymbol{\Pi}_{\boldsymbol{B}}^{\perp}\dot{\boldsymbol{B}}\boldsymbol{B}^{+} + \boldsymbol{B}\left(\boldsymbol{B}^{H}\boldsymbol{B}\right)^{-1}\left(\dot{\boldsymbol{B}}\right)^{H}\boldsymbol{\Pi}_{B}^{\perp}\end{aligned}. \tag{A.10}$$

Note that for arbitrary non-singular matrix $\boldsymbol{A}$, we have [54]

$$\left(\boldsymbol{A}^{-1}\right)^{H} = \left(\boldsymbol{A}^{H}\right)^{-1}. \tag{A.11}$$

Thus, we get

$$\left(\boldsymbol{B}^{+}\right)^{H} = \left[\left(\boldsymbol{B}^{H}\boldsymbol{B}\right)^{-1}\boldsymbol{B}^{H}\right]^{H} = \boldsymbol{B}\left[\left(\boldsymbol{B}^{H}\boldsymbol{B}\right)^{-1}\right]^{H} = \boldsymbol{B}\left(\boldsymbol{B}^{H}\boldsymbol{B}\right)^{-1}, \tag{A.12}$$

$$\left(\boldsymbol{\Pi}_{\boldsymbol{B}}^{\perp}\right)^{H} = \boldsymbol{I} - \left(\boldsymbol{B}^{+}\right)^{H}\boldsymbol{B}^{H} = \boldsymbol{I} - \boldsymbol{B}\left(\boldsymbol{B}^{H}\boldsymbol{B}\right)^{-1}\boldsymbol{B}^{H} = \boldsymbol{I} - \boldsymbol{B}\boldsymbol{B}^{+} = \boldsymbol{\Pi}_{\boldsymbol{B}}^{\perp}. \tag{A.13}$$

Based on Equation (A.12) and (A.13), Equation (A.10) can be written as

$$\frac{\partial \boldsymbol{\Pi}_{\boldsymbol{B}}}{\partial \eta_i} = \boldsymbol{\Pi}_{\boldsymbol{B}}^{\perp}\dot{\boldsymbol{B}}\boldsymbol{B}^{+} + \left(\boldsymbol{B}^{+}\right)^{H}\left(\dot{\boldsymbol{B}}\right)^{H}\left(\boldsymbol{\Pi}_{B}^{\perp}\right)^{H} = \boldsymbol{\Pi}_{\boldsymbol{B}}^{\perp}\dot{\boldsymbol{B}}\boldsymbol{B}^{+} + \left(\boldsymbol{\Pi}_{\boldsymbol{B}}^{\perp}\dot{\boldsymbol{B}}\boldsymbol{B}^{+}\right)^{H}. \tag{A.14}$$

Therefore, Equation (4.49) in Chapter 4 is proved.

Appendix B

Proof of Equation (4.75)

Note that the first order partial derivative with respect to multipath delays takes the form

$$\frac{\partial L}{\partial \tau_i} = -2\frac{N}{\sigma^2}\mathrm{Re}\left\{\frac{1}{N}\sum_{n=1}^{N} tr\left\{\left[\boldsymbol{FDg}-\boldsymbol{x}(n)\right]\boldsymbol{g}^H\left(\dot{\boldsymbol{D}}_{\tau_i}\right)^H \boldsymbol{F}^H\right\}\right\}. \tag{B.1}$$

And the second order partial derivative with respect to multipath delays is given by

