

**SPACE-TIME PROCESSING FOR
WIRELESS BASE STATIONS**

APPROVED BY
SUPERVISING COMMITTEE:

Brian L. Evans, Supervisor

Gustavo de Veciana

To my parents without whose inspiration and support this report would have
been a dream.

**SPACE-TIME PROCESSING FOR
WIRELESS BASE STATIONS**

by

SRIKANTH GUMMADI, B. Tech.

REPORT

Presented to the Faculty of the Graduate School of
The University of Texas at Austin
in Partial Fulfillment
of the Requirements
for the Degree of

MASTER OF SCIENCE IN ENGINEERING

THE UNIVERSITY OF TEXAS AT AUSTIN

December, 1998

Acknowledgments

I am deeply grateful to Dr. Brian L. Evans who, in addition to being a great advisor has been a good friend and an inspiring example. Brian accepted me into the Embedded Signal Processing Laboratory (ESPL), provided financial assistance and guided me in my research area. I will always remember my stay at ESPL as a period of learning and productivity, and this was possible only through the dedication and enthusiasm of Brian. It has been a great pleasure and privilege to work with him.

I would like to thank all the members of Motorola's Fort Worth Research Lab for giving me a chance to work on a challenging project as a summer intern. This project is a part of this report. My special thanks to Bill Alexander, Eric Schorman, Walt Rozanski, Jim O'Connor and Mark Harrison. I will be joining them after finishing my MSEE.

I want to thank all the members of ESPL for the great time I had at UT. ESPL has a collection of excellent graduate students who work on various topics complementing each other. Among others, we have the ever helping Güner, the enthusiastic Niranjan, the intelligent Wade, the busy Jong-Il and the hard working David and Magesh. I owe special thanks to our lab manager Biao for helping me with \LaTeX , proof reading my report, buying the books needed for research, and making the lab more productive.

I would also like to thank the Indian community and my friends in Austin, India, and the United States for their encouragement and friendship.

I thank my roommates Adduri, Kiran (Pondy) and KV for putting up with me for 2 years. I also thank Ravi, Praveen, Neeraj, Parminder, Kartick, Amey and Arun. Special thanks to Neeraj and Parminder who made tea (*chai*) which helped me work late night on this report. Special thanks to my cousin Nagesh and sister-in-law Bhramaramba for the great time I had during my internship.

I would like to express my gratitude to my committee member Dr. Gustavo de Veciana for taking the time to read my report. His course on communication networks is among the best courses I have taken at UT.

I am forever indebted to my parents, brother, sister-in-law, and Shreya for their love, sacrifice, and support.

SRIKANTH GUMMADI

The University of Texas at Austin

December, 1998

SPACE-TIME PROCESSING FOR WIRELESS BASE STATIONS

SRIKANTH GUMMADI, M.S.E.

The University of Texas at Austin, 1998

Supervisor: Brian L. Evans

The explosive growth in the demand for cellular communications services has driven research into improving spectrum efficiency, battery life, and link quality in wireless systems. Using space-time diversity in antenna arrays for transmission and reception is an emerging solution to all three problems. A space-time system operates simultaneously on all of the antennas by processing signal samples both in space and time. This extra dimension enables interference cancellation in a way that is not possible with single antenna systems.

At the base station, received signals suffer from cochannel signal interference from other users and degradation from the channel. This report discusses a Constant Modulus (CM) array for separating cochannel users and 2-D unitary ESPRIT for estimating channel parameters. Each stage in a CM array consists of a CM beamformer and an adaptive signal canceler that recovers one cochannel signal. The 2-D unitary ESPRIT algorithm jointly estimates the direction-of-arrival (DOA) and time-of-arrival (TOA) of each path impinging on an antenna array. By knowing the DOA and TOA, we can locate the mobile user, and transmit the signal in a narrow beam towards the mobile, which improves spectral efficiency, link quality and battery life. The key contributions of this report are (i) a modified error criterion for CM array, which makes the CM array phase sensitive, and (ii) a closed-form solution for joint angle and delay estimation, which significantly reduces computational complexity.

Table of Contents

Acknowledgments	iv
Abstract	vi
Table of Contents	vii
List of Figures	x
1. Introduction	1
1.1 Wireless Communications	1
1.2 Space-Time Processing	3
1.3 Estimation of Channel Parameters	4
1.4 Radio Propagation Model	6
1.4.1 Path Loss and Fading	7
1.4.2 Multipath	8
1.5 Characterizing Propagation in Wireless Cells	10
1.5.1 Macrocell propagation	11
1.5.2 Microcell and picocell propagation	11
1.6 An Overview of the Report	12
2. Constant Modulus Array	14
2.1 Introduction	14
2.2 Background	16

2.2.1	Constant Modulus Beamformer	16
2.2.2	Adaptive Signal Canceler	17
2.3	Channel Model for Digital Signals	17
2.4	Modified Error Criterion	19
2.4.1	Decision Directed Beamforming	20
2.4.2	The Modified Error Criterion	21
2.5	Computer Simulations	23
2.6	Conclusion	26
3.	Joint Angle and Delay Estimation (JADE) using	
	2-D Unitary ESPRIT	27
3.1	Introduction	27
3.2	Discrete Space-Time Channel Model	30
3.2.1	Transmitter Model	31
3.2.2	Receiver Model	32
3.2.3	Channel Model	33
3.3	Channel Model Transformation	35
3.4	Deconvolution	36
3.5	Joint Angle and Delay Estimate	38
3.6	Computer Simulations	39
3.7	Conclusion	41
A.	Notation and Abbreviations	43
A.1	Notation	43
A.2	Abbreviations	44

B. 2-D Unitary ESPRIT	46
B.1 Introduction	46
B.2 Real-Valued Processing with a Uniform Linear Array	47
B.3 Unitary ESPRIT for Uniform Linear Array	48
B.4 2-D Unitary ESPRIT	52
 BIBLIOGRAPHY	 58
 Vita	 62

List of Figures

1.1	In a wireless network, each base station services a cell. Each base station is connected to the Mobile Telephone Switching Office (MTSO) and the MTSO is connected to the Public Switched Telephone Network (PSTN).	2
1.2	Fading of a transmitted signal	7
1.3	Macrocell multipath propagation and scattering	10
1.4	Microcell multipath propagation and scattering	12
2.1	Multistage constant modulus array. Signals $x(k)$, $y_m(k)$, and $e(k)$ are complex-valued.	15
2.2	A decision directed beamformer shown for an array of four sensors.	20
2.3	Performance of constant modulus array for two different error criteria: (a) power of fading signals, (b) output power of first beamformer using traditional CMA, and (c) output power of first beamformer using modified error criterion. The traditional CMA locks onto the signal of highest power which alternates between the signal of interest and the interference signal. The modified CMA locks onto the signal of interest.	24
2.4	Cochannel Signals (SNR = 10 dB) separated by a CM array using the modified error criterion.	25

3.1	Location of mobile is given by intersection of (a) three TOA circles or (b) two TDOA hyperbolas or (c) two DOA lines for three-sites visibility or (d) TOA circle and DOA line for single-site visibility.	29
3.2	Parameters associated with square-root of raised cosine pulse shape: (a) time domain and (b) frequency domain.	36
3.3	The root mean square error of TOA vs. SNR for the prompt ray.	40
3.4	The root mean square error of DOA vs. SNR for the prompt ray.	40
B.1	Unitary ESPRIT algorithm for a uniform linear array of N sensors.	52
B.2	2-D unitary ESPRIT for a uniform rectangular array of $N \times M$ sensors.	56

Chapter 1

Introduction

1.1 Wireless Communications

The wireless era has its origins in the 1895 demonstration by Guglielmo Marconi of the use of radio waves to communicate over long distances [1]. In the 1960's, Bell Laboratories developed the *cellular concept*. Cellularization consists of breaking up a large geographical service area into smaller cells, each of which can reuse a portion of the available bandwidth (*frequency reuse*), thus making it possible to provide wireless links to many users despite limited spectrum. Channels are reused when there is enough distance between cells to avoid mutual interference (*cochannel interference*). In each cell, a *base station* communicates with the mobile users via a radio link on one side and on the other side with the Mobile Telephone Switching Office (MTSO), as shown in Figure 1.1. The MTSO is connected to the Public Switched Telephone Network (PSTN). Cellularization is used in most commercial mobile radio communications systems, e.g. Advanced Mobile Phone Systems (AMPS), Digital European Cordless Telephone (DECT), Global System for Mobile (GSM), and the North American Digital Cellular System (IS-54, IS-36, and IS-95) [2].

Cells have irregular shapes and dimension, and are determined largely by terrain features and man-made objects. Depending on their size, cells can be classified as macrocells (where the base station has enough transmit power

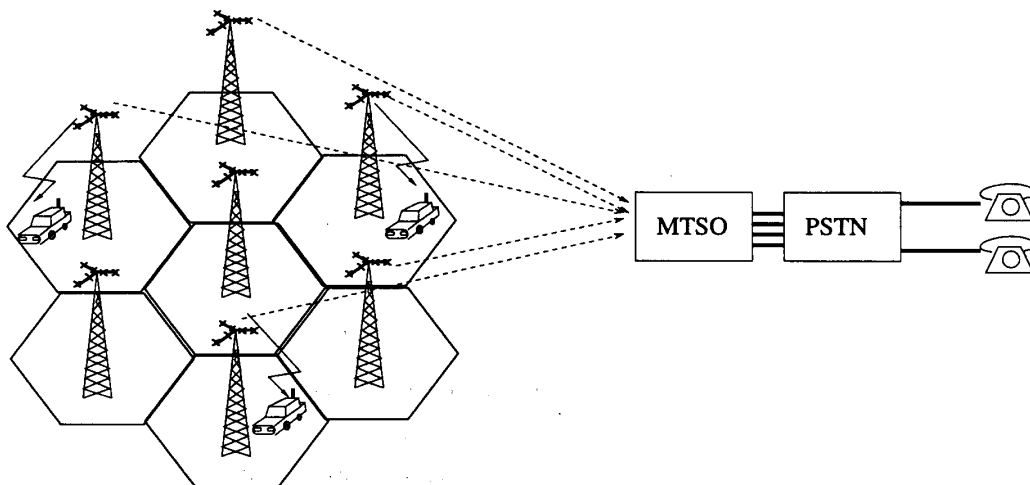


Figure 1.1: In a wireless network, each base station services a cell. Each base station is connected to the Mobile Telephone Switching Office (MTSO) and the MTSO is connected to the Public Switched Telephone Network (PSTN).

to cover a radius of 1–20 km), microcells (radius of 0.1–1 km) and picocells (indoor areas). Cellular systems operate in the 800–900 MHz and 1800–2000 MHz radio frequency bands.

Wireless communication systems, unlike their wireline counterparts, pose several unique challenges: *(i)* a variable number of users operating on the same limited allocated spectrum; *(ii)* the radio propagation environment and mobility of users give rise to signal fading and spreading in time, space, and frequency; *(iii)* the limited battery life at the mobile user imposes restrictive power constraints; and *(iv)* interference due to frequency reuse. Over the last 20 years, many technologies have been developed to offer solutions to these challenges, including multiple access, channel coding, and space-time signal processing techniques.

1.2 Space-Time Processing

In the early development of signal processing for wireless communication systems, the received signal was sampled only in time using one antenna. The goal of *space-time processing* is to combine spatial and temporal information gathered by sampling the signal with an array of antenna elements. Space-time signal processing techniques exploit the new spatial dimension to add spatial diversity among different users. These techniques increase coverage and capacity, mitigate multipath effects and cochannel interference, simplify user hand-off from one base station to another, improve security, and increase power efficiency [3]. These benefits come at the expense of increased complexity at the base station, because of the extra antennas and embedded computers placed at the base station.

The embedded processors used at the base station usually have dual-core architectures which combine an advanced programmable digital signal processor (DSP) with a reduced instruction set computer (RISC) microcontroller, e.g. Motorola's MC56652 and MC56651 and Lucent's DSP1620 and DSP16210. These processors have substantial on-chip memory, a sophisticated interface that allows its two cores to function together seamlessly, and a mature peripheral set for cellular applications. The DSP core compresses and decompresses speech, and controls reliable transmission and reception of the digital signal via radio waves. The microcontroller core generates control signals, frame headers, and power control bits.

Each signal received by an antenna array can be characterized by its direction-of-arrival (DOA) and time-of-arrival (TOA), which are often called

its *signature*. The interference and desired signals generally arrive at the antenna array with distinct and often well separated signatures, even in complex multipath environments. We can exploit this difference to reduce cochannel interference. In the transmitter, we can use space-time processing to deliver signals to the desired mobile using spatial selectivity to minimize the interference to other mobiles. Wireless channels are rapidly time-varying, thus the use of training for equalization not only consumes bandwidth but is also inefficient, as we have to transmit the training sequence often. The use of blind equalization (i.e., equalization without the use of a training sequence) gives an additional advantage when used with space-time processing. Therefore, blind channel equalization and estimation of multiple users signals can improve network capacity and performance.

1.3 Estimation of Channel Parameters

In wireless systems, a transmitted signal may arrive at the receiver along multiple propagation paths. Each path has its own direction-of-arrival (DOA), time-of-arrival (TOA), and attenuation (fading). As the transmitted signal (a radio wave) propagates through the environment, reflection, diffraction and scattering occur each time the signal collides with an object such as a building, hill, or tree. Each collision causes a change in direction (azimuth/elevation) and amplitude (attenuation). The different lengths of the propagation paths lead to different time delays.

Several motivations exist for estimating the channel parameters, esp. the DOAs and TOAs. Classical applications are source localization in sonar

systems and modeling layers of the earth in seismic survey systems. Two modern applications include mobile user localization and channel equalization and tracking.

Many methods for mobile user localization are based on DOA and/or time difference of arrival (TDOA) estimation of the received signals [4]. The time difference of arrival of a signal at two base stations traces a hyperbola, and intersection of two hyperbola gives the location of mobile. A key application of mobile positioning is personal safety, such as in emergency localization (E-911 service) and automatic location identification of cell phone users. Federal Communications Commission (FCC) regulations state that these features be implemented by the year 2001 [5]. Mobile positioning can also be used by advanced user hand-off schemes and by many user services for which a global positioning system (GPS) receiver is impractical [4]. Other applications are automatic billing and fraud detection for cellular providers, accident reporting, law enforcement, cargo tracking, and intelligent transportation systems [4].

Channel equalization and tracking aids directive transmission in the downlink. Also, knowledge of the more slowly varying channel parameters (such as the DOAs) allows for effective channel tracking in fast-changing environments. For example, an improved receiver could obtain DOA/TOA estimates and then use a Viterbi algorithm to jointly find the signals and channel by updating only the fading and signals at each state [6].

