DESIGN, DEVELOPMENT, AND EVALUATION OF
A SYSTEM FOR ANGLE-OF-ARRIVAL MEASUREMENT
AT A UHF MOBILE RADIO BASE STATION

DISSERTATION

Presented in Partial Fulfillment of the Requirements for
the Degree Doctor of Philosophy in the
Graduate School of The Ohio State University

By

Steven W. Ellingson, B.S., M.S.

* * * * *

The Ohio State University

2000

Dissertation Committee:
Prof. Walter D. Burnside, Adviser
Dr. Inder J. Gupta
Prof. Randolph L. Moses
Prof. Edward H. Newman

Approved by
Walter D. Burnside
Adviser
Department of Electrical Engineering
© Copyright by

Steven W. Ellingson

2000
ABSTRACT

A system for angle-of-arrival (AOA) measurement at a UHF mobile radio base station is described. A seven-element "Y"-shaped array of monopoles above a ground plane is used to observe the signal from a hand-held radio transmitting at 460 MHz as the user moves through a variety of terrain conditions. A novel low-cost digitizing array receiver is used to capture the signal. Array calibration is achieved by fitting a physical model of the open-circuit impedance matrix to a small number of previous AOA measurements for known transmitter locations. The calibrated data is then analyzed using conventional AOA estimators assuming discrete AOAs. The quality of the estimate is assessed using a simple "residual power" metric which provides an indication of how well the estimate describes the data obtained from the array. The system was evaluated in both line-of-sight (LOS) and non-LOS cases. It is shown that transmitter localization to within a few degrees is usually possible. In certain cases, a higher-order estimate is used to identify significant multipaths. A few examples of complex terrain scattering are also observed, including evidence of multipath reflections from buildings. In general, this multipath is found to be weak relative to the primary path. The results also suggest that angle spread was less than 5° in LOS cases, and less than 20° in non-LOS cases, which is consistent with the findings of other field studies.
To Karen, Colin, Nathan, and Ryan
ACKNOWLEDGMENTS

I would like to express my deep appreciation to my adviser, Prof. W.D. Burnside, who provided me the opportunity to pursue the research described in this dissertation. He was also a great source of inspiration and encouragement. I would also like to thank Dr. I.J. Gupta, Prof. R.L. Moses, and Prof. E.H. Newman for their service on the dissertation reading committee.

The system described in this dissertation consists of 40 printed circuit boards including 10 new designs, plus seemingly endless enclosures, cables, and connectors. Such a project would simply not have been possible without the combined effort of many members of the Ohio State University ElectroScience Laboratory. I am particularly grateful to Gary Winchell, who collaborated on the layout of the critical RFGB, mixer, WBDIF, and DFIFO boards; and who did the CAD entry. He and Scott Freeman built many of these boards, demonstrating an impressive level of craftsmanship. Scott Freeman designed and built a system power distribution shelf. Jim Moncrief, Pablo Diez, Elaine Manning, and Michelle Davies assembled the receiver subsystems and made lots of cables. Pablo Diez built additional special purpose test equipment and helped in the field testing. Russell Eckley assembled the clock generation board and the LNAs. Jim Hetrick worked out the mechanical design of the antenna array and built it on the roof. Also, the DFIFO design was modified from a single-FIFO design by Scott Yano.
Most importantly, I would like to thank my wife, Karen, and sons, Colin, Nathan, and Ryan for their support and patience, especially over the last three years.

This work was supported by the National Science Foundation under Grant ANI-9809018 and the Ohio State University. Certain aspects of the receiver design described in this study were developed as part of an unrelated project sponsored by the SETI Institute.
VITA

July 9, 1965 .................................. Born - Tonawanda, NY, USA

1987 ........................................... B.S. Electrical & Computer Engineering, Clarkson University, Potsdam, NY, USA

1987-1989 ..................................... Graduate Research Associate, The Ohio State University ElectroScience Laboratory, Columbus, OH, USA

1989 ........................................... M.S. Electrical Engineering, The Ohio State University, Columbus, OH, USA

1989-1993 ..................................... Active Duty, U.S. Army

1993-1995 ..................................... Senior Consultant, Booz-Allen & Hamilton, McLean, VA, USA

1995-1997 ..................................... Systems Engineer, Raytheon E-Systems, Falls Church, VA, USA

1997-present ................................... Senior Research Associate, The Ohio State University ElectroScience Laboratory, Columbus, OH, USA

PUBLICATIONS

Research Publications

FIELDS OF STUDY

Major Field: Electrical Engineering

Studies in:

Electromagnetics
Communications & Signal Processing

Prof. Walter D. Burnside
Prof. Randolph L. Moses
# TABLE OF CONTENTS

<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Abstract</td>
<td>ii</td>
</tr>
<tr>
<td>Dedication</td>
<td>iii</td>
</tr>
<tr>
<td>Acknowledgments</td>
<td>iv</td>
</tr>
<tr>
<td>Vita</td>
<td>vi</td>
</tr>
<tr>
<td>List of Tables</td>
<td>xi</td>
</tr>
<tr>
<td>List of Figures</td>
<td>xii</td>
</tr>
<tr>
<td>Chapters:</td>
<td></td>
</tr>
<tr>
<td>1. Introduction</td>
<td>1</td>
</tr>
<tr>
<td>1.1 Summary of this Research</td>
<td>3</td>
</tr>
<tr>
<td>1.2 Organization of this Dissertation</td>
<td>5</td>
</tr>
<tr>
<td>2. Angle Spectrum Estimation</td>
<td>8</td>
</tr>
<tr>
<td>2.1 Data Model</td>
<td>8</td>
</tr>
<tr>
<td>2.2 Model-Independent Techniques</td>
<td>10</td>
</tr>
<tr>
<td>2.3 Discrete AOA Techniques</td>
<td>12</td>
</tr>
<tr>
<td>2.3.1 Maximum Likelihood</td>
<td>13</td>
</tr>
<tr>
<td>2.3.2 Subspace Techniques</td>
<td>16</td>
</tr>
<tr>
<td>2.3.3 Modulation-Specific Techniques</td>
<td>21</td>
</tr>
<tr>
<td>2.4 Distributed AOAs</td>
<td>21</td>
</tr>
<tr>
<td>2.5 Conclusions</td>
<td>24</td>
</tr>
</tbody>
</table>

viii
3. Antenna Array .................................................. 25
   3.1 Background ............................................... 25
   3.2 Azimuth Plane Analysis .................................. 29
   3.3 Elevation Plane Design & Analysis ...................... 35
   3.4 Electromagnetic Coupling Between Elements ........... 45
   3.5 Improving Elevation Plane Performance Using Rolled Edges .... 47
   3.6 Conclusions ............................................... 55

4. Calibration ..................................................... 56
   4.1 Background ............................................... 57
   4.2 Estimating the Receiver Gains and Phases .............. 62
      4.2.1 Sinusoidal Parameter Estimation ...................... 63
      4.2.2 Precalibration ....................................... 69
      4.2.3 Improving Performance using Wideband Calibration Signals .... 72
   4.3 Estimating the Open-Circuit Impedance Matrix ........... 72
   4.4 Calibration Accuracy Requirements ..................... 77
   4.5 Conclusions ............................................... 79

5. Array Receiver ................................................ 80
   5.1 Requirements ............................................. 80
      5.1.1 Coherency, Calibration, and Bandwidth ............... 81
      5.1.2 Linearity ............................................ 82
      5.1.3 Sensitivity .......................................... 87
      5.1.4 SFDR and the Implications of a Wideband Front End .... 93
      5.1.5 Image Rejection ..................................... 94
      5.1.6 I/Q Balance ......................................... 95
      5.1.7 Requirements Summary ................................ 96
   5.2 Design Overview .......................................... 97
      5.2.1 Signal Flow: Hardware ................................ 97
      5.2.2 Signal Flow: Software ................................ 103
      5.2.3 Specifications ....................................... 107
   5.3 Design Details ........................................... 107
      5.3.1 Antenna Array Subsystem (AAS) ....................... 107
      5.3.2 Low-Noise Amplifier Array (LNAA) .................... 109
      5.3.3 Receivers (RX-1 through -4) ......................... 111
      5.3.4 PC Interface (PCIF) .................................. 122
      5.3.5 Local Oscillator and Clock Distribution (LCDIST, SG-1, SG-2, and ARB) .... 126
6. Field Measurements ........................................ 129
   6.1 Test Conditions ....................................... 129
       6.1.1 Test Area ...................................... 129
       6.1.2 Mobile Transmitter ............................ 131
   6.2 System Operation .................................... 133
       6.2.1 Receiver Operation ............................ 133
       6.2.2 Array Calibration .............................. 140
   6.3 LOS Sites ........................................... 143
   6.4 Shadowed Sites ...................................... 144
   6.5 Suburban Sites ...................................... 153
   6.6 Sites with Complex Local Structure ............. 153
   6.7 Sources of Error .................................... 160
   6.8 Conclusions ......................................... 162

7. Conclusions .............................................. 164

Bibliography ................................................. 169
# LIST OF TABLES

<table>
<thead>
<tr>
<th>Table</th>
<th>Description</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>3.1</td>
<td>Magnitude (dB) of the open circuit impedance matrix, relative to the self-impedance of the center element.</td>
<td>46</td>
</tr>
<tr>
<td>3.2</td>
<td>Phase (degrees) of the open circuit impedance matrix, relative to the self-impedance of the center element.</td>
<td>47</td>
</tr>
<tr>
<td>3.3</td>
<td>Summary of results for the monopole $1.5\lambda$ from the edge (or attachment point).</td>
<td>54</td>
</tr>
<tr>
<td>4.1</td>
<td>Calibration accuracy requirements to achieve $1^\circ$ RMS estimation error, given the experiment conditions described in the text, assuming only one type of error is present at a time.</td>
<td>79</td>
</tr>
<tr>
<td>5.1</td>
<td>Array receiver specifications, as built, compared to requirements developed in Section 5.1.</td>
<td>109</td>
</tr>
<tr>
<td>5.2</td>
<td>GNI analysis of the analog portion of the receiver.</td>
<td>119</td>
</tr>
<tr>
<td>6.1</td>
<td>LOS examples</td>
<td>143</td>
</tr>
<tr>
<td>6.2</td>
<td>Shadowing examples</td>
<td>148</td>
</tr>
<tr>
<td>6.3</td>
<td>Suburban examples. Numbers in parentheses indicate number of outliers among the trials; if outliers are indicated, then the bias is calculated as the median, as opposed to the mean, of the difference between the estimated AOA and the true bearing.</td>
<td>154</td>
</tr>
<tr>
<td>6.4</td>
<td>Complex local structure examples. Numbers in parentheses indicate number of outliers among the trials; if outliers are indicated, then the bias is calculated as the median, as opposed to the mean, of the difference between the estimated AOA and the true bearing.</td>
<td>155</td>
</tr>
</tbody>
</table>
LIST OF FIGURES

Figure                                      Page

3.1 Y array geometry. $k = 1$ for uniform spacing between elements on each   29
    leg. .................................................................

3.2 Normalized cost-per-aperture ratio $(N/Q)$. Solid: ULA $(N = 8)$, Dash:    31
    UCA, Dash-dot: L $(N = 7)$, Dot: Y $(N = 7)$. ........................

3.3 Gain patterns of the 8-element ULA for various main beam pointing        32
    directions. Broadside is $0^\circ$. ................................

3.4 Gain patterns of the 7-element Y array for various main beam pointing     33
    directions. ........................................................

3.5 CRB of the estimation variance for the fixed AOA as a function of the      34
    variable AOA. Fixed AOA is at (a) $0^\circ$ and (b) $55^\circ$. ............

3.6 CRB of the estimation variance for the variable AOA as a function of the 36
    variable AOA. Fixed AOA is at (a) $0^\circ$ and (b) $55^\circ$. ............

3.7 CRB of the estimation variance for the variable AOA as a function of the 37
    variable AOA. ULA, Y, and sparse Y (Ye) only. Fixed AOA is at (a) $0^\circ$  
    and (b) $55^\circ$. ................................................

3.8 CRB of the estimation variance for the fixed AOA as a function of SNR · L, 38
    when the fixed AOA is $0^\circ$ and the variable AOA is $3^\circ$. “Ye”  
    indicates the sparse $(k = 1.8)$ Y. ................................

3.9 Top view of array geometry. The circles indicate monopole positions       40
    and the outline of the ground plane is indicated. As built, the top edge  
    of this figure is North. .........................................
3.10 Array as built. View is slightly to the North of East. 41

3.11 Geometry for a monopole on a ground plane with "knife-edge" termination. 43

3.12 Power pattern of half-wave dipole in free space and a quarter-wave monopole on an infinite ground plane. 44

3.13 Pattern of a quarter-wave monopole mounted on a ground plane with an edge at the indicated distances from the monopole. 45

3.14 Geometry for a monopole on a ground plane with an elliptical rolled edge termination. 48

3.15 Pattern of a quarter-wave monopole 1.5\(\lambda\) from the junction between a ground plane and circular rolled edges of various radii. 52

3.16 Same as Figure 3.15, except now the rolled edge radius is fixed at 5\(\lambda\) and the distance from the monopole to the junction point is varied. 53

3.17 Pattern of a quarter-wave monopole mounted on a ground plane of radius 1.5\(\lambda\), using various elliptical terminations. 54

4.1 A technique for estimating the receiver calibration, C. 63

5.1 UAX array location. 83

5.2 0-1000 MHz, as observed through a wideband omnidirectional antenna near the array. The horizontal grid lines are 10 dBm divisions, with 0 dBm at the top. 84

5.3 460-470 MHz, as observed through the center element of the array. The horizontal grid lines are 10 dBm divisions, with 0 dBm at the top. 84

5.4 The relationship between receiver gain, 1dB compression, and IIP_3. The x axis is \(\log_{10}\) of the input power, whereas the y axis is \(\log_{10}\) of the output power. 85

5.5 Application of Lee's Model using the "suburban" standard condition. Note "Tx" refers to the transmitter; whereas "Rx" refers to the receiver. 90
5.6 SNR for an ideal receiver as a function of range, assuming free-space transmission (broken line) and Lee's Model with a "suburban" standard condition (solid line). ........................................ 93

5.7 Signal flow for one UAX receiver subsystem. ........................................ 98

5.8 Frequency plan for acquisition of the 25 MHz IF. Acquisition of the 35 MHz IF differs only in the bottom two spectra. .................. 101

5.9 Response of the digital baseband filter. ........................................ 106

5.10 UAX system block diagram. ................................................... 108

5.11 LNAA. ................................................................. 110

5.12 One of four identical receiver subsystems. .............................. 111

5.13 RFGB schematic. ....................................................... 112

5.14 RFGB construction: (a) Top side of PCB, (b) Underside of RFGB PCB, showing helical filter mounting. .................. 113

5.15 Mixer schematic. ....................................................... 114

5.16 Mixer assembly: (a) Top side of PCB, (b) As mounted in receiver chassis.115

5.17 IFGB schematic. ....................................................... 118

5.18 IFGB board mounted in enclosure. ..................................... 118

5.19 WBDIF schematic (A/D section). ................................... 120

5.20 WBDIF schematic ($F_s/4$ section). .................................. 121

5.21 WBDIF construction details. ........................................... 122

5.22 DFIFO schematic, data section. FIFO A shown, FIFO B is identical. 123

5.23 DFIFO schematic, control section. ................................... 124

5.24 DFIFO. .............................. 125

xiv
5.25 Clock and frequency reference generation. ........................................ 126

6.1 Test Area. The right edge faces East (0°), and the top edge faces North (90°). ................................................................. 130

6.2 View from the array, looking North. ................................................. 131

6.3 View from the array, looking South. ................................................ 132

6.4 Spectrum from a field measurement at various stages of software processing. The horizontal axis is RF frequency in MHz; whereas the vertical axis is received power spectral density in dBFS/bin. FFT with Blackman window. Top: Original 16K sample block. Middle: After filtering and decimation by 4. Bottom: After channelization. ................. 134

6.5 Calibrated spectrum from each of the elements of the array. The horizontal axis is RF frequency in MHz; whereas the vertical axis is received power spectral density in dBFS/bin. 4K FFT with Blackman window. 137

6.6 SNR versus range for every acquisition reported in this chapter. Circles: LOS sites, asterisks: shadowed sites, diamonds: suburban sites, x’s: complex scattering sites. Predicted path loss assuming free space (dashed) and Lee’s suburban model (solid) also shown. ................. 138

6.7 Eigenvalue spreads observed from 6 acquisitions while the mobile was at Site 2. ................................................................. 139

6.8 Bias for each of six acquisitions (corresponding to Sites 2-7) as a function of the impedance matrix phase parameter. ............................. 141

6.9 Optimization of the impedance matrix phase parameter for array calibration. The difference between the two most extreme values of bias ("o") for six acquisitions (corresponding to Sites 2-7) are shown. The mean bias ("x") is also shown. ............................. 142

6.10 Angle spectrum estimates for Site 3 (5 trials). ............................... 145

6.11 ML₁ AOAs for Sites 3, 4, 6, and 7. .............................................. 146
6.12 AOAs estimated for Site 5. (a) ML₁; (b) ML₂ scatter plot; (c) ML₂, First AOA; (d) ML₂, Second AOA. For ML₂, note there are 6 AOA pairs superimposed for first AOA equal to 135.8°. .......................... 147

6.13 ML₁ AOAs for Sites 1 and 12. .......................... 149

6.14 Angle spectrum estimates for Site 12 (6 trials). .......................... 150

6.15 AOAs estimated for Site 1. (a) ML₁; (b) ML₂ scatter plot; (c) ML₂, First AOA; (d) ML₂, Second AOA. The dashed lines in (b) indicate the optical boundaries of west and east edges of Building A (top and middle lines, respectively) and the east edge of Building B (bottom line). .......................... 152

6.16 View of the suburban sites, looking East. The light pole at the street corner is Site 22. .......................... 154

6.17 ML₁ AOAs for Site 18 (4 trials). .......................... 155

6.18 ML₁ AOAs for Sites 19–22. .......................... 156

6.19 ML₁ AOAs for Sites 9 (8 trials). .......................... 157

6.20 ML₁ AOAs for Sites 10 (12 trials). .......................... 157

6.21 ML₁ AOAs for Site 13 (5 trials). .......................... 158

6.22 ML₁ AOAs for Site 26 (13 trials). .......................... 159
CHAPTER 1

INTRODUCTION

In ultra-high frequency\textsuperscript{1} (UHF) land mobile radio systems, a signal transmitted by a mobile user may be scattered by the surrounding environment and arrive at the base station along multiple paths. Only in the case where the mobile is far from the base station, and the environment does not scatter the signal, will all the power arrive at the base station along a single path, or “angle of arrival” (AOA). If scattering of the signal occurs close to the mobile, then the power which arrives at the base station will appear to be spread over a small range of angles. If there is scattering from objects which are distant from both the mobile and the base station, then there may be multiple AOAs, each of which may appear to be spread. Measurement of the angular distribution of this signal power is a current problem of great theoretical and practical interest. Specific applications include “smart antennas” ([1] and references); cellular site planning and channel modeling ([2] and references); and various location-based services such as intelligent transportation systems, wireless E-911, and location-specific billing ([3] and references).

Despite the current intense commercial and academic interest in this problem, progress in developing practical systems for angle spectrum estimation has been slow.

\textsuperscript{1}300 MHz – 3000 MHz
One reason is that systems suitable for the acquisition of data in support of the necessary research are not commonly available. Several groups have reported custom-built systems designed to perform measurements in stationary conditions, using for example rotating dishes or aperture probing techniques [4, 5, 6, 7, 8, 9]. These approaches lead to simple and effective hardware and have been shown to be quite useful for "off-line" characterization of angle spectrum. However, they are not suitable for use in the goal systems, which will use fixed arrays of a few antenna elements and will be making observations in dynamically changing propagation channels. In fixed-element arrays, one is faced with numerous additional problems which complicate the process of angle spectrum estimation; for example, discrete sampling of the aperture, electromagnetic coupling between elements, and calibration and coherency among multiple receivers. A few groups have reported practical fixed-array AOA estimation systems for UHF [10, 11], and a few have reported field results in live mobile radio systems [12, 13, 14, 15]. These efforts have provided some cause for optimism that certain applications of angle spectrum estimation – e.g., beam pointing, Wireless E-911 – are practical using fixed arrays.

However, the application-specific nature of these studies leaves unanswered some of the important questions that pertain to fixed arrays generally. For example, if one has the freedom to select a geometry for the fixed array, what is the best choice? What aspects of calibration are important in practice? What aspects of receiver design should be reconsidered when designing arrays as opposed to single antennas? These are various issues associated with the same general question: How does one design a practical fixed-array angle spectrum measurement system for UHF mobile radio? This is currently a difficult question to answer; not only for the reasons
indicated above, but also because it involves relationships between many disciplines – propagation, antennas, RF design, digital signal processing, and spectral estimation – that are traditionally considered separately. The premise in this research is that the desired insights can best be obtained by going through the processes of designing and building a complete system, and then observing its performance in actual field conditions. Along the way, the issues which come to bear on the questions posed above can be identified and analyzed.

1.1 Summary of this Research

This dissertation describes the design, development, and evaluation of a system for measurement of AOA at mobile radio base stations. The system has been developed from scratch, and each aspect of the design has been optimized for the task of AOA measurement using existing AOA estimation algorithms. A seven-element "Y"-shaped array of monopole antenna elements on a ground plane is used. The "Y" offers nearly-uniform performance with low cost per unit aperture. The ground plane reduces errors due to scattering from structures close to the array. To capture the signals from the array, a novel wideband digital receiver system is developed. The central feature of the system is a "dual receiver" which accepts inputs from two elements of the array, and outputs a digital representation of those signals in a form which makes them easy to manipulate in post-processing software. The receiver uses a single analog frequency conversion to a 25 MHz or 35 MHz intermediate frequency (IF), which is digitized and further manipulated using a series of multiply-free \( F_S/4 \) downconversions and downsampling. Individual narrowband signals are extracted from the
2.5 MHz passband using a successive estimation-subtraction technique to remove all undesired signals, followed by a spectral shift to zero-IF.

Before using AOA estimation algorithms, system calibration is required. In the approach implemented here, the three main contributions to the system calibration – the array response (neglecting mutual coupling), the mutual coupling, and the receiver complex gains – are computed separately. Because the array described above is rigid, its response can be computed in advance. The receiver complex gains are computed “on-the-fly”; i.e., simultaneously with the data collection, using CW signals injected at the antenna terminals. Finally, the mutual coupling is estimated using a novel procedure in which a model of the open-circuit impedance matrix, computed using the Moment Method [16], is varied using a single parameter to optimize the bias for a small number of AOA estimates corresponding to known transmitter locations. An estimate of orientation error (bias due to an unknown rotation of the array with respect to the experimental coordinate system) is generated as side information.

To perform the necessary experiments, regular access to a suitable block of frequency spectrum was required. In this project, the United States’ 462 MHz “Family Radio Service” (FRS) band is selected. This is an unlicensed band for which inexpensive hand-held mobile radio equipment is commercially available. Each FRS channel is allocated for two-way (half-duplex) voice-only communications using FM modulation with 12.5 kHz maximum bandwidth. Although this band is quite far removed from the 800–900 MHz and 1800–1900 MHz bands which are currently generating the most commercial interest, it is known that the fundamental propagation mechanisms
are basically the same, with predictable dependence on wavelength [17]. Furthermore, this band displays other features common to all UHF mobile radio bands, most notably cochannel and adjacent-channel interference.

This system has been evaluated in field conditions, and the results are presented here. Both line-of-sight (LOS) and non-LOS cases were tested. It is shown that transmitter localization to within a few degrees is possible in both LOS and non-LOS conditions using traditional AOA estimators. A few examples of complex terrain scattering are observed, including evidence of multipath reflections from buildings.

A tight constraint on this study is cost. It is important that the hardware developed for this study be built within a small fixed budget. While this necessarily forces certain compromises on system capabilities, this emphasis is well-matched to the overall objectives of this study. Thus, there is considerable effort expended in this study to determine the specifications required to minimize cost while maintaining a useful level of performance.

1.2 Organization of this Dissertation

The organization of this dissertation is as follows. Chapter 2 ("Angle Spectrum Estimation") provides a brief review of the relevant theory and descriptions of candidate AOA estimators. From this review, three well-known estimators are selected. To gain additional insight into the true nature of the AOA distribution, a simple procedure is proposed for calculating the "residual power" associated with a discrete AOA spectrum estimate. This metric provides an indication of how well the estimate describes the data obtained from the array.
Chapter 3 ("Antenna Array") discusses the design and analysis of the antenna array used in this study. Linear, Circular, “L”-shaped, and “Y”-shaped arrays are considered from the perspective of signal processing performance and cost. From this analysis, a 7-element “Y”-shaped array of vertically-polarized monopoles over a ground plane is selected. An electromagnetic analysis of this array shows that the ground plane provides the desired suppression of signals from low elevation angles, but also degrades the gain along the horizon, from where the desired signals originate. An improved design using convex rolled edge terminations is proposed and analyzed. Also, a separate electromagnetic analysis is performed to estimate the mutual coupling expected in the original (untreated edge) design, which was actually built and used in this study.

Chapter 4 ("Calibration") addresses the problem of calibrating an AOA measurement system, including the antenna array and receivers. Various known calibration strategies are considered, but each is found to have limitations which make these techniques undesirable in this application. The calibration approach described above – computing the array response (neglecting mutual coupling), the mutual coupling, and the receiver complex gains separately – is described. The necessary procedures and algorithms are developed.

Chapter 5 ("Array Receiver") documents the design of the array receiver. First, requirements are developed based on the findings in Chapters 2–4, plus additional findings related to coherency, linearity, range, and various other considerations. High-level and detailed descriptions of the array receiver follow.
Chapter 6 ("Field Measurements") provides the results of field experiments conducted using the instrumentation and techniques developed in Chapters 2–5. Results, analysis, and sources of error are reported.

Finally, Chapter 7 ("Conclusions") summarizes the findings of this dissertation and offers some suggestions for future research in this direction.
CHAPTER 2

ANGLE SPECTRUM ESTIMATION

This section provides a short survey of techniques for azimuthal angle spectrum estimation. First, the data model used in this project is presented. This is followed by descriptions of three groups of techniques: those which are model-independent, those which assume discrete AOAs, and those which consider parametrically-distributed AOAs. This chapter concludes with a summary of the techniques found to be most suitable for this study.