$$\begin{aligned}
\frac{\partial^2 L}{\partial \tau_i \partial \tau_j} &= -2\frac{N}{\sigma^2}\mathrm{Re}\left\{\frac{1}{N}\sum_{n=1}^{N} tr\left\{\boldsymbol{F}\,\dot{\boldsymbol{D}}_{\tau_j}\,\boldsymbol{g}\boldsymbol{g}^H\left(\dot{\boldsymbol{D}}_{\tau_i}\right)^H \boldsymbol{F}^H + \left[\boldsymbol{FDg}-\boldsymbol{x}(n)\right]\boldsymbol{g}^H\left(\ddot{\boldsymbol{D}}_{\tau_i\tau_j}\right)^H \boldsymbol{F}^H\right\}\right\} \\
&= -2\frac{N}{\sigma^2}\mathrm{Re}\left\{tr\left\{\boldsymbol{F}\,\dot{\boldsymbol{D}}_{\tau_j}\,\boldsymbol{g}\boldsymbol{g}^H\left(\dot{\boldsymbol{D}}_{\tau_i}\right)^H \boldsymbol{F}^H\right\} + \frac{1}{N}\sum_{n=1}^{N} tr\left\{\left[\boldsymbol{FDg}-\boldsymbol{x}(n)\right]\boldsymbol{g}^H\left(\ddot{\boldsymbol{D}}_{\tau_i\tau_j}\right)^H \boldsymbol{F}^H\right\}\right\} \\
&= -2\frac{N}{\sigma^2}\mathrm{Re}\left\{tr\left\{\boldsymbol{F}\,\dot{\boldsymbol{D}}_{\tau_j}\,\boldsymbol{g}\boldsymbol{g}^H\left(\dot{\boldsymbol{D}}_{\tau_i}\right)^H \boldsymbol{F}^H\right\}\right\} - 2\frac{N}{\sigma^2}\mathrm{Re}\left\{\frac{1}{N}\sum_{n=1}^{N} tr\left\{\left[\boldsymbol{FDg}-\boldsymbol{x}(n)\right]\boldsymbol{g}^H\left(\ddot{\boldsymbol{D}}_{\tau_i\tau_j}\right)^H \boldsymbol{F}^H\right\}\right\}
\end{aligned}$$

. (B.2)

Take the statistical expectation at both sides of Equation (B.2), we get

$$\begin{aligned}
E\left[\frac{\partial^2 L}{\partial \tau_i \partial \tau_j}\right] &= -2\frac{N}{\sigma^2}\mathrm{Re}\left\{tr\left\{\boldsymbol{F}\,\dot{\boldsymbol{D}}_{\tau_j}\,\boldsymbol{g}\boldsymbol{g}^H\left(\dot{\boldsymbol{D}}_{\tau_i}\right)^H \boldsymbol{F}^H\right\}\right\} \\
&- 2\frac{N}{\sigma^2}\mathrm{Re}\left\{\frac{1}{N}\sum_{n=1}^{N} tr\left\{E\left[\left[\boldsymbol{FDg}-\boldsymbol{x}(n)\right]\boldsymbol{g}^H\left(\ddot{\boldsymbol{D}}_{\tau_i\tau_j}\right)^H \boldsymbol{F}^H\right]\right\}\right\} \\
&= -2\frac{N}{\sigma^2}\mathrm{Re}\left\{tr\left\{\boldsymbol{F}\,\dot{\boldsymbol{D}}_{\tau_j}\,\boldsymbol{g}\boldsymbol{g}^H\left(\dot{\boldsymbol{D}}_{\tau_i}\right)^H \boldsymbol{F}^H\right\}\right\} \\
&- 2\frac{N}{\sigma^2}\mathrm{Re}\left\{\frac{1}{N}\sum_{n=1}^{N} tr\left\{E\left[\left(\boldsymbol{FDg}-\boldsymbol{x}(n)\right)\right]\boldsymbol{g}^H\left(\ddot{\boldsymbol{D}}_{\tau_i\tau_j}\right)^H \boldsymbol{F}^H\right\}\right\}
\end{aligned} \tag{B.3}$$

Note that, in the second term of Equation (B.3), $\left[\boldsymbol{FDg}-\boldsymbol{x}(n)\right]$ corresponds to an AWGN noise vector, therefore,

$$E\left[\left(\boldsymbol{FDg}-\boldsymbol{x}(n)\right)\right] = \boldsymbol{0}. \tag{B.4}$$

Substituting Equation (B.4) back into Equation (B.3) yields

$$E\left\{\frac{\partial^2 L}{\partial\tau_i\partial\tau_j}\right\} = -2\frac{N}{\sigma^2}\mathrm{Re}\left\{tr\left\{\boldsymbol{F}\,\dot{\boldsymbol{D}}_{\tau_j}\,\boldsymbol{g}\boldsymbol{g}^H\left(\dot{\boldsymbol{D}}_{\tau_i}\right)^H\boldsymbol{F}^H\right\}\right\}. \tag{B.5}$$

Thus, Equation (4.75) in Chapter 4 is proved. Likewise, the other expectations of second order partial derivatives in Section 4.3.3, Chapter 4 can be proved in this way.