Techniques for estimating channel parameters can be classified into two categories: spectral-based and parametric. Spectral-based techniques such as MUSIC [7] form a spectrum (a function) of the desired parameters and search

the spectrum for peaks, which determine the value of the desired parameters. A parametric approach such as ESPRIT [8] involves optimization of a highly non-linear function of the parameters and yields more accurate estimates. The complexity of spectral-based techniques increases exponentially in the number of parameters while that of parametric techniques increases linearly. This report develops closed-form parametric solutions that rely on models of radio propagation which are described next.

1.4 Radio Propagation Model

Understanding the physics of radio frequency (RF) wave propagation is crucial to the development of realistic models for use in space-time wireless signal processing systems. Radio wave propagation is a complex phenomenon which cannot be completely described by a single model. We describe propagation models that are largely empirical. They are drawn from field measurements and often contain simplifying assumptions. Due to the random fluctuations of the various parameters associated with RF wave propagation, the models are often statistical rather than deterministic. As a radio signal propagates through the environment, it experiences (*i*) a decrease in power level due to path loss and fading as described in Section 1.4.1 and (*ii*) spreading in space, time, or frequency, due to multipath effects and user mobility as described in Section 1.4.2.

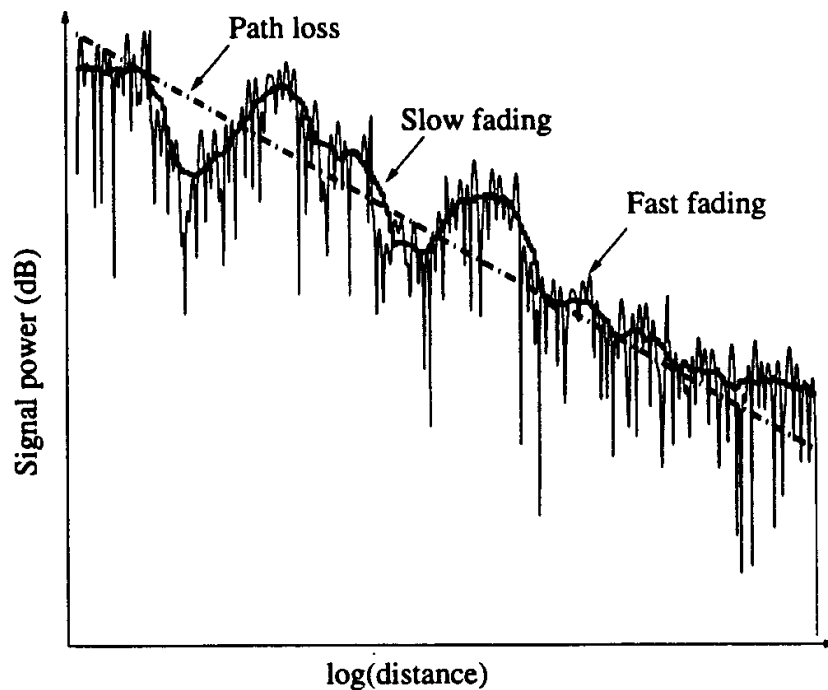


Figure 1.2: Fading of a transmitted signal

1.4.1 Path Loss and Fading

An important measure of link quality in mobile communications is path loss. It is defined as the ratio between the received and transmitted power. Path loss increases with distance, carrier frequency, and the number of large obstructions. The mean received signal level varies with distance d as d^{-n} , where n is a parameter in the range of 2-5. The value of n depends on the type of environment [9], e.g. $n = 2$ is realistic for free space propagation. The more built-up and obstructed the environment, the larger the value of n .

In addition to path loss, the signal exhibits fluctuations in power level. These fluctuations, called fading, also depend on distance. The two types of fading are *slow* (or *long-term*, or *shadowing*) and *fast* (or *short-term*) fading, as

shown in Figure 1.2. A signal experiences slow fading when it is shadowed by obstructions between the transmitter and receiver such as hills and buildings. The envelope of a slow-fading signal is determined statistically by the local mean of the fast-fading signal, which is the average signal level for few tens of wavelengths.

Fast fading is caused by multiple reflections of the transmitted wave by objects around the mobile such as houses and trees. Since the waves scattered by these objects have different attenuation and phases, they may add constructively or destructively, thereby causing fast fluctuations in the signal level. The received signal power may change by a few orders of magnitude (e.g. 20-40 dB) within a few wavelengths (50 ns at a carrier frequency of 1.8 GHz). When the mobile is completely obstructed from the base station, i.e., when there is no direct *line-of-sight*, the envelope of the received signal is best modeled as a Rayleigh distribution [9]. When we have direct line-of-sight, the received signal can be modeled as a Rician distribution. Usually in an urban environment, we do not have a line-of-sight for about 75% of the time.

1.4.2 Multipath

Multipath phenomena are caused by objects (*scatterers*) lying in the environment in which the radio signal is propagating. Multipath causes the spreading of signals in time and space (and also in frequency if the source is moving). So, the received signal consists of multiple time-delayed replicas of the transmitted signal that have arrived from various directions. Multipath effects are caused by *reflection*, *diffraction*, and *scattering* of the propagating waves [10]. Reflection occurs when a propagating wave impinges upon an ob-

struction with dimensions that are very large relative to its wavelength, e.g. the earth surface and buildings. In refraction, components of the radio wave travel into the obstruction medium. Most buildings are made of materials that absorb most of the energy of the wave, so that the refracted wave is not significant in strength when compared to the reflected wave. Reflection and refraction occur according to Snell's laws. Diffraction occurs when the radio path between the transmitter and receiver is obstructed by an impenetrable object; then, according to Huyghen's principle, secondary waves form behind this object [10]. This phenomenon explains how radio waves arrive at the receiver even though there is no direct line-of-sight, as is the case in many urban environments. Scattering occurs when the wave impinges upon objects of dimensions that are on the order of the wavelength (or less), such as street signs and lamp posts. Scattering causes the energy of the wave to be radiated in many directions.

The relative importance of these propagation mechanisms depends on the particular environment. If there is a direct line-of-sight between the mobile and base station, then reflection dominates the propagation. If the mobile is in a heavily build-up area with no line-of-sight to the base station, then diffraction and scattering will play a major role.

Doppler spread is caused by the movement of the mobile or other objects in the environment. The maximum value of this frequency shift is given by $f_D = \frac{2V_m f_c}{c}$, where f_D is the maximum Doppler frequency, V_m is the velocity of the mobile, f_c is the communication carrier frequency, and c is the velocity of propagation.

To summarize, multipath propagation results in signal spreading in

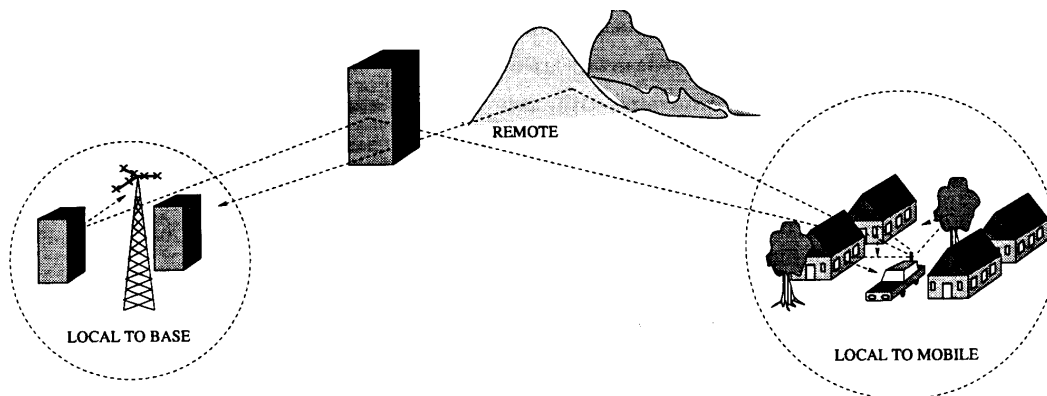


Figure 1.3: Macrocell multipath propagation and scattering

time (*delay spread*), space (*angle spread*), and frequency (*Doppler spread*). In typical outdoor cellular systems, the delay spread is on the order of 0-10 microseconds, angle spread ranges from 2 to 60 degrees, and Doppler spread varies from 5 to 200 Hz [3].

1.5 Characterizing Propagation in Wireless Cells

The propagation of signal in wireless cells depends on the type of scatterers, the size of cell, and the elevation of the antenna at the base station. Propagation can be classified based on the size of the cell as macrocell, microcell and picocell propagation. Both microcell and picocell have similar characteristics as the antenna for both of them is below the roof top. Macrocell propagation is discussed in Section 1.5.1, and microcell and picocell propagation is discussed in Section 1.5.2.

1.5.1 Macrocell propagation

In the context of macrocell mobile communications, the scattering objects in the environment can be classified as *scatterers local to base*, *scatterers local to mobile*, and *remote scatterers*, as shown in Figure 1.3. An object is called a local scatterer if it is (i) it is greater in size than the operating wavelength and (ii) higher than the antenna [11]. Scatterers local to base are local structures (e.g., buildings and trees) in the proximity of the base station. They can cause large angle spread of the incoming signal, but small delay spread and no additional Doppler spread. Scatterers local to mobile are structures within a few meters of the mobile. Because of the relatively low height of the mobile antenna, these scatterers can cause severe angular spread, as well as Doppler spread, when the vehicle is in motion. From the point of view of the base station, however, these scatterers cause only small delay and angle spread. Finally, remote scatterers, also known as dominant reflectors, are terrain features (e.g. hills) or high-rise buildings located far (hundreds of meters) from the base or mobile. They give rise to *specular multipath*, and cause large delay and angle spreads, but no additional Doppler spread.

1.5.2 Microcell and picocell propagation

In microcells and picocells, the base station antenna is usually below roof top level. The numerous scatterers, such as building and cars, cause short delay spread, large angle spread, and some Doppler spread of the signal arriving at the base station, as shown in Figure 1.4. When a line-of-sight exists, a Rician model for fast signal fading is more suitable and scattering is non-isotropic [9]. Compared to macrocell environments, the number of paths may be very large.

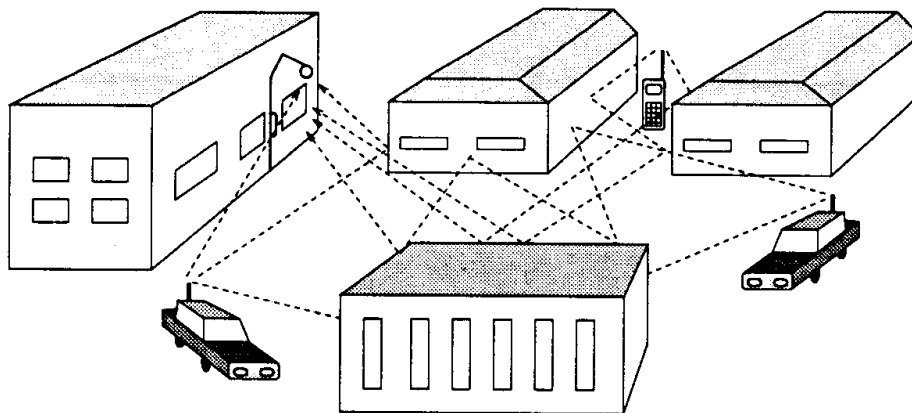


Figure 1.4: Microcell multipath propagation and scattering

Indoor environments exhibit scatterers (e.g., walls and furniture) as well. Base Station antennas are at roughly the same height as mobile antennas, such as on the same floor of a building. This causes the large angle spread found in microcells. A line-of-sight may or may not exist. In addition, the signal can also experience the “hallway effect”, in which a hallway acts as a wave guide [2].

Because microcells and picocells often exhibit a large number of paths and insufficient delay spread (as the propagation paths are small), the techniques for joint angle and delay estimation derived in this work may not readily apply. Therefore, we focus our attention to macrocell environments.

1.6 An Overview of the Report

This report address two key issues in wireless macrocell base stations: (i) cochannel interference rejection and (ii) joint estimation of spatial and temporal parameters of the channel based on space-time processing. We make the following assumptions:

- Source signals are received by an array consisting of at least two antennas;
- The environment exhibits specular (as opposed to diffuse) multipath propagation;
- Doppler shifts and residual carriers of sources are not appreciable in small time periods (200 ns); and
- Source signals are linearly modulated with a known pulse shaping function.

Based on the introduction to wireless communications systems in this chapter, Chapter 2 discusses the Constant Modulus (CM) Array for cochannel interference rejection. Recovery of multiple cochannel signals based on sequential application of CM array is also shown. A class of algorithms for joint angle-delay estimation (referred to as JADE) are discussed in Chapter 3. Chapter 3 also introduces a closed-form solution for JADE using a uniform linear array and 2-D unitary ESPRIT. These ESPRIT-like algorithms take advantage of uniform sampling in space and time; thus, the estimation of DOAs/TOAs reduces to solving a generalized eigenvalue problem. A discrete receiver data model (sampled in space and time) is also presented. Appendix A describes the notation and abbreviations used in this report. Appendix B derives 2-D unitary ESPRIT.

Chapter 2

Constant Modulus Array

2.1 Introduction

In mobile wireless communication systems, the transmitted signal may be unintentionally reflected, refracted, or scattered on its way to the receiver. The received signal is a linear combination of delayed, scaled, and attenuated versions of the transmitted signal. Other users transmitting at the same frequency cause cochannel interference. Vehicle motion affects the received frequency because of Doppler shift and creates standing waves. Standing waves produce regions of high and low amplitudes, which is known as fading. Fading causes both attenuation and phase shift, which in turn cause errors at the receiver. Attenuation decreases SNR and phase shifts rotate the signal constellation which cause errors at the receiver.

One way to reduce errors at the receiver is to use a smart antenna system. Smart antenna systems improve signal recovery in severe cochannel signal environments. One smart antenna system, the multistage constant modulus (CM) array [12], is capable of separating cochannel signals. Figure 2.1 shows the stages in a CM array. Each stage consists of two components: (i) a weight-and-sum beamformer adapted by the constant modulus algorithm (CMA) [13] that captures one source, and (ii) a signal canceler adapted by the least-mean-squares (LMS) algorithm [14] that removes the captured source from the array

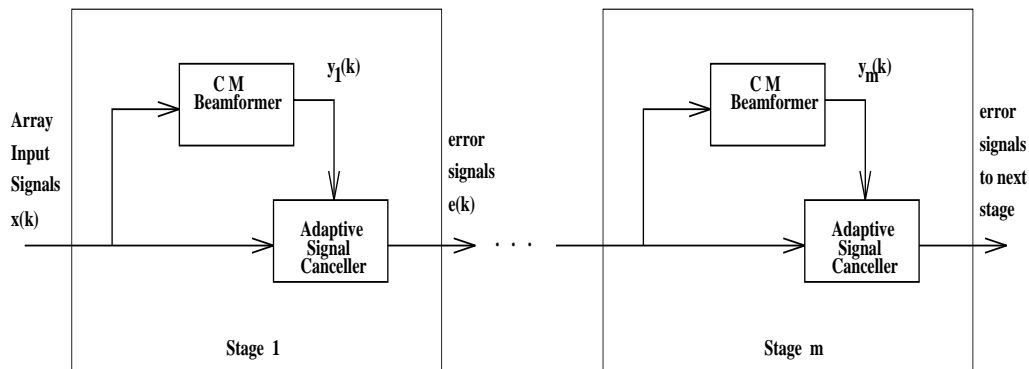


Figure 2.1: Multistage constant modulus array. Signals $x(k)$, $y_m(k)$, and $e(k)$ are complex-valued.

input.