2.1 Data Model

Consider an antenna array consisting of $N$ identical elements, with each element having a pattern which, in the absence of the other elements, is isotropic in azimuth. Let us further assume that the narrowband assumption applies. That is, any signals incident on the array are assumed to have bandwidth sufficiently small such that the array response can be accurately described in terms of gains and phase shifts, as opposed to gains and time delays. For a single narrowband signal $s(t)$ incident on this array from angle $\phi$ at time $t$, the complex baseband receiver outputs are given by $x(t) = [x_1(t) \ x_2(t) \ .. \ x_N(t)]^T$ where

$$x(t) = CZa(\phi)s(t) + n(t) \quad \text{(2.1)}$$
In this model, \( a(\phi) \) is an \( N \times 1 \) “steering vector” which describes the currents induced on each element by the incident signal in the absence of the other elements; i.e., neglecting the electromagnetic mutual coupling between the elements. \( Z \) is the \( N \times N \) open-circuit impedance matrix, in this case normalized to a unitless symmetric matrix that describes the coupling between elements. Thus, \( Za(\phi) \) is proportional to the voltages at the terminals of each element in the array. \( \mathbf{C} \) is a \( N \times N \) diagonal matrix which represents the gain/phase errors introduced due to difference among feedlines and receivers. It is assumed that cross-talk between elements within the array receiver, if any, is negligible. Finally, \( \mathbf{n}(t) \) is an \( N \times 1 \) vector representing the noise present in the data, which is assumed to be thermal in nature, and thus is well-modeled as stationary, ergodic, complex circular Gaussian white noise which is identically distributed and uncorrelated between elements.

In the more general case where power is incident on the array from \( M \) discrete directions, the model can be compactly expressed as

\[
\mathbf{x}(t) = \mathbf{CZA}s(t) + \mathbf{n}(t) \tag{2.2}
\]

where \( \mathbf{A} \) is an \( N \times M \) matrix composed of \( M \) steering vectors associated with the incident AOAs, and \( s(t) = [ s_1(t) \ s_2(t) \ldots \ s_M(t) ]^T \).

AOA estimation requires knowledge of the matrix \( \mathbf{B} = \mathbf{CZA} \). The process of finding \( \mathbf{B} \) is part of system calibration, discussed further in Chapter 4. Recall that the data model given by Equation 2.2 is expressed in terms of discrete AOAs. However, the true angle spectrum must in fact be continuous. For situations in which the true angle spectrum is dominated by a few discrete AOAs, Equation 2.2 is an excellent approximation. However, in more complex situations, this model may not be valid. Nevertheless, the matrix operator formalism in Equation 2.2 is convenient and will be
retained. This can be justified by noting that the number of signals $M$ can be made arbitrarily large, so that the model can be interpreted as a decomposition of the true incident fields into a spectrum of plane waves, for example. This concept will be put to use in the next section.

Given the above formulation, the following additional notation is introduced to facilitate the discussion of angle spectrum estimators which follows. The set of steering vectors $\mathbf{a}(\phi) \ \forall \ \phi \in [0, 2\pi)$ is referred to as the array manifold, $\mathcal{A}$. Similarly, the set of system response vectors $\mathbf{b}(\phi) \ \forall \ \phi \in [0, 2\pi)$ is referred to as the system manifold, $\mathcal{B}$. Thus, $\mathcal{A}$ and $\mathcal{B}$ are composed from vectors which lie in $\mathcal{A}$ and $\mathcal{B}$, respectively.

### 2.2 Model-Independent Techniques

If the true nature of the angle spectrum is of interest, then an angle spectrum estimator which is “model independent” may be preferred. Specifically, the estimator should not assume that the spectrum is composed of a finite number of discrete, resolvable AOAs. In this section, three such estimators are considered.

Perhaps the simplest approach that meets this criterion is the classical beamforming (CBF) method.\(^2\) In this approach, one attempts to measure the power for a given AOA by selecting a linear combination of the antenna outputs that results in maximum gain in the specified direction, subject to a unit norm constraint on the coefficients. The AOA spectrum is then computed by repeating the procedure for each AOA of interest. One obtains the angle power spectrum estimate:

$$P_{\text{cbf}}(\phi) = \mathbf{b}^H(\phi) \mathbf{R} \mathbf{b}(\phi), \text{ where}$$

\[\text{(2.3)}\]

\(^2\)Attributed by Krim and Viberg [18] to Bartlett [19]; See also [20, 21] for authoritative references.
the superscript “$H$” denotes the conjugate transpose, and $\mathbf{R}$ is the spatial covariance matrix, defined as the expected value of the outer product $\mathbf{x}(t)\mathbf{x}^H(t)$ and approximated from $L$ samples using

$$\mathbf{R} \approx \hat{\mathbf{R}} = \frac{1}{L} \sum_{l=1}^{L} \mathbf{x}(t_l) \mathbf{x}^H(t_l).$$  \hfill (2.4)

This method is simple and robust, but has severe performance limitations. First, the best possible resolution is about $57^\circ$ degrees per wavelength of physical aperture, assuming the AOA is close to broadside [16]. Second, this method can be badly biased, especially if multiple AOAs exist within a beamwidth. Finally, this method is affected by high sidelobes, with the effect that AOAs well outside the beamwidth for the AOA of interest can contribute significantly to the observed power of the desired AOA. This also manifests itself here as a dynamic range problem. In other words, it may be difficult to detect weak signals in the presence of stronger signals, when the difference is greater than the sidelobe level.

It should be noted that the CBF and all other techniques discussed in this chapter require that the set of all possible steering vectors $\mathbf{a}(\phi)$ be linearly independent in the field of view. In other words, each $\phi$ should be associated with a unique $\mathbf{a}(\phi)$. This is in effect a requirement on the geometry of the antenna array. For an example where this condition is not true, consider the linear array, which experiences a front-to-back ambiguity in steering vector for a field of view of $2\pi$ radians. For this reason, the field of view of a linear array must be restricted to one of the half-spaces delimited by its endfire directions. The issue of array geometry is addressed further in Chapter 3.

An alternative to the CBF estimator is the Capon estimator:

$$P_{\text{Capon}}(\phi) = \left[ \mathbf{b}^H(\phi) \mathbf{R}^{-1} \mathbf{b}(\phi) \right]^{-1} \hfill (2.5)$$
which in fact is the maximum likelihood (ML) estimate of power incident from angle $\phi$ [20, 22, 23]. In the absence of multipath, the Capon estimator yields low sidelobes and resolution which is superior to CBF. Unlike the CBF, it is degraded in the presence of correlated signals (multipath) and fails completely in the presence of widely-spaced, perfectly-correlated multipath of nearly equal strength.

Another approach to model-independent angle spectrum estimation is deconvolution. In this approach, one recognizes that $P_{\text{CBF}}$ represents the convolution of the true incident fields with the CBF array pattern as it is swept in azimuth. If $B$ is known, then the true angle spectrum can in principle be recovered by deconvolution. Unfortunately, the deconvolution tends to be ill-conditioned unless the main lobe of the CBF pattern is vary narrow, which requires that the physical aperture be very large. To avoid grating lobes, the number of elements must then be very large as well. Some headway can be gained using model-based deconvolution, such as the CLEAN procedure [24]. If one must resort to model-dependent estimation, however, the model-based AOA estimators discussed in the following sections are preferred.

2.3 Discrete AOA Techniques

In this section, estimators which assume discrete AOAs are considered. Among the simplest of these are the discrete AOA interpretations of the CBF and Capon estimators, presented above. For these methods, the discrete AOAs are chosen as the global and possibly some local peak values of $P_{\text{CBF}}(\phi)$ and $P_{\text{Capon}}(\phi)$. Although common, this procedure leads to some problems in practical use. First, there is the problem of deciding how many peaks to interpret as bona fide discrete AOAs. A
related problem is discriminating between *bona fide* peaks and sidelobe maxima, although this is easier for the Capon estimator than for the CBF. One might arbitrarily decide to restrict attention to the largest peak; in other words, adopt an $M = 1$ model. In this case, the estimate is likely to be badly biased if the true $M$ is greater than 1 or if the true angle spectrum is not discrete. In fact, these problems are common to various degrees for all model-based AOA estimators.

The motivation for considering more sophisticated discrete AOA estimators is to gain some advantage over CBF and Capon in terms of resolution and statistical performance. Considered below are two well-known classes of discrete AOA estimators, namely Maximum Likelihood (ML) techniques and Subspace techniques. A feature of both classes is that the maximum number of AOAs that can be found is less than $N$, the number of elements in the array.$^3$

2.3.1 Maximum Likelihood

In the deterministic ML method of Ziskind and Wax [25], the concept is to conduct an $M$-dimensional search for the set of AOAs which maximize a certain likelihood function. The likelihood function is derived from the following observation. Since $\mathbf{n}(t)$ is Gaussian, Equation 2.2 infers that the joint probability density function for all unknowns in Equation 2.2 is

$$\prod_{l=1}^{L} \frac{1}{\pi \det (\sigma^2 \mathbf{I})} \exp \left\{ -\frac{1}{\sigma^2} \| \mathbf{x}(t_l) - \mathbf{B}(\Phi)\mathbf{s}(t_l) \|^2 \right\}$$

(2.6)

where $\sigma^2 \mathbf{I}$ is the spatial covariance matrix for the noise alone, $\mathbf{B}(\Phi)$ represents values of the system response matrix generated assuming various trial values for the discrete AOAs given by the set $\Phi = \{\phi_1, \phi_2, \ldots, \phi_M\}$, and one considers a set of $L$ samples taken

$^3$Assuming narrowband signals and using only the spatial covariance, $\mathbf{R}$. 

13
at times \( t_l, l = 1, 2, ... L \). The unknowns can be estimated by maximizing Equation 2.6 over the range of the unknown parameters. Ziskind and Wax show that the AOAs can be decoupled from the other unknown parameters to the extent that this optimization reduces to the estimator

\[
\arg \max_{\phi} \text{Tr}\{P_{\phi}R\} \tag{2.7}
\]

where \( P_B \) is the linear operator

\[
P_B = B(B^H B)^{-1} B^H \tag{2.8}
\]

which projects its argument into the column span of \( B \), \( \text{span}\{B\} \). "Tr" in Equation 2.7 indicates the trace operator, which is the sum of the eigenvalues of the matrix argument, or equivalently, the sum of its diagonal elements. Thus, Equation 2.7 finds the AOAs that maximize the power resulting when \( R \), representing the array output, is projected into the column span of the trial system response matrix \( B \). In other words, this estimator seeks the AOAs (represented by \( B \)) which are most consistent with the second-order statistics of the observed data, \( R \). This approach is optimal in the sense that no other estimator yields lower variance when the model conditions are satisfied and the estimate of \( R \) is perfect.

For the special case of \( M = 1 \) the ML estimator simplifies to

\[
\arg \max_{\phi_1} \; b^H(\phi_1) \; R \; b(\phi_1) \tag{2.9}
\]

Note that the function to be optimized in the above equation is exactly \( P_{\phi_1}(\phi) \). Hence this estimator, which shall be referred to as "ML1", is exactly the CBF discrete AOA estimator for \( M = 1 \). Similarly, the ML estimators for \( M > 1 \) can be interpreted as multidimensional generalizations of CBF-based discrete AOA estimator. Like the CBF, the ML estimator is robust to the presence of multipath. Unlike the CBF, the
ML estimator remains unbiased for any \( M \), as long as the correct value of \( M \) is used in the estimator. If \( M \) is selected to be too small, the results may be biased; whereas, if \( M \) is selected to be too large, unnecessary computational effort is expended and the additional AOA estimates may not correspond to physical signals. The problem of selecting \( M \) correctly is known as model order estimation. Considerable theoretical effort has been applied to this problem. Perhaps most applicable among these are the multipath-robust minimum description length (MDL)-based methods of Wax and Ziskind [26] and Wax [27]. The approach in [26] jointly estimates \( M \) and the AOAs using a procedure similar to the AOA-only method described above. The latter method is generalized to accommodate the presence of noise which does not fit the \( \sigma^2 I \) model. Such techniques are not employed in this study. For simplicity, let us instead concentrate on the use of fixed-order estimators with \( M = 1 \) or \( 2 \), and use other means to assess the probability of additional significant AOAs.

One such method, suitable for exposing strong, resolvable, but unmodeled signals after AOA estimation is as follows. Given the order-\( M \) AOA estimate \( \hat{\Phi} \), a spatially-filtered covariance matrix can be computed as follows:

\[
R_F = P_{B(\hat{\Phi})}^\perp R (P_{B(\hat{\Phi})}^\perp)^H , \quad \text{where}
\]

\[
P_{B(\hat{\Phi})}^\perp = I - P_{B(\hat{\Phi})} . \quad (2.11)
\]

Assuming \( \hat{\Phi} \) is a perfect estimate of the true incident signals, the power associated with those signals is eliminated from \( R_F \). Otherwise, the residual power relative to the total power is greater than zero and can be estimated using:

\[
\frac{\text{Tr}\{R_F\}}{\text{Tr}\{R\}} . \quad (2.12)
\]
Thus, the residual indicates how much of the total power collected by the array can be attributed to the estimate \( \hat{\Phi} \). If \( \hat{\Phi} \) is accurate, there are no other signals, and the incident signal-to-noise ratio (SNR) is high, then the residual is approximately equal to the reciprocal of the signal-to-noise ratio, \((\text{SNR})^{-1}\). If, on the other hand, other AOAs of significant power are present, the residual will increase toward a value of unity. Again assuming the SNR is high, a value close to unity indicates that none of the incident signal power has been identified.

Once \( M \) and \( \Phi \) have been estimated in a discrete AOA model, one can solve for other parameters which were decoupled from the AOAs in the derivation of the ML estimator. To solve for the power of each signal and the correlations between signals, one can calculate the source signal covariance matrix [10]:

\[
R_{ss} = (B^H B)^{-1} B^H (R - \sigma^2 I) B (B^H B)^{-1},
\]  

(2.13)

where the diagonal elements of \( R_{ss} \) are the powers associated with each discrete AOA, and the off-diagonal elements are the correlations between those signals.

### 2.3.2 Subspace Techniques

An alternative to the optimal ML approach described above is a subspace-based approach. Subspace-based approaches exploit the fact that \( R \) can be decomposed as

\[
R = U \Lambda U^H
\]  

(2.14)

where \( U \) is an \( N \times N \) matrix whose columns are the \( N \) eigenvectors of \( R \), and \( \Lambda \) is an \( N \times N \) diagonal matrix whose elements are the corresponding eigenvalues of \( R \). Since \( R \) is hermetian, it has \( N \) real eigenvalues which are greater than or equal to \( \sigma^2 \). It is assumed that \( U \) and \( \Lambda \) are ordered such that the largest eigenvalue is in the top left
position of $\Lambda$, and all other eigenvalues are arranged in descending order from top left to bottom right. If there are no signals incident on the array, $R = \sigma^2 I$ and therefore $\Lambda = \sigma^2 I$. In the presence of signals, $k \leq N$ eigenvalues of $R$ will become greater than $\sigma^2$. $R$ can then be partitioned into the sum of "signal" and "noise" covariance matrices as follows:

$$R = U_S \Lambda_S U_S^H + U_N \Lambda_N U_N^H$$  \hspace{1cm} (2.15)

where $U_S$ is the $N \times k$ matrix whose columns are the leftmost $k$ columns of $U$, $\Lambda$ is an $k \times k$ diagonal matrix whose elements are the eigenvalues corresponding to $U_S$, $U_N$ is the $N \times (N - k)$ matrix whose columns are the rightmost $N - k$ columns of $U$, and $\Lambda_N$ is the $(N - k) \times (N - k)$ diagonal matrix whose elements are the eigenvalues corresponding to $U_N$. In this context, the signal subspace is defined as $\text{span}\{U_S\}$ and the noise subspace is defined as $\text{span}\{U_N\}$. Note that selecting $k$ is a problem in model order estimation, and is related to the problem of selecting $M$ for ML.

**MUSIC**

The "Multiple Signal Classification" (MUSIC) estimator exploits the fact that the noise subspace is orthogonal to the system response vector associated with each incident signal [28]. If a candidate response vector $b(\phi)$ is orthogonal to the noise subspace, then it must correspond to one of the AOAs. MUSIC uses this fact to identify the AOAs as the peaks of the "pseudospectrum"

$$P_{\text{MUSIC}}(\phi) = \|b^H(\phi) U_n U_n^H b(\phi)\|^{-1}.$$  \hspace{1cm} (2.16)

This amounts to a one-dimensional search regardless of the true $M$, and therefore has a compelling advantage over ML if $M$ is known to be large. However, MUSIC is
not robust to multipath and in fact fails in the presence of widely-spaced, perfectly-correlated signals of nearly equal strength. In this case \( b(\phi) \) may not be orthogonal to the noise subspace, even if \( \phi \) is one of the actual AOAs.

There are two well-known strategies for overcoming this problem. One approach is spatial smoothing [29]. Spatial smoothing uses subarrays within a large array, combining the covariance matrices in a way that restores the dimension of the signal subspace. In this study, spatial smoothing is undesirable because one wants to minimize the number of elements in the array in order to reduce the cost. The second approach is Generalized MUSIC [30]. In this approach, one uses test vectors drawn from \( \text{span}\{B\} \), as opposed to \( B \), in place of \( b(\phi) \) in Equation 2.16. The modified estimator yields the desired peaks when the correct linear combination of system response vectors is used. This allows one to identify coherent sets of signals, but unfortunately increases the dimensionality of the search by two for each additional system response vector considered in the linear combination used to form the test vectors. Thus, Generalized MUSIC may not have an advantage over ML in terms of computational effort.

At least one previous field study [12] has indicated that naturally-occurring resolvable multipaths are unlikely to be perfectly coherent. In this case, the original MUSIC estimator yields peaks for the desired AOAs, but they are not infinitately large. In practice, this occurs even for perfectly uncorrelated AOAs since the system calibration is unlikely to be perfect. Thus, it may not be strictly necessary or even worthwhile to use spatial smoothing or Generalized MUSIC. This situation will be observed in the field measurements presented in Chapter 6.
ESPRIT

"Estimation of Signal Parameters by Rotational Invariance Techniques" (ESPRIT) is another subspace-based parameter estimation technique [31]. In ESPRIT, the tactic is to constrain the array geometry. For example, if the array is linear with uniform spacing, the matrix $A$ has a special structure which allows it to be represented by a diagonal matrix $\Upsilon$ whose elements are invertible functions of the spatial frequency associated with each of the AOAs. Also, a related matrix $\Psi$ is associated with $R$ and has the same eigenvalues as $\Upsilon$. Since the eigenvalues of a diagonal matrix are simply the diagonal elements, $\Upsilon$ is completely determined from an eigenanalysis of $\Psi$.

Thus, the ESPRIT algorithm can be summarized as follows: (1) Obtain an estimate of $R$ from the measurements, (2) Perform an eigendecomposition of $R$, (3) Compute $\Psi$ from the eigenvectors of $R$, (4) Find the eigenvalues of $\Psi$ (giving the diagonal elements of $\Upsilon$), and finally (5) Compute the spatial frequencies of the AOA estimates from the diagonal elements of $\Upsilon$. Note that Step 3 involves the solution of a linear system of equations that can be solved either in a least squares (LS) sense (traditional ESPRIT) or Total LS (TLS) sense (TLS-ESPRIT).

ESPRIT is popular because it does not require a search and because it has been shown to have good statistical performance. However, it only works with linear arrays and certain other geometries that have "shift invariant" structure; that is, one must be able to divide the sensors into a pair of displaced but otherwise identical subarrays. ESPRIT has the convenient feature that the individual subarrays do not need to be calibrated; it is only required that the displacement between the subarrays be fixed.
ESPRIT, like MUSIC, fails in the presence of correlated signals unless the rank of the signal subspace is first restored using spatial smoothing.

**Weighted Subspace Fitting**

A third class of subspace techniques is *subspace fitting*. An attractive technique from this class is the Weighted Subspace Fitting technique of Viberg *et al.* [32]. The WSF estimator is given as:

$$
\arg \min_{\hat{\Phi}} \text{Tr}\{P_{\mathcal{B}(\hat{\Phi})}^{-1} U_S W U_S^H \}
$$

where $W = (\Lambda_S - \hat{\sigma}^2 I)^2 \Lambda_S^{-1}$. Neglecting the weighting function $W$, this has the simple geometrical interpretation that one seeks to minimize the power resulting when the signal subspace (obtained from an eigendecomposition of an estimate of $R$) is projected into the orthogonal complement of $\text{span}\{\mathcal{B}(\hat{\Phi})\}$. The choice of $W$ ensures the lowest asymptotic variance of the resulting AOA estimates.

Note the similarity to the ML estimator, given by Equation 2.7. Like ML, WSF is robust to the presence of coherent multipath but requires knowledge of $M$. Whereas ML must be combined with MDL to obtain a reasonable estimator of model order, WSF gives rise to its own model order estimator that appears to have favorable statistical properties [32]. However, robustness to non-white noise does not appear to be one of these properties. Also, WSF requires an eigendecomposition of $R$ and a consistent estimate of $\sigma^2$, whereas ML does not. Estimating $\sigma^2$ in particular poses some difficulties in practice. For these reasons, WSF does not appear to have a compelling advantage over ML for this study.
2.3.3 Modulation-Specific Techniques

The analog FM signals employed in this study have the property of constant modulus (CM); that is, the magnitude of the complex baseband (zero-IF) version of the signal is constant. CM is a strong property that is often sufficient to separate uncorrelated signals received through an antenna array [33, 34]. Once the CM property is used to separate signals, the task of attributing system response vectors to AOAs is straightforward. Recent findings indicate that discrete AOA estimation algorithms based on the CM property, simply by virtue of exploiting this knowledge of the waveform, may be competitive with other techniques described above [35]. Unfortunately, this approach has the drawback that highly-correlated signals from different AOAs will be identified with a single response vector, so this approach is of limited use in this study.

Other modulation-based techniques exist, exploiting properties such as cyclostationarity, or other more specific features of the waveform. However, because the source signal is analog FM with voice input, it is expected that the carrier will be indistinguishable from an unmodulated carrier much of the time. Thus, there seems to be no advantage in this approach.

2.4 Distributed AOAs

The techniques described in Section 2.2 make no assumptions about the angle spectrum, but yield only low-resolution spectrum estimates. The techniques described in Section 2.3 assume the spectrum can be parameterized in terms of a small number of resolvable discrete AOAs. This has the advantage of yielding higher resolution, albeit at the cost of increased complexity and model sensitivity. Since the true AOA
spectrum must be smooth and continuous, it is important to consider the effect of applying a discrete AOA estimator in this case.

A few researchers have considered this problem, including Jäntti [36], Moses et.al. [37], and Astély and Ottersten [15]. Jäntti tested MUSIC and ESPRIT in the presence of rectangular and triangular angle spread distributions of less than 4°. He found that they gave estimates which were biased, but generally within the extent of the spreading. This seems to be confirmed by the work of Moses et.al., who developed a first order performance analysis of MUSIC, ESPRIT, and WSF in the presence of AOA distributions parameterized in terms of the mean AOA $\bar{\phi}$ and AOA variance $\sigma_{\phi}^2$. They found that the effect on the estimation variance was small, but that effect on the bias could be significant. For a uniform linear array, they found the magnitude of first-order bias term was $\frac{\sigma_{\phi}^2}{2} \tan \bar{\phi}$, where $\phi = 0$ corresponds to broadside. This suggests that the geometry of the array may have a significant effect in the presence of spread AOAs. However, this analysis is limited in that it assumes that the samples are independent discrete-AOA realizations drawn from a known distribution of $\phi$. In practice, however, it may be that all samples correspond to a single spread steering vector, and vary simply due to noise. Astély and Ottersten [15] take this approach assuming a uniform distribution and, in another first-order analysis, find that the estimates are in fact not biased. They further find the expected absolute error\(^4\) to be an approximately linear function of the angular distribution of the scatterers. This error is found to about $\frac{1}{\sqrt{12}}^\circ$ for each degree of angular extent.

\(^4\)I.e., the mean of the absolute values of the error; as opposed to the bias, which is simply the mean of the error.
Recognizing that the angle spread in the mobile radio environment can be perceived, it is reasonable to ask if the spreading can be included as additional parameter(s), thereby recovering some of the advantages of the model-based techniques. One possible approach is to replace the discrete AOA model with a general parameterization, such as an autoregressive moving average (ARMA) model [38]. However, this considerably increases the complexity and further aggravates the model order selection problem. A less ambitious scheme has been developed by Valaee et al. [39] which parameterizes each incident signal in terms of a AOA plus a spreading parameter, and modifies the MUSIC estimator to account for the effect of a specified angle distribution on the array manifold. AOA estimation then consists of a two-dimensional search for the peaks of the modified MUSIC cost function. Note that this is precisely the same principle as Generalized MUSIC. Trump and Ottersten [40] developed a joint AOA-angle spread estimator based on the ML principle assuming a single distributed AOA with Gaussian angle distribution. A concern with these approaches is that the angle spread distributions are arbitrary (i.e., not experimentally derived) and is therefore not necessarily representative of the true form of possible AOA distributions.

There have also been several attempts to derive angle spread parameters from the CBF spectrum. This includes Mogensen et al. [14], who estimated the angle spread by comparing measured CBF spectrum to the ideal (discrete AOA) CBF spectrum, assuming a single discrete AOA. Kalliola and Vainikainen [11] have described joint AOA-angle spread estimators based on conventional beamforming principles assuming uniform angle distribution. Again, the assumptions made in these approaches are quite bold and difficult to justify.
As a final observation, note that the simple model-verification method embodied in Equations 2.10 through 2.12 can be used to assess the quality of any angle spectrum estimate that can be expressed in terms of a small number of response vectors. For example, that method gives a rough indication of the validity of the discrete AOA model when only one signal is present: If the model holds exactly and the incident SNR is high, then the residual is approximately equal to (SNR)$^{-1}$. If, on the other hand, the AOA is not discrete but is rather distributed, the residual will increase toward a value of unity. If the discrete AOA model is completely invalid then the residual will be close to unity. Note that this latter condition is true even if the discrete AOA estimate is reasonable in some "center of mass" sense. The quality and validity of a distributed AOA estimate can thus be assessed in a simple way by using the appropriate response vectors in Equation 2.10.