References

[1] J. G. Proakis and M. Salehi, “Communication systems engineering”, Prentice-Hall Inc., Englewood Cliffs, New Jersey, 1994.

[2] T. S. Rappaport, “Wireless communications”, Prentice-Hall Inc., Englewood Cliffs, New Jersey, 1996.

[3] J. G. Proakis, “Digital communications” 4th edition, Mc-Graw Hill, 2000.

[4] R. O. Schmidt, “Multiple emitter location and signal parameter estimation”, *IEEE Trans. on Antennas and Propagation*, vol. AP-34, No. 3, March 1986.

[5] R. Roy and T. Kailath, “ESPRIT-Estimation of signal parameters via rotational invariance techniques”, *IEEE Trans. on Acoustics, Speech, and Signal Processing*, vol. 37, pp. 984-995, July 1986.

[6] R. Kumaresan and D. W. Tufts, “Estimating the angles of arrival of multiple plane waves”, *IEEE Trans. on Aerospace and Electronic Systems*, AES-19(1), Jan. 1983.

[7] M. Viberg and B. Ottersten, “Sensor array processing based on subspace fitting”, *IEEE Trans. on Signal Processing*, vol. 39, No. 5, May 1991.

[8] P. Stoica and K. C. Sharman, “Maximum likelihood methods for direction of arrival estimation”, *IEEE Trans. on Signal Processing*, vol. 38, No. 7, July 1990.

[9] T. J. Shan, M. Wax and T. Kailath, “On spatial smoothing for directions of arrival estimation of coherent signals”, *IEEE Trans. on Acoustics, Speech and Signal Processing*, vol. ASSP-33, No. 4, pp. 806-811, April 1985.

[10] R. Muhamed and T. S. Rappaport, “Direction of arrival estimation using antenna arrays”, in *Technical Report MPRG-TR-96-03*, Mobile & Portable Radio Research Group, Virginia Tech, Blackburg, VA, Jan. 1996.

[11] A. G. Jaffer, “Maximum likelihood direction finding of stochastic sources: a separable solution”, in *IEEE Proc. ICASSP 88,* vol. 5, pp. 2893-2896, April 1988.

[12] Y. Bresler and A. Macovski, “Exact maximum likelihood parameter estimation of superimposed exponential signals in noise*”, IEEE Trans. on Signal Processing*, vol. ASSP-34, No. 5, Oct. 1986.

[13] J. Sheinvald, M. Wax and A. J. Weiss, “On maximum likelihood localization of coherent signals”, *IEEE Trans. on Signal Processing*, vol. 44, No. 10, Oct. 1996.

[14] D. Starer and A. Nehorai, “Maximum likelihood estimation of exponential signals in noise using a Newton algorithm”, in *Proc. The Fourth Annual ASSP Workshop on Spectrum Estimation and Modeling*, Aug. 1988, pp. 240 – 245.

[15] J. Li and R. T. Compton, “Maximum likelihood angle estimation for signals with known waveforms”, *IEEE Trans. on Signal Processing*, vol. 41, No. 9, Sep. 1993.

[16] A. N. Lemma, A. J. Van der Veen, and E.F. Deprettere, "Joint angle frequency estimation for slow frequency hopping signals", in *IEEE Proc. Circuits, System Signal processing Workshop,* Nov. 1998, pp. 363-370.

[17] K. T. Wong, "Blind beamforming /geolocation for wideband-FFHs with unknown hop-sequences", *IEEE Trans. on Aerospace and Electronic Systems*, vol. 37, No.1, pp. 65-76, Jan. 2001.