The CMA is a *blind equalization* algorithm (i.e., does not require a training or pilot signal) that makes use of the property that the transmitted signals are of constant amplitude. The CMA in [13] is insensitive to the phase of received signals because the error which is used to update the weights comprises of only amplitude differences. In this chapter, we propose a modification to the CM array which makes the error dependent on both phase and amplitude of the output. In digital communications, the phase of the received signal is important because a shift in the phase would rotate the constellation of the received signal thus causing decision errors at the decoder.

Section 2.2 describes the operation of a CM array and the CM algorithm in [13]. Section 2.3 describes our channel model. Section 2.4 discusses decision-directed equalization and introduces our modification to the CMA error criterion. Section 2.5 presents computer simulations to show the robustness of signal tracking using the modified error criterion in a fading environment. Section 2.6 concludes the chapter. This chapter is an expanded version of [15].

2.2 Background

Each stage of the constant modulus (CM) array consists of a constant modulus beamformer and an adaptive signal canceler. The purpose of the beamformer is to lock onto and track a particular user. Section 2.2.1 describes the weight-and-sum beamformer which is adapted by the constant modulus algorithm (CMA). Section 2.2.2 describes the adaptive signal canceler which uses a least-mean-squares (LMS) algorithm [14] to remove the captured source from the array input.

2.2.1 Constant Modulus Beamformer

The CM beamformer with complex input $\mathbf{x}(k)$ and complex output $y(k) = \mathbf{w}^H(k) \mathbf{x}(k)$ is shown in Figure 2.1. The complex weight vector $\mathbf{w}(k) = [w_1(k) \ w_2(k) \ \cdots \ w_N(k)]^T$ is updated using the CM algorithm according to

$$\mathbf{w}(k+1) = \mathbf{w}(k) + 2 \mu_{cma} \mathbf{x}(k) \epsilon_{cma}^*(k) \quad (2.1)$$

where $\mu_{cma} > 0$ is the step size and $\epsilon_{cma}(k)$ is the CMA error

$$\epsilon_{cma}(k) = \frac{y(k)}{|y(k)|} - y(k) = y(k) \left(\frac{1}{|y(k)|} - 1 \right) \quad (2.2)$$

where $|y(k)| = \sqrt{y(k) y^*(k)}$. The term $\frac{y(k)}{|y(k)|}$ is called the instantaneous modulus (amplitude) of the received signal. The error criterion in (2.2) does not contain any phase information; thus, the update in (2.1) is phase insensitive.

The update of the weights in (2.1) is similar to the update used in the Least Mean Squares (LMS) algorithm [14] where the instantaneous modulus $\frac{y(k)}{|y(k)|}$ acts as the “desired response” signal. The CMA tries to make the

instantaneous modulus constant. Thus, it can be used only with signals of constant modulus (amplitude) such as FSK and QPSK [16].

The value of step size μ_{cma} has to be chosen appropriately (as is the case with most of the adaptive algorithms) so that the algorithm converges fast. If the value is too small, the algorithm takes a long time to converge; if it is too large, then the algorithm may diverge. If $0 < \mu_{cma} < \frac{2}{\lambda_{max}}$, where λ_{max} is the maximum eigenvalue of input autocorrelation matrix, then the algorithm converges [17].

2.2.2 Adaptive Signal Canceler

Every stage of a CM array contains an adaptive signal canceler. The output $y(k)$ is weighted by the canceler weights $\mathbf{u}(k) = [u_1(k) \ u_2(k) \ \cdots \ u_N(k)]^T$ which is subtracted from $\mathbf{x}(k)$ to generate the error vector $\mathbf{e}(k)$. This error vector serves as the input to the next stage and is also used to update the canceler weights according to

$$\mathbf{u}(k+1) = \mathbf{u}(k) + 2 \mu_{lms} y^*(k) \mathbf{e}(k) \quad (2.3)$$

where $\mu_{lms} > 0$ is the step size. The weights of canceler estimate the columns of the array response matrix which in turn gives an estimate of the directions of arrival of various signals. We define the array response matrix in the next section.

2.3 Channel Model for Digital Signals

We assume that the transmitted signals are narrowband and that the receiver antenna array is in the far field of the transmitter. The baseband

analog waveform transmitted by the l^{th} source is

$$s_l(t) = \sum_{n=0}^{\infty} d_l(n) g(t - nT), \quad l = 1, \dots, L \quad (2.4)$$

where $d_l(n)$ are the digital symbols, T is the symbol period, and $g(t)$ is the pulse shape. The $d_l(n)$ terms are symbols of constant modulus (amplitude) such as BPSK and QPSK so that points in the constellation for the $d_l(n)$ terms lie on a circle of the same radius. We assume that the l^{th} source propagates along M_l paths, where each multipath has a different attenuation and propagation delay [18].

The received baseband signal at the m^{th} antenna element of a uniform linear array is

$$x_m(t) = \sum_{l=1}^L \sum_{i=1}^{M_l} \alpha_{li}(k) e^{j2\pi f_c \tau_{li}} e^{j\phi_{mli}} s_l(t - \tau_{li}) + n_m(t) \quad (2.5)$$

where f_c is the carrier frequency, τ_{li} is the propagation delay of the i^{th} multipath of the l^{th} signal, $\alpha_{li}(k)$ is the corresponding attenuation, $n_m(t)$ is white Gaussian noise and

$$\phi_{mli} = 2\pi \frac{d}{\lambda} (m - 1) \sin(\theta_{li})$$

where d is the inter-sensor spacing, λ is the carrier wavelength, and θ_{li} is the signal angle of arrival (AOA) for path i of the l^{th} source. The array input can be compactly written as

$$\mathbf{x}(k) = \mathbf{A}(k) \mathbf{s}(k) + \mathbf{n}(k) \quad (2.6)$$

where \mathbf{A} is called the array response matrix. We assume that $\mathbf{s}(k)$ and $\mathbf{n}(k)$ have zero mean and are uncorrelated with each other. We model the attenuation coefficients by using a Rayleigh fading channel [9] with Doppler shift. The

attenuation terms are given by

$$\alpha_{li}(k) = \frac{1}{\sqrt{M}} \sum_{m=1}^M e^{j(\Omega_i k \cos(\gamma_m) + \phi_m)} \quad (2.7)$$

where $\Omega_i = 2\pi \frac{f_i}{f_s}$, such that f_i is the Doppler shift of the i^{th} source and f_s is the symbol rate, M is the number of scattering paths received, $\gamma_m = \frac{2\pi m}{M}$ which assumes that the scatterers are uniformly distributed, and ϕ_m is a random variable uniformly distributed on $[-\pi, \pi]$ representing the initial phase of the m th scattering path.

2.4 Modified Error Criterion

In this section, we convert a constant modulus (CM) array into a robust smart antenna by modifying the error criterion to be a weighted sum of the conventional CM array error and decision-directed equalization error. The new error criterion enables the CM array to (1) separate digital cochannel signals with multipath and inter-symbol interference and (2) track fading signals. The key contribution is that the modified error criterion adds phase sensitivity to the otherwise phase insensitive CM error criterion. Phase is important for constellations which are based on phase like BPSK and QPSK. Section 2.4.1 explains decision directed equalization and shows the possibility of error propagation if the decisions are in error. Section 2.4.2 derives the proposed modified error criterion.

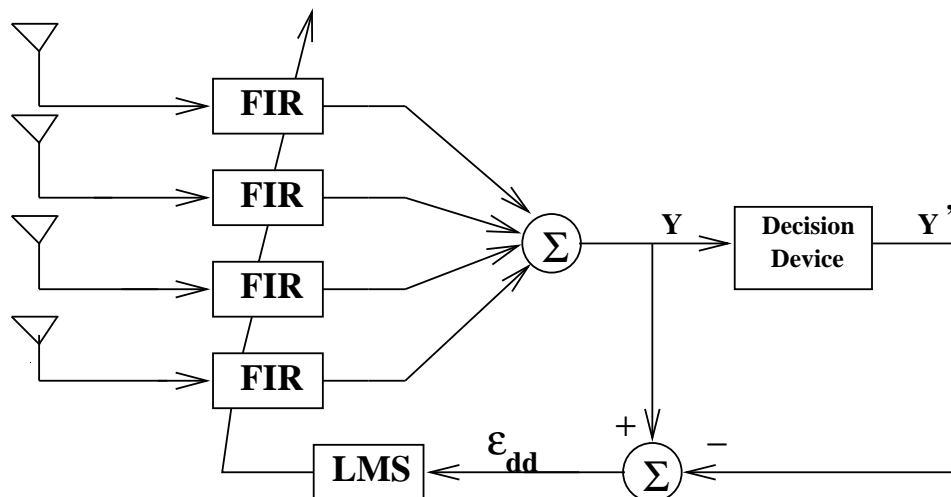


Figure 2.2: A decision directed beamformer shown for an array of four sensors.

2.4.1 Decision Directed Beamforming

If the output of a communication channel were the correct transmitted sequence, then the output may be used as the “desired” response for the purpose of adaptive equalization. This method of equalization, called *Decision Directed Equalization* [17], can only be used if the output is free of errors; otherwise, an error in output will propagate through the receiver. The error ϵ_{dd} for updating the weights using decision directed equalization is

$$\epsilon_{dd} = \mathbf{Y} - \mathbf{Y}' \quad (2.8)$$

where \mathbf{Y} is the output of the beamformer and \mathbf{Y}' is the output of the decision device.

The primary disadvantage of using decision directed equalization is error propagation when wrong decisions are made. In a real-time application, pure decision directed equalization cannot be performed. When we start the

receiver, we get errors as the weights of the receiver filter are not set correctly. Since the decisions are in error, we cannot use decision directed equalization to update the weights of the receiver filter; thus, an initial training signal is needed.

2.4.2 The Modified Error Criterion

The primary advantages of CM beamformers are that they do not require a reference signal (because they perform blind equalization) and they can be implemented in real time (even on fixed-point processors). Its close resemblance to the LMS algorithm means that a hardware or software subsystem configured to use the complex LMS algorithm could be used for the CM algorithm.

From (2.2), ϵ_{ema} does not contain any phase information in it; i.e., the update in (2.1) is insensitive to phase shifts. Therefore, a conventional CM beamformer would not be able to give the desired response for a wireless communication system with fading effects. For signals like QPSK and FSK, the decision depends on the phase of the output wave and not on the amplitude. Thus, a phase shift in the output could result in many wrong decisions. In addition, the CM array captures the source having the maximum power. When deep fading occurs, the CM beamformer captures the interfering source, which causing erroneous decisions to be made. Thus, the need arises for a modified error criterion that would keep the advantages of CM array while removing its disadvantages.

The key advantage in decision directed equalization is its sensitivity to

phase in updating the weights, which prevents an error if we have a phase offset. Decision directed equalization also has the ability to track small frequency offsets because it would adjust its weights to track the change in frequency. For a conventional CM array, however, a small frequency offset would result in a phase offset that cannot be corrected by CMA due to its insensitivity to phase. The decision directed equalization would also latch onto a captured source even in the case of deep fading. All of the above predictions are based on the assumption that the decisions being made are correct, and this assumption may not be valid in practice.

To overcome the drawbacks of the CM array, we propose a new error criterion which is a weighted mean of the CM error ϵ_{cma} and the decision directed error ϵ_{dd}

$$\epsilon(k) = \alpha_{cma} \epsilon_{cma}(k) + \alpha_{dd} \epsilon_{dd}(k) \quad (2.9)$$

where α_{cma} is the weight of CM error and α_{dd} is the weight of decision directed error and

$$\alpha_{cma} + \alpha_{dd} = 1 \quad (2.10)$$

During the initial stage when the weights of the beamformer have not converged, i.e., when the decisions may not be correct, we use only the CM error by setting $\alpha_{cma} = 1$ and $\alpha_{dd} = 0$. As the weights converge and the output becomes stable and decisions become correct, we move from pure CM error to the modified error given in (2.9) with both α_{cma} and α_{dd} being non-zero. We determine this transition based on the absolute value of the CM error. When the error ϵ_{cma} becomes less than a threshold (1% of the maximum CM error in our case), then we switch from a CM error criterion to modified error crite-

tion [19]. The modified error builds phase sensitivity into the update equation and improves the performance of the CM array in the following ways:

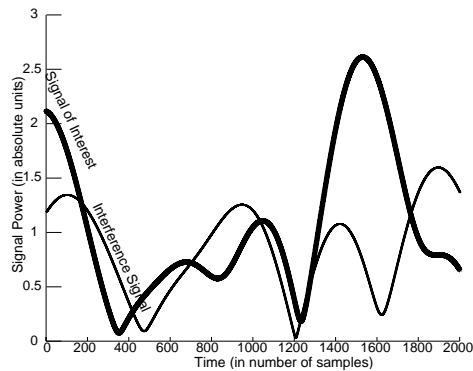
- it latches onto a captured signal irrespective of the power of the signal,
- it overcomes small frequency offsets, and
- it reduces phase offsets in the CM array.

The first property implies that the modified error criterion gives the correct output even when the captured signal has destructive fading, which occurs when the power of the captured signal is less than that of the interference. The new error criterion would perform better than the CM error by itself and make decisions more reliable.

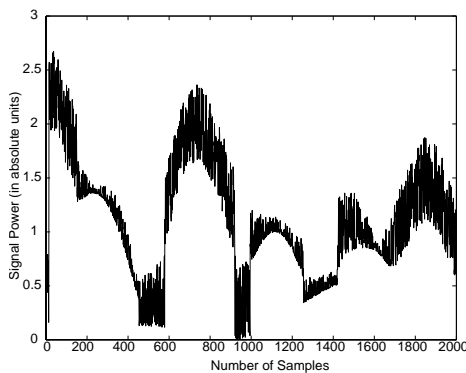
2.5 Computer Simulations

We present the results of a computer simulation for the case where the modified error could separate two cochannel sources undergoing fading but the traditional CM array fails to latch onto one signal. The simulation uses $L = 2$ users, $N = 4$ antenna elements, and $M_1 = M_2 = 12$ multipaths. The symbol rate f_s is 24,300 baud and the fading frequency f_i is 72 Hz to correspond to a vehicle traveling at approximately 45 MPH with a transmitting frequency of 1.8 GHz. For the pulse shape $g(t)$ we used a square-root raised-cosine spectrum [20] with a roll-off parameter of $\beta = 0.35$. We transmit QPSK signals.