2.5 Conclusions

This chapter has provided a brief survey of angle spectrum estimators applicable to this study. Of these estimators, the following are employed in the field study presented in Chapter 6. The discrete ML estimator is selected because it is optimal under the stated assumptions, and through its relationship to the CBF, is simple to analyze. The model-independent Capon estimator and the popular MUSIC estimator will be used for comparison. The use of more complex estimators, such as those which assume distributed AOAs, will not be used. This is due to uncertainty in how the AOA distribution should be parameterized. However, the method summarized by Equations 2.10 through 2.12 can be used to sense significant departures from the discrete AOA model, and possibly also errors in model order.
CHAPTER 3

ANTENNA ARRAY

This chapter addresses the problem of antenna array design for angle spectrum estimation, as described in the previous chapter. The objective is to obtain a reasonable antenna array design for this study. The first section provides some background on this problem and identifies some important requirements. Based on these requirements, a 7-element "Y" shaped array of monopoles mounted on a ground plane turned out to be a good choice. The next two sections provide an analysis of the azimuth- and elevation-plane aspects of this design. This is followed by an analysis of the electromagnetic coupling between the array elements, which is an important consideration for calibration, further discussed in Chapter 4. This chapter concludes with proposed enhancement for future arrays, such as the use of rolled edges to improve the elevation-plane performance.

3.1 Background

The primary consideration in the design of an azimuth-plane array for this study is the geometry of the array; in particular, the number and arrangement of the antenna elements. The issue of array geometry has received very little attention in the past. As evidence, recent papers summarizing the state of research in array processing for
spectral estimation [18], wireless communications [1], and spatial channel studies [2] make no mention of this issue from the performance aspect. A short investigation of this topic, as it applies to this study, is presented here.

Consider the following desired properties of an array geometry for this study. First, the array geometry should have uniform performance in all azimuth directions. Elevation coverage is a secondary consideration, but can become important depending on the design of the elements and the potential for scattering from below the array. It is desirable to keep the spacing between at least some of the elements at $\lambda/2$ or less, in order to avoid grating lobes and minimize sidelobe levels. It is also desirable to have as large a physical aperture as possible, in order to improve resolution. Given the dynamic nature of the mobile radio environment, it is not desirable to mechanically rotate or translate the array to achieve this aperture. From this perspective, an ideal DOA spectrum estimation instrument would be an antenna array with a physical aperture greater than 60 wavelengths ($\lambda$) in all directions, with Nyquist-spaced ($\lambda/2$) elements. With such a system, $1^\circ$ resolution could be obtained using conventional beamforming (CBF), which is robust and model-independent. A volumetric distribution of elements would be even better, giving the ability to discriminate in elevation as well as azimuth without front-back or up-down ambiguity. This may be an important feature to prevent experimental errors due to ground bounce and reflections from low-flying aircraft. Further, a polarimetric distribution of elements would be desirable, as polarization diversity could then be used to further improve angle spectrum estimation. With respect to this study, however, any of these aspects would make the system prohibitively expensive to construct. In fact, the total number of
antennas in the array must be aggressively minimized in order to bear the cost of the associated electronics.

Spectral estimation research has traditionally focussed on the uniformly-spaced linear array (ULA) and the uniformly-spaced circular array (UCA). The ULA is popular because it is simple to analyze and maximizes the physical aperture given a fixed number of elements, albeit for the broadside direction only. A well-known drawback of the ULA is poor AOA estimation performance near endfire. One explanation for this is that the effective physical aperture of a ULA goes to zero at endfire. For the same reason, the ability of the ULA to resolve closely-spaced sources also severely degrades near endfire. Further, it was pointed out in Section 2.4 that discrete AOA estimators may be biased when applied to a ULA in the presence of non-zero angle spread, and that the effect worsens towards endfire. These considerations rule out the ULA when consistent performance over a broad field-of-view is desirable.

When consistent performance over 360° is desired, the UCA is an obvious candidate. The UCA has constant usable aperture as a function of azimuth. However, the UCA requires about twice as many elements to achieve the same usable aperture as the ULA at broadside. Since it is desired to use a dedicated receiver for each sensor, the system cost is proportional to the number of elements. Thus, the UCA is a relatively expensive alternative in practice.

A related class of geometries which has received some attention in the past is thinned, or sparse, arrays. A sparse array is one in which element spacing is irregular, with some spacings allowed to be greater than \( \lambda/2 \). The main design challenge for such arrays is minimizing the ambiguity resulting from subsampling the aperture. Pillai et.al. [41] and Gavish and Weiss [42] provide techniques for characterizing the
ambiguity. Recently, Wong & Zoltowski [43] described a sparse grid (not linear) array for AOA estimation. However, their motivation was not to find a better geometry \textit{per se}, but rather to identify a very sparse array suitable for use with the ESPRIT algorithm, which requires identical subarrays. In general, sparse arrays are attractive from the perspective of the performance-cost tradeoff, since a reduced number of elements is required to achieve a specified aperture. Also, there is potential simplification in calibration if all elements are sufficiently far apart that $Z$ can be approximated as being proportional to $I$. Disadvantages of sparse arrays include increased sidelobe level and difficulty maintaining element positions.

Returning to compact arrays, two additional papers deserve note here. A 1991 paper by Hua \textit{et al.} [44] compares the Cramer-Rao Bound (CRB) for 2-D AOA estimation performance of several arrays, including an “L”-shaped array and a “+” (cross) -shaped array. Because the effective physical aperture of the “L” is always larger than the “+” the CRB is found to be significantly lower for the L array. A 1995 paper by Liang and Paulraj [45] addresses the relationship between array geometry and range, using basic signal-to-noise ratio arguments. This paper introduces a fairly wide variety of geometries, and seems to imply that there may be advantages to these alternative geometries. However, the concept is not developed beyond the application to coverage extension, which is irrelevant to the angle spectrum estimation problem.

The Y geometry, shown in Figure 3.1, is proposed for this study. It is not new, but neither has it received much attention. The Y is mentioned in [46] as a subarray concept; that is, as an option for selecting a small set of elements out of a much larger array for further processing. No justification or evaluation of this geometry is provided. The Y has also appeared in recent array signal processing papers by
Gönen and Mandel [47, 48]. In [47], it simply serves as a test geometry for a blind copy algorithm. In [48], the Y is formed by augmenting an existing ULA with two additional uncalibrated linear arrays, intended in this case as a test geometry for AOA estimation with a mix of calibrated and uncalibrated subarrays.

3.2 Azimuth Plane Analysis

In this section, let us consider the performance of the Y shaped array in the azimuth plane; i.e., in terms of its angle spectrum estimation performance. To demonstrate the advantages of the Y with respect to some other popular geometries, several
aspects of performance will be considered. This study considers only the array manifold \( \mathcal{A} \), under the assumption that \( \mathbf{C} \) and \( \mathbf{Z} \) can be calibrated out in the final system.

First, there is the important issue of cost. To compare array geometries fairly, it is desired to know the cost per “unit of performance”. Cost is easy to quantify; it is proportional to the number of elements \( N \). Performance can be quantified as proportional to the usable aperture \( Q \), that is, the physical aperture that can actually be applied to a specific AOA estimation problem. Note that \( Q \), in general, varies with \( \phi \). A reasonable cost metric is \( N/Q \); that is, the number of elements required per unit of useful aperture. For some popular array geometries and the Y, one finds that:

\[
N/Q = \frac{N}{N-1} |\cos \phi|^{-1}, \text{ for the ULA } (\phi = 0 \text{ for broadside}) \tag{3.1}
\]

\[
N/Q \approx \pi, \quad \text{for the UCA,} \tag{3.2}
\]

\[
N/Q = \frac{2N}{N-1} (|\cos \phi| + |\sin \phi|)^{-1}, \text{ for the L, and} \tag{3.3}
\]

\[
\frac{1.73N}{N-1} \leq N/Q \leq \frac{2N}{N-1}, \text{ for the Y.} \tag{3.4}
\]

Above, \( Q \) is figured in half-wavelength units. Also note that the expression for the UCA is a large \( N \) approximation; in fact, \( N/Q \) for the UCA is higher for finite \( N \). Finally, note that that \( N/Q \) for the Y has 60° symmetry, so the limits given above are achieved every 60°.

A plot of the above results is shown in Figure 3.2. Note that the Y has a lower overall normalized cost than the UCA and is superior to the ULA for \( |\phi| > 50° \) or so. The L is better over most of the field of view, but the Y has less variation.

Next, consider CBF performance. Array patterns for an 8-element ULA are shown in Figure 3.3. Note the front-to-back ambiguity. Also note that the beamwidth widens toward endfire (±90°), and the main beam becomes distorted towards endfire.
Figure 3.2: Normalized cost-per-aperture ratio \((N/Q)\). Solid: ULA \((N = 8)\), Dash: UCA, Dash-dot: \(L\) \((N = 7)\), Dot: \(Y\) \((N = 7)\).

At broadside, the main beam half-power beamwidth (HPBW) is 14°; at ±60° it has degraded to 29°. The same test is shown for the 7-element Y array in Figure 3.4. In this case, the HPBW = 30° and is nearly uniform for all directions. The beamwidth is greater than that of the ULA over much of the field of view, because there is one less element and two of the remaining elements do not contribute to widening the aperture. The peak sidelobe level is quite a bit higher, but is nevertheless a favorable tradeoff, since there is no front-back ambiguity and the shape of the main beam is independent of \(\phi\).

Next, the AOA estimation performance of various geometries including the Y will be investigated in terms of the most optimistic estimation variance, as predicted by the CRB. A suitable formulation is given by Equation 7.6 (with a few changes in
Figure 3.3: Gain patterns of the 8-element ULA for various main beam pointing directions. Broadside is 0°.

\[
\text{var}_{CR}(\hat{\phi}_i) = \frac{\sigma^2}{2L} \left[ \left\{ (D^H P_A^\perp D) \odot R_{ss}^T \right\}_{ii}^{-1} \right], \quad \text{where}
\]
\[
P_A^\perp = I - A(A^H A)^{-1} A^H, \quad \text{and}
\]

\( D \) is the matrix formed by taking \( d/d\phi \) of each column of \( A \). Note that \( \odot \) denotes the Hadamard matrix product, defined as \( [F \odot G]_{ij} = F_{ij} G_{ij} \). In this experiment, there are two equal-strength discrete-\( \text{AOA} \) signals which are 99\% correlated. One \( \text{AOA} \) is fixed, and the other is moved to generate each plot. In each case, the "large \( L \)" result is shown, but is normalized for \( \text{SNR} = 0 \) dB and \( L = 1 \) snapshot.\(^5\) The

\(^5\)Note that "\( L \)" (math font) refers to the number of snapshots, where as "\( L \)" (text font) refers to the "\( L \)"-shaped array geometry.
Figure 3.4: Gain patterns of the 7-element Y array for various main beam pointing directions.

The actual value of the RMS error can be obtained by multiplying the result shown by (SNR)^{-\frac{1}{2}} L^{-\frac{1}{2}}.

Figure 3.5 shows the CRB of the estimation variance for the fixed AOA, when the fixed AOA is at broadside with respect to the ULA, and 55° off broadside. As expected, the variance becomes large as the two AOAs approach each other. In the broadside case, the ULA is clearly superior, which is also expected. The Y and L geometries are about the same, and both are significantly better than the UCA. In the off-broadside case, however, the Y is superior to the UCA and L, and superior to the ULA after 65° or so. The Y is also not affected by the difficulties experienced by the ULA and L when the variable AOA is left of -45°.
Figure 3.5: CRB of the estimation variance for the fixed AOA as a function of the variable AOA. Fixed AOA is at (a) $0^\circ$ and (b) $55^\circ$. 
Figure 3.6 shows the CRB of the estimation variance for the variable AOA in the same test. The results are similar, except that the ULA performance is somewhat worse; in fact, the estimation variance for the ULA is unbounded at endfire.

Figure 3.7 shows the same data as Figure 3.6, except now only for the ULA, the Y, and a sparse version the Y with the spacing of the outer elements increased by a factor of 1.8. Note that the variance improves significantly for the sparse Y. Figure 3.8 shows the CRB of the estimation variance for the fixed AOA as a function of SNR-L, when the fixed AOA is at 0° and the variable AOA is at 3°. This gives an idea of the SNR or number of snapshots required to obtain a desired estimation variance when the AOAs are extremely close. Note that the aperture of the array has a far more significant impact than either the SNR or the number of snapshots in this case.

Based on the above study, it appears that the Y has a favorable balance between usable aperture, field of view, and cost. A sparse version of the Y was considered, and was shown to have even better performance. However, it was difficult to implement the sparse Y due to mechanical and space limitations. For these reasons, the Y array with half-wavelength spacing was selected and built for this study. Additional details of the design are described in the following sections.

### 3.3 Elevation Plane Design & Analysis

The previous section provided an analysis of the Y array in the azimuth plane. In this section, the elevation plane design and analysis is presented.

UHF-band mobile antennas are typically vertically polarized (e.g., for vehicle rooftop use), or are intended to be used in a manner that favors vertical polarization (e.g., hand-held radios). Therefore, if the base station is limited to one polarization,
Figure 3.6: CRB of the estimation variance for the variable AOA as a function of the variable AOA. Fixed AOA is at (a) 0° and (b) 55°.
Figure 3.7: CRB of the estimation variance for the variable AOA as a function of the variable AOA. ULA, Y, and sparse Y (Ye) only. Fixed AOA is at (a) 0° and (b) 55°.
Figure 3.8: CRB of the estimation variance for the fixed AOA as a function of SNR·L, when the fixed AOA is 0° and the variable AOA is 3°. “Ye” indicates the sparse (k = 1.8) Y.
it should be vertical. A natural choice for an array element in this case is the vertical
dipole. However, a concern in this project is that the electromagnetic scattering and
interference local to the array may corrupt the results. This was a particular concern
since the array intended for this study was to be mounted on the roof of a laboratory
building. Since a vertical dipole has considerable gain at low elevations, it is not a
good choice.

To mitigate the effects of local scattering and interference, the final design of the
antenna array consists of a Y-shaped array of monopoles on a ground plane, as shown
in Figure 3.9 and 3.10. The ground plane helps to suppress signals scattered from the
mounting structure or transmitted from locations which are too close to the array to
be interpreted correctly as azimuth-only AOAs. Each monopole is constructed from
brass rod 157 mm long (about one quarter wavelength at 465 MHz) and approximately
2 mm in diameter, and welded to the center conductor of panel mount style “N”
coaxial connector. The N connectors are mounted to a single section of aluminum
sheet which forms the central portion of the ground plane. The ground plane is then
extended to the dimensions indicated in Figure 3.9 using aluminum screen material
supported by a wooden frame.

While it is obvious that the ground plane should provide deep suppression of sig-
nals arriving from directly below the array, the ground plane also affects the elevation
pattern close to the horizon. As will be shown below, the fact that the ground plane
is truncated significantly reduces gain close to the horizon, where the desired sig-
nals originate. This unfortunately reduces the effective range of the system, but is
preferable to the possibility of misleading results due to spurious local signals.
Figure 3.9: Top view of array geometry. The circles indicate monopole positions and the outline of the ground plane is indicated. As built, the top edge of this figure is North.
Figure 3.10: Array as built. View is slightly to the North of East.
To fully understand the tradeoff involved in this design feature, an expression for the elevation plane pattern is derived below, followed by some pattern calculations. A few preliminaries: All structures are assumed to be perfect electrical conductors, and an $e^{j\omega t}$ time dependence is assumed and suppressed. Since it is only the elevation plane which is of interest, this analysis will be done in two dimensions.

Consider a dipole centered at the origin and oriented along the $y$ axis. The far-zone radiated field can be obtained from the three-dimensional result for a $z$-directed dipole, which is given by

$$H_{\phi,3D}^{hwd}(r, \theta) \approx \frac{jI_0}{2\pi} \left[ \frac{\cos \left( \frac{\pi}{2} \cos \theta \right)}{\sin \theta} \right] \frac{e^{-jkr}}{r}$$ (3.7)

where $I_0$ is the complex magnitude of the source current [50, p.131].

Since this has the form of a geometrical optics (GO) field, the two-dimensional ($z$-invariant) version of this field can be obtained by rotating the (3D) dipole into the $y$ axis and replacing the spherical (3D) spread factor with the appropriate cylindrical (2D) spread factor:

$$H_z^{hwd}(\rho, \phi) \approx \frac{jI_0}{2\pi} \left[ \frac{\cos \left( \frac{\pi}{2} \sin \phi \right)}{\cos \phi} \right] \frac{e^{-jk\rho}}{\sqrt{\rho}}$$ (3.8)

Let $H_i^z$ be the magnetic field radiated by a quarter-wave monopole over an infinite ground plane. By image theory, $H_i^z = H_z^{hwd}$ for $0 < \phi < \pi$, and is identically zero below the ground plane ($-\pi < \phi < 0$). If the $x > 0$ end of the ground plane is truncated as shown in Figure 3.11, the total field $H_z^t$ is given by $H_i^z + H_e^z$, where $H_e^z$ is the due to the diffraction of $H_i^z$ from the edge. A good high-frequency approximation for $H_e^z$ is the Uniform Geometrical Theory of Diffraction (UTD) which is given by

$$H_e^z(P) = H_i^z(Q_E)D_h \frac{e^{-jksd}}{\sqrt{s_d}}$$ (3.9)
where $P$ is the field point, $Q_E$ is the edge, $s^d$ is the distance from $Q_E$ to $P$, and $D_h$ is the diffraction coefficient for grazing incidence which is given by [51, Eq.52]

$$D_h = \frac{-e^{-j\pi/4}}{2\sqrt{2\pi k}} \left[ \frac{F[kL_\alpha(\psi)]}{\cos \frac{\psi}{2}} \right]$$  \hspace{1cm} (3.10)

where $\psi$ is the angle measured from the ground plane to the edge-diffracted ray. Note the additional factor of $\frac{1}{2}$ in the above equation, which is required due to grazing incidence in this case [51, p.1458]. $L$ is given by [51, Eq.32]$^6$

$$L = \frac{s^i s^d}{s^i + s^d}$$  \hspace{1cm} (3.11)

where $s^i = d/2$ is the distance from the base of the monopole to the edge, and (from [51, Eq.50])

$$a(\psi) = 2 \cos^2 \frac{\psi}{2}.$$  \hspace{1cm} (3.12)

$^6$Note $L$ is used in other chapters to represent the number of samples. Nevertheless, $L$ is used to represent the UTD distance parameter here in order to be consistent with applicable literature on this topic.
Figure 3.12: Power pattern of half-wave dipole in free space and a quarter-wave monopole on an infinite ground plane.

Also, $F(x)$ is the usual UTD transition function [51, Eq.26].

This completes the formulation of the elevation plane pattern. For results, let us begin with the case of an infinite ground plane as shown in Figure 3.12. The pattern of a half-wave dipole in free space is also computed for comparison. Note that the monopole pattern is zero below the ground plane, and therefore is 3 dB greater than the dipole pattern above the ground plane.

Figure 3.13 shows the result when the ground plane is terminated in a "knife edge" at various distances from the monopole. This configuration has the advantage that the close-in coverage (i.e., elevations less than $-10^\circ$ or so, depending on antenna height) is greatly attenuated with respect to the dipole. However, note that the monopole pattern is 3 dB below the dipole at the incident shadow boundary (in this
Figure 3.13: Pattern of a quarter-wave monopole mounted on a ground plane with an edge at the indicated distances from the monopole.

In this case, the horizon) regardless of the distance between the edge of the ground plane and the monopole. Thus, this configuration is not as good as a dipole in the primary coverage zone, corresponding to elevations between zero and $-10^\circ$ or so.

A remedy for this problem, recommended for future arrays using this design approach, is described in Section 3.5.

### 3.4 Electromagnetic Coupling Between Elements

This section presents a computation of the open-circuit impedance matrix $Z$ for the antenna array design described above. The mutual coupling information contained in $Z$ is a factor of the all-important system manifold $\mathcal{B}$, which describes the
<table>
<thead>
<tr>
<th></th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0.0000</td>
<td>-7.5023</td>
<td>-6.7710</td>
<td>-6.9131</td>
<td>-11.7435</td>
<td>-11.8313</td>
<td>-11.6182</td>
</tr>
<tr>
<td>4</td>
<td>-6.9131</td>
<td>-10.7743</td>
<td>-10.8473</td>
<td>-0.6121</td>
<td>-15.6405</td>
<td>-14.1261</td>
<td>-7.6907</td>
</tr>
</tbody>
</table>

Table 3.1: Magnitude (dB) of the open circuit impedance matrix, relative to the self-impedance of the center element.

A relationship between incident plane waves and captured data snapshots. Having detailed information about the nature of \( Z \) will be exploited in Chapter 4 to enable accurate array calibration.

A reasonable technique to calculate \( Z \) is the moment method [16]. The Electromagnetic Surface Patch code (ESP5) is a particularly convenient way to use the moment method [52]. In this study, the antenna array described above was modeled as an arrangement of perfectly-conducting plates and wires. ESP5 provides as output (among other things) its calculation of \( Z \). For the array described above, the result is given in Tables 3.1 and 3.2. In these figures, the elements are numbered as follows: The center element is 1. The inner and outer elements are 2–3–4 and 5–6–7, respectively. The three arms of the Y are defined by elements 1–2–5, 1–3–6, and 1–4–7.

Table 3.1 shows that the mutual coupling is related to the distance between elements, as expected. The strongest coupling exists between adjacent elements along each arm of the array, and is about −7 dB. Other coupling coefficients range between
<table>
<thead>
<tr>
<th></th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0.0000</td>
<td>-138.1843</td>
<td>-135.5026</td>
<td>-139.0885</td>
<td>50.2623</td>
<td>52.4188</td>
<td>49.6645</td>
</tr>
<tr>
<td>2</td>
<td>-138.1843</td>
<td>-1.8959</td>
<td>103.2372</td>
<td>100.6559</td>
<td>-134.8560</td>
<td>-56.9560</td>
<td>-58.8381</td>
</tr>
<tr>
<td>3</td>
<td>-135.5026</td>
<td>103.2372</td>
<td>-2.9665</td>
<td>97.7325</td>
<td>-65.1102</td>
<td>-141.1965</td>
<td>-73.4746</td>
</tr>
<tr>
<td>4</td>
<td>-139.0885</td>
<td>100.6559</td>
<td>97.7325</td>
<td>-2.0369</td>
<td>-72.2133</td>
<td>-71.5850</td>
<td>-139.2167</td>
</tr>
<tr>
<td>5</td>
<td>50.2623</td>
<td>-134.8560</td>
<td>-65.1102</td>
<td>-72.2133</td>
<td>-0.5956</td>
<td>154.2619</td>
<td>155.3350</td>
</tr>
<tr>
<td>6</td>
<td>52.4188</td>
<td>-56.9560</td>
<td>-141.1965</td>
<td>-71.5850</td>
<td>154.2619</td>
<td>-6.1152</td>
<td>141.9007</td>
</tr>
<tr>
<td>7</td>
<td>49.6645</td>
<td>-58.8381</td>
<td>-73.4746</td>
<td>-139.2167</td>
<td>155.3350</td>
<td>141.9007</td>
<td>-5.5676</td>
</tr>
</tbody>
</table>

Table 3.2: Phase (degrees) of the open circuit impedance matrix, relative to the self-impedance of the center element.

-11 dB and -17 dB. Because this coupling is significant, it will be important to account for it in the calibration of the array.

### 3.5 Improving Elevation Plane Performance Using Rolled Edges

In Section 3.3, it was shown that the untreated truncation of the ground plane results in a 3 dB loss of gain along the horizon. This section presents a study on the use of a convex rolled edge treatment to improve the gain close to the horizon while maintaining suppression of interference and scattering from below the array. It is shown in this section that this approach yields horizon gain which is actually about 2 dB greater than a dipole, without much degradation of the desirable close-in attenuation. The height of the proposed configuration is limited to the height of a monopole plus the vertical dimension of the rolled edge, which can be as small as one wavelength.

The following analysis is an extension of the analysis presented in Section 3.3. Consider the case where the edge of the ground plane is treated by attaching a rolled
edge with elliptical cross section as shown in Figure 3.14. The rolled edge is attached with its semi-major axis parallel to the ground plane, such that the surface normal is continuous across the junction. This greatly reduces the diffraction from the edge (now, a junction). However, there will still be a discontinuity in curvature across the junction, so diffraction from $Q_E$ may still be significant. Let $H^j_2(P_L)$ be this diffracted field contribution, where "$P_L$" is used to indicate a field point in the lit region; i.e., $0 < \phi < \pi$.

Furthermore, a "creeping wave" will cling to the surface after the incident field crosses the junction point. This energy will eventually be shed toward field points below the ground plane, and so may be a significant contribution to the total fields.
there. Let \( H_z^e(P_s) \) be this diffracted field contribution, where "\( P_s \)" is used to indicate a field point in the shadow region.