[18] X. Liu, N. D. Sidiropoulos and A. Swami, "Blind high resolution localization and tracking of multiple frequency hopped signals", *IEEE Trans. on Signal Processing*, vol. 50, No. 4, pp. 889-901, April 2002.

[19] J. Sheinvald and M. Wax, "Localization of multiple signals using subarrays data", in *Advances in Spectrum Analysis and Array Processing*, S. Haykin, Editor, Englewood cliffs, NJ: Prentice-Hall, 1995, Chapter 7, vol.3, pp. 324-351.

[20] J. Sheinvald and M. Wax, "Direction finding with fewer receivers via time-varying preprocessing", *IEEE Trans. on Signal Processing*, vol. 47, No.1, Jan. 1999.

[21] S. M. Kay, "Fundamentals of Statistical Signal Processing: Estimation Theory", Prentice Hall, Englewood Cliffs, New Jersey, 1994.

[22] I. J. Gupta and A. A. Ksienski, "Effect of mutual coupling on the performance of adaptive arrays", *IEEE Trans. on Antenna and Propagation*, vol. AP-31, No.5, pp. 785-791, March. 1991.

[23] Y. Rockah and P.M. Schultheiss, “Array shape calibration using sources in unknown locations-Part I: Far field sources,” *IEEE Trans. on Acoust., Speech, Signal Processing*, vol. ASSP-35, pp.286-299, March 1987.

[24] _____, “Array shape calibration using sources in unknown locations-Part II: Near-field sources and estimator implementation,” *IEEE Trans. on Acoust., Speech, Signal Processing*, vol. ASSP-35, pp.724-735, June 1987.

[25] A. Paulraj and T. Kailath, “Direction of arrival estimation by eigenstructure methods with unknown sensor gain and phase”, in *IEEE Proc. Int. Conf. Acoust., Speech, Signal Processing (ICASSP) 85*, Tampa, Florida, March 1985.

[26] G. C. Brown, J. H. McClellan and E. J. Holder, “A phased array calibration technique using eigenstructure methods”, in *IEEE Proc.International Radar Conference,* 1990.

[27] J. Pierre and M. Kaveh, “Experimental performance of calibration and direction finding algorithm”, in *IEEE Proc. ICASSP,* vol. 2, pp. 1365-1368, April 1991.

[28] A. G. Jaffer, “Constrained mutual coupling estimation for array calibration”, *IEEE Trans. on Aerospace and Electronic Systems*, vol. 13, No.2, pp. 336-342, Oct. 1996.

[29] H. M. Aumann, A. J. Fenn and F. G. Willwerth, “Phased array antenna calibration and pattern prediction using mutual coupling measurements”, *IEEE Trans. on Antennas Propagation,* vol 37, pp 844-851, 1989.

[30] C. Roller and W. Wasylkiwskyj, “Effect of mutual coupling on super-rsolution DF in liner arrays”, in *IEEE Proc. ICASSP,* No.5, pp. 257-260, 1992.

[31] K. M. Pasala and E. M. Friel, “Mutual coupling effects and their reduction in wideband direction of arrival estimation”, *IEEE Trans. Aerospace and Electronic Systems*, vol. 30, No.4, pp. 1116-1122, Oct. 1994.

[32] K. R. Dandekar, H. Ling and G. Xu, “Experimental study of mutual coupling compensation in smart antenna applications”, *IEEE Trans. on Wireless Communications,* vol. 1, No. 3, pp. 480-487, July 2002.

[33] Y. Inoue, K. Mori and H. Arai, “DOA estimation in consideration of the array element pattern”, in *IEEE Proc. Vehicular Technology conference (VTC)*, vol. 2, pp. 745-748, May 2002.

[34] R. S. Adve and T. K. Sarkar, “Compensation for the effects of mutual coupling on direct data domain adaptive algorithms”, *IEEE Trans. on Antennas and Propagation,* vol. 48, No.1, pp. 86-94, Jan. 2000.

[35] T. K. Sarkar, M. C. Wicks, M. S. Palma and R. J. Bonneau, “Smart antennas”, John Wiley&Sons, Inc., Hoboken, New Jersey, 2003.