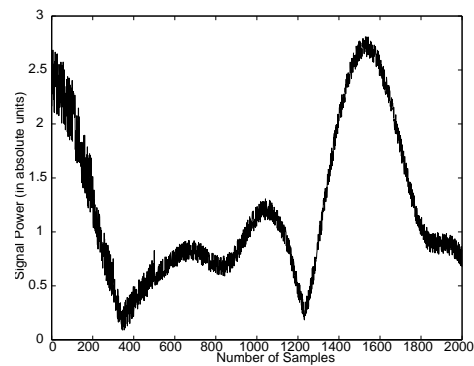
Figure 2.3(a) shows the amplitude of the received signal in a fading channel. Initially, the amplitude of the first signal is greater than that of the second. The first signal fades with time, and eventually, the power of the second



(a)



(b)



(c)

Figure 2.3: Performance of constant modulus array for two different error criteria: (a) power of fading signals, (b) output power of first beamformer using traditional CMA, and (c) output power of first beamformer using modified error criterion. The traditional CMA locks onto the signal of highest power which alternates between the signal of interest and the interference signal. The modified CMA locks onto the signal of interest.

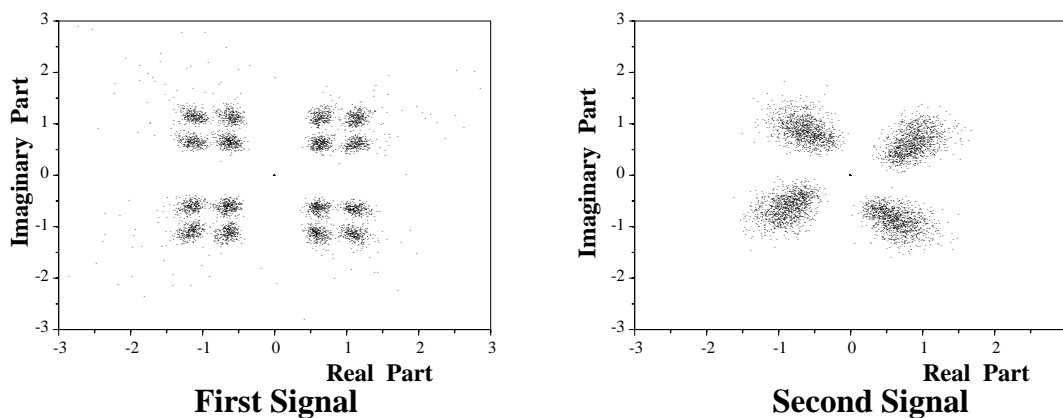


Figure 2.4: Cochannel Signals (SNR = 10 dB) separated by a CM array using the modified error criterion.

signal becomes greater than that of the first. We refer to the crossover point as the critical point. A traditional CM array would capture the first signal until the critical point. After the critical point, the CM array would be forced to capture the second signal, thereby causing errors at the decision device because the decision device thinks that it is still receiving the first signal. Figure 2.3(b) plots the output signal power using the traditional CMA. The CM beamformer clearly captures the interference signal when interference becomes stronger than the signal of interest.

We overcome the problem of capturing different sources by using the modified error criterion. Figure 2.3(c) shows the output signal power using the modified error criterion. Using the modified error criterion, the CM array latches onto one signal even during deep fades. Figure 2.4 displays the received constellation of the separated signals using the modified error criterion. The signals have been clearly separated without any errors after we switch from pure CM error to the modified error.

2.6 Conclusion

We present a new error criterion, which is a weighted mean of the Constant Modulus and Decision Directed Error criteria, to improve the performance of CM beamformer. The new error criterion allows the beamformer to latch onto a signal even during deep fades because it adds phase sensitivity to the beamformer. The use of a modified CM error criterion requires an additional comparison and two additions, which is insignificant with respect to the complexity of the CM array. The cochannel signal separator has been shown to work in simulations of severe signal conditions, e.g. when the cochannel interference is stronger than the signal of interest. Additional advantages of using the CM cochannel signal separator are that (a) its canceler weights can be used to estimate the angles of arrival, and (b) it can be used to estimate the number of cochannel signals.

The step size μ_{cma} for the beamformer and μ_{lms} for the canceler could change with the situation. They should be appropriately chosen so that both the algorithms converge and do so in a few iterations. A typical value for both μ_{cma} and μ_{lms} is 0.01. Another decision that the designer has to make is the transition from CM mode to modified error mode. We perform the shift when the error from the CM mode falls below a threshold of 1% to the maximum CM error. This shift should be performed once the receiver is sure that the decisions on the output are error free. It could be made either empirically or when the mean CM error reaches a steady value.

Chapter 3

Joint Angle and Delay Estimation (JADE) using 2-D Unitary ESPRIT

3.1 Introduction

In wireless cellular communications, locating the position of mobile user has many applications. Federal communication commission regulations state that the wireless emergency 911 (E-911) feature be implemented by the year 2001 [5]. Additional personal safety applications include emergency road-side service, early warning evacuation, and emergency locator service. Other applications include location-sensitive billing, mobile yellow pages, and driving directions.

Adding the global positioning system (GPS) to each new cell phone is a solution that would increase cost and power consumption but not support the 58 million current cellular customers. Thus, a cost-effective solution should not be subscriber-based but should be based on existing cellular infrastructure. Infrastructure-based solutions depend on computing one or more of the following:

- Time-of-arrival (TOA)
- Time difference-of-arrival (TDOA)
- Direction-of-arrival (DOA)

Infrastructure-based solutions are classified into two groups: (i) three-site visibility and (ii) single-site visibility, based on the number of base stations with which the mobile can communicate.

With three-site visibility, the infrastructure-based solutions can use triangulation to locate the mobile unit. The location of mobile units may be given by the

- intersection of the three TOA circles as shown in Fig. 3.1(a),
- intersection of any two TDOA hyperbolas as shown in Fig. 3.1(b), or
- intersection of any two DOA lines as shown in Fig. 3.1(c).

In practical situations, we do not use DOAs by themselves because the error in estimating DOAs is usually larger than the error in estimating TOAs. For single-site visibility, a combination of the TOA and DOA can be used to locate the mobile unit. The intersection of TOA circle with the DOA line gives the location of mobile, as shown in Figure 3.1(d).

This chapter presents an algorithm for the joint high-resolution estimation of multipath DOAs and delays that uses the shift-invariance properties of a special sensor array geometry. This algorithm gives a closed-form low-complexity solution for jointly estimating and automatically pairing the DOAs and TOAs (delays). The first step of this algorithm consists of transforming the data (i.e., channel estimates) by a Discrete Fourier Transform (DFT), which maps delay in the time domain into phase shifts in the frequency domain, followed by a deconvolution by the known pulse shaping function. Deconvolution

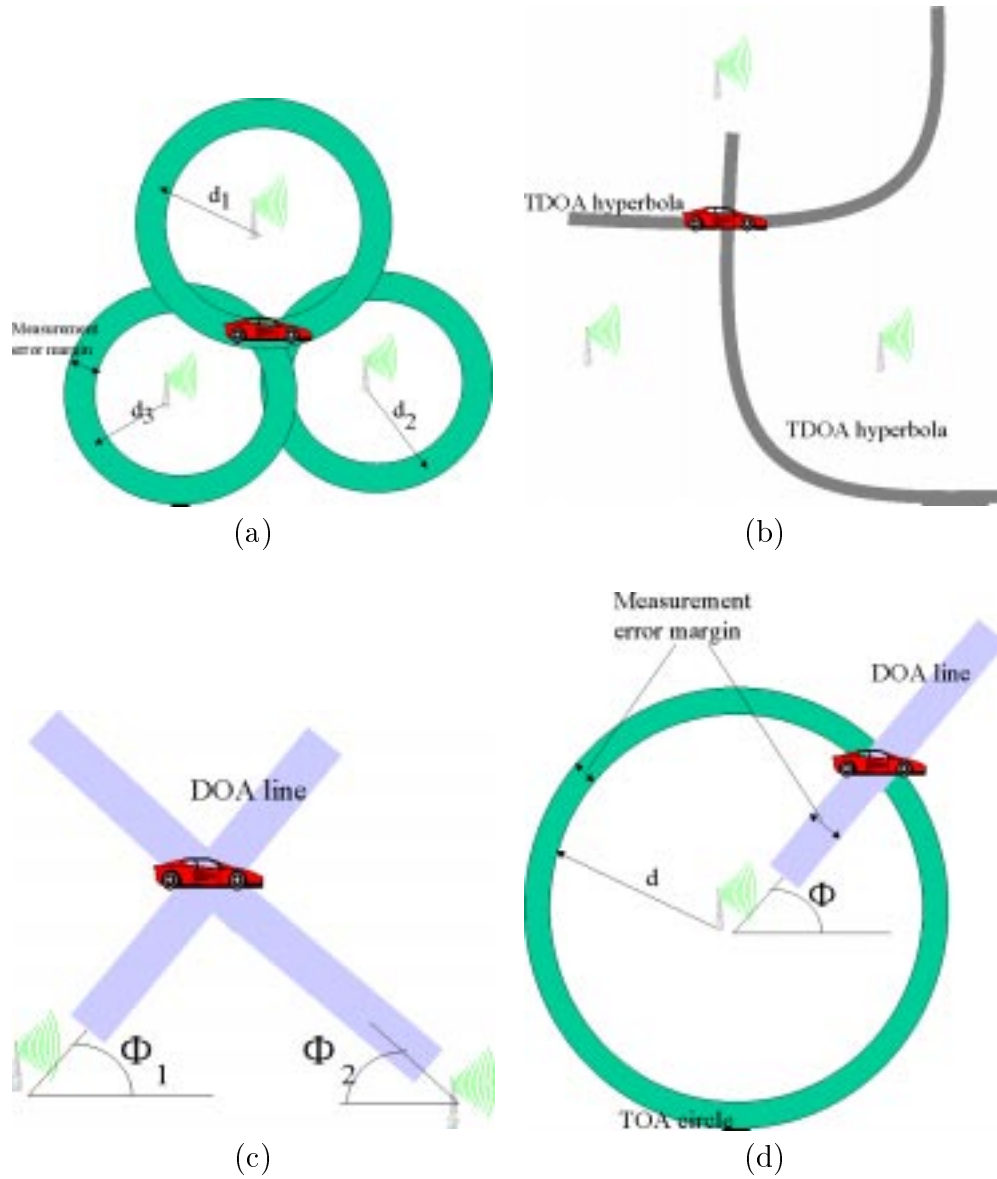


Figure 3.1: Location of mobile is given by intersection of (a) three TOA circles or (b) two TDOA hyperbolas or (c) two DOA lines for three-sites visibility or (d) TOA circle and DOA line for single-site visibility.

is done via the DFT, as it is more accurate if the number of samples is small, because there are no edge effects in the frequency domain. The algorithm incorporates many techniques developed for ESPRIT and DOA estimation, such as spatial smoothing [21] and forward-backward averaging [22]. The number of paths may be more than the number of antennas, which overcomes a limitation of the sequential estimation of TOA and DOA [23], which assumes that the number of paths are no more than the number of antenna elements.

Section 3.2 describes the space-time transmitter, receiver, and channel model. Section 3.3 derives the channel model transformation that is required for TOA estimation. Section 3.4 discusses the deconvolution via the DFT method. Section 3.5 presents JADE-ESPRIT and Section 3.6 gives computer simulations of JADE-ESPRIT. Section 3.7 concludes this chapter.

3.2 Discrete Space-Time Channel Model

We derive a channel model for wireless communications that is sampled in space and in time. This section is less concerned with the details of radio propagation, but instead adopts a space-time signal processing point of view. We make the following assumptions:

- The sensor array has a doublet (pairwise equal antennas) structure, e.g., uniform linear array (ULA) whose sensors are spaced at a half-wavelength or closer.
- The signal received by the antennas is sampled above the Nyquist rate.
- Source signals are received by an array of at least two antennas;

- Source signals are linearly modulated with a known pulse shaping function.

Section 3.2.1 describes the space-time transmitter model. The receiver model is discussed in Section 3.2.2, and the channel model is derived in Section 3.2.3.

3.2.1 Transmitter Model

We focus on the case of a single user transmitting a modulated digital signal in a specular multipath environment. Digital modulation is the process by which a digital baseband signal is converted into an bandpass signal for RF transmission. The digital sequence $\{s_k\}$ is modulated by a pulse shaping function $g(t)$, such that, assuming linear modulation for simplicity, the baseband transmitted signal is the convolution of the digital signal with the modulation waveform

$$s(t) = \sum_l s_l g(t - lT_s) \quad (3.1)$$

where T_s is the symbol period. The digital signal is described by a sequence of Dirac delta impulses $\sum s_l \delta(t - lT_s)$. The data points s_l come from a signal constellation (a set of k -point vectors). In Binary Phase Shift Keying (BPSK), $s_l = +1$ or -1 . The modulated waveform (i.e., the output of the transmitter) is a function of the pulse shape. For example, the GSM system uses binary signals with Gaussian Minimum Shift Keying (GMSK) which is a type of nonlinear modulation [2]. This is a modulation scheme in which the phase of the carrier is instantaneously varied by the “modulating” signal (i.e. the information to transmit).

3.2.2 Receiver Model

Let the number of receiving antennas be M . If we assume for simplicity that the source and sensors are located in the same plane, and that the source is in the far-field of the array (i.e., the impinging wave appears as a plane wave rather than a spherical wave), then only the DOA θ characterizes the location of the source. The m th antenna response (i.e., the measured power) is proportional to

$$A_m(\theta) = G_m(\theta)e^{-j2\pi f_c \tau_m(\theta)}$$

where $G_m(\theta)$ is the complex amplitude or gain of the m th sensor towards direction θ (for isotropic sensors, $G_m(\cdot) = 1$), and $\tau_m(\theta)$ is the propagation delay between the reference point and the m th sensor for a wavefront impinging from θ . If one of the sensors (say, the first one) is taken as a reference point, then the array response is

$$\mathbf{a}(\theta) = \begin{bmatrix} G_1(\theta) \\ G_2(\theta)e^{-j2\pi f_c \tau_2(\theta)} \\ \vdots \\ G_M(\theta)e^{-j2\pi f_c \tau_M(\theta)} \end{bmatrix} \quad (3.2)$$

Let the received signal at the antenna array be arranged into an $M \times 1$ vector $\mathbf{x}(t)$. The continuous-time received signal can be modeled from (3.2) as

$$\mathbf{x}(t) = \sum_{i=1}^Q \mathbf{a}(\theta_i)\beta_i s(t - \tau_i) + \mathbf{n}(t) \quad (3.3)$$

where $\mathbf{n}(t)$ is additive noise (e.g., thermal and measurement noise), Q is the number of discrete paths that are received by the antenna array, and β_i is the

path attenuation of i th path. The output of antenna array can be represented as the convolution

$$\mathbf{x}(t) = s(t) * \mathbf{c}(t) + \mathbf{n}(t) \quad (3.4)$$

where $\mathbf{c}(t)$ is the channel impulse response,

$$\mathbf{c}(t) = \sum_{i=1}^Q \mathbf{a}(\theta_i) \beta_i \delta(t - \tau_i) \quad (3.5)$$

$$\mathbf{x}(t) = \sum_l s_l \mathbf{h}(t - lT) + \mathbf{n}(t) \quad (3.6)$$

$h(t)$ is the *channel pulse response*

$$\mathbf{h}(t) = \sum_{i=1}^Q \mathbf{a}(\theta_i) \beta_i g(t - \tau_i) \quad (3.7)$$

which is the channel response to the pulse shaping function $g(t)$. We assume that $g(t)$ has finite support; i.e., it is non-zero only for $t \in [0, L_g T)$, where L_g is the number of non-zero samples of $g(t)$. Then, the (integer) channel length is L , where $LT = L_g T + \Delta\tau$ seconds and $\Delta\tau$ is the delay spread. Thus, $\mathbf{h}(t)$ is non-zero only for $t \in [0, LT)$.