The theory described in [51] does not apply to convex surface scattering, as considered in this case. Pathak \textit{et. al.} developed a solution for convex surface scattering in [53]; however, this solution does not apply to surfaces which have a discontinuity in surface curvature. In [54], Pathak \textit{et. al.} presented a diffraction coefficient for a discontinuity in surface curvature; however, this solution does not account for the creeping wave. A method for calculating the total fields which is applicable to the scenario in Figure 3.14 was suggested by Heedy and Burnside [55]. Their method uses the latter diffraction coefficient with the former convex surface scattering solution along with a correction factor to ensure the continuity of fields across the shadow boundary. This heuristic approach shows good agreement with a hybrid UTD/moment method technique developed by Chuang and Burnside for the analysis of "aperture matched" horns [56, 57]. Since the geometry of that problem is quite similar to the geometry of the present problem, it seems reasonable to apply the method of Heedy and Burnside here. The solution for the diffracted field in the lit region is as follows [58, Eq. A.20 through A.27]:

\[
H_z^d(P_L) = H_z^s(Q_E) D_h^{lit} e^{-jks^d}; \quad \text{where} \quad (3.13)
\]

\[
D_h^{lit} = \frac{-e^{-j\pi/4}}{\sqrt{2\pi k}} \left[ \frac{\cos\left(\frac{\psi}{2}\right)F(X_1) + C_h(\xi)F(X_2)}{1 + \cos \psi} \right] \quad (3.14)
\]

\[
C_h(\xi) = \sqrt{\frac{2\cos\left(\frac{\psi}{2}\right)}{\rho_g(Q_E)}} \left[ \frac{1}{\pi X} F(X) + m(Q_E) \sqrt{\frac{2}{k}} q^*(\xi) \right] e^{-j\pi/4} e^{-j\xi^2/12} \quad (3.15)
\]

\[
\xi = -2m(Q_E) \cos \frac{\psi}{2} \quad (3.16)
\]

\[
m(Q_E) = \left[ \frac{1}{2} k \rho_g(Q_E) \right]^{1/3} \quad (3.17)
\]
\[ X = 2kL \cos^2 \frac{\psi}{2} \]  \hspace{1cm} (3.18)

\[ X_1 = \frac{kL(1 + \cos \psi)^2}{8 \cos^2 \frac{\psi}{2}} \]  \hspace{1cm} (3.19)

\[ X_2 = \frac{k \rho_y(Q_E)(1 + \cos \psi)^2}{4 \cos \frac{\psi}{2} \left[ 1 + \left( \frac{1}{s^2} + \frac{1}{s_t^2} \right) \frac{1}{2} \rho_y(Q_E) \cos \frac{\psi}{2} \right]} \]  \hspace{1cm} (3.20)

\( q^*(x) \) is a special function defined in [53], and \( \rho_y(Q) \) is the surface radius of curvature (on the rolled edge side, for \( Q = Q_E \)) in the plane of diffraction.

The solution for the diffracted field in the shadow region is [58, Eq. A.28 through A.34]:

\[ H_2^t(P_S) = H_2^t(Q_E) D_h^{sha} e^{-jks^d} \]  \hspace{1cm} (3.21)

where \( s^d \) is now measured from the surface tangent point \( Q_T \) to the shadow zone field point \( P_S \). Also, one finds that

\[ D_h^{sha} = CT_h \]  \hspace{1cm} (3.22)

where \( T_h \) is the diffraction coefficient describing the effect of the wave creeping from \( Q_E \) to \( Q_T \) [53] and is given by

\[ T_h = \sqrt{m(Q_E)m(Q_T)} e^{-jkx} \sqrt{\frac{2}{k}} e^{-\frac{j\pi}{4}} \left[ \frac{F(X^d)}{2 \xi^d \sqrt{\pi}} - q^*(\xi^d) \right] \]  \hspace{1cm} (3.23)

Note that \( C \) is a constant which enforces the continuity of fields across the shadow boundary:

\[ C = \left[ \begin{array}{c} H_2^t(P_L) \\ H_2^t(P_S) \end{array} \right]_{SB} = \left[ \begin{array}{c} H_2^t(P_L) + H_2^t(Q_E) D_h^{lit} e^{-jks^d} \\ H_2^t(Q_E) T_h e^{-jks^d} \end{array} \right] \]  \hspace{1cm} (3.24)

The remaining terms are given by

\[ m(Q_T) = \left[ \frac{k \rho_y(Q_T)}{2} \right]^{\frac{1}{3}} \]  \hspace{1cm} (3.25)

\[ X^d = \frac{kL^d(\xi^d)^2}{2m(Q_E)m(Q_T)} \]  \hspace{1cm} (3.26)
where $L^d$ is the same as $L$ in Equation 3.11 except for using the shadow-zone $s^d$ (i.e., the distance from $Q_T$ to $P_S$).

For an ellipse parameterized in $v$ with semi-major dimension $a_1$ and semi-minor dimension $a_2$ [59, p.135], one finds that

$$\rho_d(v) = \frac{(a_1^2 \cos^2 v + a_2^2 \sin^2 v)^{\frac{3}{2}}}{a_1 a_2}$$  \hspace{1cm} (3.27)

$$t = \int_{Q_E}^{Q_T} \sqrt{a_1^2 \cos^2 v + a_2^2 \sin^2 v} \, dv, \quad \text{and}$$  \hspace{1cm} (3.28)

$$\xi^d = \left(\frac{ka_1^2 a_2^2}{2}\right)^{\frac{1}{2}} \int_{Q_E}^{Q_T} \frac{dv}{\sqrt{a_1^2 \cos^2 v + a_2^2 \sin^2 v}}.$$  \hspace{1cm} (3.29)

Finally, it should be noted that this solution may not be valid for $|\rho_d(Q_E)| < 0.5\lambda$ or $d/2 < 1.5\lambda$; i.e., if the rolled edge is very small or the the distance between the monopole and the rolled edge-to-ground plane junction is small.

This formulation allows one to determine the elevation pattern of a vertical quarter-wave monopole on a ground plane with various edge treatments. Let us begin with circular terminations. Figure 3.15 shows the result for an edge $1.5\lambda$ from the monopole, terminated using cylinders of various radii. Note that the primary zone gain is now better than the dipole, regardless of the size of terminating cylinder used. The close-in attenuation is not as great as with the knife-edge, but is still mostly better than the dipole. Figure 3.16 shows that this performance is not sensitive to the size of the ground plane.

A disadvantage of circular rolled edges is that they become very large with respect to the original monopole-plus-ground-plane configuration. An elliptical rolled edge offers performance which is nearly as good, but is much more compact. The performance of various elliptical rolled edges is shown in Figure 3.17. The elliptical termination with parameters $a_1 = a_2 = 5.00\lambda$ is identical to the circular rolled edge.
Figure 3.15: Pattern of a quarter-wave monopole 1.5\(\lambda\) from the junction between a ground plane and circular rolled edges of various radii.

of radius 5\(\lambda\) considered above. The elliptical termination with parameters \(a_1 = 2.32\lambda\) and \(a_2 = 1.08\lambda\) is designed to have a radii of curvature of 5\(\lambda\) at the junction point (same as the 5\(\lambda\) circular rolled edge) and 0.5\(\lambda\) at the extreme end. Note that this rolled edge has performance which is nearly identical to the 5\(\lambda\) circular rolled edge for elevations within 40\(^\circ\) of the horizon.

To show the effect of further reduction in the size of the elliptical rolled edge, Figure 3.17 also shows the result for an elliptical termination with parameters \(a_1 = 1.26\lambda\) and \(a_2 = 0.79\lambda\). This rolled edge has a radius of curvature of 2\(\lambda\) at the junction point. Note that the degradation follows the same trend observed for circular rolled edge. These results indicate that the performance of an elliptical rolled edge is strongly influenced by the radius of curvature at the point of attachment.

52
Figure 3.16: Same as Figure 3.15, except now the rolled edge radius is fixed at $5\lambda$ and the distance from the monopole to the junction point is varied.

Table 3.3 provides a summary of the results from this section, illustrating the tradeoff between rolled edge size versus gain in the primary coverage zone ($-5^\circ$ taken as an example) and close-in gain (more negative elevations). Recall that $a_1$ contributes to the radial size of the structure, while $a_2$ contributes to the height of the structure. It can be seen from this table that the elliptical rolled edges provide a very favorable tradeoff compared to the other configurations.

In this section, a monopole array on a ground plane with rolled edges was proposed as an alternative to the “knife-edge” termination of the original ground plane. Knife-edge, circular, and elliptical rolled edge terminations were compared. It was found that elliptical rolled edges provided a good tradeoff between primary coverage zone gain, close-in attenuation, and overall size. Using an elliptical rolled edge with a
Figure 3.17: Pattern of a quarter-wave monopole mounted on a ground plane of radius 1.5\lambda, using various elliptical terminations.

<table>
<thead>
<tr>
<th>Termination</th>
<th>$a_1$</th>
<th>$a_2$</th>
<th>Pattern at: $-5^\circ$</th>
<th>$-20^\circ$</th>
<th>$-40^\circ$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Knife-edge</td>
<td>0.00\lambda</td>
<td>0.00\lambda</td>
<td>-4.5 dB</td>
<td>-8.3 dB</td>
<td>-12.3 dB</td>
</tr>
<tr>
<td>Circular</td>
<td>0.50\lambda</td>
<td>0.50\lambda</td>
<td>+1.0 dB</td>
<td>-0.9 dB</td>
<td>-4.6 dB</td>
</tr>
<tr>
<td>Circular</td>
<td>5.00\lambda</td>
<td>5.00\lambda</td>
<td>+1.6 dB</td>
<td>-1.1 dB</td>
<td>-9.3 dB</td>
</tr>
<tr>
<td>Elliptical</td>
<td>2.32\lambda</td>
<td>1.08\lambda</td>
<td>+1.6 dB</td>
<td>-1.4 dB</td>
<td>-8.2 dB</td>
</tr>
<tr>
<td>Elliptical</td>
<td>1.26\lambda</td>
<td>0.79\lambda</td>
<td>+1.6 dB</td>
<td>-0.7 dB</td>
<td>-6.2 dB</td>
</tr>
<tr>
<td>Dipole</td>
<td>-</td>
<td>-</td>
<td>-0.0 dB</td>
<td>-0.8 dB</td>
<td>-3.2 dB</td>
</tr>
</tbody>
</table>

Table 3.3: Summary of results for the monopole 1.5\lambda from the edge (or attachment point).
vertical dimension of about a wavelength (depending on where the rolled edge itself is terminated) and slightly longer in the horizontal dimension, one obtains horizon gain which is significantly better than a half-wave dipole with close-in attenuation that is 3-6 dB better.

3.6 Conclusions

In this Chapter, various antenna array options were considered, and a 7-element "Y"-shaped array of monopoles on a ground plane was selected. This configuration seems to yield the most favorable combination of cost and performance for an array with a 360° field of view. However, it was shown that the abrupt truncation of the ground plane results in a 3 dB loss of gain along the horizon, with respect to a dipole in free space. A method to fix this problem, which in fact improves the horizon gain by about 2 dB with respect to a dipole, is to use rolled edge termination. Although not implemented in this project, results in following chapters will show the value of including this feature in future work.
CHAPTER 4

CALIBRATION

In the previous chapters, techniques for angle spectrum estimation have been described and the issues associated with antenna array design have been identified. An assumption so far has been that the system manifold $B$ is either known or depends only on the array manifold $A$, which in turn depends only on the array geometry and element patterns. In practice, the problem is not so simple. It was shown in Section 3.4 that the coupling between the antenna elements (represented by the off-diagonal elements of $Z$) was weak, but significant. It will be further demonstrated in this chapter that even this weak coupling must be taken into account for reasonable AOA estimation performance. Also, there will be differences among the cabling and receivers used to process the signal from each element in a real system, making $C$ important. Finally, it is possible that there will be errors in one's knowledge of the element positions or patterns. All of these effects have the potential to vary in both time and frequency, and taken together may be sufficient to render the system unusable. The process of compensating for these effects, so that the theory from previous chapters can be applied, is calibration.

This chapter is organized as follows. First, some known approaches to the problem are presented. An argument is made for the approach of computing the factors $C$,
\( \mathbf{Z} \), and \( \mathcal{A} \) independently. Specifically, the selected approach is to assume \( \mathcal{A} \) is known, measure \( \mathbf{C} \) on-the-fly at the moment of data acquisition, and estimate \( \mathbf{Z} \) off-line using disjoint measurements for known AOAs. In order to determine calibration accuracy requirements, a brief study of the effects of the various errors parameterized in the model is presented.

\section{4.1 Background}

A simple and direct method to estimate the system manifold \( \mathcal{B} \) is to measure a set of response vectors \( \mathbf{b}_m \) for plane waves incident from various AOAs within the field of view. In this case, the response vectors are estimated as the mean over \( L \) samples

\[
\mathbf{b}_m = \frac{1}{L} \sum_{l=1}^{L} \mathbf{x}(t_l, \phi_m),
\]

where \( t_l \) is the time of the \( l^{th} \) sample, and \( \phi_m \) is the AOA. It is assumed that the noise \( \mathbf{n}(t) \) has zero mean (as discussed in Chapter 2) and that the situation is stationary over the time that the samples are taken.

One version of this approach is simply to sample at the highest resolution of interest over the entire field of view. Then, the values of \( \mathbf{b}_m \) represent all the necessary samples from \( \mathcal{B} \). However, if the system response varies slowly with angle, a more practical alternative is to interpolate between widely-spaced values of \( \mathbf{b}_m \) to estimate the intermediate values. Even in this case, this method of calibration may require a very large number of measurements to achieve reasonable resolution, and a source must be available for every AOA resolution increment.

Alternatively, one can rotate the array in the presence of a smaller number (as few as one) source. This option might be reasonable for estimating the product \( \mathbf{Z} \mathcal{A} \) for small, rigid arrays, which could be measured once after manufacture and assumed
constant thereafter. The calibration problem then reduces to that of finding the angle-invariant receiver calibration matrix $C$, which can be done as a separate "on-line" measurement using techniques described later in this chapter. A limitation of this approach is that the environment in which the system is tested may be different from the environment in which the system is used. For example, if the system is tested in a indoor antenna range, the effect of local scattering in the intended operating site is absent and errors due to limitations of the test range may be introduced.

It is most desirable to have an approach suitable for large, non-rigid arrays, which allows the calibration to be done in situ, and uses a small number of calibration sources. Pierre and Keveh [60] described a simple method for estimating $CZ$ in this case, using as few as $N$ far-field sources. Their approach is as follows. Let the results of the source measurements be summarized by the $N \times M$ matrix $B_M = [b_1, b_2, \ldots b_M]$, $M \geq N$. Similarly, let $A_M = [a_1, a_2, \ldots a_M]$ where $a_m = a(\phi_m)$. Then:

$$B_M = (CZ)A_M,$$  \hspace{1cm} (4.2)

which in general is an over-determined set of linear simultaneous equations. The least squares solution is given by

$$CZ = B_M A_M^H (A_MA_M^H)^{-1}.$$  \hspace{1cm} (4.3)

This is applicable under generally mild assumptions, the most important being that the $a_m$ are linearly independent. This implies that each AOA corresponds to a unique vector in span{$A$}, and the measurements are not so close together in angle that the quantity $(A_MA_M^H)^{-1}$ approaches singularity, becoming numerically ill-conditioned. It is further understood that the $M$ measurements are temporally disjoint, such that
they do not interfere with each other. Alternatively, the $M$ measurements may be simultaneous if they are distributed in a small frequency range (i.e., spectrally disjoint) over which the system response does not vary significantly.

A weakness in this technique is that it does not take into account the effect of errors in one's knowledge of the antenna element locations. Position errors are problematic because the element positions are used to calculate the columns of $A_M$. However, an ML solution to the problem of jointly estimating $\mathbf{CZ}$ and the true element positions in this case was developed by Ng and See [61]. They formulate the problem as:

$$\arg \min_{\mathbf{CZ}, \psi} \left| (\mathbf{CZ}) \mathbf{A}_M(\psi) - \mathbf{B}_M \right|^2_F,$$

in which one seeks the set of element positions $\psi$ and $\mathbf{CZ}$ which minimizes the error between trial values $\mathbf{CZ} \mathbf{A}_M(\psi)$ and measurements $\mathbf{B}_M$. This problem is linear in the elements of the product $\mathbf{CZ}$, but is nonlinear in $\psi$, and thus cannot be solved in closed form. A numerical solution is found to be

$$\hat{\psi} = \arg \min_{\psi} \operatorname{Tr} \left\{ \mathbf{P}_{\mathbf{A}_M(\psi)}^\perp \mathbf{B}_M^H \mathbf{B}_M \right\}.$$

This requires a multidimensional search over the coordinates of the sensor positions; i.e., a $2N$-dimensional search for an $N$-element planar array. The resulting ML estimate of $\psi$ is then substituted back into Equation 4.3, where it is used to obtain the correct value of $\mathbf{A}_M$, yielding the ML estimate of $\mathbf{CZ}$.

For this study, the array presented in Chapter 3 is rigid and thus the element positions are known with high accuracy all the time. Thus, the computational complexity of the search over $\psi$ is probably not warranted. Also, there is the issue that measurements at $M \geq N$ separate AOAs are required, and the corresponding $\phi_M$
must be known \textit{a priori}, which presents numerous practical difficulties in a mobile radio environment.

A different approach to this problem has been suggested by Weiss & Friedlander [62]. This method assumes that the element positions are known. The procedure consists of estimating $C$ and $Z$ separately by an iterative technique using a small number of resolvable sources, which need not be temporally or spectrally disjoint. Their procedure begins by assigning initial values to $C$ and $Z$ based on the best available information. The first step in the iteration is to use MUSIC to estimate the AOAs. Since $C$ and $Z$ are in error, the AOAs are likely to be badly biased. However, as long each AOA generates a resolved peak in the MUSIC pseudo-spectrum, the procedure can continue. The second step in the iteration is to compute an improved value of $C$, which is a linear problem in the available parameters and so does not require a search. The final step in the iteration is to compute an improved value of $Z$. The difficulty of this step depends on the array geometry. For perfect ULAs and UCAs, the form of the mutual coupling is known and $Z$ can be characterized in one parameter, pertaining to the element spacing. Otherwise, more parameters may be required. The quality of the estimates of $C$ and $Z$ obtained from this three-step process improves on each iteration. To know when to stop, one observes the values of MUSIC pseudo-spectrum at the estimated AOAs. When the sum of these values is maximized, there is no advantage in additional iterations.

This procedure was extended by Solomon \textit{et.al.} [63]. Their approach increases the number of parameters that can be estimated by exploiting the temporal or spectral disjointness of the signals, when it exists. They also incorporate other prior work by
Weiss and Friedlander [64] to allow joint estimation of position errors as an additional step in the above iterative procedure.

These procedures have the advantage that the AOAs do not need to be known a priori. This is useful in the mobile radio environment because it facilitates the use of existing signals of opportunity (e.g., mobile radio users) as calibration sources. Furthermore, the number of sources $M$ can in principle be less than $N$ if $Z$ can be described in sufficiently few parameters. However, this is only practical if the mutual coupling can be parameterized in a useful way. As stated above, this is true for perfect ULAs and UCAs. For arbitrary arrays, arrays in the presence of complex structures, and ULAs and UCAs with otherwise unknown mutual coupling, $Z$ cannot be simply parameterized. Another problem with the techniques in [62] and [63] is that $C$ and $Z$ must be stationary over the observation time, which may be a problem if multiple measurements are required. The problem here is likely to be the receiver calibration, $C$. Finally, because the technique uses MUSIC, the calibration quality is degraded in the presence of strong multipath and in the presence of model errors such as angle spread. In fact, the failure of real world propagation to conform to the discrete AOA model causes problems for the ML estimator of Ng and See (described above) as well.

The key idea in the above strategy was to estimate the factors $C$, $Z$, and $A$ of the system manifold $B$ separately. Consider a different approach using the same idea of estimating the factors separately. As discussed above, $A$ for the array designed in Chapter 3 can be assumed to be known exactly, since it is rigid and the element positions can be measured with high accuracy. $C$ can be estimated independently of $A$ and $Z$ by injecting calibration signals at the antenna terminals. Such a procedure does not depend on external sources, and so stationarity need not be a concern. Once
C and \( A \) are known, the calibration problem is reduced to finding \( Z \). \( Z \) is apparently the least straightforward to estimate, but has the advantage of being time-invariant if \( A \) is time-invariant. Further, \( Z \) does not need to be estimated jointly with other calibration-sensitive parameters. The following sections present suitable methods for estimating \( C \) and \( Z \).

4.2 Estimating the Receiver Gains and Phases

A straightforward method to estimate the receiver gains and phases (diagonal elements of \( C \)) is illustrated in Figure 4.1. At the output of each antenna element is a directional coupler. A directional coupler has the properties that (1) received signals pass through it with only a small insertion loss (on the order of 1 dB), (2) signals applied to the coupled port are added to the received signals from the antenna, albeit with more loss (e.g. 10 dB), and (3) signals traveling from the coupled port to the input port are suppressed (e.g., 40 dB attenuation). A signal from a single source is split \( N \) ways through a power divider, and applied to each coupler. If there are no other signals in the passband, the elements of \( C \) are given by the cross-correlations of the receiver outputs \( x_n(t) \). In the case where the calibrating signal is CW (zero bandwidth), the problem reduces to that of estimating the relative magnitudes and phases between the \( x_n(t) \) and some other reference. This is a straightforward problem in sinusoidal parameter estimation. Since the measurements are all relative, \([C]_{11}\) is set equal to unity by convention in this study, and all other elements of \( C \) are computed with respect to \([C]_{11}\).

In practice, there are a few difficulties with this approach. First, it is assumed that the complex gains from the power divider through each coupler are identical.
This is unlikely to be true. Depending on the design of the receiver, it is also possible that there may be other signals in the passband. (This is true for the design described in Chapter 5.) Assuming all signals in the passband are spectrally disjoint, the latter issue poses only a slightly more complex problem in sinusoidal parameter estimation. The former issue can be managed by "precalibrating" the portion of the calibration signal path which is not part of the desired signal path. Once the complex gains associated with those paths are known, the calculated calibration coefficients can be corrected using those gains. Techniques for solving the parameter estimation problem and performing precalibration are presented in Sections 4.2.1 and 4.2.2 below.

4.2.1 Sinusoidal Parameter Estimation

The calibration procedure described above requires estimates of magnitude and phase for a sinusoid in the presence of noise and perhaps also interference. This section describes suitable techniques. Let us begin with by considering a scenario in
which there is a single complex sinusoid in noise:

\[ y(t) = Se^{j(\omega t + \theta)} + n(t). \]  

(4.6)

Where \( S \) is a real-valued positive scalar, \( \omega \) is the frequency, \( \theta \) is the phase, and \( n(t) \) is spectrally white, ergodic, complex, circular Gaussian noise with zero mean and variance \( \sigma^2 \). In this case, the optimal estimate of the parameters given \( L \) samples of \( y(t) \) is known to be [65]:

\[ \hat{\omega} = \arg \max_{\omega} |Y(\omega)|^2 \]  

(4.7)

\[ \hat{S} = \left| \frac{1}{L} Y(\hat{\omega}) \right|, \text{ and} \]  

(4.8)

\[ \hat{\theta} = \arctan \left[ \frac{\text{Im}\{Y(\hat{\omega})\}}{\text{Re}\{Y(\hat{\omega})\}} \right]; \text{ where} \]  

(4.9)

\[ Y(\hat{\omega}) = \sum_{t=1}^{L} y(t_l)e^{-j\hat{\omega}t_l}. \]  

(4.10)

One interpretation of this estimator is as follows. If \( \omega = 0 \), then \( y(t) = Se^{j\theta} + n(t); \) i.e., a complex constant in noise. The quantity \( Se^{j\theta} \) can then be estimated simply as the mean value of the \( L \) samples of \( y(t) \). If \( \omega <> 0 \), one can transform the problem to the simpler \( \omega = 0 \) case simply by downconverting or upconverting as necessary (as indicated in Equation 4.10) and then taking the mean. This works because \( S \) and \( \theta \) are unchanged in the spectral shift. The resulting estimates are unbiased, with variance lower-bounded by the CRB as follows [65]:

\[ \text{var}(\hat{\omega}) \geq \frac{6}{(2\pi)^2 (\text{SNR})L(L^2 - 1)} \]  

(4.11)

\[ \text{var}(\hat{A}) \geq \frac{\sigma^2}{2L}, \text{ and} \]  

(4.12)

\[ \text{var}(\hat{\theta}) \geq \frac{2L - 1}{(\text{SNR})L(L + 1)} \]  

(4.13)
At this point, let us identify two issues which may emerge in practical use. First, it is assumed that the noise in the \( L \) samples of \( y(t) \) is uncorrelated. However, the noise component of a receiver output is typically white (rather, approximately so) only within the analog passband of the receiver. The noise within the Nyquist bandwidth of the receiver, but outside the analog passband, is colored by the passband response. If the Nyquist bandwidth is significantly larger than the passband, then the noise will appear to be correlated. In fact, in the limit as the relative bandwidth of the passband shrinks to zero, the noise appears sinusoidal and thus is perfectly correlated. The danger in this phenomenon is that the parameter estimates are not statistically consistent in this case; i.e., the estimation variance is not guaranteed to improve as \( L \) is increased.

A second practical issue is that additional signals may be present in the data. This may be either by design (e.g., desired source and calibration signals), interference coming through the antenna, or perhaps spurious signals generated within the receiver. In this case, the model in Equation 4.6 is no longer valid, and estimation performance is degraded. For example, if the extra signals are included as part of \( n(t) \), the noise is no longer white. Even if the additional signals are much weaker and "noise-like", their presence again destroys the statistical consistency of the estimator. If the additional signals can also be modeled as complex sinusoids, the appropriate revision to Equation 4.6 is

\[
y(t) = \sum_{m=1}^{M} S_m e^{i(\omega_m t + \theta_m)} + n(t) .
\]  

(4.14)

As before, a reasonable strategy is to estimate frequencies first. Note that there is a dual relationship between frequency and angle (actually, "spatial frequency"). Thus,
any of the discrete AOA estimators discussed in Section 2.3 can also be used to estimate the \( M \) frequencies in Equation 4.14, with the same caveats. Since \( L \) is typically large, it is practical to substitute the Fast Fourier Transform (FFT) for the CBF to obtain high-resolution frequency estimates with low computational complexity. Given the frequencies, the joint least squares solution for the magnitudes and phases is [65]

\[
z = (W^H W)^{-1} W^H y
\]  

(4.15)

where \( y \) is the \( L \times 1 \) sample vector \([y(t_1) \ldots y(t_L)]^T\), \( W \) is an \( L \times M \) matrix whose columns are the sampled sinusoids \([\exp(j\omega_1 t_1) \ldots \exp(j\omega_1 t_L)]^T\) through \([\exp(j\omega_M t_1) \ldots \exp(j\omega_M t_L)]^T\), and \( z \) is the desired \( M \times 1 \) vector of complex coefficients; i.e., \( z_m = S_m e^{j\theta_m} \).

In extreme conditions, \( M \) may be very large and some signals present may not be well-modeled as sinusoids. This is likely to occur if the signals themselves have significant bandwidth, and the ratio of \( L \) to the sample rate is low. In this case, Equation 4.15 provides degraded estimates of the desired sinusoidal parameters in Equation 4.14. Such conditions are described by Miller et.al [66] pertaining to an ultra-wideband radar system. Their approach is to decompose \( y(t) \) through a multistage process consisting of estimation, synthesis, and subtraction for each class of signals. The widest bandwidth signals are removed first, followed by FM signals and finally by signals which can be modeled simply as complex sinusoids. This strategy is attractive for this study as well, but same changes are required. First, it is desirable to achieve a guaranteed minimum level of suppression for any signal in the passband. This is difficult using the direct parametric approach of [66], because the sample rate-to-bandwidth ratio for this study is not as high as in ultra wideband radar. To compensate, one can use any fast, suboptimal technique that does not
degrade the desired signals. To ensure a minimum level of suppression, the procedure can be repeated as necessary. Then, the remaining CW signals (calibration signals or unmodulated mobile user signals) can be estimated jointly using the LS approach described above.