[36] B. Friedlander and A. J. Weiss, “Direction finding in the presence of mutual coupling”, *IEEE Trans. on Antenna and Propagation,* vol. 39, No.3, pp. 271-284, March 1991.

[37] E. K. L. Hung, “A critical study of a self-calibrating direction-finding method for Arrays”, *IEEE Trans. on Signal Processing*, vol. 42, No.2, pp. 471-474, Feb. 1994.

[38] T. Svantesson, “Mutual coupling compensation using subspace fitting”, in *IEEE Proc. Sensor Array and Multichannel Signal Processing Workshop,* 2000, pp. 494-498.

[39] J. Mao, B. Champagne, M. O’Droma and K. Kwiat, “Separable dimension subspace method for joint signal frequencies, DOAs and sensor mutual coupling estimation”, in C*onf. Rec. 34th Asilomar Conf. on Signals, Systems and Computers*, vol. 1, pp. 605-609, Nov. 2000.

[40] A. G. Jaffer, “Sparse mutual coupling matrix and sensor gain/phase estimation for array auto calibration”, in *IEEE Proc. Radar Conference* 2002. pp. 294-297.

[41] H. Krim and M. Viberg, “Two decades of array signal processing research”, *IEEE Signal processing magazine*, pp. 67-94, July 1996.

[42] I. Ziskind and M. Wax, “Maximum likelihood localization of multiple sources by alternating projection”, *IEEE Trans. on Acoustics, Speech and Signal Processing*, vol. 36, No. 10, pp. 1553-1560, Oct. 1988.

[43] M. S. Bazaraa, H. D. Sherali and C. M. Shetty, “Nonlinear programming – theory and algorithms ”, John Wiley & Sons Inc., New York, 1993.

[44] H. H. Jenkins, “Small-aperture radio direction finding”, Norwood, MA: Artech House, 1991.

[45] J. Sheinvald and M. Wax, “Localization of multiple signals using subarrays data”, *Advances in Spectrum Analysis and Array Processing*, S. Haykin, Editor, Englewood cliffs, NJ: Prentice-Hall, 1995, Chapter 7, vol.3, pp. 324-351.

[46] A. Zeira and B. Friedlander, "Direction finding with time-varying arrays", *IEEE Trans. on Signal Processing*, vol. 43, No. 4, April 1995.

[47] B. Friedlander and A. Zeira, "Eigenstructure-based algorithms for direction finding with time-varying arrays", *IEEE Trans. on Aerospace and Electronics Systems*, vol 32, No. 2, April 1996.

[48] J. Sheinvald and M. Wax, "Direction finding with fewer receivers via time-varying preprocessing", *IEEE Trans. on Signal Processing*, vol. 47, No.1, Jan. 1999.

[49] J. G. Worms, "RF direction finding with a reduced number of receivers by sequential sampling", *IEEE conference on Phased Array Systems and Technology*, May 2000, pp.165-168.

[50] E. Fishler and H. Messer, "Multiple source direction finding with an array of m sensors using two receivers", *IEEE Proceeding of the 10th Workshop on Statistical Signal and Array Processing*, 2000, pp. 86-89.

[51] J. Tabrikian and A. Faizakov, "Optimal preprocessing for source localization by fewer receivers than sensors", *IEEE Proceedings of the 11th Signal Processing Workshop on Statistical Signal Processing*, pp. 213 - 216, Aug. 2001.

[52] Q. Bao, C. C. Ko and W. Zhi, "DOA estimation in the presence of unknown mutual coupling and multipath propagation in a frequency hopping communication system", *IEEE International Conference on Communications (ICC)*, vol. 5, pp. 2502-2506, June 2004, Paris, France.

[53] Q. Bao, C. C. Ko and W. Zhi, “DOA estimation in the presence of unknown mutual coupling and multipath propagation in a frequency hopping communication system”, accepted by *IEEE Trans. on Aerospace and Electronic Systems*.

[54] H. Lutkepohl, “Handbook of matrices”, John Wiley & Sons Inc., New York, 1996.