3.2.3 Channel Model

Let the data be collected over N symbol periods. Let the oversampling factor be P , i.e., at each symbol period we take P samples of $\mathbf{x}(t)$. Without loss of generality, it can be assumed that the sampling is perfectly synchronized with the transmission, so we sample exactly at instants $t = kT$. Then, (3.6) can be rewritten as

$$\mathbf{x}(kT) = \sum_l s_l \mathbf{h}((k-l)T) + \mathbf{n}(kT), \quad k = 0, \frac{1}{P}, \dots, N - \frac{1}{P} \quad (3.8)$$

Given that the channel is FIR of length L taps,

$$\mathbf{x}(kT) = \sum_{l=\lfloor k-L+1 \rfloor}^{\lfloor k \rfloor} s_l \mathbf{h}((k-l)T) + \mathbf{n}(kT), \quad k = 0, \frac{1}{P}, \dots, N - \frac{1}{P} \quad (3.9)$$

If we collect the channel vectors into an $M \times PL$ matrix

$$\mathbf{H} = \left[\mathbf{h}(0) \ \mathbf{h}\left(\frac{T}{P}\right) \ \dots \ \mathbf{h}\left(\left(L - \frac{1}{P}\right)T\right) \right] \quad (3.10)$$

then based on the channel model in (3.7),

$$\mathbf{H} = [\mathbf{a}(\theta_1) \ \dots \ \mathbf{a}(\theta_Q)] \begin{bmatrix} \beta_1 & & 0 \\ & \ddots & \\ 0 & & \beta_Q \end{bmatrix} \begin{bmatrix} \mathbf{g}^T(\tau_1) \\ \vdots \\ \mathbf{g}^T(\tau_Q) \end{bmatrix} = \mathbf{A}(\boldsymbol{\theta}) \mathbf{D} \mathbf{G}^T(\boldsymbol{\tau}) \quad (3.11)$$

where $\mathbf{g}(\tau_i) = [g(kT - \tau_i)]_{k=0, \frac{1}{P}, \dots, L - \frac{1}{P}}$ is a LP -dimensional column vector containing the samples of $g(t - \tau_i)$, $\boldsymbol{\theta} = [\theta_1 \ \dots \ \theta_Q]$, $\boldsymbol{\tau} = [\tau_1 \ \dots \ \tau_Q]$, $\mathbf{D} = \text{diag}(\boldsymbol{\beta})$ and $\boldsymbol{\beta}^T = [\beta_1 \ \dots \ \beta_Q]$. Note that $\boldsymbol{\theta}$, $\boldsymbol{\tau}$, and $\boldsymbol{\beta}$ vary with time and are stationary only for a short intervals.

For notation simplicity, we drop the dependence of \mathbf{A} on $\boldsymbol{\theta}$ and \mathbf{G} on $\boldsymbol{\tau}$. Assuming that the antenna array gain is omni-directional, \mathbf{A} can be written as

$$\mathbf{A} = [\mathbf{a}'(\theta_1) \ \dots \ \mathbf{a}'(\theta_Q)] \quad (3.12)$$

where

$$\mathbf{a}'(\theta_i) = \begin{bmatrix} 1 \\ e^{-j2\pi f_c \tau_2(\theta_i)} \\ \vdots \\ e^{-j2\pi f_c \tau_M(\theta_i)} \end{bmatrix}$$

The derivation of 2D-ESPRIT given in Appendix B uses the above form of \mathbf{A} .

3.3 Channel Model Transformation

Let $\mathbf{g}(0)$ be the vector samples of the known waveform $g(t)$ and let $\tilde{\mathbf{g}}$ be its DFT, $\tilde{\mathbf{g}}^T = \mathbf{g}^T F$, where F denotes the DFT matrix of size $LP \times LP$ defined by

$$F = \begin{bmatrix} 1 & 1 & \cdots & 1 \\ 1 & \phi & \cdots & \phi^{LP-1} \\ \vdots & \vdots & \vdots & \vdots \\ 1 & \phi^{LP-1} & \cdots & \phi^{(LP-1)^2} \end{bmatrix} \quad \phi = e^{-j\frac{2\pi}{LP}} \quad (3.13)$$

If τ is an integer multiple of $\frac{T}{P}$, then the Fourier Transform of $\mathbf{g}(\tau)$ is given by

$$\tilde{\mathbf{g}}_\tau = [1 \ \phi^{\tau P} \ (\phi^{\tau P})^2 \ \cdots \ (\phi^{\tau P})^{LP-1}] \text{diag}[\tilde{\mathbf{g}}] \quad (3.14)$$

The same holds for any τ if $g(t)$ is bandlimited and sampled above the Nyquist rate. This is not in full agreement with the FIR assumption of the channel impulse response, which requires that $g(t)$ be of finite length L_g , so that L is finite. Because of this truncation, the spectrum of $g(t)$ widens and sampling at a rate of TP introduces aliasing due to spectral folding. This gives extra terms in (3.10) that will eventually lead to a bias in the delay estimate. In typical situations, however, the extra terms are small. For example, for $P = 2$, a raised-cosine pulse with excess bandwidth $\alpha = 0.10$ truncated at a length $L_g = 4$ ($T = 1$) leads to a model mismatch (relative error) in (3.14) of a maximum of 6% for any τ , with a corresponding delay estimation error of less than 0.002%. The error becomes even smaller for larger P , L_g , or α . Hence, in comparison to estimation errors that will occur in the presence of noise, this bias will not be significant [24].

Based on the channel model in (3.11),

$$\mathbf{H} = \mathbf{ADG} \quad (3.15)$$

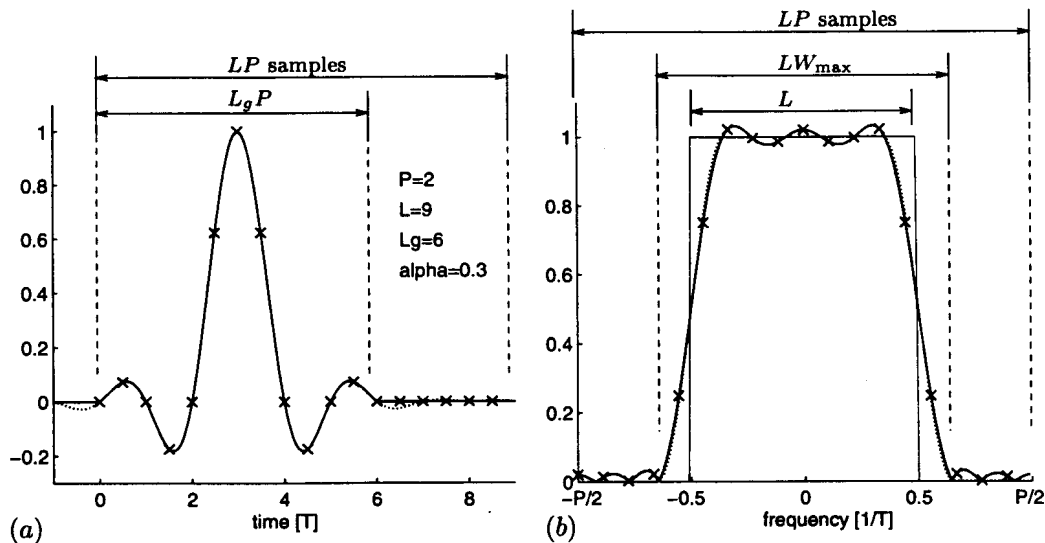


Figure 3.2: Parameters associated with square-root of raised cosine pulse shape: (a) time domain and (b) frequency domain.

Performing the DFT on the rows of \mathbf{H} , we obtain

$$\tilde{\mathbf{H}} = \mathbf{H}\mathbf{F} = \mathbf{A}\mathbf{D}\mathbf{F}^T \text{diag}[\tilde{\mathbf{g}}] \quad (3.16)$$

where \mathbf{F} is a Vandermonde matrix [25] of the form

$$\mathbf{F}^T = \begin{bmatrix} 1 & \phi_1 & \phi_1^2 & \cdots & \phi_1^{LP-1} \\ 1 & \phi_2 & \phi_2^2 & \cdots & \phi_2^{LP-1} \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ 1 & \phi_Q & \phi_Q^2 & \cdots & \phi_Q^{LP-1} \end{bmatrix}, \quad \phi_i = \phi^{\tau_i P} = e^{-j\frac{2\pi}{L}\tau_i} \quad (3.17)$$

3.4 Deconvolution

We perform a deconvolution of $g(t)$ by dividing $\tilde{\mathbf{H}}$ by $\text{diag}[\tilde{\mathbf{g}}]$. This can be performed only on intervals for which $\tilde{\mathbf{g}}$ is non-zero. To illustrate this, assume that $g(t)$ is bandlimited with normalized bandwidth W_{max} in rad/sample (that is, its Fourier transform is non-zero only for frequencies $|\omega| \leq W_{max}$),

and assume that $P > W_{max}$. Then $\tilde{\mathbf{g}}$ has at most LW_{max} nonzero entries. A raised-cosine pulse with excess bandwidth α has $W_{max} = (1 + \alpha)B$, where B is the bandwidth of the transmitted signal. However, instead of taking LW_{max} entries of $\tilde{\mathbf{g}}$, we would select a somewhat smaller number, e.g., LW , since the entries at the border can be relatively small as well, and their inversion can amplify the noise, as shown in Figure 3.2. In the case of a raised cosine pulse, it is advisable to set $W = 1$ and select only the L center frequency samples as the other values are close to zero.

Let the $LP \times LW$ matrix $\mathbf{J}_{\tilde{\mathbf{g}}}$ be the corresponding selection matrix for $\tilde{\mathbf{g}}$, such that $\tilde{\mathbf{g}}\mathbf{J}_{\tilde{\mathbf{g}}}$ has only nonzero entries. For later use, we require that the selected frequencies appear in increasing order, which with the definition of the DFT in (3.13) usually means that the final $\left\lceil \frac{LW}{2} \right\rceil$ samples of $\tilde{\mathbf{g}}$ should be moved to the front. Thus, $\mathbf{J}_{\tilde{\mathbf{g}}}$ performs the function of DFT shift matrix and has the form

$$\mathbf{J}_{\tilde{\mathbf{g}}} = \begin{bmatrix} \mathbf{0} & \mathbf{I}_{\lceil \frac{LW}{2} \rceil} \\ \mathbf{0} & \mathbf{0} \\ \mathbf{I}_{\lceil \frac{LW}{2} \rceil} & \mathbf{0} \end{bmatrix}_{LP \times LW} \quad (3.18)$$

Multiplying both sides of (3.17) by $\mathbf{J}_{\tilde{\mathbf{g}}}$ given by (3.18),

$$\tilde{\mathbf{H}}\mathbf{J}_{\tilde{\mathbf{g}}} = \mathbf{A}\mathbf{D}\mathbf{F}^T \{\text{diag}[\tilde{\mathbf{g}}] \mathbf{J}_{\tilde{\mathbf{g}}}\} \quad (3.19)$$

The post-processed channel model can be represented as

$$\check{\mathbf{H}} = \tilde{\mathbf{H}}\mathbf{J}_{\tilde{\mathbf{g}}}\{\text{diag}[\tilde{\mathbf{g}}] \mathbf{J}_{\tilde{\mathbf{g}}}\}^{-1}, \quad (M \times LW) \quad (3.20)$$

$$\check{\mathbf{H}} = \mathbf{A}\mathbf{D}\mathbf{F}^T, \quad (3.21)$$

where \mathbf{F} is as in (3.17) but with LW columns rather than LP columns.

3.5 Joint Angle and Delay Estimate

At each time index n , we obtain the channel estimate, perform the discrete Fourier transform and deconvolve with the known pulse, as described in Section 3.3. If we have a uniform linear array of M omni-directional antennas uniformly spaced at distance d , then the array response

$$\mathbf{A} = \begin{bmatrix} 1 & \cdots & 1 \\ \psi_1 & \cdots & \psi_Q \\ \vdots & & \vdots \\ \psi_1^{M-1} & \cdots & \psi_Q^{M-1} \end{bmatrix} \quad (3.22)$$

where $\psi_i = e^{-j2\pi\frac{d}{\lambda} \sin\theta_i}$. The amplitude matrix \mathbf{D} is a function of the snapshots (a snapshot is the samples at the sensors at an instant of time)

$$\mathbf{D}(n) = \begin{bmatrix} \beta_1(n) & & 0 \\ & \ddots & \\ 0 & & \beta_Q(n) \end{bmatrix} = \text{diag}(\boldsymbol{\beta}(n)) \quad n = 1, 2, \dots, N_s \quad (3.23)$$

where N_s is the number of snapshots available. The channel model is

$$\check{\mathbf{H}}(n) = \mathbf{A} \mathbf{D}(n) \mathbf{F}^T \quad (3.24)$$

If \mathbf{Y} is a diagonal matrix, then

$$\text{vec}(\mathbf{XYZ}) = (\mathbf{Z}^T \circ \mathbf{X}) \text{diag}(\mathbf{Y}) \quad (3.25)$$

where \circ represents the Khatri-Rao product of matrices [25]. Applying vec on both sides of (3.24), we use (3.25) to obtain

$$\check{\mathbf{Y}}(n) = (\mathbf{F} \circ \mathbf{A}) \boldsymbol{\beta}(n) \quad (3.26)$$

Stacking all the vectors $\boldsymbol{\beta}(n)$ to form matrix \mathbf{B} , we can rewrite (3.26) as

$$\check{\mathbf{Y}} = \mathbf{UB} \quad (3.27)$$

where $\mathbf{U} = \mathbf{F} \circ \mathbf{A}$.

\mathbf{U} has structure similar to the array manifold matrix in the 2-D DOA estimation given by (B.29). Thus, we can use the 2-D unitary ESPRIT derived in Section B.4 on $\check{\mathbf{Y}}$ and estimate ψ_i and ϕ_i , to obtain TOA and DOA. By using 2-D estimators, we automatically pair the TOA and DOA for each multipath. This estimation of TOA and DOA using ESPRIT is called JADE-ESPRIT [26].