For the first stage of the algorithm, the following simple algorithm is proposed:

1. Make a copy of the input data and perform the following steps on the copy.

2. Perform a length-$N_{FFT}$ FFT of the data and identify the $N_P$ largest peaks. Sort them from strongest to weakest.

3. Starting with the strongest signal identified, perform the single-tone parameter estimation defined by Equations 4.7–4.9. Synthesize the complex sinusoid with those parameters and subtract from the input signal copy. Repeat for all $N_P$ frequencies. Remember the frequency, magnitude, and phase of each signal.

4. Spectrally shift the input signal copy up or down by one-half of an FFT bin.

5. Repeat Steps 2 and 3.

6. Apply the opposite spectral shift to restore the original frequency alignment.

7. Repeat Steps 2 through 6 until the power in all bins is below some threshold, $P_{\text{thresh}}$.

If $P_{\text{thresh}}$ is chosen to be close to the noise power per bin, the above procedure has the effect of decomposing the input data into the sum of many sinusoids which collectively describe all the non-noiseful signals in the passband. $N_{FFT}$ and $N_P$ are selected based on the characteristics of the data to be input to the algorithm. The
purpose of the half-bin spectral shift is to ensure that signals straddling two or more bins are suppressed as aggressively as signals which are contained primarily within one bin.

Note that a single-tone estimator is used in Step 3, despite the presence of many other tones. Yet, this method has been found to be surprisingly effective. This result can be explained as follows. When the estimator operates on a sinusoid which is not at \( \omega = 0 \), the periodogram consists of two parts. The first part is the integration over an integral number of periods, whereas the second part is the integration over the remaining fractional periods at the beginning and the end of the data record. Integration over a single period results in a zero value, so integration over an integer number of periods is also zero. If the sinusoid’s frequency is sufficiently different from \( \omega = 0 \), the remaining fractional period is very small compared to the length of the total record. Thus, the contribution of the fractional part, even in the worse case (half a period) is also very small. Furthermore, any signal whose bandwidth is small relative to the Nyquist bandwidth will appear to be “approximately sinusoidal” and will exhibit similar behavior. The conclusion is that for highly-oversampled data with interferers which are not too close to \( \omega = 0 \), the simple estimators in Equations 4.7–4.9 are quite robust even in the presence of other signals. Note that the above algorithm fully exploits this behavior by estimating and subtracting the signals from strongest to weakest, so that the parameter estimation can never be biased by a stronger signal.

For best performance, a window should be applied to the data before the FFT. The purpose of the window is reduce the endpoint contributions and also to reduce the likelihood that signals can be hidden underneath the sidelobes of nearby stronger signals. The latter is a problem only because one FFT is used per \( N_P \) signal removals;
and it is not required if \( N_P = 1 \). The window should be selected based on the type of signals expected. If all signals are expected to be very close to sinusoidal, a triangular window is a reasonable choice so as to maintain resolution for accurate frequency estimation. On the other hand, it may be that many signals occupy several adjacent bins. In this case, the issue of frequency resolution becomes moot and a more aggressive window is appropriate.

The second stage of the proposed algorithm is simply to synthesize and subtract all sinusoids identified using the above procedure whose frequencies are outside the bandwidth of the signal(s) of interest. The result is unaltered spectrum for the channels of interest, plus noise only in the remainder of the passband. The threshold criterion of Step 7 ensures that no undesired signal can remain which is greater than \( 10 \log_{10} N_{\text{FFT}} \text{ dB/bin} \) below the total noise power. At this point, proper single-tone or joint LS parameter estimation of calibration signals can be performed as described earlier.

4.2.2 Precalibration

Recall that the receiver calibration scheme illustrated in Figure 4.1 requires either that the complex gains from the power divider through each coupler are identical, or that they are known. As the former is unlikely to be true, this section addresses the problem of finding the complex gains.

Let \( D_n \) be the complex gain for the calibration path associated with antenna \( n \). Note that it is only the relative values of the complex gains which are important for angle spectrum estimation, and so \( D_1 \) is set to unity and all other values are computed with respect to \( D_1 \). Also, note that any measurement of \( D_n \) using the
receivers will also include a factor of $C_n$, namely, the complex gain associated with feedlines and receivers. This is a problem because the purpose of the calibration system is to measure these coefficients (C); therefore they are unknown at this point. This problem can be bypassed by simultaneously observing two sinusoids: one injected through the calibration system, and a second at a slightly different frequency received through the array. By making each complex gain measurement relative to the second sinusoid, each coefficient $C_n$ should cancel. The second signal is radiated by an additional antenna local to the base station, and doubles as a “system monitor”; that is, a signal with known characteristics which is always present and hence can be used to diagnose system problems.

Let $g_n$ be the complex magnitude of the sinusoid injected through the calibration system and measured at receiver $n$. Similarly, let $h_n$ be the complex magnitude of the sinusoid received through the antenna array and simultaneously measured at receiver $n$. Then one finds:

$$g_n = C_n D_n g_0 \quad \text{and} \quad (4.16)$$

$$h_n = C_n H_n h_0 \quad , \quad (4.17)$$

where $g_0$ is the complex magnitude of the sinusoid at the input of the power divider, $h_0$ is the complex magnitude of the sinusoid at the system monitor antenna, and $H_n$ is the complex gain for the path from the system monitor antenna to the terminals of element $n$. It is assumed that the relative response (main path relative to coupled path) is the same for each coupler. Next we can define:

$$v_n = \frac{g_n}{h_n} = \frac{D_n g_0}{H_n h_0} \quad \text{and} \quad (4.18)$$

$$\tilde{v}_n = \frac{v_n}{v_1} = \frac{D_n H_1}{D_1 H_n} . \quad (4.19)$$
Now the $C_n$'s have been eliminated from the calculation, but unfortunately a factor of $H_1/H_n$ remains. One can deal with this factor as follows. Let us define $\tilde{v}_n^{(1)}$ as equal to $\tilde{v}_n$ as given above. Then let us define $\tilde{v}_n^{(p)}$ as the exact same measurement for a modified configuration in which the calibration signal inputs for antennas 1 and $p$ are swapped. Now:

$$\tilde{v}_n^{(p)} = \frac{v_n^{(p)}}{v_1^{(p)}} = \frac{D_1 H_1}{D_p H_n},$$

where $i = p$ if $n = 1$, $i = 1$ if $n = p$, and $i = n$ otherwise. Thus we find that

$$\frac{D_n}{D_1} = \frac{\tilde{v}_n^{(p)}}{\tilde{v}_1^{(p)}}$$

which yields $N - 2$ useful estimates of $D_n/D_1$ corresponding to values of $p$ between 2 and $N$, excluding $p = n$.

The precalibration procedure is summarized as follows:

1. Estimate $g_n^{(p)}$ and $h_n^{(p)}$ for each configuration $p$ using the sinusoidal parameter estimation technique from the previous section. (Recall that "configuration $p$" refers to a swapping of the calibration injection cables between the couplers associated with antennas 1 and $p$.)

2. Compute $\tilde{v}_n^{(p)} = [g_n^{(p)} h_1^{(p)}] / [g_1^{(p)} h_n^{(p)}]$ for $n = 2, 3, \ldots N$ and $p = 1, 2, \ldots N$ excluding $p = n$.

3. Using the convention that $D_1 = 1$, compute the $N - 2$ estimates of $D_n$ given by $\tilde{v}_n^{(p)}/\tilde{v}_1^{(p)}$.

4. Average the estimates of $D_n$ for each $n$, excluding outliers if desired.
4.2.3 Improving Performance using Wideband Calibration Signals

There is one additional problem which can emerge in the use of the receiver calibration technique described above. For the case where the array is electrically large, it is possible that phase ambiguities may emerge with the use of CW signals. This occurs when the differences in propagation delays through the system correspond to phase shifts in excess of $2\pi$ radians. If the electrical length is not too long, it may be practical to resolve the ambiguity using simple approaches, such as constructing all cables to be the same length. If this is not practical, a wideband calibrating signal, such as a chirp or a noise source, can be used to resolve the ambiguity. In this case the relative delays are estimated, which is generally a more complex task. On the other hand, wideband signals have the advantage that many frequencies can be calibrated simultaneously. In fact, using spread spectrum techniques, it is possible to apply low-level calibration signal continuously, even in the presence of a desired signal, with negligible interference to either signal.

For the receiver designed for this study (described in Chapter 5) the total instantaneous bandwidth of interest is 150 kHz. Thus, the use of wideband calibration signals was determined not to be necessary.

4.3 Estimating the Open-Circuit Impedance Matrix

This section addresses the issue of how to estimate the open circuit impedance matrix $Z$, suitable for use in calculating the system manifold $B = CZA$. Three candidate methods are considered.
The most straightforward method is by direct measurement. This is a two step process. First, the array’s \( N \)-port scattering matrix \( S \) is estimated. \( S \) is defined by

\[
[S]_{ij} = \frac{V_{j}^{\text{out}}}{V_{i}^{\text{in}}}
\]  

(4.22)

where \( V_{j}^{\text{out}} \) is the voltage measured at the terminals of element \( j \) when \( V_{i}^{\text{in}} \) is applied to the terminals of element \( i \). \( S \) can be measured simply by injecting a known signal into element \( i \) and measuring the result at all other elements \( j \), for all \( i \). This is possible using CW sources, as described in Section 4.2.1. Of course, the receiver calibration matrix \( C \) must be calculated and removed from the measurement to obtain the true value of \( S \). Then, \( Z \) can be calculated from \( S \) using

\[
S = (Z - I)(Z + I)^{-1},
\]

(4.23)

which assumes \( Z \) has been normalized by the feedline characteristic impedance. Thus, one can calculate the impedance matrix using

\[
Z = -(S - I)^{-1}(S + I).
\]

(4.24)

This method has the advantage of providing an accurate on-line, potentially real-time, estimate of \( Z \). However, it has the disadvantage that it requires a considerable increase in the number of components required in the antenna array structure. First, \( N \) additional directional couplers are required to inject signals. Second, \( N \) additional frequency channels are required for simultaneous injection at all antennas. At a center frequency spacing of 25 kHz, this consumes another 175 kHz over which the array and receivers must have a flat frequency response. Alternatively, the measurements could be performed sequentially using one channel, by means of an RF switch to route injection signals. This means either \( N \) additional cables must be run between
the array and the control location, or power and control signals must be routed to the antenna array. In either case, the RF switch plus additional couplers amounts to a significant increase in the cost and complexity of the antenna system.

A second approach is simply to use the moment method computation of $\mathbf{Z}$ from Section 3.4. This has the compelling advantage that it requires no additional hardware, and provides answers which must at least have the same general structure as the true value of $\mathbf{Z}$. The problem with this approach is that it is difficult to know if the model used to compute $\mathbf{Z}$ accurately conveys all the effects which may have a significant impact on the computation. For example, it may be that local scattering plays a role, or that the feedline impedance was not correctly modeled, or that there is sensitivity to the model used to describe the element feeds. Therefore, it is difficult to be confident that a reasonable value of $\mathbf{Z}$ is obtained.

A third approach, selected for use in this study, is to fit a physical model of the impedance matrix to AOA observations made while the system is in operation. This technique was inspired by an observation of Fenn [67] in a study of an $11 \times 11$ array of quarter-wave monopoles with half-wavelength spacing, on a flat ground plane. When comparing the multiport scattering coefficients computed using the moment method to the same coefficients computed from measurements, he observed that the magnitudes in each case were in good agreement, but with significant error in the phases. However, the phases were in error by about the same amount for every coefficient. This was attributed to model assumptions in the moment method calculation; in particular, the nature of the monopole feeds and the assumed current modes for each monopole. Note that the impedance matrix is related to the scattering coefficients by Equation 4.23. This suggests that it may be reasonable to parameterize
the difference between an impedance matrix for an array computed with element-wise model errors, and the true impedance matrix, using a single parameter – namely a phase shift in the scattering coefficients.

To exploit this observation, a method is needed to determine the “correctness” of the impedance matrix estimates obtained by varying the phase parameter. One can measure the array output due to sources at \( N_C \) sites distributed over the field of view. \( N_C \geq N \) with half-beamwidth spacing is desirable; however this procedure is feasible for smaller \( N_C \) depending on the accuracy with which the “prototype” impedance matrix is computed. The AOA is estimated in each case and compared to the true bearing to determine the estimation bias, using as many trials as necessary to accurately determine the bias. The desired impedance matrix is the one which minimizes the overall variation among the biases.

Ideally, the optimal value of the phase parameter would reduce the bias for each of the test sites to zero. However, it is possible that there is some error in the orientation of the array. In other words, the element positions may be known exactly with respect to other array elements, but the array itself may be rotated by some previously undetected amount. In this case, the optimal phase parameter is that which yields the smallest variation around some constant bias, and this bias is an estimate of the orientation error.

To summarize, the technique consists of the following steps:

1. A “prototype” value of \( Z \) is obtained. This could be computed using the moment method or possibly some other electromagnetic modeling technique. Alternatively, it could be based on measurements of the same or similar array under different conditions.
2. $S$ is computed using Equation 4.23. A new scattering matrix, $\hat{S}$, is defined as $Se^{i\theta}$, where $\theta$ is the parameter to be optimized. For a given value of $\theta$, the corresponding impedance matrix is obtained by substituting $\hat{S}$ for $S$ in Equation 4.23 and solving for $Z$.

3. Data is obtained for transmitters positioned at $N_C$ sites. The data for each site is processed using any appropriate AOA estimator.

4. The AOA estimates in Step 3 are recalculated using various values of $\theta$. The value which minimizes the variation in among the biases is taken to be the optimum value. $Z(\theta)$ corresponding to this value is taken to be the true impedance matrix.

5. The mean bias is interpreted as a rotational error in the orientation of the array, and is corrected as a post precessing step after the AOAs are estimated.

Results obtained using this technique in field conditions are presented in Chapter 6.

It should be noted that this technique is limited by a number of factors. First, and most important, is that the assumption in Step 2 must be reasonable. This was found to be the case for the moment method computation in [67]. In Chapter 6, this assumption will be shown to be useful – although not necessarily verified to be strictly true – for the array developed for this study. Second, the prototype $Z$ must be “good enough”; i.e., an inaccurate prototype may not yield completely unreasonable AOAs. Third, it is assumed that the AOA estimates themselves are unbiased for a perfectly-calibrated array. If the propagation is complex or an inappropriate estimator is used, this will not be the case, and the result may not be useful. Fourth, the mutual coupling must be weak enough in both the prototype and actual impedance matrices.
that the AOA estimates are simply biased, and not completely wrong. However, this is usually true for arrays of wire antennas with half-wavelength spacing between elements.

4.4 Calibration Accuracy Requirements

In the previous sections of this chapter, a calibration strategy has been developed. Now, a brief study will be conducted to determine the effects of various calibration errors on AOA estimation accuracy. The purpose is to determine calibration accuracy requirements. This is a consideration in the design of the system described in the next chapter, and also provides some indication as to the relative importance of various calibration parameters.

Considerable research effort has been applied to the problem of assessing calibration requirements recently; see for example [68] and [69]. In this study, however, it is sufficient to identify some loose bounds on the tolerable error in each calibration parameter. The parameters to be considered are:

- Magnitude and phase of elements of the receiver calibration matrix $C$,
- Magnitude and phase of elements of the open circuit impedance matrix $Z$, and
- Antenna element position errors.

In this case, direct simulation is a simple means to identify these bounds. The performance metric in this study will be the root mean square error (RMSE) of AOA estimates for the following simulation scenario:

- Y-shaped array developed in Chapter 3.
• 2 uncorrelated signals, each 10 dB SNR, incident from 90° (North as seen in Figure 3.9) and 110° (West of North).

• $\mathbf{R}$ estimated from $L = 3025$ snapshots (This number is derived from the receiver design considerations discussed in Chapter 5).

• $\text{ML}_2$ estimator. The Alternating Projections search algorithm of Ziskind & Wax [25] is used, initializing with the true AOAs, incrementing in 0.1° steps, and halting upon the first repeated AOA set.

• 100 trials.

This study will focus on the effect on estimation of the signal from 90°.

First, let us consider a “baseline” scenario: $\mathbf{C} = \mathbf{I}$, true $\mathbf{Z}$ equal to the nominal $\mathbf{Z}$ (computed in Section 3.4), and true element positions are equal to the nominal element positions. $\mathbf{Z}$ is assumed known and is used in the trial system matrix $\mathbf{B}(\hat{\phi})$. In this case, it was found that the mean AOA estimate $\hat{\phi} = 0.0°$; hence, the estimate is approximately unbiased. The RMS estimation error $\sigma_\hat{\phi} = 0.0°$; in other words, less than the 0.1° resolution granularity. The associated CRB is also less than 0.1°.

Next, each of the calibration parameters listed above is randomly varied according to a Gaussian distribution, followed by estimation of the AOAs using the new values. In this step, it is assumed that all calibration parameters are known exactly except for the one being varied. The variance of the parameter is increased until the AOA estimation variance over 100 trials reaches 1° RMSE.

The resulting calibration accuracy requirements are summarized in Table 4.1. These should be interpreted as upper bounds, with goal values being perhaps an order of magnitude less. This is because these requirements are computed assuming
<table>
<thead>
<tr>
<th>Error in...</th>
<th>Requirement</th>
</tr>
</thead>
<tbody>
<tr>
<td>Magnitude of elements of $C$</td>
<td>$&lt; 0.06 \ (0.3 \ dB)$ RMS</td>
</tr>
<tr>
<td>Phase of elements of $C$</td>
<td>$&lt; 4^\circ \ RMS$</td>
</tr>
<tr>
<td>Magnitude of elements of $Z$</td>
<td>$&lt; 0.02 \ (0.1 \ dB)$ RMS</td>
</tr>
<tr>
<td>Phase of elements of $Z$</td>
<td>$&lt; 6^\circ \ RMS$</td>
</tr>
<tr>
<td>Element position errors</td>
<td>$&lt; 0.018\lambda \ (1.2 \ cm)$ RMS</td>
</tr>
</tbody>
</table>

Table 4.1: Calibration accuracy requirements to achieve $1^\circ \ RMS$ estimation error, given the experiment conditions described in the text, assuming only one type of error is present at a time.

only one type of error at a time. Thus, the goal values should be considerably less to maintain the desired RMS accuracy in the presence of multiple error types.

4.5 Conclusions

In this chapter, a strategy for calibration of the array system was developed. With respect to the system manifold $\mathcal{B} = CZ\mathcal{A}$, the strategy is to compute $C$, $Z$, and $\mathcal{A}$ separately. The array manifold $\mathcal{A}$ is determined directly from the array geometry, and is never updated. The receiver calibration $C$ is determined at the same time the data is collected, by injecting additional signals at the antenna terminals. The open-circuit impedance matrix $Z$ is estimated by optimizing a single-parameter model of $Z$, using as a cost function the bias from AOA observations made while the system is in operation. This optimization also reveals orientation errors in the array, if any. Finally, a brief study was conducted to bound the calibration accuracy required to achieve AOA bias under $1^\circ$ for a realistic test scenario.
CHAPTER 5

ARRAY RECEIVER

The measurements and techniques proposed in previous chapters require a 7-element array receiver. The first section of this chapter further develops these requirements. In addition to the issues raised in previous chapters, requirements associated with cost, available technology, and the radio frequency environment at the proposed measurement site are identified. The second section provides a high-level description of the receiver that was developed from these requirements, and summarizes its performance. For the sake of brevity, this system is referred to as “UHF Array Experimental” (UAX). The final section provides additional details pertaining to the receiver, thereby providing complete documentation of the design.

5.1 Requirements

From previous chapters, the following requirements are inferred. First, seven receivers are required. Each receiver must convert the signal of interest at one set of antenna terminals to a digital form suitable for input to an array signal processing algorithm implemented on a PC. In addition to some hard limits on cost, it was also important that the receiver met the minimum essential requirements for this study. These requirements are discussed below.
5.1.1 Coherency, Calibration, and Bandwidth

The receivers must be coherent, which means that the phase difference between any two receivers, given the same input, is time-invariant for the duration of the measurement, at each frequency in the passband. Since virtually all practical receivers also convert the input signal to a lower frequency, the coherency requirement also means that the frequency conversion error should be the same for all receivers. Of course, no practical design is exactly time-invariant over a finite period. Thus, a more practical statement of the coherency requirement is that variations in the phase differences between receivers be sufficiently small so as not to noticibly affect the measurement. To achieve this, the receiver must either be extremely stable, or each measurement must be extremely short. The latter option leads to a less expensive design. Furthermore, the findings in [15] (discussed in Section 2.4) indicate the potential advantage of estimating channels which are invariant over the observation time: namely, that the MUSIC estimates are approximately unbiased even in the presence of small amounts of angle spread.

Another requirement is that the receivers must be capable of calibration, as described in Section 4.2. The findings summarized in Figure 4.1 can be interpreted as upper bounds on the required calibration accuracy. Since receiver stability is not a priority, it is important to ensure that the calibration be known precisely at the time of the measurement. A simple way to accomplish this is to perform the calibration signal injection/acquisition process simultaneously with the measurement. If the receiver calibration is simultaneous, calibration signals should not be applied at exactly the same frequency as the measurement. Instead, the calibration signal should be placed at a frequency which is close enough to accurately represent the calibration at
the measurement frequency, yet far enough away to be separated from the signal of interest. This provides incentive to make the passband response as flat as possible, suggesting an analog signal chain which has instantaneous bandwidth that is much greater than the difference between the source and calibration frequencies.

Another motivation for a wideband front end is that it is desirable to capture all 7 frequency channels within a 150 kHz band simultaneously. This provides 7 simultaneous channels which can be allocated among the tasks of receiver calibration and array calibration in addition to observing multiple mobile test sources. Furthermore, it allows monitoring of activity on nearby channels to detect conditions which could affect the measurement. This was found to be essential in the measurement campaign, as will be discussed below. An instantaneous bandwidth of 150 kHz is required to capture all 7 channels in one band. For sufficient passband flatness, the narrowest analog filter in the receiver should be perhaps 10 times wider, or 1.5 MHz. In the process of researching components for this design, excellent low-cost filters with 2.5 MHz bandwidth were identified. Thus, the instantaneous bandwidth was set at 2.5 MHz.

5.1.2 Linearity

Further requirements stem from the intended test location and the antenna array design. The antenna array was constructed on the roof of the Ohio State University ElectroScience Laboratory, a three-story building located at 1320 Kinnear Road, Columbus, OH, USA, as shown in Figure 5.1. An important aspect of receiver design is to ensure that it is robust to the various signal levels – both in-band and out-of-band – that can be expected at the input. To determine the characteristics of the radio spectrum over a broad range of frequencies, the output of a 25-1300 MHz discone
antenna\textsuperscript{7} was mounted near the array and observed using a spectrum analyzer. A typical result is shown in Figure 5.2. The strongest signal present is a broadcast TV station at about 504 MHz, which transmits continuously. When this experiment was repeated using the center monopole of the array, this signal was received at about \(-36\) dBm. Several mobile radio signals unrelated to this research were almost always present in the 460-470 MHz band. Figure 5.3 shows an example. Typically, at least one of these signals was received between \(-50\) dBm and \(-40\) dBm. All other signals observed through the center monopole were received at \(-40\) dBm or less.

One of the functions of the analog portion of a receiver is to boost signals from the level at which they are received to a level which makes the best possible use of an A/D's dynamic range. Full-scale for a typical A/D is typically between 0 dBm

\textsuperscript{7}Radio Shack Model No. 20-403
Figure 5.2: 0-1000 MHz, as observed through a wideband omnidirectional antenna near the array. The horizontal grid lines are 10 dB divisions, with 0 dBm at the top.

Figure 5.3: 460-470 MHz, as observed through the center element of the array. The horizontal grid lines are 10 dB divisions, with 0 dBm at the top.
Figure 5.4: The relationship between receiver gain, 1dB compression, and IIP₃. The $x$ axis is $\log_{10}$ of the input power, whereas the $y$ axis is $\log_{10}$ of the output power.

and +10 dBm at 50Ω. Thus, a minimum of 30 dB of gain is required in the analog front end. The challenge is to ensure this gain is linear. The observations above suggest that the receiver must be utterly linear for input powers up to $-36$ dBm at 505 MHz, and up to $-40$ dBm in the 460-470 MHz band. A common metric from describing linearity is the 1-dB compression point. The compression point is defined as the signal level at which the gain departs from linearity by 1 dB. This is illustrated in Figure 5.4. To be conservative, the 1-dB compression point should be at least $-30$ dBm in this case.
Since there is the potential for multiple strong signals in the passband, it is also
prudent to impose a requirement on the third-order intermodulation products (IM₃).
IM₃ are among the various spurious signals which arise when signals at different
frequencies are passed through a non-linearity. If the two frequencies are ω₁ and ω₂,
IM₃ appears at 2ω₁ − ω₂ and 2ω₂ − ω₁. These products are of particular interest in this
design because they are the strongest to appear within the analog passband of the
receiver. A useful metric when evaluating IM₃ is the “two-tone spurious free dynamic
range” (SFDR₂), defined as follows. Two CW signals are applied to the input of
the receiver. The frequency spacing between the signals is several channels, and is
chosen such that the third-order intermodulation products appear in the passband.
The power of the two input tones is increased until the power in each intermodulation
product is equal to the noise power in one channel bandwidth. SFDR₂ is then defined
as the difference between the power in one input signal and one intermodulation
product. SFDR₂ is useful to know because it indicates how strong two signals can be
before the associated IM₃ becomes significantly greater than the noise in any given
channel.

To calculate SFDR₂, one must know the noise power spectral density in the pass-
band, referenced to the input; and the input third-order intercept point (IIP₃). The
latter is illustrated in Figure 5.4. Referring back to the definition of SFDR₂, IIP₃ is
the defined as the input power per signal for the case in which the power in the two
input signals is sufficiently strong to generate IM₃ power equal to the output power
of the desired signals. Since the log power of the IM₃, referenced to the input, de-
creases at three times greater than that of log input signal power [70], IIP₃ provides a
convenient reference point from which to compute the IM₃ power for any other input signal level.

The SFDR₂ also depends on the noise power spectral density, which in turn depends on the noise figure of the receiver. This is a complex issue and is discussed in the next section. The SFDR₂ requirement is revisited at the end of this section and a value is assigned.

5.1.3 Sensitivity

A practical requirement for this system is that the effective range must be long enough to conduct meaningful experiments. Range is determined by transmit power, antenna and propagation characteristics, and receiver sensitivity. All of these parameters have been constrained based on other considerations, except sensitivity. The primary consideration in determining sensitivity is noise figure. Noise figure is traditionally defined as the ratio of the output SNR to the input SNR [70]. This ratio is greater than one for practical receivers, because real analog devices inevitably contribute noise to the signals they process.

Managing the noise figure in a receiver design is difficult because actions which improve noise figure tend to degrade linearity, and vice versa. Thus, one must typically constrain one while the other is optimized under the constraint. The nature of the measurements desired in this study demand that linearity be constrained (as discussed above). Thus, noise figure must be a secondary consideration. Since the simplest (hence lowest-cost) receiver design is desired, the noise figure requirement should be relaxed to be the maximum possible value that yields an acceptable system range.
To characterize the noise figure-versus-range tradeoff, one can use the following procedure:

1. For a given range, estimate the signal power $P_r$ at the terminals of one element in the array due to the mobile transmitter.

2. Estimate the total noise power $N_0$ for the receiver, referenced to the antenna terminals and the instantaneous bandwidth.

3. Calculate the SNR at the receiver output assuming an ideal noise figure $F = 1$, which is simply $P_r/N_0$.

4. The ratio between this ideal SNR and the minimum acceptable SNR is the receiver noise figure requirement for the given system range.

The method used in Step 1 must be chosen with care. If the mobile transmitter has line of sight (LOS) with the array and the environment is simple (negligible multipath scattering), a free space propagation model is appropriate. If the mobile transmitter is far from the array and does not have LOS, the free space model does not apply and an alternative method is required. In the LOS case, the following form of the Friis range equation applies [50]:

$$P_r = G_r \left( \frac{\lambda}{4\pi R} \right)^2 G_t P_t$$  \hspace{1cm} (5.1)

where $G_r$ is the receive antenna gain (approximately $-3$ dB below the gain of a dipole for grazing incidence on the ground plane, or $-1.2$ dBi), $G_t$ is the transmit antenna gain (typically 0 dBi), $P_t = +27$ dBm, $\lambda$ is the wavelength (0.6 m to be conservative), and $R$ is the range. Thus, one obtains

$$P_r = -0.6 - 20 \log_{10} R \text{ dBm, with } R \text{ in meters.}$$  \hspace{1cm} (5.2)
In the non-LOS case, the meaning of the term "path loss" is slightly different. In the narrowband mobile radio environment, the instantaneous signal level undergoes rapid variations with distance [71]. This is due to multipath scattering, typically local to the mobile, and is often modeled as "Rayleigh fading" [72]. In addition, there is a more subtle variation due to terrain shadowing, often modeled as "Lognormal fading" [71]. Rayleigh fading can vary the signal level by as much as 35 dB over as little as one-half wavelength of movement. Lognormal fading introduces another 8 dB or so of variation over a distance of hundreds of wavelengths. In the presence of these mechanisms, "path loss" is defined as the mean value of the instantaneous signal level over 20\(\lambda\) to 40\(\lambda\) of mobile motion [71]. Empirical studies have shown that the range dependence of path loss under these conditions varies between \(R^{-3}\) and \(R^{-5}\), as opposed to \(R^{-2}\) for free space conditions [71]. A representative estimate can be obtained using the terrestrial propagation model of Lee [71]. In Lee's Model, one selects from a set of empirically-derived "standard conditions", and then applies corrections to account for the difference between the standard condition and the situation of interest. A standard condition appropriate to this study is the "suburban" case described in [71], which assumes path loss equal to \(R^{-3.8}\). The details are summarized in Figure 5.5.8

When applied to the conditions of this study, one obtains

\[
P_r = 27.1 - 38 \log_{10} R \text{ dBm, with } R \text{ in meters.}
\]  

(5.3)

Several caveats apply to the above analysis. First, the LOS result should be regarded as an absolute lower bound; in other words, the path loss cannot be expected

8Note that the definitions of "transmitter" and "receiver" have been swapped with respect to Lee's definitions.
### Table

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Standard Condition</th>
<th>Actual Value</th>
<th>Correction</th>
</tr>
</thead>
<tbody>
<tr>
<td>Range</td>
<td>$R' = 1600$ m</td>
<td>$R$</td>
<td>$-38 \log_{10} R/R'$ dB</td>
</tr>
<tr>
<td>Tx Power</td>
<td>$P_t' = +40$ dBm</td>
<td>$P_t = +27$ dBm</td>
<td>$P_t - P_t' = -13$ dB</td>
</tr>
<tr>
<td>Tx Ant. Height</td>
<td>$h_t' = 3$ m</td>
<td>$h_t = 1$ m</td>
<td>$10 \log_{10} \left{h_t/h_t'\right}$ dB</td>
</tr>
<tr>
<td>Tx Ant. Gain</td>
<td>$g_t = 0$ dB</td>
<td>$g_t = 0$ dB</td>
<td>$g_t - g_t' = 0$ dB</td>
</tr>
<tr>
<td>Rx Ant. Height</td>
<td>$h_r' = 30$ m</td>
<td>$h_r = 15$ m</td>
<td>$20 \log_{10} \left{h_r/h_r'\right}$ dB</td>
</tr>
<tr>
<td>Rx Ant. Gain</td>
<td>$g_r' = 6$ dB</td>
<td>$g_r = -3$ dB</td>
<td>$g_r - g_r' = -9$ dB</td>
</tr>
<tr>
<td>Rx Power</td>
<td>$-61.7$ dBm</td>
<td>$27.1 - 38 \log_{10} R$ dBm</td>
<td></td>
</tr>
</tbody>
</table>

Figure 5.5: Application of Lee’s Model using the “suburban” standard condition. Note “Tx” refers to the transmitter; whereas “Rx” refers to the receiver.

To ever be less than this. On the other hand, the results from Lee’s model should be interpreted as “representative”, and not as a upper bound. In other words, it is possible that under certain situations the instantaneous path attenuation can be considerably greater than the path loss predicted by Lee’s model. A good example would be shadowing by a large building. Furthermore, the Rayleigh fading mechanism present in the non-LOS case can be expected to drop the instantaneous SNR by 10 dB or more 9% of the time [71]. This effect can be mitigated only by moving the mobile through the fading pattern during testing to ensure that not all of the samples are so impaired. As indicated previously, moving the mobile along a path of length $20\lambda-40\lambda$ ensures that the sampled signal powers are representative of the true distribution. Moving the mobile though the fading pattern during testing allows one to also benefit from the Rayleigh fading mechanism, since the signal level in a
Rayleigh fading environment is greater than the mean 40% of the time, and is 10 dB
greater 6% of the time [71].

There are other reasons to believe that multipath-induced fast fading will not be
a serious problem. First, true Rayleigh fading conditions are actually something of
a worse case. Typically, the fading statistics observed in field studies more closely
correspond to a mixture of LOS and Rayleigh fading mechanisms, known commonly
as Ricean fading [17]. In the region beyond optical LOS but short of the Lee Model
region-of-validity (given by the standard condition), it can be expected that the LOS
component will be relatively strong. This will significantly reduce the multipath-
induced variance in the instantaneous signal level. Thus, multipath fading is not
expected to be a serious problem.

Next, let us consider $N_0$. In the mobile radio environment, there are two significant
components to $N_0$: thermal and man-made noise. Thermal noise is an unavoidable
property of nature, and is given by $kTB$, where $k$ is Boltzmann’s Constant ($1.3806 \times
10^{-23}$ $J/\circ K$) and $T$ is temperature [70]. At room temperature, the thermal noise
power spectral density $kT$ is about $-174$ dBm/Hz. For an instantaneous bandwidth
of 2.5 MHz, the resulting thermal noise power is about $-110$ dBm referenced to the
input of the receiver.

It should be noted that for a traditional receiver design, the noise outside the
channel bandwidth of 25 kHz would normally be suppressed by additional filtering.
In this case, the intercepted thermal noise power would decrease by 20 dB. Due to the
short capture time and the wide bandwidth, such filtering is difficult to implement in
this design. Instead, the iterative subtraction technique described in Section 4.2.1 is
used for channelization in this design. Since this technique removes only interference and not noise, the 20 dB SNR improvement is not realized.

The second contribution to $N_0$, man-made noise, is attributable to devices which generate broadband noise, such as automobiles, lawn mowers, and electric motors. In suburban and urban areas, the large number of sources collectively generates spectrum which appears approximately white, and thus hard to distinguish from thermal noise. For this reason, it is convenient to express man-made noise in terms of an environmental noise figure, $F_{env}$, whose value indicates the ratio of the actual noise power spectral density to the thermal noise power spectral density. Experimental findings reported in [72] indicate that $F_{env}$ ranges from 10 dB to 25 dB in densely-populated areas. This is consistent with measurements performed at the array location on the roof of ESL, which indicated a value of about 12 dB at 462 MHz. A reasonable estimate of $N_0$ is therefore about $-98$ dBm.

Based on these findings, Figure 5.6 shows the best possible SNR (assuming a receiver with 0 dB noise figure) as function of range. Let us assume that a minimum acceptable SNR is 10 dB. From Figure 5.6, it can be seen that range is likely to be much less than the standard condition of 1.6 km using an ideal receiver under non-LOS conditions. For LOS conditions, the range is likely to be greater than 4 km. As the receiver noise figure increases from 0 dB, the SNR for any given range will decrease. As the receiver noise figure approaches $F_{env}$, the SNR degradation will approach 3 dB and the range degradation will become significant. This analysis makes it clear that there is not much to be gained by designing for a noise figure much below $F_{env}$, since then the range dependence will be dominated by man-made

92
Figure 5.6: SNR for an ideal receiver as a function of range, assuming free-space transmission (broken line) and Lee’s Model with a “suburban” standard condition (solid line).

noise. Based on these considerations, a non-strict noise figure requirement of about 12 dB seems appropriate.

5.1.4 SFDR and the Implications of a Wideband Front End

Let us now return to the topic of SFDR. Using $N_0 = -95$ dBm as an estimate of the total power in the receiver passband, one finds that the total noise power in one 25 kHz channel, referenced to the input, is $-115$ dBm. Based on the spectrum analyzer measurements described in Section 5.1.2, let us assume two signals at a maximum level of $-45$ dBm and within the 2.5 MHz passband are present at the
input to the receiver. To make the IM$_3$ less than the expected channel noise power, SFDR$_2$ must be at least 70 dB, and IIP$_3$ must be at least $-10$ dBm.

At this point, it should be noted that by requiring the receiver to accept all channels through a single wideband analog signal chain, IIP$_3$ has become the main design problem in the performance of the receiver. To make this clear, consider the alternative narrowband architecture, in which each channel is received through a separate analog signal chain. In this case, the IIP$_3$ can be controlled by analog filtering, and is therefore an issue only for the first few components of the receiver, before the channels are separated. By requiring a single wideband analog signal chain for all channels, the only defense against odd-ordered intermodulation products is to maintain high linearity at every stage of the receiver. The tradeoff is that only one signal chain is required for all channels; whereas, the narrowband approach requires a number of signal chains equal to the number of simultaneous channels desired.

### 5.1.5 Image Rejection

Another important consideration should be pointed out, namely image rejection. Images are signals which originate outside the passband, but are translated into it as a result of the frequency conversion process. For example, consider a single-conversion analog receiver which uses an LO at frequency $\omega_{LO}$ to convert the desired frequency $\omega_{RF}$ to the intended IF frequency $\omega_{IF}$. In a “low-side” conversion, this is accomplished by mixing (multiplying) the input and LO signals and then low-pass filtering, with the result that $\omega_{IF} = \omega_{RF} - \omega_{LO}$. However, this process also allows signals at $\omega_{LO} - \omega_{IF}$ to be translated into the same IF, superimposed on the desired signals. To prevent this, the image band $\omega_{LO} - \omega_{IF}$ must be suppressed before the conversion. The amount
of suppression is called “image rejection”. To be perfectly safe, the image rejection should be equal to the difference between the maximum signal power expected in the image band and the channel noise power. The image band is not yet defined; however, if one assumes the strongest signals observed in Figure 5.2, the “perfectly safe” image rejection requirement is about 80 dB. This is quite difficult to achieve in practice and is not strictly necessary unless the image band actually contains the worst-case signal levels. To ease this requirement, the receiver should be designed so that the image band falls in regions in which the signal levels are weak.

5.1.6 I/Q Balance

Ultimately, the algorithms envisioned in Chapters 2 and 4 will be implemented in complex baseband form; that is, operating on signals which are centered at 0 Hz. A natural representation for such signals is Cartesian form, in which the components are traditionally referred to as “in-phase” (I) and “quadrature” (Q). However, all physical signals are real-valued. The conversion from real to complex must be performed accurately, or the subsequent algorithm performance can be expected to suffer.

A common approach in wideband receiver design is to perform the real-to-I/Q conversion as part of frequency downconversion. In this scheme, the real signal is divided, and the two parts are downconverted separately using LOs which are 90° out of phase. The resulting analog I and Q signals are then digitized separately. In practical receiver systems, this method introduces numerous problems. First, it is difficult to design the divider, mixers, and associated RF components such that both the I and Q channels have exactly the same gain and phase response. Second, it is difficult to design a phase shifting scheme for the LO which maintains exactly 90°
over a wide frequency range. The resulting I/Q imbalance not only distorts the final signal, but is also likely to introduce a strong spurious component at DC – precisely in the center of the desired channel. The process of correcting the I/Q imbalance to make the signal usable is complex and not always successful.

For this reason, it is highly desirable to digitize the receiver passband while it is still in real form; i.e. at an IF stage. Then, the I/Q conversion can be implemented digitally, with performance which is limited only by quantization effects. Therefore, real IF sampling is considered a requirement for this design.

5.1.7 Requirements Summary

This section has identified the requirements for the array receiver. A minimum bandwidth of 2.5 MHz was set and linearity, in particular IIP3, was identified as a key specification. It was noted that sensitivity must be traded off to satisfy the bandwidth and linearity requirements at low cost. However, it was noted that the environmental noise power spectral density was also very high, which implies that it is not required to seek a receiver noise figure which is much less than 12 dB or so. It should be noted that the limited sensitivity affects spatial dynamic range as well; i.e., multipaths may not be visible unless the signals are very strong. The easiest way to improve the system range would be to use higher transmit power. However, the 0.5 W assumed here is the maximum allowed in the proposed frequency band. The most efficient way to improve the system sensitivity would be to implement one of the rolled edges described in Section 3.5, which could be expected to improve the system SNR by as much as 5 dB. Unfortunately, this was not implemented in the present project.
5.2 Design Overview

This section provides a functional description of the UAX system as built. First, the hardware signal flow is described. Since a significant portion of the functionality of UAX lies in post-processing software, this is described separately. The specifications are given at the end of this section. Details of the subsystems described here are provided in the Section 5.3.

5.2.1 Signal Flow: Hardware

UAX consists of four identical receiver subsystems, each responsible for processing two antenna elements; and various peripheral subsystems. The signal flow for one pair of antenna elements is shown in Figure 5.7. First, the output of each antenna is delivered to a low-noise amplifier (LNA), whose main purpose is to provide gain while bounded the noise figure of the system. This is followed by an RF gain block (RFGB), consisting of a bandpass filter and additional amplification. The output of the RFGB is downconverted to a new center frequency. One antenna is downconverted to 25 MHz; whereas the other antenna is downconverted to 35 MHz. These signals are then bandlimited to 2.5 MHz and combined.

The motivation for combining two elements into a single analog signal is that it simplifies the rest of the receiver; half the number of components are required in the IF and subsequent stages. Also, this approach eliminates cross-talk between elements in the IF stage, which greatly reduces the effort (hence cost) involved in signal isolation and shielding. This approach also makes efficient use of the large bandwidth offered by modern A/D and digital signal processing components occurring later in the signal chain; that is, digitizing 5 MHz of spectrum costs virtually the same as digitizing
Figure 5.7: Signal flow for one UAX receiver subsystem.
2.5 MHz of spectrum, and the difference in achievable performance is negligible. Furthermore, having multiple signals in each digital IF simplifies the transfer of data to PC, as will be seen below.

The dual-IF approach has the disadvantage that two local oscillator (LO) synthesizers are required. However, two suitable synthesizers were already available (benchtop signal generators were used), so no cost was involved. Another potential disadvantage of the dual-IF approach is that it is more difficult to maintain coherency; however, this problem can be addressed in post-processing software, as discussed below.

The selection of 25 MHz and 35 MHz as the center frequencies for the analog IF stage is based on several considerations. First, this choice places the image bands for low-side conversion between 392 MHz and 418 MHz, which is committed mostly to satellite- and aircraft-based services and is therefore exceptionally quiet spectrum. In the wideband spectrum measurements discussed in Section 5.1.2, no signals were observed above $-70 \text{ dBm}$ (the sensitivity limit of that measurement) in this band. This reduces the worst-case image rejection requirement to a relatively modest 45 dB, which can be achieved with inexpensive filters.

Other considerations driving the selection of the IFs are as follows. Higher IFs place the even-order intermodulation products at higher frequencies, making them easier to filter out. Modern high-dynamic range A/Ds have sample rates greater than 20 MSPS with analog bandwidths on the order of 100 MHz, so such IFs can be directly acquired using IF undersampling techniques. No additional analog downconversions are necessary if the IFs can be placed in the analog passband. Second, generation of the A/D sample clock is typically an expensive part of the receiver, due to the desire
for high stability and precise frequency control. Several suitable 40 MHz oscillators were already on-hand at the start of this project, as well as suitable 40 MSPS A/Ds, so selecting this sample rate significantly reduced the cost of new equipment. Given a 40 MSPS sample rate, center frequencies of 25 MHz and 35 MHz are convenient, because their second harmonics alias to gaps between the two IF bands. For example, the second harmonic of 25 MHz (50 MHz), sampled at 40 MSPS, aliases to 10 MHz, 30 MHz, 50 MHz, and so on. Similarly, the second harmonic of 35 MHz (70 MHz) aliases to 30 MHz, 50 MHz, and so on. Since these harmonics cannot be superimposed into one of the two passbands by aliasing, the filtering requirements in the analog IF strip are greatly simplified. In fact, the UAX design actually uses no antialiasing filter before the A/D; this function is provided solely by the bandlimiting/isolation filters at the mixer outputs.

The level of the combined signal is adjusted using a digitally-controlled variable attenuator. The purpose is to adjust the signal level to a specific value for the rest of the signal chain, which serves to optimize the dynamic range through the IF strip and A/D. Following the attenuator, an IF gain block (IFGB) provides the additional gain needed to bring the output close to the maximum encodable level of the A/D. The A/D digitizes the combined-element signal to 12 bits at 40 MSPS. As a result, the 25 MHz and 35 MHz element center frequencies are aliased to new center frequencies of 15 MHz and 5 MHz, respectively. Because these signals are aliased from the second Nyquist zone, they are spectrally reversed in the A/D output; however, this is easily corrected in software processing.

After digitization, a commercially-available integrated circuit (IC), the Harris HSP43216, is used to perform quadrature demodulation. Specifically, the HSP43216
Figure 5.8: Frequency plan for acquisition of the 25 MHz IF. Acquisition of the 35 MHz IF differs only in the bottom two spectra.

uses "$F_s/4$" processing, which converts the real-valued input to a complex-valued output which is translated in frequency by one-fourth the input sample rate \(^9\). This process is illustrated in the top four spectra shown in Figure 5.8.

A brief explanation of the $F_s/4$ technique follows. Consider a signal $s(t)$ which is "lowpass"; i.e., complex and centered at zero frequency. Multiplication by a factor of $\exp\{j\omega_{IF}t\}$ causes a spectral shift to a new center frequency, $\omega_{IF}$. Let us identify this version of the signal as $s_{IF}(t)$. If one samples $s_{IF}(t)$ at $F_S = 4\omega_{IF}$, then $\exp\{j\omega_{IF}t\}$ becomes $\exp\{j\frac{\pi}{2}l\}$, where $l$ is the sample index. Writing $s(t)$ and $s_{IF}(t)$ in I-Q form,

\(^9\) See [73] for an example of this technique in the literature.
one obtains:

\[ s'(lT) + js''(lT) = [s'_{IF}(lT) + js''_{IF}(lT)] e^{-j\frac{2\pi}{S} l} \]  

(5.4)

where \( T \) is the sample period \((1/F_s)\), the primed values indicate the in-phase component, and the double-primed values indicated the quadrature component. By setting \( l = 0, l = 1, l = 2, \) and \( l = 3 \) in the above equation, and solving for \( s'(lT) \) and \( s''(lT) \) for each \( l \), one finds:

\[ s'(0) = +s'_{IF}(0) \quad \text{and} \quad s''(0) = +s''_{IF}(0) \]  

(5.5)

\[ s'(T) = +s''_{IF}(T) \quad \text{and} \quad s''(T) = -s'_{IF}(T) \]  

(5.6)

\[ s'(2T) = -s'_{IF}(2T) \quad \text{and} \quad s''(2T) = -s''_{IF}(2T) \]  

(5.7)

\[ s'(3T) = -s''_{IF}(3T) \quad \text{and} \quad s''(3T) = +s'_{IF}(3T) \]  

(5.8)

and so on, with a pattern that repeats every 4 samples. Therefore, the spectral shift is accomplished simply by rearranging the I and Q components and changing signs as indicated above. For real-IF sampling, \( s''_{IF}(t) = 0 \), which introduces many zeros into the above equations and further simplifies the computations.

The spectral output of the HSP43216 is the pair of element signals centered at \(-5 \text{ MHz}\) (corresponding to the 35 MHz IF) and \(+5 \text{ MHz}\) (corresponding to the 25 MHz IF). The HSP43216 also lowpass filters the output, allowing the sample rate to be decimated by 2 without aliasing. This results in separate in-phase (I) and quadrature (Q) signal components, each 16 bits, at 20 MSPS. The A/Ds and HSP43216’s are placed together on a single circuit board known as the Wideband Digital IF (WBDIF).

The output of the WBDIF is captured by a length-16,384 (16K) first-in first-out (FIFO) memory buffer, allowing about 802 \( \mu s \) of contiguous capture. This requires
two 16K × 18 FIFO ICs, which are located on another circuit board known as the Dual FIFO (DFIFO). Using a FIFO memory buffer eliminates the need for a high-speed digital backplane, which would certainly dominate the system cost if used in this design. Since continuous real-time operation is not strictly required for the current study, it makes sense to avoid it in this design.

The DFIFO includes an interface circuit that allows it to be accessed using commercially-available PC-hosted data acquisition cards. In this way, data can be transferred from the DFIFO to a PC for subsequent processing and analysis. The data blocks are simply archived on the PC; all further processing occurs off-line on this or some other PC. This proved to be a simple and low-cost method of transferring data from the receiver subsystems to the PC. To minimize this cost even further, a 96-bit digital I/O card was used in the PC. The large width of the data path simplified the design of the interface; however, such cards have limited low I/O bandwidth. In this design, the transfer rate is limited to about 10 KB/s, so it requires about 24 s to transfer one 802 μs acquisition.

5.2.2 Signal Flow: Software

A significant portion of the functionality of UAX lies in software. This is necessary because the hardware section described above provides only blocks of data describing 2.5 MHz swaths of spectrum, with element outputs multiplexed into groups of two. However, this is also the strength of the system, as this leaves considerable flexibility in how the signals are processed. Above all, signal processing on PCs is inexpensive, and facilitates the use of convenient high-level software tools such as C or MATLAB.
The standard procedure for processing a 16K sample block is as follows. First, the two signals present in the output of each dual receiver subsystem are separated. The choice of center frequencies and sample rates makes this easy to do. Since ±5 MHz is one-fourth of the sample rate of $F_s = 20$ MSPS, “$F_s/4$” processing can again be used. This technique exploits the property that a spectral shift of ±$F_s/4$ can be implemented simply by rearranging I and Q components and possibly changing signs. This processing is the same as the technique used by the WBDIF’s HSP43216, except in this case the IF signal is complex. The signal in the upper sideband can be recovered in zero-IF form by $F_s/4$ downconversion followed by lowpass filtering; whereas, the signal in the lower sideband can be recovered in zero-IF form using $F_s/4$ upconversion followed by filtering. In fact, given either of the signals in zero-IF form (before lowpass filtering), the other can be obtained simply by changing the sign of every other sample. This is easily verified by substituting $\exp\{j\frac{\pi}{2}l\}$ for $\exp\{-j\frac{\pi}{2}l\}$ in Equation 5.4 and noting that the only differences are sign changes in Equations 5.6 and 5.8.

The two-sided bandwidth of the final digital lowpass filters is chosen to be equal to the instantaneous bandwidth of the analog section of the receiver, 2.5 MHz. A wider bandwidth would make the subsequent processing vulnerable to spurious analog signals generated by the receiver. A much narrower filter with slow roll-off is possible, but there are two reasons for choosing instead to preserve the entire passband with a fast roll-off. First, it is desired to monitor the spectrum well outside the channel of interest in order to identify signals that are strong enough to drive the front end
into compression, or "splatter" spurious signals into the channel of interest.\textsuperscript{10} Second, it is desired to maintain the maximum possible number of independent samples for accurate phase estimation. As shown in Equations 4.11 through 4.13, the variance of the sinusoidal frequency estimates is proportional to $L^{-3}$ and the variance of the sinusoidal amplitude and phase estimates is proportional to $L^{-1}$, where $L$ is the number of samples. A narrower filter with sharp roll-off would not only reduce the number of independent samples, but would also require more taps and thereby produce a longer transient response, rendering a greater fraction of the fixed number of samples unusable.