3.6 Computer Simulations

We use computer simulations to assess the performance of JADE-ESPRIT. We assume that one source emits signals that arrive at an array of $M = 2$ sensors separated by a half-wavelength along $Q = 3$ paths. The additive noise at the array has an SNR that varies from 10 dB to -40 dB. Signal power is calculated by averaging over the frame of bits (192 bits in case of IS-95). The angles-of-arrival are $\{-5^\circ, 0^\circ, 10^\circ\}$ and time delays are $\{1, 1.1, 2\} T_s$, where T_s is the symbol rate. The path fading is generated from a complex Gaussian distribution with zero mean and variance $[1, 0.9, 0.8]$, respectively, for the three paths. The pulse shaping function is a raised cosine with excess bandwidth $\beta = 0.35$, truncated to a length of $L_g = 6$ symbols. After truncation 2% of the energy is lost. The oversampling rate is $P = 2$ and the symbol rate is normalized to $T = 1$.

The multipath which arrives first at the array is called the prompt ray. Our aim is to find the TOA and DOA of the prompt ray. The JADE-ESPRIT is compared against a total least squares TLS-ESPRIT [8] sequential estimator. As we have more samples in the time domain than in the spatial domain, TOA

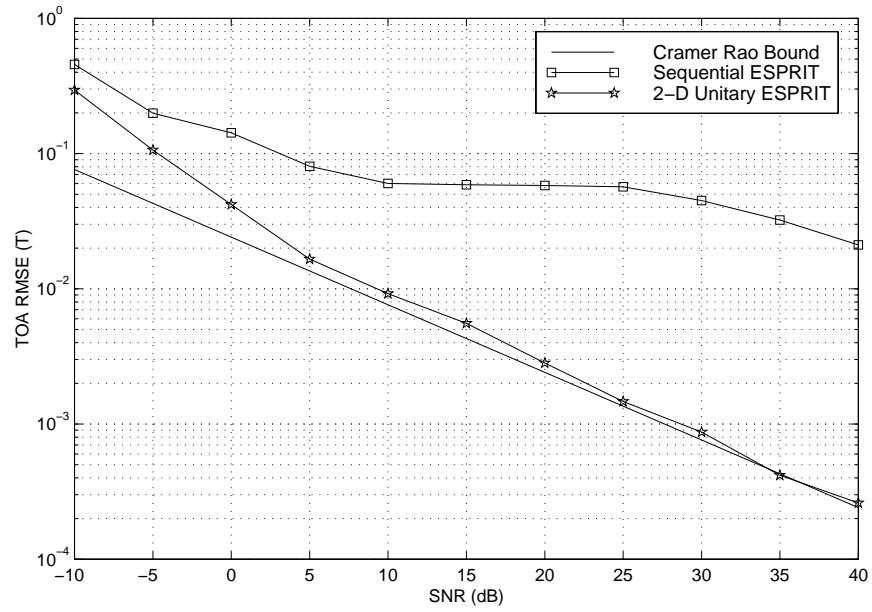


Figure 3.3: The root mean square error of TOA vs. SNR for the prompt ray.

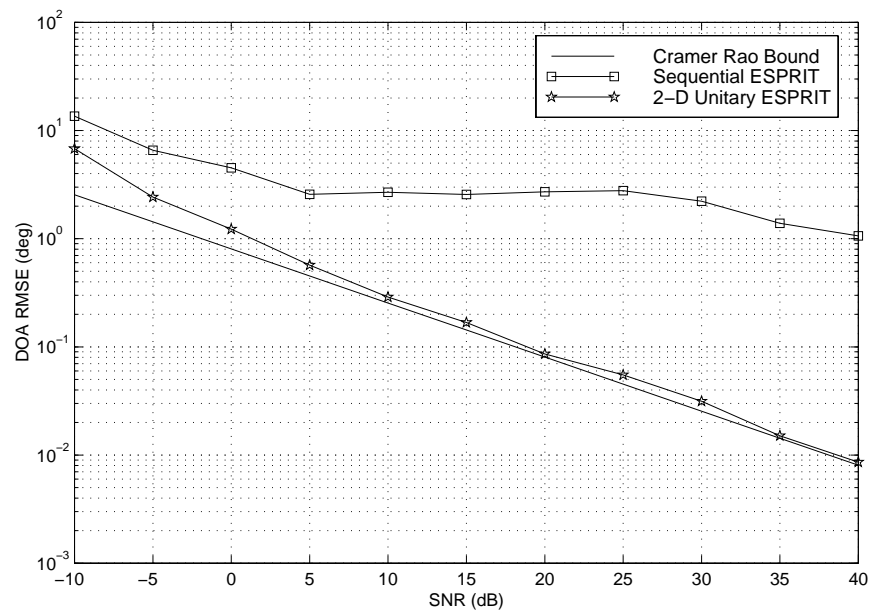


Figure 3.4: The root mean square error of DOA vs. SNR for the prompt ray.

is estimated first and then the DOA is calculated from the TOA estimate. The sequential estimator is not capable of separating rays when they are very close (less than $0.3T$) in the time domain, while the JADE-ESPRIT has no problem separating the rays. Figs. 3.3 and 3.4 compare the two ESPRIT techniques and show the root mean square error (RMSE) of the estimate of TOA and DOA as a function of the noise power σ^2 . JADE-ESPRIT asymptotically approaches the Cramer-Rao Bound as SNR goes to infinity [8]. To obtain a DOA RMSE of 2.5° , TLS-ESPRIT requires 10 dB higher SNR than JADE-ESPRIT, which in turn requires 5 dB higher SNR than the Cramer-Rao Bound.

3.7 Conclusion

In this chapter, we develop a discrete space-time channel model for an array of receivers and a transformation that converts the time delays to phase shifts in the frequency domain. This facilitates the use of 2-D ESPRIT algorithms that were originally developed to find 2-D angles of arrival to the joint angle and delay estimation (JADE). We verify that the JADE-ESPRIT algorithm is capable of jointly estimating the TOA and DOA estimates of the rays. JADE-ESPRIT is a closed-form solution that can estimate more paths than the number of antenna elements (discussed in detail in Appendix B).

Signal subspace estimation is the most computational expensive step in the JADE-ESPRIT. For an $M \times M$ matrix a singular value decomposition (SVD) requires $\mathcal{O}(M^3)$ operations to be performed. In traditional ESPRIT, SVD has to be taken on complex data, but in JADE-ESPRIT, it has to be taken on real data, thereby reducing the computational complexity. The channel

estimates obtained using JADE-ESPRIT are more reliable because it improves the underlying signal subspace estimation. Another key advantage of JADE-ESPRIT is that the number of paths could be more than the number of antenna elements, while in traditional ESPRIT, the number of paths should be less than the number of antenna elements.

Appendix A

Notation and Abbreviations

A.1 Notation

This report follows the IEEE conventions. Vectors are denoted by lowercase boldface letters, matrices by boldface capital or calligraphic capital letters, scalar parameters by capital or lower case Latin or Greek letters. Other symbols that are used often follow:

T	matrix or vector transpose
$*$	matrix or vector complex conjugate
H	Hermitian of the matrix or vector (complex conjugate transpose)
\mathbf{I}_m	identity matrix of size $m \times m$
$\mathbf{0}_m$	a zero matrix of m columns and an appropriate number of rows
$*$	convolution operator
\otimes	Kronecker or outer product [25]
\circ	Khatri-Rao product [25] which is a column-wise Kronecker product $\mathbf{A} \circ \mathbf{B} = [\mathbf{a}_1 \otimes \mathbf{b}_1 \quad \mathbf{a}_2 \otimes \mathbf{b}_2 \quad \cdots]$
$\mathcal{Re}\{\cdot\}$	real part of complex scalar, vector or matrix
$\mathcal{Im}\{\cdot\}$	imaginary part of complex scalar, vector or matrix
$\text{vec}\{\cdot\}$	column vector obtained by stacking the columns of a matrix
$\text{diag}\{\cdot\}$	column vector of diagonal elements of a matrix or a diagonal matrix whose diagonal elements are the column vector
$\lceil \cdot \rceil$	smallest integer greater than real number (ceiling)
$\lfloor \cdot \rfloor$	largest integer smaller than real number (floor)
$\delta()$	Dirac delta functional
δ_{kl}	Kronecker delta which has value of 1 if $k = l$ and 0 otherwise
$[\mathbf{A} \mathbf{B}]$	concatenate matrices \mathbf{A} and \mathbf{B} into a new matrix.

A.2 Abbreviations

The following are the primary abbreviations (in alphabetical order) that are used in this report.

BPSK	Binary Phase Shift Keying
CCI	Co-channel Interference
CDMA	Code Division Multiple Access
CM	Constant Modulus
CMA	Constant Modulus Algorithm
DFT	Discrete Fourier Transform
DOA	Direction of Arrival
ESPRIT	Estimation of Signal Parameters using Rotational Invariance Techniques
FFT	Fast Fourier Transform
FIR	Finite Impulse Response
GMSK	Gaussian Minimum Shift Keying
GSM	Global System for Mobile
GPS	Global Positioning System
ISI	Inter-Symbol-Interference
JADE	Joint Angle and Delay Estimation
LOS	Line-of-sight
MIMO	multiple-input, multiple-output
ML	Maximum Likelihood
NLOS	Non Line-of-sight
MSE	Mean Square Error
RMSE	Root Mean Square Error
TDMA	Time Division Multiple Access
TDOA	Time Difference of Arrival
TLS	Total Least Squares
TOA	Time of Arrival
ULA	Uniform linear Array
URA	Uniform Rectangular Array

Some wireless systems and standards:

- IS-54, 136, 95 (Interim Standard 54, 136, 95) - North American digital wireless system. IS-54 and IS-136 are TDMA while IS-95 is CDMA.

- PHS (Personal Handyphone System) - Japanese personal wireless system
- GSM (Global System for Mobile) - European digital cellular system
- DECT (Digital European Cordless Telecommunications) system
- DCS 1800 (Digital Cellular System 1800) - Europe
- AMPS (Advanced Mobile Phone System) - North America

Appendix B

2-D Unitary ESPRIT

B.1 Introduction

Estimation of Signal Parameters via Rotational Invariance Techniques (ESPRIT) [8] is a high-resolution parameter estimation method used in DOA estimation, system identification, and time series analysis. ESPRIT exploits an underlying *rotational invariance* among signal subspaces induced by an array of sensors with a *translational invariance* structure. That is, we decompose an array into two overlapping subarrays containing the same number of sensors whose sensor locations are related by a single phase shift. Unitary ESPRIT finds DOA with increased accuracy and reduced computational complexity [27]. All computations in unitary ESPRIT are real-valued.

In Section B.2, we discuss the real-valued processing using a uniform linear array (ULA). This real-valued processing forms the basis of unitary ESPRIT which is derived in the Section B.3. Section B.4 discusses the extension of 1-D unitary ESPRIT to 2-D unitary ESPRIT. Most of the material in this chapter has been borrowed from [28].

B.2 Real-Valued Processing with a Uniform Linear Array

The development of unitary ESPRIT relies on aspects of real-valued processing with a ULA [22] [27] [29]. Employing the center of the ULA as a phase reference, the array manifold is conjugate centrosymmetric. For example, if the number of elements N in the ULA is odd, then a sensor is located at the array center and the array response vector is

$$\mathbf{a}_N(\mu) = [e^{-j(\frac{N-1}{2})\mu}, \dots, e^{-j\mu}, 1, e^{j\mu}, \dots, e^{j(\frac{N-1}{2})\mu}]^T \quad (\text{B.1})$$

where $\mu = \frac{2\pi}{\lambda} \Delta_x u$ with λ as the wavelength, Δ_x as the inter-element spacing, and u as the direction cosine (sine of the DOA) of the received ray relative to the array axis. The conjugate symmetry of $\mathbf{a}_N(\mu)$ can be expressed as

$$\mathbf{\Pi}_N \mathbf{a}_N(\mu) = \mathbf{a}_N^*(\mu) \quad (\text{B.2})$$

where

$$\mathbf{\Pi}_N = \begin{bmatrix} \mathbf{0} & & & 1 \\ & & 1 & \\ & \ddots & & \\ 1 & & & \mathbf{0} \end{bmatrix} \in \Re^{N \times N}$$

The inner product between any two conjugate centrosymmetric vectors is real-valued. Any matrix whose rows are each conjugate centrosymmetric may be employed to transform the complex-valued element space manifold $\mathbf{a}_N(\mu)$ into a real-valued manifold. As noted in [22] [27] [29], the simplest matrices for accomplishing this are

$$\mathbf{Q}_{2K} = \frac{1}{\sqrt{2}} \begin{bmatrix} \mathbf{I}_K & j\mathbf{I}_K \\ \mathbf{\Pi}_K & -j\mathbf{\Pi}_K \end{bmatrix} \quad (\text{B.3})$$

if N is even, or

$$\mathbf{Q}_{2K+1} = \frac{1}{\sqrt{2}} \begin{bmatrix} \mathbf{I}_K & \mathbf{0} & j\mathbf{I}_K \\ \mathbf{0}^T & \sqrt{2} & \mathbf{0}^T \\ \mathbf{\Pi}_K & \mathbf{0} & -j\mathbf{\Pi}_K \end{bmatrix} \quad (\text{B.4})$$

if N is odd. \mathbf{Q}_N^H is a sparse unitary matrix that transforms $\mathbf{a}_N(\mu)$ into an $N \times 1$ real-valued manifold,

$$\mathbf{d}_N(\mu) = \mathbf{Q}_N^H \mathbf{a}_N(\mu) \quad (\text{B.5})$$

If the number of elements comprising the ULA is odd, then

$$\begin{aligned} \mathbf{d}_N(\mu) &= \mathbf{Q}_N^H \mathbf{a}_N(\mu) \\ &= \sqrt{2} [\cos(\frac{N-1}{2}\mu), \dots, \cos(\mu), \frac{1}{\sqrt{2}}, -\sin(\frac{N-1}{2}\mu), \dots, -\sin(\mu)]^T \end{aligned} \quad (\text{B.6})$$

If \mathbf{X} denotes the $N \times N_s$ element data matrix containing N_s snapshots in time as columns, then the signal eigenvectors may be computed as the “largest” left singular vectors of the real-valued matrix $[\mathcal{R}e\{\mathbf{Y}\} \mid \mathcal{I}m\{\mathbf{Y}\}]$, where $\mathbf{Y} = \mathbf{Q}_N^H \mathbf{X}$.

B.3 Unitary ESPRIT for Uniform Linear Array

In this section, we develop the unitary ESPRIT method which uses only real computations. The element space manifold in (B.1) satisfies the invariance relation [8]

$$e^{j\mu} \mathbf{J}_1 \mathbf{a}_N(\mu) = \mathbf{J}_2 \mathbf{a}_N(\mu) \quad (\text{B.7})$$

where \mathbf{J}_1 and \mathbf{J}_2 are the $(N-1) \times N$ selection matrices

$$\mathbf{J}_1 = \begin{bmatrix} 1 & 0 & 0 & \cdots & 0 & 0 \\ 0 & 1 & 0 & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & 1 & 0 \end{bmatrix} \in \mathfrak{R}^{(N-1) \times N} \quad (\text{B.8})$$

$$\mathbf{J}_2 = \begin{bmatrix} 0 & 1 & 0 & \cdots & 0 & 0 \\ 0 & 0 & 1 & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & 0 & 1 \end{bmatrix} \in \mathfrak{R}^{(N-1) \times N} \quad (\text{B.9})$$

\mathbf{J}_1 and \mathbf{J}_2 select the first and last $(N - 1)$ components of an $N \times 1$ vector, respectively. \mathbf{J}_1 and \mathbf{J}_2 satisfy the relationship

$$\mathbf{\Pi}_{N-1} \mathbf{J}_2 \mathbf{\Pi}_N = \mathbf{J}_1 \quad (\text{B.10})$$

Since \mathbf{Q}_N is unitary, it follows that

$$\mathbf{Q}_N \mathbf{Q}_N^H = \mathbf{I}_N \quad (\text{B.11})$$

Substituting (B.11) into (B.7),

$$e^{j\mu} \mathbf{J}_1 \mathbf{Q}_N \mathbf{Q}_N^H \mathbf{a}_N(\mu) = \mathbf{J}_2 \mathbf{Q}_N \mathbf{Q}_N^H \mathbf{a}_N(\mu) \quad (\text{B.12})$$

Substituting (B.5) into (B.12),

$$e^{j\mu} \mathbf{J}_1 \mathbf{Q}_N \mathbf{d}_N(\mu) = \mathbf{J}_2 \mathbf{Q}_N \mathbf{d}_N(\mu) \quad (\text{B.13})$$

Multiplying both sides by \mathbf{Q}_{N-1}^H on the left yields the invariance relationship

$$e^{j\mu} \mathbf{Q}_{N-1}^H \mathbf{J}_1 \mathbf{Q}_N \mathbf{d}_N(\mu) = \mathbf{Q}_{N-1}^H \mathbf{J}_2 \mathbf{Q}_N \mathbf{d}_N(\mu) \quad (\text{B.14})$$

Since \mathbf{Q}_N satisfies the relationship

$$\mathbf{\Pi}_N \mathbf{Q}_N = \mathbf{Q}_N^* \quad (\text{B.15})$$

and $\mathbf{\Pi}_N$ satisfies the relationship

$$\mathbf{\Pi}_N \mathbf{\Pi}_N = \mathbf{I}_N \quad (\text{B.16})$$

We combine (B.10), (B.15), and (B.16) to obtain

$$\begin{aligned}\mathbf{Q}_{N-1}^H \mathbf{J}_2 \mathbf{Q}_N &= \mathbf{Q}_{N-1}^H \mathbf{\Pi}_{N-1} \mathbf{\Pi}_{N-1} \mathbf{J}_2 \mathbf{\Pi}_N \mathbf{\Pi}_N \mathbf{Q}_N \\ &= \mathbf{Q}_{N-1}^T \mathbf{J}_1 \mathbf{Q}_N^* \\ &= (\mathbf{Q}_{N-1}^H \mathbf{J}_1 \mathbf{Q}_N)^*\end{aligned}\quad (\text{B.17})$$

Substituting (B.17) in (B.14) we obtain

$$e^{j\mu} (\mathbf{Q}_{N-1}^H \mathbf{J}_2 \mathbf{Q}_N)^* \mathbf{d}_N(\mu) = \mathbf{Q}_{N-1}^H \mathbf{J}_2 \mathbf{Q}_N \mathbf{d}_N(\mu) \quad (\text{B.18})$$

Let \mathbf{K}_1 and \mathbf{K}_2 be the real and imaginary parts of $\mathbf{Q}_{N-1}^H \mathbf{J}_2 \mathbf{Q}_N$, respectively,

$$\mathbf{K}_1 = \mathcal{R}e\{\mathbf{Q}_{N-1}^H \mathbf{J}_2 \mathbf{Q}_N\} \quad (\text{B.19})$$

$$\mathbf{K}_2 = \mathcal{I}m\{\mathbf{Q}_{N-1}^H \mathbf{J}_2 \mathbf{Q}_N\} \quad (\text{B.20})$$

\mathbf{K}_1 and \mathbf{K}_2 are real-valued $(N-1) \times N$ matrices. Based on these definitions, (B.18) may be expressed as

$$e^{j\frac{\mu}{2}} (\mathbf{K}_1 - j\mathbf{K}_2) \mathbf{d}_N(\mu) = e^{-j\frac{\mu}{2}} (\mathbf{K}_1 + j\mathbf{K}_2) \mathbf{d}_N(\mu) \quad (\text{B.21})$$

Rearranging terms, we have

$$(e^{j\frac{\mu}{2}} - e^{-j\frac{\mu}{2}}) \mathbf{K}_1 \mathbf{d}_N(\mu) = j(e^{j\frac{\mu}{2}} + e^{-j\frac{\mu}{2}}) \mathbf{K}_2 \mathbf{d}_N(\mu) \quad (\text{B.22})$$

which becomes an expression involving only real-valued quantities

$$\tan\left(\frac{\mu}{2}\right) \mathbf{K}_1 \mathbf{d}_N(\mu) = \mathbf{K}_2 \mathbf{d}_N(\mu) \quad (\text{B.23})$$

For d sources, we define the $N \times d$ real-valued DOA matrix as

$$\mathbf{D} = [\mathbf{d}_N(\mu_1), \mathbf{d}_N(\mu_2), \dots, \mathbf{d}_N(\mu_d)]$$

The real-valued manifold in (B.23) translates into the real-valued DOA matrix relation

$$\mathbf{K}_1 \mathbf{D} \mathbf{\Omega} = \mathbf{K}_2 \mathbf{D} \quad (\text{B.24})$$

where

$$\mathbf{\Omega} = \text{diag} \left\{ \tan\left(\frac{\mu_1}{2}\right), \dots, \tan\left(\frac{\mu_d}{2}\right) \right\} \quad (\text{B.25})$$

As discussed in the previous section, if \mathbf{X} denotes the $N \times N_s$ element space data matrix containing N_s snapshots as columns, then the signal eigenvectors for unitary ESPRIT may be computed as the largest left singular vectors of the real-valued matrix $[\mathcal{R}e\{\mathbf{Y}\} \mid \mathcal{I}m\{\mathbf{Y}\}]$, where $\mathbf{Y} = \mathbf{Q}_N^H \mathbf{X}$. As the number of snapshots becomes infinitely large, the $N \times d$ real-valued matrix of signal eigenvectors \mathbf{E}_s is related to the real $N \times d$ real-valued DOA matrix \mathbf{D} as

$$\lim_{N_s \rightarrow \infty} \mathbf{E}_s = \mathbf{D}\mathbf{T} \quad (\text{B.26})$$

where \mathbf{T} is an unknown $d \times d$ non-singular, real-valued matrix. Substituting $\mathbf{D} = \mathbf{E}_s \mathbf{T}^{-1}$ into (B.24) yields,

$$\begin{aligned} \mathbf{K}_1 \mathbf{E}_s \mathbf{T}^{-1} \mathbf{\Omega} &= \mathbf{K}_2 \mathbf{E}_s \mathbf{T}^{-1} \\ \mathbf{K}_1 \mathbf{E}_s \mathbf{T}^{-1} \mathbf{\Omega} \mathbf{T} &= \mathbf{K}_2 \mathbf{E}_s \\ \mathbf{K}_1 \mathbf{E}_s \mathbf{\Psi} &= \mathbf{K}_2 \mathbf{E}_s \end{aligned} \quad (\text{B.27})$$

where $\mathbf{\Psi} = \mathbf{T}^{-1} \mathbf{\Omega} \mathbf{T}$. The eigenvalues of the $d \times d$ solution $\mathbf{\Psi}$ to the $(N - 1) \times d$ matrix equation above are $\tan\left(\frac{\mu_i}{2}\right)$, $i = 1, 2, \dots, d$. If $\Delta_x = \frac{\lambda}{2}$ so that $\mu_i = \frac{2\pi}{\lambda} \Delta_x u_i = \pi u_i$. In this case, there is a one-to-one mapping between $-1 \leq u_i \leq 1$, corresponding to the range of possible values for a direction cosine. The eigenvalues of $\mathbf{\Psi}$ give $\tan\left(\frac{\mu_i}{2}\right)$; thus we know μ_i so u_i (the desired direction cosine) is $\frac{\mu_i}{\pi}$.

Figure B.1 gives the unitary ESPRIT algorithm for a uniform linear array of sensors. We do not address the problem of estimating the number of sources d . We assume that an estimate is available via a procedure such as that described in [30], which explicitly exploits the conjugate centrosymmetry of the array manifold for a ULA.

- Compute \mathbf{E}_s via the d “largest” left singular vectors of $[\mathcal{R}e\{\mathbf{Y}\}, \mathcal{I}m\{\mathbf{Y}\}]$, where $\mathbf{Y} = \mathbf{Q}_N^H \mathbf{X}$. \mathbf{X} contains the $N \times N_s$ snapshots of the array sample values and \mathbf{Q}_N is defined by either (B.3) or (B.4) as appropriate.
- Compute $\mathbf{\Psi}$ as the solution to $(N - 1) \times d$ matrix equation $\mathbf{K}_1 \mathbf{E}_s \mathbf{\Psi} = \mathbf{K}_2 \mathbf{E}_s$ where \mathbf{K}_1 and \mathbf{K}_2 are defined in (B.19) and (B.20).
- Compute $\omega_i, i = 1, 2, \dots, d$, as the eigenvalues of the $d \times d$ real-valued matrix $\mathbf{\Psi}$.
- Compute the spatial frequency estimates as $\mu_i = 2 \arctan(\omega_i)$, $i = 1, \dots, d$.

Figure B.1: Unitary ESPRIT algorithm for a uniform linear array of N sensors.

B.4 2-D Unitary ESPRIT

In this section, we develop 2-D unitary ESPRIT for uniform rectangular array (URA) as an extension of unitary ESPRIT for ULA. The URA consists of $N \times M$ elements lying in the $x - y$ plane and equispaced by Δ_x in the x direction and Δ_y in the y direction. In addition to $\mu = \frac{2\pi}{\lambda} \Delta_x u$, where u is the direction cosine relative to the x -axis, we define the spatial frequency variable $\nu = \frac{2\pi}{\lambda} \Delta_x v$, where v is the direction cosine variable relative to the y -axis.

The array response is denoted as an $NM \times 1$ vector $\mathbf{a}(\mu, \nu)$ or as an $N \times M$ matrix denoted by $\mathbf{A}(\mu, \nu)$. The two forms are related through the operators, $\text{vec}(\cdot)$ and $\text{mat}(\cdot)$ as $\mathbf{a}(\mu, \nu) = \text{vec}(\mathbf{A}(\mu, \nu))$ and $\mathbf{A}(\mu, \nu) = \text{mat}(\mathbf{a}(\mu, \nu))$. The operator $\text{vec}(\cdot)$ maps a $N \times M$ matrix to an $NM \times 1$ vector by stacking the columns of the matrix. The operator $\text{mat}(\cdot)$ performs the inverse mapping, mapping a $NM \times 1$ vector to an $N \times M$ matrix such that $\text{mat}(\text{vec}(\mathbf{X})) = \mathbf{X}$. An important property of the vec operator that will be used heavily in the

development of 2-D unitary ESPRIT is

$$\text{vec}(\mathbf{ABC}) = (\mathbf{C}^T \otimes \mathbf{A})\text{vec}(\mathbf{B}) \quad (\text{B.28})$$

where \otimes denotes the Kronecker matrix product [25].

In matrix form, the array manifold may be expressed as

$$\mathbf{A}(\mu, \nu) = \mathbf{a}_N(\mu)\mathbf{a}_M^T(\nu) \quad (\text{B.29})$$

where $\mathbf{a}_M(\nu)$ is defined by (B.1) with N replaced by M and μ replaced by ν . Similar to the 1-D case, multiplying $\mathbf{A}(\mu, \nu)$ by \mathbf{Q}_N^H on the left and multiplying by \mathbf{Q}_M^* on the right, creates the $N \times M$ real-valued manifold

$$\begin{aligned} \mathbf{D}(\mu, \nu) &= \mathbf{Q}_N^H \mathbf{A}(\mu, \nu) \mathbf{Q}_M^* \\ &= \mathbf{Q}_N^H \mathbf{a}_N(\mu) \mathbf{a}_M^T(\nu) \mathbf{Q}_M^* \\ &= \mathbf{d}_N(\mu) \mathbf{d}_M^T(\nu) \end{aligned} \quad (\text{B.30})$$

where $\mathbf{d}_M(\nu)$ is defined in (B.5) with N replaced by M and μ replaced by ν .