The lowpass filter currently used is a 256-tap equiripple FIR filter, designed using the Remez algorithm, that achieves 54 dB stopband suppression. The response is shown in Figure 5.9. The output of this filter is decimated by 4, yielding 4096 samples per block at a sample rate of 5 MSPS. To eliminate the transient response of the filter, the first and last one-eighth of the data samples are removed, which reduces the number of samples to 3025.

The next step is to extract the individual 25 kHz channels corresponding to mobile transmitters and calibration signals. A conventional method to achieve this is to shift the channel of interest to zero-IF and then to apply a lowpass filter with bandwidth commensurate to the channel width. As noted above, this is not practical for the UAX system. This is because a sufficiently narrow filter would generate a transient response that would render most of the samples in the record unusable. Instead, the algorithm described in Section 4.2.1 is used to identify and suppress all but the desired signal in the 2.5-MHz passband. This algorithm models all undesired signals

\textsuperscript{10}In the actual field use, this rarely happens. When it does, however, the situation is easier to diagnose if the wideband information is available.

105
in the passband as complex sinusoids, and iteratively estimates their parameters and subtracts them from the data until only the desired signal(s) remains. The algorithm ensures that all signals which are not of interest are suppressed to a level at least 36 dB below the total noise power. The algorithm is prevented from injecting signals within the channels of interest, so the narrowband signal(s)-of-interest is in no way distorted by this process. This technique also has the advantage that the passband response (both magnitude and phase) remains flat across the channels of interest, which would not be the case if the the channels were separated using traditional FIR filters. While this technique has been found to be quite effective as an alternative to FIR-filter channelization, it has the disadvantage that it does nothing to reduce the total noise power in the 2.5-MHz passband.
The final stage in the software portion of the receiver is to spectrally shift the remaining desired signal to zero-IF. Because separate synthesizers are used for the 25 MHz and 35 MHz IFs, it is not assumed that this shift has exactly the same value for each element. Furthermore, the synthesizers by themselves experience considerable "pulling"; that is, short departures from the desired center frequency due to phase-locked loop design limitations. To account for this, the exact frequency of the calibration signal is estimated in one of the receiver paths using a 25 MHz IF, and one using a 35 MHz IF. Equation 4.7 is used to estimate the calibration tone's instantaneous frequency to within 10 Hz. This value is then used to generate the desired digital LO signals for conversion to zero IF for each individual receive path.

5.2.3 Specifications

The specifications of the UAX array receiver design, as built, are given in Table 5.1. Note that the linearity specifications (compression point and IIP3), as well as image rejection, vary considerably due to the wide instantaneous bandwidth and the use of multiple IFs. For these specifications, the worst case values are given.

5.3 Design Details

In this section, a detailed description of the UAX design is presented. Figure 5.10 shows a block diagram of the system. Subsystems identified in this figure are described in the following sections.

5.3.1 Antenna Array Subsystem (AAS)

The AAS includes the antenna array described in Chapter 3 plus the calibration system described in Section 4.2. The entire AAS is installed on the roof of ESL,
Figure 5.10: UAX system block diagram.
<table>
<thead>
<tr>
<th></th>
<th>Actual</th>
<th>Requirement</th>
</tr>
</thead>
<tbody>
<tr>
<td>Number of receivers</td>
<td>8</td>
<td>7</td>
</tr>
<tr>
<td>Tuning Range</td>
<td>455-470 MHz</td>
<td>462-468 MHz</td>
</tr>
<tr>
<td>Instantaneous Bandwidth</td>
<td>2.5 MHz</td>
<td>1.5 MHz</td>
</tr>
<tr>
<td>IIP₃</td>
<td>-11 dBm</td>
<td>-10 dBm</td>
</tr>
<tr>
<td>Noise Figure</td>
<td>13 dB</td>
<td>12 dB</td>
</tr>
<tr>
<td>SFDR₂ (25 kHz noise bandwidth)</td>
<td>70 dB</td>
<td>70 dB</td>
</tr>
<tr>
<td>Image Rejection</td>
<td>44 dB</td>
<td>45 dB</td>
</tr>
<tr>
<td>Output Format</td>
<td>16-bit I × 16-bit Q</td>
<td>(none)</td>
</tr>
<tr>
<td>Output Sample Rate</td>
<td>20 MSPS</td>
<td>(none)</td>
</tr>
<tr>
<td>Block Acquisition Time</td>
<td>802 µs</td>
<td>Short as possible</td>
</tr>
<tr>
<td>Block Transfer Time</td>
<td>24 s</td>
<td>(none)</td>
</tr>
<tr>
<td>Calibration Accuracy, Magnitude</td>
<td>0.1 dB</td>
<td>0.3 dB</td>
</tr>
<tr>
<td>Phase</td>
<td>0.3°</td>
<td>4.0°</td>
</tr>
</tbody>
</table>

Table 5.1: Array receiver specifications, as built, compared to requirements developed in Section 5.1.

as shown in Figure 5.1. Everything except for the monopoles is installed directly below the ground plane to reduce local scattering and to provide some protection from weather.

### 5.3.2 Low-Noise Amplifier Array (LNAA)

The LNAA is illustrated in Figure 5.11. It is collocated with the rest of UAX in a small office directly underneath the array. The LNAA consists of 7 Hamtronics LNW-450 low-noise amplifiers (LNA), each built from a kit and installed in a custom-built connectorized enclosure for improved RF isolation. These units are intended for use in the 420-450 MHz amateur radio band, but can be used over a much wider range with negligible degradation in performance. This approach was selected because it resulted in the lowest possible cost (approximately US$250 in materials for the entire

11 65 Moul Road, Hilton, NY, USA.
LNAA subsystem). Each LNA was tuned for optimum performance at 465 MHz. In this configuration, each LNA provides 19 dB gain, about 25 MHz 3-dB bandwidth, and worst-case 13 dB rejection of the image band. The 1-dB compression point was observed to be about -10 dBm referenced to the input. The rated noise figure and IIP₃ of this design is 1.5 dB and 0 dBm, respectively. The entire LNAA assembly draws 200 mA at +15 VDC.

During system integration, it was decided to add 10 dB coaxial attenuators behind each LNA. As will be shown below (Table 5.2), this has a negligible effect on the system noise figure, while providing an additional measure of protection from strong out-of-band signals entering the receiver front end. These were installed at the LNAA outputs (barely visible in Figure 5.11).
5.3.3 Receivers (RX-1 through -4)

A single receiver subsystem is shown in Figure 5.12. Four identical receiver subsystems were constructed and installed in a standard 19-in wide by 72-in high equipment rack.

Analog Portion of the Receiver

Each subsystem receives input from two antenna elements via the LNAA. To achieve additional image rejection and gain prior to downconversion, each input is processed by a radio frequency gain block (RFGB). The schematic of the RFGB is shown in Figure 5.13, and construction details are shown in Figure 5.14. The input to the RFGB is filtered using a helical filter, which has approximately 24 MHz 3-dB
bandwidth and provides about 28 dB worst-case image rejection. A cascade of 3 Mini-Circuits MAV-11SM MMIC amplifiers provides about 35 dB gain. The measured 1 dB compression point for this circuit was found to be about \(-17\) dBm at the input. Each RFGB consumes 200 mA at 15 VDC.

The next step is downconversion. A common approach to this part of the design is to use a passive double-balanced mixer (DBM). Unfortunately, DBMs provide poor linearity. The typical 1 dB compression point for a DBM is 6 dB below the LO power; whereas, typical LO power is between 7 dBm and 23 dBm [74]. The typical IIP\(_3\) of a DBM is 10-15 dB above its 1-dB compression point. Therefore, IIP\(_3\) is in the range \(+11\) dBm to \(+34\) dBm for a DBM. Furthermore, achieving the stated level of performance for a DBM depends on careful impedance matching combined with high suppression of spurious products. A particularly difficult aspect of this problem is bleed-through of the LO into the IF. Typically, LO-IF isolation is about 30 dB. Since the log IIP\(_3\) of a DBM is proportional to the log LO power, there is strong incentive to use a high LO power. However, this increases the amount of LO power in the IF output. Unless the output match is perfect, some portion of this is reflected
back into mixer, and degrades its linearity. As a result, an elaborate diplexer design is required to use a DBM in a wideband receiver.

An analysis of the UAX design (Table 5.2) shows that a minimum $I_{IP3}$ of +38 dBm is required to achieve a system $I_{IP3}$ of -10 dBm; therefore a DBM is not an attractive choice. A safer, simpler approach in this case is to use an active FET mixer, which is considerably less sensitive to output matching. Active FET mixers actually tend to cost somewhat less than comparable DBMs, but are not widely used because they require DC power. UAX uses the Watkins-Johnson HMJ5 active FET mixer, which has a rated $I_{IP3}$ of +38 dBm using +21 dBm LO power, and consumes 35 mA at 3 VDC. The complete mixer design is illustrated in Figure 5.15, with construction and installation details shown in Figure 5.16. An $I_{IP3}$ of greater than +28 dBm has been verified; higher values cannot be measured using available test equipment.

\[12\] Specifically, it was not possible to arrange the available signal generators and spectrum analyzers in such a way that the mixer, as opposed to the test equipment, dominated the linearity of the test setup.
However, since the system IIP₃ was found to be with 1 dB of the requirement, the actual IIP₃ of the mixer must have been very close to the goal of +38 dBm.

In this mixer, an +8 dBm LO input is amplified to +21 dBm and applied to the HMJ5. Since the LO nominally consists of a single CW signal, linearity in the LO amplification is normally not an issue. In practice, however, synthesizers generate power at other frequencies, including artifacts which are difficult to filter out, such as phase noise and reference sidetones. This is characteristic behavior for synthesizers using analog phase-locked loops. In an array receiver, there is the additional issue that a single LO synthesizer may be shared by multiple receivers. The simplest way to do this is to divide the synthesizer output into separate LO signals for each mixer, thereby ensuring coherency among the receivers. A difficulty in this approach is that spurious signals can be propagated from the individual receivers back into the LO divider, and can propagate from there into all other mixers. Therefore, it is not safe
Figure 5.16: Mixer assembly: (a) Top side of PCB, (b) As mounted in receiver chassis.

to assume that the LO path will be free of spurious signals, and the linearity of the LO amplifier chain on each mixer board becomes an issue.

To prevent this from becoming a problem, it was decided to use high-dynamic range power amplifiers in the mixer board's LO chain. Two Mini-Circuits HELA-10 balanced power amplifier ICs were used. Each HELA-10 provides about 12 dB gain, +35 dBm IIP₃, and a very reasonable noise figure of 3.5 dB. A big disadvantage of this approach is that each HELA-10 consumes 525 mA at 12 VDC; however, since no requirement was placed on system power, this was not a discriminator. A more serious issue is that each HELA-10 heats up to greater than 100°C during normal operation. Special measures were required to transfer heat away from the amplifiers to prevent damage. In this design, this was accomplished by placing all the circuitry on one side of the circuit board, and bonding the back side to a mounting bracket formed from \( \frac{1}{8} \)-in aluminum sheet. The heat slug underneath each amplifier makes direct contact with the mounting bracket, and the mounting bracket provides a path for
the excess heat to be transferred to the chassis, where it is more efficiently dissipated. Mounting the boards on end (as shown in Figure 5.16) improves air circulation, which in turn reduces heat buildup in the chassis. Since the entire face of one side of the circuit board is fastened to the mounting bracket, the board cannot flex in response to temperature changes, reducing mechanical stress on the components and traces. This approach has been found to be completely reliable, with no failures over hundreds of hours of operation.

The output of each mixer is an IF signal centered at 25 MHz or 35 MHz. In each case the output is bandlimited to 2.5 MHz using bandpass filters with 4-pole Chebychev response. These were manufactured to our specifications by TTE\textsuperscript{13}. The 60-dB bandwidth of the 25 MHz filter is from 17.7 MHz to 30.3 MHz; whereas, the 60-dB bandwidth of the 35 MHz filter is from 27.1 MHz to 41.6 MHz. The magnitude response of these filters is flat to better than 1 dB within the central 1.5 MHz of the passband. These filters serve several purposes, in addition to defining the passband. First, they prevent nearby out-of-band outputs from each mixer from being superimposed into the output of the other mixer. They also suppress the strong LO bleed-through from each mixer. Finally, they serve as anti-aliasing filters for the A/D.

After the filter outputs are combined, they are routed through a 4-bit digital step attenuator, which allows the level to dropped between 0 dB and 32 dB in 2 dB steps under PC control. As discussed earlier, this feature was added to ensure that that the total power could be adjusted to ensure optimum use of the A/D's limited dynamic

\textsuperscript{13}11652 Olympic Blvd., Los Angeles, CA, USA.
range. In practice, this attenuator is set once at the beginning of a field experiment and is not controlled dynamically.

The output of the attenuator is amplified by the IFGB, as shown in Figure 5.17. The IFGB provides 25 dB gain to move the total signal level close to the +4 dBm maximum encodable level of the A/D with minimum distortion of the signal. Construction details of the IFGB are shown in Figure 5.18. It consists of a cascade of two Burr-Brown OPA643U operational amplifiers, each providing 12 dB of gain at 15.9 dB noise figure. The worst case IIP$_3$ of each amplifier is +32 dBm at 35 MHz; at 25 MHz it improves to about +37 dBm. The final design exhibited an input 1 dB compression point of −17 dBm. At this level, the IM$_2$ is quite high; about −23 dBc. This intermodulation product is in effect a significant new signal which is outside the 25 MHz and 35 MHz IFs. Fortunately, the first appearance of this product, at 50 MHz and 70 MHz respectively, aliases to 10 MHz in each case; Since the desired IF aliases are at 5 MHz and 15 MHz, this product has no effect on spurious performance.

This completes the description of the analog portion of the receiver. Provided in Table 5.2 is a summary analysis of gain, noise figure, and IIP$_3$ (GNI), presented in stage/cascade format [70].

**Digital Portion of the Receiver**

The output of the IFGB is delivered to the WBDIF, which consists of an Analog Devices AD9042 12-bit A/D and a Harris HSP43216 IC, which accomplishes quadrature conversion to zero-IF, filtering, and decimation by 2. The schematic of the WBDIF is shown in Figures 5.19 and 5.20, with construction details shown in Figure 5.21. The AD9042 is one of several suitable A/Ds available, and was selected
Figure 5.17: IFGB schematic.

Figure 5.18: IFGB board mounted in enclosure.
Table 5.2: GNI analysis of the analog portion of the receiver.

Based mainly on familiarity with the part. It has a total SNR (including quantization and other spurious contributions) of about 67 dB relative to full scale. Individual spurious signals generated by the A/D are normally less than −80 dBc. The use of the HSP43216 was motivated by its unique features, namely its ability to deal with wide bandwidths without high decimation. This is due to the use of the $F_s/4$ approach. The IC is very simple to use, requires no microprocessor interface for control purposes, and costs less than US$50. The more traditional device used in this application, commonly known as a “digital downconverter” (DDC), typically costs in excess of US$200, requires a microprocessor interface for control, and has a minimum decimation rate of 16 or 32.

The output of the WBDIF is delivered to the DFIFO, whose function was explained in Section 5.2.1. The schematic of the DFIFO is shown in Figures 5.22
and 5.23; construction details are shown in Figure 5.24. This design uses IDT72265 16K×18 FIFOs, which are somewhat expensive (about US$120 each). Smaller FIFOs would be significantly less expensive; however, 16K was judged to be the practical lower limit due to other considerations, such as channelization. To help defray the cost, the FIFOs were purchased in pin grid array (PGA) form. This allowed them to be socketed, and therefore available for reuse on other projects.

The method of operation of the DFIFO is as follows. Control of the acquisition is done by the PC, through the PC interface (PCIF). To acquire a block of data, each FIFO is commanded to perform a master reset (MRS*). This erases existing data and resets both FIFOs, so that the first sample to arrive after triggering is placed in the first memory location. When the trigger (WEN) is asserted, samples from the WBDIF are accepted into the DFIFO. Note that the WBDIF runs continuously; the
Figure 5.20: WBDIF schematic \((F_s/4\) section).
only difference during an acquisition is that the output samples are captured. "I" and “Q” samples are saved to different FIFOs. 16K samples (802 µs) later the DFIFOs are filled. At this point, they automatically disable themselves, such that they stop accepting samples, but maintain the samples currently in memory. The PC waits for 1 second after triggering the acquisition, and then reads the data in each FIFO, one at a time. This process takes about 24 s. At this point, the next acquisition can begin.

5.3.4 PC Interface (PCIF)

The PCIF provides the interface between the DFIFOs and data acquisition card in the PC and provides the interface for the PC to control the variable attenuator
Figure 5.22: DFIFO schematic, data section. FIFO A shown, FIFO B is identical.
Figure 5.23: DFIFO schematic, control section.
in each receiver. The DFIFO control consists of two identical circuit boards, each interfacing to two receivers. Each board connects to two DFIFOs one side, and to one of the two data acquisition cables from the PC on the other side. This board partitions the data and control lines into separate sets for each DFIFO. The board receives a dedicated acquisition trigger \textit{WEN*} for each DFIFO. Typically, only one of these is used, and is jumpered to all other DFIFOs, on both boards. This is to ensure that all DFIFOs begin acquisition at precisely the same time. This also allows the option (not so far used) to extend the total observation time by sampling a smaller number of elements using multiple receivers with appropriate trigger delays.
Figure 5.25: Clock and frequency reference generation.

5.3.5 Local Oscillator and Clock Distribution (LCDIST, SG-1, SG-2, and ARB)

The ultimate frequency reference for the UAX system is a 40 MHz temperature-compensated crystal oscillator (TCXO). Figure 5.25 shows how the sine-wave output of the TCXO is buffered and distributed by the LCDIST subsystem to create four separate coherent 40 MHz TTL clock signals. One clock signal is provided to each of the four receiver subsystems, where it is used by the WBDIF as the A/D clock.
The clock generation circuit also creates a fifth TTL output at 10 MHz which is used as a frequency reference for a Stanford Research Systems DS-345 arbitrary function generator (ARB). The purpose of the ARB is to generate the calibration signal which is injected at the antenna terminals, and also to provide various diagnostic signals as required. This reference is retransmitted to SG-1, which is a Hewlett-Packard HP-8656A signal generator, which serves as the LO generator SG-1. Similarly, SG-1 retransmits the 10 MHz reference to the second LO generator SG-2, which is another HP-8656A. This scheme ensures that the ARB, SG-1, and SG-2 are all coherent with the A/D sample clock and each other.

The output from SG-1 is split four ways using a simple power divider within the LCDIST subsystem. Each output is provided to one of the four receiver subsystems, where it is used as an LO in one of the two mixers. The output from SG-2 is processed in the same way, and is used as the LO for the other mixers.

5.3.6 PC

The PC used in this system was 450 MHz Pentium-III with 256 MB RAM running the Microsoft Windows 98 operating system. Two National Instruments PCI-DIO-96 data acquisition cards were installed on the PC’s PCI bus. Each data acquisition card supports 96 bits of digital I/O. One card was dedicated to data transfer from the UAX receiver FIFOs, whereas the other was used for gain control (5 bits per receiver subsystem times 4 subsystems).

All software used on this PC was developed in C using the National Instruments LabWindows/CVI development system. Two separate programs were developed. The first program managed data acquisition, data logging and archiving, system control,
and performance monitoring. The output of this code is four streams of hex-valued sample data (one per receiver subsystem), which is convenient for storage even in ASCII format. A second program was used after the experiment to recover the raw data and demultiplex the four streams into eight element outputs in complex baseband floating-point format.

All subsequent processing was performed using the Mathworks MATLAB interpreter on a Linux PC. The processing was split between two PCs because the Windows-based system was judged to be better suited to hardware tasks, whereas the Linux-based system was judged to be better suited to long-term data archiving and iterative signal processing tasks. MATLAB programs were developed for channelization, receiver (C) calibration, computation of covariance matrices, computing various AOA estimators, Array (Z) calibration, and performing statistical analysis of the data.
CHAPTER 6

FIELD MEASUREMENTS

In this chapter, the system developed in Chapters 2–5 is put into operation. The conditions of the experiment are described in Section 6.1, which describes the test area, the mobile transmitter, and certain relevant procedures. Section 6.2 describes the operation of the UAX array receiver and the system calibration process. Sections 6.3–6.6 describe the results of angle spectrum estimation for various classes of sites within the test area: “LOS”, “shadowed”, “suburban”, and “complex”. This chapter concludes with a discussion of error mechanisms at work in this experiment, and finally with a summary of the results.

6.1 Test Conditions

This section describes the area in which the field experiments were conducted and describes the operation of the mobile transmitter.

6.1.1 Test Area

The experiments were performed in November 1999 on the west campus of the Ohio State University and the adjacent suburban neighborhood of Upper Arlington, OH. The antenna array was mounted on the roof of the 3-story ElectroScience
Laboratory building (1320 Kinnear Road, Columbus OH), with the ground plane approximately 15 m above ground level, as shown in Figures 3.10 and 5.1.

The test area is illustrated in Figure 6.1, with the test sites identified by number. A total of 26 transmit sites were identified, with 19 selected for evaluation. The position of each site is known to within about 10 m, which means the true bearing measured from the array site to the closest test site (about 500 m away) is known to within 1.2°. To the north of the array site is a large open field, as shown in Figure 6.2. This field is bordered to the west by a two-lane road with single-family homes, and to the north by a four-lane road with single-level small business offices. A very large
4-story office building, which shall be called "Building A", is located at the northwest corner of the field. Building A spans about 11° of the field of view as seen from the array site. "Building B" is a bowling alley which extends another 7° beyond the east edge of Building A. Visible in Figure 6.2 (also Figure 3.10) is a cluster of industrial buildings, including a prominent 30 m high structure (Building C) that spans about 4°. The area immediately south of the array is shown in Figure 6.3. There is a two-lane road and across this road is an apartment complex consisting mainly of two-story townhouses.

6.1.2 Mobile Transmitter

In all cases, the transmitter is an unmodified commercially-available ICOM IC-4008A hand-held transceiver carried by a person on foot. The radio transmits 500 mW at 462.6625 MHz using analog FM modulation of about 12.5 kHz maximum bandwidth. This frequency corresponds to one of 7 contiguous channels with 25 kHz
Figure 6.3: View from the array, looking South.

spacing. Interference (in fact, legitimate use) was observed on all 7 channels, including the test channel. Separate observations indicated that this interference was visible with an interference-to-noise ratio greater than 3 dB less than 10 percent of the time. No special effort was made to account for cochannel interference in the data, except to exclude those acquisitions in which the cochannel interference was obvious. Thus, some of the data reported here includes weak cochannel interference.

At each site, the user was instructed to talk normally into the radio. To obtain a more representative distribution of local fading conditions, the user was also instructed to walk in a circle of diameter 6 m (≈9λ) while transmitting. The channel coherence time $\tau = \frac{9\lambda}{16\pi\nu}$, where $\nu$ is mobile speed, was estimated to be at least 20 ms. Therefore the scenario is effectively stationary over the 802-µs acquisition time. No special attempt was made to hold the radio to maintain any particular polarization; in fact, it was observed that most users tend to hold the radio in a position about half way between vertical and horizontal polarizations. Because the handheld’s antenna
has a dipole-like pattern, and the user is walking in circles, the SNR at the array
tends to vary over a few dB from acquisition to acquisition, even in LOS conditions.

6.2 System Operation

This section provides a brief description of the operation of the system in the field
environment. First, some observations of the receiver performance are presented.
Next is a discussion of array calibration.

6.2.1 Receiver Operation

The purpose of this chapter is to provide some insight into the behavior of the
system in field conditions. Figure 6.4 shows a typical acquisition from one element of
the array (the center element, in this case). The top figure is a Blackman-windowed
FFT of the 16K sample block, after the $F_s/4$ shift to complex baseband, but before
filtering or channelization. Note that various narrowband signals are apparent above
the noise floor. Further note that there is no apparent IM$_3$ in band.

The strongest signal is a “system monitor” signal, which is transmitted during the
802 $\mu$s acquisition from a yagi antenna mounted about 30 m away from the array. All
other signals are normal users unrelated to this research. The purpose of the system
monitor signal is to facilitate an end-to-end check of the system, by providing a signal
which is always present from the same AOA, with known magnitude and frequency
characteristics. If this signal is correctly processed, it provides additional confidence
that other signals processed in the course of the experiment are also being correctly
processed. This signal is also used in precalibration (calibration of the calibration
system, so to speak) as discussed in Section 4.2.2.
Figure 6.4: Spectrum from a field measurement at various stages of software processing. The horizontal axis is RF frequency in MHz; whereas the vertical axis is received power spectral density in dBFS/bin. FFT with Blackman window. Top: Original 16K sample block. Middle: After filtering and decimation by 4. Bottom: After channelization.
Note that the system monitor signal appears at a level of about $-30 \text{ dB}$ with respect to the full-scale input of the A/D (0 dBFS). Since no IM$_3$ is apparent above the noise floor, this plot demonstrates the realized SFDR in this case is greater than 60 dB.

The middle spectrum in Figure 6.4 is the same data block after the 2.5-MHz digital filter is applied and decimation by 4. At this point the effective sample rate has been reduced to 5 MSPS, and the filtering is apparent as the suppression of noise at the edges of the plot. Most importantly, note that there is no apparent distortion of the spectrum as a result of the filtering and decimation process.

The bottom spectrum in Figure 6.4 shows the result after channelization. In this case, the goal was to remove all but the system monitor signal. Note that all other signals have been suppressed to the level of the noise floor.

Figure 6.5 shows the equivalent of the middle spectrum of Figure 6.4, now for each of the 7 elements in the array. In this figure, each spectrum has been calibrated with respect to C. Note that the system monitor signal appears at the same level in each spectrum. However, note that certain other signals do not appear at the same level in each figure. For example, note that one of the stronger signals to the left of the system monitor signal disappears beneath the noise floor in the sixth plot from the top. This is attributable to fast fading. Even though neither the mobile nor the array may be moving, the array is in effect sampling different combinations of multipath phases. Element 6 apparently is close to a point of destructive interference for that signal. Such behavior is prevalent throughout this experiment and represents no fault of the system. To the contrary, it confirms the presence of fast fading in the propagation channels observed in this study. Figure 6.6 was prepared to better visualize the fading
statistics encountered in the course of this study. This figure is essentially the same as Figure 5.6, except that the path loss curves have been corrected to reflect the actual noise figure of the receiver. In addition, SNR versus range has been plotted for every acquisition reported in this chapter. It is interesting to note that the analysis of Section 5.1.3 accurately bounded the measurements.