Given that $\mathbf{d}_N(\mu)$ satisfies the invariance relationship in (B.23), it follows that $\mathbf{D}(\mu, \nu)$ satisfies

$$\tan\left(\frac{\mu}{2}\right)\mathbf{K}_1\mathbf{D}(\mu, \nu) = \mathbf{K}_2\mathbf{D}(\mu, \nu) \quad (\text{B.31})$$

where \mathbf{K}_1 and \mathbf{K}_2 are defined in (B.19) and (B.20), respectively. Using the vec on both sides of (B.31),

$$\tan\left(\frac{\mu}{2}\right)\text{vec}(\mathbf{K}_1 \mathbf{D}(\mu, \nu) \mathbf{I}_M) = \text{vec}(\mathbf{K}_2 \mathbf{D}(\mu, \nu) \mathbf{I}_M) \quad (\text{B.32})$$

Using the property of vec operator in (B.28), we find that the $NM \times 1$ real-valued manifold in vector form, $\mathbf{d}(\mu, \nu) = \text{vec}(\mathbf{D}(\mu, \nu))$ satisfies

$$\tan\left(\frac{\mu}{2}\right)\mathbf{K}_{\mu 1}\mathbf{d}(\mu, \nu) = \mathbf{K}_{\mu 2}\mathbf{d}(\mu, \nu) \quad (\text{B.33})$$

where $\mathbf{K}_{\mu 1}$ and $\mathbf{K}_{\mu 2}$ are $(N - 1)M \times NM$ matrices defined as

$$\mathbf{K}_{\mu 1} = \mathbf{I}_M \otimes \mathbf{K}_1 \quad \text{and} \quad \mathbf{K}_{\mu 2} = \mathbf{I}_M \otimes \mathbf{K}_2 \quad (\text{B.34})$$

Similarly, the 1-D real-valued manifold $\mathbf{d}_M(\nu)$ satisfies

$$\tan\left(\frac{\nu}{2}\right)\mathbf{K}_3\mathbf{d}_M(\nu) = \mathbf{K}_4\mathbf{d}_M(\nu) \quad (\text{B.35})$$

where \mathbf{K}_3 and \mathbf{K}_4 are defined as (B.19) and (B.20) but with N replaced by M such that they are the $(M - 1) \times M$ matrices defined as

$$\mathbf{K}_3 = \mathcal{R}e\{\mathbf{Q}_M^H \mathbf{J}_2 \mathbf{Q}_M\} \quad (\text{B.36})$$

$$\mathbf{K}_4 = \mathcal{I}m\{\mathbf{Q}_M^H \mathbf{J}_2 \mathbf{Q}_M\} \quad (\text{B.37})$$

It follows that

$$\tan\left(\frac{\nu}{2}\right)\mathbf{D}(\mu, \nu)\mathbf{K}_3^T = \mathbf{D}(\mu, \nu)\mathbf{K}_4^T \quad (\text{B.38})$$

Using the vec operator, we find that $\mathbf{d}(\mu, \nu)$ satisfies

$$\tan\left(\frac{\nu}{2}\right)\mathbf{K}_{\nu 1}\mathbf{d}(\mu, \nu) = \mathbf{K}_{\nu 2}\mathbf{d}(\mu, \nu) \quad (\text{B.39})$$

where $\mathbf{K}_{\nu 1}$ and $\mathbf{K}_{\nu 2}$ are the $N(M - 1) \times NM$ matrices,

$$\mathbf{K}_{\nu 1} = \mathbf{K}_3 \otimes \mathbf{I}_N \quad \text{and} \quad \mathbf{K}_{\nu 2} = \mathbf{K}_4 \otimes \mathbf{I}_N \quad (\text{B.40})$$

Consider the $NM \times d$ real-valued DOA matrix

$$\mathbf{D} = [\mathbf{d}(\mu_1, \nu_1), \dots, \mathbf{d}(\mu_d, \nu_d)]$$

where $\mathbf{d}(\mu_i, \nu_i) = \text{vec}(\mathbf{D}(\mu_i, \nu_i))$. From (B.33), \mathbf{D} satisfies

$$\mathbf{K}_{\mu 1}\mathbf{D}\boldsymbol{\Omega}_\mu = \mathbf{K}_{\mu 2}\mathbf{D} \quad (\text{B.41})$$

where

$$\mathbf{\Omega}_\mu = \text{diag} \left\{ \tan\left(\frac{\mu_1}{2}\right), \dots, \tan\left(\frac{\mu_d}{2}\right) \right\} \quad (\text{B.42})$$

From (B.39), \mathbf{D} satisfies

$$\mathbf{K}_{\nu 1} \mathbf{D} \mathbf{\Omega}_\nu = \mathbf{K}_{\nu 2} \mathbf{D} \quad (\text{B.43})$$

where

$$\mathbf{\Omega}_\nu = \text{diag} \left\{ \tan\left(\frac{\nu_1}{2}\right), \dots, \tan\left(\frac{\nu_d}{2}\right) \right\} \quad (\text{B.44})$$

For a given snapshot, the array output is an $N \times M$ matrix. The data matrix \mathbf{Y} can be generated as follows: multiply the array output by \mathbf{Q}_N^H on the left and by \mathbf{Q}_M^* on the right, apply the vec operator, and place the resulting $NM \times 1$ vector as a column of an $NM \times N_s$ data matrix \mathbf{Y} (N_s is the number of snapshots). If \mathbf{X} denotes the $NM \times N_s$ complex-valued element space data matrix, then the relationship between \mathbf{Y} and \mathbf{X} may be expressed as $\mathbf{Y} = (\mathbf{Q}_M^H \otimes \mathbf{Q}_N^H) \mathbf{X}$ where we have again used the property in (B.38). The appropriate $NM \times d$ matrix of signal eigenvectors \mathbf{E}_s may be computed as the d largest left singular vectors of the real-valued matrix $[\mathcal{R}e\{\mathbf{Y}\} \mid \mathcal{I}m\{\mathbf{Y}\}]$. As the number of snapshots becomes infinitely large, \mathbf{E}_s approaches $\mathbf{D}\mathbf{T}$ where \mathbf{T} is an unknown $d \times d$ non-singular, real-valued matrix. Substituting $\mathbf{D} = \mathbf{E}_s \mathbf{T}^{-1}$ into (B.41) and (B.43) yields the eigenvector relations

$$\mathbf{K}_{\mu 1} \mathbf{E}_s \mathbf{\Psi}_\mu = \mathbf{K}_{\mu 2} \mathbf{E}_s \quad \text{where} \quad \mathbf{\Psi}_\mu = \mathbf{T}^{-1} \mathbf{\Omega}_\mu \mathbf{T}. \quad (\text{B.45})$$

$$\mathbf{K}_{\nu 1} \mathbf{E}_s \mathbf{\Psi}_\nu = \mathbf{K}_{\nu 2} \mathbf{E}_s \quad \text{where} \quad \mathbf{\Psi}_\nu = \mathbf{T}^{-1} \mathbf{\Omega}_\nu \mathbf{T}. \quad (\text{B.46})$$

The automatic pairing of μ and ν spatial frequency estimates is facilitated by the fact that all the quantities in (B.45) and (B.46) are real-valued. Thus

- Compute \mathbf{E}_s via the d “largest” left singular vectors of $[\mathcal{R}e\{\mathbf{Y}\}, \mathcal{I}m\{\mathbf{Y}\}]$, where $\mathbf{Y} = (\mathbf{Q}_M^H \otimes \mathbf{Q}_N^H)\mathbf{X}$. \mathbf{X} contains the $NM \times N_s$ snapshots of the array sample values, \mathbf{Q}_M and \mathbf{Q}_N are defined by either (B.3) or (B.4) as appropriate.
- Compute Ψ_μ as the solution to $(N - 1)M \times d$ matrix equation $\mathbf{K}_{\mu 1}\mathbf{E}_s\Psi_\mu = \mathbf{K}_{\mu 2}\mathbf{E}_s$ where $\mathbf{K}_{\mu 1}$ and $\mathbf{K}_{\mu 2}$ are defined in (B.34).
- Compute Ψ_ν as the solution to $N(M - 1) \times d$ matrix equation $\mathbf{K}_{\nu 1}\mathbf{E}_s\Psi_\nu = \mathbf{K}_{\nu 2}\mathbf{E}_s$ where $\mathbf{K}_{\nu 1}$ and $\mathbf{K}_{\nu 2}$ are defined in (B.40).
- Compute $\omega_i, i = 1, 2, \dots, d$, as the eigenvalues of the $d \times d$ complex-valued matrix $\Psi_\mu + j\Psi_\nu$.
- Compute the spatial frequency estimates as $\mu_i = 2 \arctan(\mathcal{R}e\{\omega_i\})$, $\nu_i = 2 \arctan(\mathcal{I}m\{\omega_i\})$, $i = 1, 2, \dots, d$.

Figure B.2: 2-D unitary ESPRIT for a uniform rectangular array of $N \times M$ sensors.

$\Psi_\mu + j\Psi_\nu$ may be spectrally decomposed as

$$\Psi_\mu + j\Psi_\nu = \mathbf{T}^{-1}\{\Omega_\mu + j\Omega_\nu\}\mathbf{T} \quad (\text{B.47})$$

The maximum number of sources d that 2-D unitary ESPRIT can locate is $d = \min\{M(N - 1), N(M - 1)\}$ assuming that at least $\frac{d}{2}$ snapshots are available. This is because forward-backward averaging is applied in unitary ESPRIT, which effectively doubles the number of snapshots [22]. If only a single snapshot is available, then one can extract $\frac{d}{2}$ or more identical rectangular sub-arrays from the master (overall) array to obtain the equivalent of multiple snapshots, thereby increasing the maximum number of sources that can be located.

Figure B.2 gives the summary of closed-form 2-D unitary ESPRIT for

a uniform rectangular array. Note that 2-D unitary ESPRIT provides closed-form, automatically paired 2-D angle estimates as long as the spatial frequency coordinates pairs $(\mu_i, \nu_i), i = 1, \dots, d$ are distinct.

BIBLIOGRAPHY

- [1] A. J. Paulraj, C. B. Papadias, V. U. Reddy, and A. J. van der Veen, *Signal Processing for Wireless Communications*, ch. 4, Space-Time Blind Signal Processing for Wireless Communication Systems. Prentice Hall, 1997.
- [2] W. C. Y. Lee, *Mobile Cellular Telecommunications Systems*. McGraw-Hill, 1989.
- [3] A. Paulraj and C. B. Papadias, “Space-time processing for wireless communication systems,” *IEEE Signal Processing Magazine*, vol. 16, pp. 49–83, Nov. 1997.
- [4] T. S. Rappaport, J. H. Reed, and B. D. Woerner, “Position location using wireless communications on highways of the future,” *IEEE Communications Magazine*, vol. 34, pp. 33–41, Oct. 1996.
- [5] “Wireless E911 Legislative Issue: Questions & Answers.” Bell South Corporate, May 1997. <http://www.bellsouthcorp.com/issues/911>.
- [6] J. T. Chen and A. Paulraj, “Multi-channel MLSE equalizer for GSM using a parametric channel model,” in *Proc. IEEE Int. Conf. on Acoustics, Speech, and Signal Processing*, vol. 5, (Munich, Germany), pp. 3885–3888, May 1997.
- [7] R. O. Schmidt, “Multiple Emitter location and Signal Parameter Estimation,” *IEEE Trans. on Acoustics, Speech, and Signal Processing*, vol. 34,

- pp. 276–280, Mar. 1986.
- [8] 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 1989.
- [9] W. C. Jakes, *Microwave Mobile Communications*. New York: John Wiley & Sons, 2 ed., 1974.
- [10] G. Stuber, *Principles of Mobile Communications*. Boston: Kulwer, 2 ed., 1996.
- [11] W. C. Lee, *Mobile Communications Design Fundamentals*. New York: John Wiley & Sons, 2 ed., 1993.
- [12] J. J. Shynk, A. V. Keerthi, and A. Mathur, “Steady state analysis of the multistage CM array,” *IEEE Trans. on Signal Processing*, vol. 44, pp. 948–962, May 1996.
- [13] J. R. Treichler and B. G. Agee, “A new approach to multipath correction of constant modulus signals,” *IEEE Trans. on Acoustics, Speech, and Signal Processing*, vol. 31, pp. 459–471, April 1983.
- [14] B. Widrow and S. D. Stearns, *Adaptive Signal Processing*. Prentice-Hall, N.J, 1985.
- [15] S. Gummadi and B. L. Evans, “Chochannel signal separation in fading channels using a modified constant modulus array,” in *Proc. IEEE Asilomar Conf. on Signals, Systems, and Computers*, (Pacific Grove, CA), Nov 1-4 1998. to appear.

- [16] E. A. Lee and D. G. Messerschmitt, *Digital Communication*. Kulwer Academic Publishers, 2 ed., 1994.
- [17] S. Haykin, *Adaptive Filter Theory*. Englewood Cliffs, N.J. 07632: Prentice Hall, 2 ed., 1991.
- [18] A. V. Keerthi, A. Mathur, and J. J. Shynk, “Analysis of the Multistage CM Array for Digital Communication Signals,” in *Proc. IEEE Asilomar Conf. on Signals, Systems and Computers*, vol. 1, pp. 285–289, Nov. 1996.
- [19] K. Hilal and P. Duhamel, “Blind equalizer allowing soft transition between the constant modulus and the decision-directed algorithm for PSK modulated signals,” in *Proc. IEEE Int. Conf. on Communications*, (Geneva, Switzerland), pp. 1144–1148, May 1993.
- [20] J. Proakis, *Digital Communication*. McGraw-Hill, 2 ed., 1989.
- [21] T. J. Shan, M. Wax, and T. Kailath, “On spatial smoothing for direction-of-arrival estimation of coherent signals,” *IEEE Trans. Acoustics, Speech and Signal Processing*, vol. 33, pp. 806–811, Apr. 1985.
- [22] D. A. Linebarger, R. D. DeGroat, and E. M. Dowling, “Efficient direction finding methods employing forward/backward averaging,” *IEEE Trans. on Signal Processing*, vol. 42, pp. 2136–2145, Aug. 1994.
- [23] M. Wax and A. Leshem, “Joint estimation of time delays and directions of arrival of multiple reflections of a known signal,” in *Proc. IEEE Int. Conf. Acoustic Speech and Signal Processing*, vol. 5, (Atlanta, GA), pp. 2622–2625, May 1996.

- [24] M. C. Vanderveen, *Estimation Of Parametric Channel Models In Wireless Communication Networks*. PhD thesis, Stanford University, May 1998.
- [25] G. Golub and C. V. Loan, *Matrix Computations*. The Johns Hopkins University Press, 3 ed., 1996.
- [26] A. J. van der Veen, M. C. Vanderveen, and A. Paulraj, "Joint angle and delay estimation using shift-invariance properties," *IEEE Trans. on Signal Processing*, vol. 46, pp. 142–145, Feb. 1998.
- [27] M. Haardt and J. A. Nossek, "Unitary ESPRIT: How to obtain increased estimation accuracy with a reduced computational burden," *IEEE Trans. on Signal Processing*, vol. 43, pp. 1232–1242, May 1995.
- [28] M. D. Zoltowski, M. Haardt, and C. P. Mathews, "Closed-form 2-D angle estimation with rectangular arrays in element space or beamspace via unitary ESPRIT," *IEEE Trans. on Signal Processing*, vol. 44, pp. 316–328, Feb. 1996.
- [29] K. C. Hurang and C. C. Yeh, "A unitary transformation method for angle-of-arrival estimation," *IEEE Trans. on Signal Processing*, vol. 39, pp. 975–977, Apr. 1991.
- [30] G. Xu, R. H. Roy, and T. Kailath, "Detection of number of sources via exploitation of centro-symmetry property," *IEEE Trans. on Signal Processing*, vol. 42, pp. 102–112, Jan. 1994.

Vita

Srikanth Gummadi was born in Hyderabad, India, on May 30, 1976 to Sujatha and Veerabhadra Rao. He received a Bachelor of Technology degree in Electrical Engineering from the Indian Institute of Technology, Madras, India, in May of 1997, and was then admitted to The University of Texas at Austin where he has been pursuing the degree of Master of Science in Engineering. Since June of 1997, he has been a Research Assistant in the Embedded Signal Processing Laboratory working with Dr. Brian L. Evans and has enjoyed every moment of it. His current research interests include smart antennas, signal processing for mobile communication systems, and embedded processing on microcontrollers.

Permanent address: 2910 Medical Arts St.
Apt. #302
Austin, Texas 78705

This report was typeset¹ with L^AT_EX by author.

¹L^AT_EX document preparation system was developed by Leslie Lamport as a special version of Donald Knuth's T_EX program for computer typesetting. T_EX is a trademark of the American Mathematical Society. The L^AT_EX macro package for The University of Texas at Austin report format was written by Khe-Sing The.