Figure 6.5 also reveals a limitation of the UAX system for angle spectrum estimation. Note that the level of the noise floor changes slightly for each receiver. This is perhaps most easily seen by comparing the third and seventh plots from the top. The reason for this is as follows: Each receiver inevitably varies in gain, which is expected and is calibrated out. Also, examination of Figure 5.2 shows that both the LNA and the IFGB contribute significantly to the noise figure. Just as the gain varies from component to component, so does the noise figure. Thus, it is possible for a specific LNA or IFGB to have a relatively low gain and high noise figure compared to the other receivers, or vice versa. Since the gain per receiver is normalized, this difference manifests itself as an apparent variation in the noise floor from receiver to receiver. This results in a violation of the noise model described in Chapter 2, since the noise covariance is now significantly different from $\sigma^2 I$. To assess the effect of this problem, the following procedure was used. With the mobile stationed at Site 2, six acquisitions were performed. Site 2 is the closest site to the array, has optical LOS, and has nearby structures to generate scattering. The covariance matrix was estimated from each acquisition. The eigenvalues were then computed from the covariance matrix estimates and sorted from largest to smallest. The results are shown in Figure 6.7. If one assumes a single discrete AOA and noise covariance $\sigma^2 I$, one would expect to see one strong eigenvalue (representing the signal subspace) followed
Figure 6.5: Calibrated spectrum from each of the elements of the array. The horizontal axis is RF frequency in MHz; whereas the vertical axis is received power spectral density in dBFS/bin. 4K FFT with Blackman window.
Figure 6.6: SNR versus range for every acquisition reported in this chapter. Circles: LOS sites, asterisks: shadowed sites, diamonds: suburban sites, x’s: complex scattering sites. Predicted path loss assuming free space (dashed) and Lee’s suburban model (solid) also shown.
Figure 6.7: Eigenvalue spreads observed from 6 acquisitions while the mobile was at Site 2.

by six weaker eigenvalues of equal strength (representing the noise subspace). In Figure 6.7, however, note that the six noise eigenvalues are not equal. Based on measurements presented later, the propagation from Site 2 is believed to be dominated by a single strong AOA which is either discrete or has very limited spread. Thus, the variation of the noise eigenvalues is attributable mainly to the receiver’s contribution to the noise covariance. This effect can also be verified by simulation or theoretical analysis.

The main effect of this behavior is to limit one’s ability to make conclusions about single-source phenomena which involve rank greater than one. For example, instantaneous angle spread tends to increase apparent rank, and would be expected to generate the same kind of eigenvalue spread observed in Figure 6.7. A second limitation imposed by this behavior is that it limits one’s ability to properly estimate the
true number of signals, \( M \) (for ML), or the signal subspace rank, \( k \) (for MUSIC). For example, techniques based on MDL principles will fail in these conditions. Assuming CCI and multipaths are not too weak, however, the non-white noise covariance should not seriously limit one’s ability to estimate AOAs for the those signals, as long as the SNR (represented by the signal subspace eigenvalues) is relatively high.

The only practical way that has been identified to correct this problem is to buy large quantities of each component and attempt to match the components in such a way that each receiver has nearly identical gain versus noise figure characteristics. Unfortunately, such an approach is prohibitively expensive from the perspective of this study. Of course, the cost of any practical system developed for this task will be similarly effected, so it remains of interest to observe performance under this limitation. For future work, Wax’s technique for estimating AOA in the presence of unknown noise covariance [27] may prove useful.

### 6.2.2 Array Calibration

Calibration of the antenna array was performed using the technique proposed in Section 4.3. Data was collected while the user was transmitting from Sites 2-7. Separately, an estimate of the open-circuit impedance matrix was computed using the ESP5 computer code [52], which is based on the moment method. Using the \( \text{ML}_1 \) estimator, the matrix \( Z \) was found which minimized the variation in the bias of one AOA estimate from each of the 6 sites. The block of data selected to represent each site was chosen on the basis of SNR.

Using \( Z = I \) (no calibration), the bias varied between 2.4° and 6.6°, for a total range of 4.2°. The results of the calibration procedure are shown in Figures 6.8 and
Figure 6.8: Bias for each of six acquisitions (corresponding to Sites 2-7) as a function of the impedance matrix phase parameter.

6.9. Using \( Z \) from ESP5 with phase parameter equal to 0°, the bias varied between -1.8° to +11.4°. Using a phase parameter equal to 210°, the observed biases are limited to +3.7° ± 1.4°. Subsequent examination of the array revealed that it was in fact installed off-center by about 3°, which fortunately could be accounted for in the post-processing. The apparent second point of minimum variation near 50° is unusable, as it gives rise to response vectors (\( b \)) corresponding to incident signals which do not appear originate from far-field sources. From these results, it is inferred that the antenna array calibration error contributes less than 2° of error to the bias measurements described in this Chapter for Sites 2–7. Also, the calibrated array was observed to reduce the residual power after ML\( _4 \) AOA estimation by about 3 dB with respect to the uncalibrated array, to the values shown in Table 6.1.

141
Figure 6.9: Optimization of the impedance matrix phase parameter for array calibration. The difference between the two most extreme values of bias ("o") for six acquisitions (corresponding to Sites 2-7) are shown. The mean bias ("x") is also shown.
<table>
<thead>
<tr>
<th>ID</th>
<th>Range</th>
<th>Bearing</th>
<th>Trials</th>
<th>Mean SNR</th>
<th>ML Bias/RMSE</th>
<th>MUSIC Bias/RMSE</th>
<th>Capon Bias/RMSE</th>
<th>Residual Mean/Max</th>
<th>2Δ from [15]</th>
</tr>
</thead>
<tbody>
<tr>
<td>2</td>
<td>0.5 km</td>
<td>90°</td>
<td>6</td>
<td>32 dB</td>
<td>+1.8°/2.1°</td>
<td>+1.7°/2.1°</td>
<td>+1.0°/2.0°</td>
<td>-7.2 dB/-5.7 dB</td>
<td>6.3°</td>
</tr>
<tr>
<td>3</td>
<td>0.5 km</td>
<td>105°</td>
<td>5</td>
<td>30 dB</td>
<td>+1.3°/1.0°</td>
<td>+1.2°/1.0°</td>
<td>+1.0°/1.4°</td>
<td>-8.3 dB/-6.6 dB</td>
<td>5.5°</td>
</tr>
<tr>
<td>4</td>
<td>0.6 km</td>
<td>120°</td>
<td>6</td>
<td>31 dB</td>
<td>-0.3°/1.2°</td>
<td>-1.0°/1.1°</td>
<td>-0.7°/1.3°</td>
<td>-5.1 dB/-1.0 dB</td>
<td>3.3°</td>
</tr>
<tr>
<td>5</td>
<td>0.7 km</td>
<td>135°</td>
<td>22</td>
<td>27 dB</td>
<td>-1.6°/2.1°</td>
<td>-1.7°/2.1°</td>
<td>-0.5°/2.2°</td>
<td>-6.1 dB/-3.5 dB</td>
<td>5.7°</td>
</tr>
<tr>
<td>6</td>
<td>1.2 km</td>
<td>115°</td>
<td>9</td>
<td>20 dB</td>
<td>+0.6°/1.9°</td>
<td>+0.4°/2.0°</td>
<td>+1.6°/2.1°</td>
<td>-5.2 dB/-3.3 dB</td>
<td>5.4°</td>
</tr>
<tr>
<td>7</td>
<td>1.1 km</td>
<td>100°</td>
<td>6</td>
<td>25 dB</td>
<td>+0.9°/1.1°</td>
<td>+0.8°/1.1°</td>
<td>-0.5°/1.3°</td>
<td>-7.0 dB/-5.0 dB</td>
<td>3.0°</td>
</tr>
</tbody>
</table>

Table 6.1: LOS examples

6.3 LOS Sites

AOA estimation results have been organized into four sets of measurements: “LOS”, “Shadowed”, “Suburban”, and “Complex”. Sites 2-7 were classified LOS, with results as summarized in Table 6.1. (In this and the following tables, “MUSIC\(_1\)” refers to MUSIC with \( k = 1 \).) Note that the bias and RMS error (RMSE) in each case is less than 2°, which in light of the known errors in true bearing and calibration-induced bias noted earlier, indicates that the performance of the AOA estimators is very good. The reasonable performance of MUSIC and the Capon estimator relative to ML indicates that perfectly-correlated wide-angle multipath, if present, must be very weak. This is further supported by the residuals, which indicate that any multipath separated by more than a beamwidth is probably at least 5 dB below the LOS component. Shown in the last column is an angle spread estimate obtained using the method of Astély and Ottersten [15]. As discussed in Chapter 2, their method is based on the statistics of the MUSIC estimates. Angle spreads in the range 3° to 6° are estimated.

Figure 6.10 shows the “raw” angle spectrum using estimators which yield continuous spectrum: CBF, Capon, and MUSIC. In this case, Site 3 is used as an example representative of the LOS sites. Note that despite the accuracy of the estimates as
indicated in Table 6.1, the continuous spectrum estimates are not impressive. Although all three methods suggest a single dominant AOA, in the Capon and MUSIC spectra this peak is weaker than might be expected. In ideal conditions, the MUSIC peak should go to infinity and the Capon peak (measured from the peak to the spectral “noise floor”) should be equal to the SNR. The fact that this does not occur indicates either an error in the calibration or angle spectrum that is not well-modeled by discrete AOAs. In fact, it is likely that both play a role.

Figure 6.11 is provided to aid in visualizing the LOS results. In this figure, each AOA estimated by ML$_1$ is shown as a line originating from the array site.

Due to the relatively large number of acquisitions for Site 5, results for this site were further analyzed. Figure 6.12 shows the AOA pairs estimated by ML$_1$ and ML$_2$; that is, ML with $M = 1$ and $M = 2$. An interesting finding is that the second AOA is usually estimated to be about 150°. The power associated with this AOA is estimated to be about 7 dB (mean) below that of the first AOA. A possible explanation is scattering from houses and perhaps vehicles traveling along the nearby road; in fact, the second AOA is close to the expected angle for specular reflection from these features. However, increasing the model order has no apparent effect on the bias or variance of the first AOA in this case.

### 6.4 Shadowed Sites

Two cases of shadowing by buildings were considered. The first case involves Sites 1 and 12, which are summarized in Table 6.2 and Figure 6.13. Site 1 is blocked from LOS with the array by a cluster of trees; it is also located relatively close to a cluster of industrial buildings to the southeast and a cluster of 8 satellite dishes
Figure 6.10: Angle spectrum estimates for Site 3 (5 trials).
Figure 6.11: ML\(_1\) AOAs for Sites 3, 4, 6, and 7.
Figure 6.12: AOAs estimated for Site 5. (a) ML₁; (b) ML₂ scatter plot; (c) ML₂, First AOA; (d) ML₂, Second AOA. For ML₂, note there are 6 AOA pairs superimposed for first AOA equal to 135.8°.
<table>
<thead>
<tr>
<th>ID</th>
<th>Range</th>
<th>Bearing</th>
<th>Trials</th>
<th>Mean SNR</th>
<th>ML1 Bias/RMSE</th>
<th>MUSIC1 Bias/RMSE</th>
<th>Capon Bias/RMSE</th>
<th>Residual Mean/Max</th>
<th>$2\Delta$ from [15]</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0.5 km</td>
<td>76°</td>
<td>26</td>
<td>27 dB</td>
<td>+0.6°/6.1°</td>
<td>+0.6°/6.1°</td>
<td>-0.3°/7.9°</td>
<td>-5.2 dB/-5.1 dB</td>
<td>15.4°</td>
</tr>
<tr>
<td>12</td>
<td>0.3 km</td>
<td>69°</td>
<td>9</td>
<td>14 dB</td>
<td>-3.7°/6.4°</td>
<td>-4.0°/6.4°</td>
<td>-3.3°/9.3°</td>
<td>-3.3 dB/-1.5 dB</td>
<td>17.7°</td>
</tr>
</tbody>
</table>

Table 6.2: Shadowing examples

(10-15 ft diameter) to the northwest (at the position marked “SCF”). It is interesting to note that the bias for Site 1 is low relative to the nearby LOS Site 2, but that the RMSE is much higher. Furthermore, the residual indicates that propagation from this site includes a strong LOS-like component. Site 12 is about 20 m to the east of Site 1, but is blocked from the array by a two-story industrial building in the Northwest corner of the industrial park. In this case, the shadowing results in about 13 dB attenuation. It is also observed that AOA estimates are biased an additional 2°-4° to the east, but that the RMSE is about the same. Again, this is somewhat unexpected, as it seems that diffraction around the west side of the building, or over the roof, would be more likely. A better explanation is that the signal from Site 12 is scattered by each of the buildings to the south of it, and not just the closest building. This would create a bias to the East and would greatly increase the RMSE. This is consistent with the fact that the residual is large, indicating that the AOA estimates, although reasonable, do not properly account for the signals incident on the array. The increased angle spread estimates seem to further support this, suggesting a complex AOA distribution.

Figure 6.14 shows the “raw” continuous angle spectrum from CBF, Capon, and MUSIC for Site 12. It is interesting to note that the Capon and MUSIC spectra suggest a single dominant AOA, whereas in the CBF spectrum this is not as clear.
Figure 6.13: ML₁ AOAs for Sites 1 and 12.

However, the CBF spectrum is noticeably skewed to the left. This supports the argument that the large bias and RMSE is in fact due to scattering from buildings to the south of Site 12.

Due to the relatively large number of acquisitions for Site 1, results for this site can be further analyzed. Figure 6.15 shows the AOA pairs estimated by ML₂ for Site 1. In this case, two types of behavior are evident. One cluster of AOA pairs shows significant RMSE for the first AOA, and with a second AOA in the range 93°–110°. The power associated with the second AOA for these pairs is estimated to be about 8 dB (mean) below that of the first AOA. As shown in Figure 6.15, the distribution of the second AOA in this case strongly suggests that it is mainly due to scattering from Buildings A and B. The second cluster of AOA pairs shows very low bias for the first AOA, and with a second AOA of about 250°. The power associated with the second AOA in this case is also estimated to be about 8 dB (mean) below that
Figure 6.14: Angle spectrum estimates for Site 12 (6 trials).
of the first AOA, but no terrain feature can be associated with this AOA. Therefore, it is believed to be spurious in the sense that the best model order for those trials may actually be 1. For this site, increasing the model order from 1 to 2 reduces the RMSE for the first AOA from 6.1° to 5.0°, which can also be observed in Figure 6.15. This was the only case encountered in this field study in which increasing the model order significantly improved the quality of the first AOA estimate.

It is also interesting to note that some of the apparent outliers in Figure 6.15(d) intersect the SCF, the Building C, and the building south of Site 12. Thus, it is likely that these and possibly other estimated AOAs in this figure represent legitimate scattering, and are not simply outliers.

Also with regard to shadowing, Sites 7 and 25 were considered. Note that Site 7 is a LOS site on the near side of Building A; whereas, Site 25 is directly opposite Building A. The mean SNR for Site 25 was found to be 7 dB, indicating about 15 dB attenuation due to shadowing. All three AOA estimators completely failed for this case; the estimates are more-or-less randomly distributed over 360°. In fact, the continuous CBF and Capon AOA spectra are effectively featureless as well. This is interesting because the building itself spans only about 10° of the field of view. Thus, one might expect AOA estimates that are generally in the proper direction, but perhaps with RMSE on the order of 10° due to diffraction around the building. Since this is not the case, the signals incident on the array must be spread over a range of angles which is much greater than the 30° beamwidth of the array. This suggests that signals reach the array primarily via wide-angle multipaths, probably from the row of houses along the west side of the field and the row of industrial buildings along the east side of the field.
Figure 6.15: AOAs estimated for Site 1. (a) $ML_1$; (b) $ML_2$ scatter plot; (c) $ML_2$, First AOA; (d) $ML_2$, Second AOA. The dashed lines in (b) indicate the optical boundaries of west and east edges of Building A (top and middle lines, respectively) and the east edge of Building B (bottom line).
6.5 Suburban Sites

Table 6.3 and Figures 6.17–6.18 summarizes the data obtained from Sites 18–22. These sites are located along one of the suburban side streets that extends from the west edge of the north field, as shown in Figure 6.16. Site 18 is at the opening to the field, and Sites 19–22 run east to west along a sidewalk. Site 18 has LOS to the array, but is close to homes on the North and South sides of the street. It exhibits the low bias, low RMSE, and low residual expected from an LOS site. Sites 19-22 are blocked from LOS by two-story single-family homes and foliage. This blockage contributes at least 14 dB attenuation relative to Site 18. Sites 19-21 exhibit RMSE of 3°-4° and the high residual suggests scattering in addition to a strong LOS component. The biases for Sites 19-21 suggest that at least some part of the signal may be directed by the street back toward the open field; having the effect of making these sites appear to be further to the east than they actually are. At Site 22, the situation changes dramatically: First, the bias returns to near-LOS values, suggesting that perhaps the postulated street-guided propagation path is now weaker than the direct path to the array. However, the residual is still large, indicating that even this more direct propagation path may not be well-modeled as a discrete AOA. Furthermore, all three AOA estimators are beginning to generate outliers.

6.6 Sites with Complex Local Structure

Finally, three sites with complex local structure were examined, as summarized in Table 6.4 and Figures 6.19–6.22. Site 10 is situated between two sections of two-story townhouses, with a single-story parking shelter, foliage, and a chain-link fence blocking LOS to the array. The parking shelter has metal siding, extends for hundreds
Figure 6.16: View of the suburban sites, looking East. The light pole at the street corner is Site 22.

<table>
<thead>
<tr>
<th>ID</th>
<th>Range</th>
<th>Bearing</th>
<th>Trials</th>
<th>Mean SNR</th>
<th>MLL Bias/RMSE</th>
<th>MUSIC Bias/RMSE</th>
<th>Capon Bias/RMSE</th>
<th>Residual Mean/Max</th>
<th>2Δ from [15]</th>
</tr>
</thead>
<tbody>
<tr>
<td>18</td>
<td>0.8 km</td>
<td>132°</td>
<td>4</td>
<td>27 dB</td>
<td>-0.3°/0.6°</td>
<td>-0.4°/0.6°</td>
<td>-0.4°/0.8°</td>
<td>-6.2 dB/1.1 dB</td>
<td>2.1°</td>
</tr>
<tr>
<td>19</td>
<td>0.9 km</td>
<td>137°</td>
<td>7</td>
<td>12 dB</td>
<td>-3.5°/3.7°</td>
<td>-3.4°/3.8°</td>
<td>-3.7°/3.4°</td>
<td>-3.2 dB/2.3 dB</td>
<td>12.4°</td>
</tr>
<tr>
<td>20</td>
<td>0.9 km</td>
<td>130°</td>
<td>5</td>
<td>11 dB</td>
<td>-3.0°/3.2°</td>
<td>-3.6°/1.4°</td>
<td>-3.0°/2.8°</td>
<td>-2.6 dB/0.9 dB</td>
<td>17.3°</td>
</tr>
<tr>
<td>21</td>
<td>1.1 km</td>
<td>146°</td>
<td>6</td>
<td>9 dB</td>
<td>-8.4°/3.1°</td>
<td>-8.5°/2.0°</td>
<td>-8.7°/2.6°</td>
<td>-2.8 dB/1.3 dB</td>
<td>9.0°</td>
</tr>
<tr>
<td>22</td>
<td>1.2 km</td>
<td>150°</td>
<td>19</td>
<td>8 dB</td>
<td>-0.2° (7)</td>
<td>-2.8° (8)</td>
<td>-1.6° (1)</td>
<td>-1.6 dB/1.3 dB</td>
<td>10.5°</td>
</tr>
</tbody>
</table>

Table 6.3: Suburban examples. Numbers in parentheses indicate number of outliers among the trials; if outliers are indicated, then the bias is calculated as the median, as opposed to the mean, of the difference between the estimated AOA and the true bearing.
Figure 6.17: ML\textsubscript{1} AOAs for Site 18 (4 trials).

<table>
<thead>
<tr>
<th>ID</th>
<th>Range</th>
<th>Bearing</th>
<th>Trials</th>
<th>Mean SNR</th>
<th>ML\textsubscript{1} Bias/RMSE</th>
<th>MUSIC\textsubscript{1} Bias/RMSE</th>
<th>Capon Bias/RMSE</th>
<th>Residual Mean/Max</th>
<th>2Δ from [15]</th>
</tr>
</thead>
<tbody>
<tr>
<td>10</td>
<td>0.5 km</td>
<td>227°</td>
<td>12</td>
<td>29 dB</td>
<td>+3.3°/2.3°</td>
<td>+4.2°/2.3°</td>
<td>+2.8°/2.3°</td>
<td>-7.4 dB/-1.2 dB</td>
<td>6.9°</td>
</tr>
<tr>
<td>13</td>
<td>0.5 km</td>
<td>24°</td>
<td>5</td>
<td>18 dB</td>
<td>-10.0°/3.0°</td>
<td>-9.9°/3.8°</td>
<td>-9.2° (2)</td>
<td>-3.0 dB/-1.6 dB</td>
<td>12.7°</td>
</tr>
<tr>
<td>26</td>
<td>1.7 km</td>
<td>101°</td>
<td>13</td>
<td>8 dB</td>
<td>-8.2° (5)</td>
<td>-11.8° (4)</td>
<td>-7.4° (7)</td>
<td>-1.0 dB/-0.8 dB</td>
<td></td>
</tr>
</tbody>
</table>

Table 6.4: Complex local structure examples. Numbers in parentheses indicate number of outliers among the trials; if outliers are indicated, then the bias is calculated as the median, as opposed to the mean, of the difference between the estimated AOA and the true bearing.

of meters to the east and west of Site 10, and is one car-length deep. In this case, note that the bias and RMSE are much larger than for LOS cases, as might be expected. However, it is interesting that the residual is very low, indicating that a significant portion of the incident power is well-modeled by a single discrete AOA in this case. The intermittent occurrences of high residual and RMSE are probably related to the movement of traffic on the road which runs east-west between Site 10 and the array.
Figure 6.18: ML₁ AOAs for Sites 19–22.
Figure 6.19: ML$_1$ AOAs for Sites 9 (8 trials).

Figure 6.20: ML$_1$ AOAs for Sites 10 (12 trials).
Figure 6.21: ML₁ AOAs for Site 13 (5 trials).

Site 13 is situated at the eastern edge of a parking area to the northeast of the array, and is surrounded on the north, west, and south by industrial buildings. Here, one can see that the AOA estimates are severely biased to the south, the RMSE is small relative to the bias, and that the residual is high. There is no structure (for example, a multistory building) close enough to the estimated AOA path to explain the large bias in terms of specular multipath. Thus, the propagation from this site may be quite complex.

Note that Sites 9, 10, and 13 all seem to exhibit high bias relative to Sites 1-7. As sites which are suspected to generate complex propagation, this is not surprising. However, it should be noted that the array calibration did not use any AOAs from the angular vicinity of these sites, so it is possible that calibration error also contributes to the apparent bias.
Figure 6.22: ML_1 AOAs for Site 26 (13 trials).

Site 26 is inside a sandwich shop located on the north side of the four-lane road that runs along the north edge of the test area. Propagation to the array is obstructed by furnishings within the shop, moving vehicles on the road, Building B, and the east corner of Building A. In this case a bias to the East of about 10° is observed, a high residual indicating complex propagation, and a significant number of estimation failures. As noted for Site 22, it is difficult to attribute these failures completely to the low SNR in this case. In fact, it seems likely that the buildings to the East and Northeast of ESL are actually scattering as indicated, and that this scattering is now apparent due to the occlusion of the LOS component in this case.
6.7 Sources of Error

In this section, the known sources of error and other limitations in this field study are summarized in preparation for summary conclusions, presented in the next section.

- **Z estimate.** Recall that the array calibration procedure – in particular, the computation of $Z$ – is based on heuristic arguments. As pointed out earlier, the parameterization of $Z$ using the phase of the scattering coefficients, and subsequent attempt to estimate that phase using a small number of AOA observations, is likely to be contributing an error in the measurements. A particular concern is that the AOA measurements used to estimate $Z$ were drawn from the LOS sites in this study, which are clustered between $70^\circ$ and $140^\circ$. Since no suitable LOS sites were available outside this range, it is difficult to know precisely how well the calibration algorithm is working elsewhere. This may explain the large bias observed in the results for Sites 9, 10, and 13, which lie well outside this sector. On the other hand, it was interesting to note that the array calibration procedure correctly identified the $3^\circ$ (rotational) orientation error in the array to within a fraction of a degree.

- **Precalibration.** Recall that precalibration is the process of calibrating the signal paths that are later used to inject the needed signals for calibration of $C$. Also recall that this requires an off-line procedure involving a series of measurements, modifying cable connections each time. Through a series of off-line experiments, it was found that the estimates provided by this procedure tended to remain
within 0.5 dB and 0.05° over a 24-hour period. Nevertheless, this does not preclude the possibility that momentary weather conditions – for example, strong wind gusts – could flex certain cables and thereby cause a significant departure from this range. Furthermore, this would be difficult to detect. However, all measurements presented in this chapter were performed on days of relatively mild wind conditions, so this error is judged not to be a significant factor.

- **Receiver-Induced Noise Covariance.** As pointed out earlier in this chapter, the differences in noise level for each receiver represent a model violation, and are conceivably affecting the bias and variance of the AOA estimates.

- **Position Uncertainty.** Recall that the precise position of each site is known only to about 10 m, leading to a bias uncertainty of 1.2° for the closest sites.

- **Cochannel Interference.** Recall that CCI greater than 3 dB above the noise level was observed 10% of the time, and that these acquisitions were discarded. However, weaker cochannel is almost certain to be present. Since it is not correctly accounted for in ML₁ or MUSIC₁, this factor is also a potential source of error.

- **Limited number of acquisitions.** Using this system, it requires 24 s to acquire 802 µs of signal capture. Since the total time to perform the experiment was limited, the number of acquisitions per site was too low to compute accurate statistics. Thus, caution is required when evaluating the bias and variance data computed for each site.
- **Low SNR.** Only the LOS sites experienced SNR greater than 20 dB or so. For the other sites, especially the shadowed sites, the SNR was on the order of 10 dB. This implies that the error mechanisms cited above are more likely to affect the non-LOS sites. Thus, it may not be appropriate to assign physical meaning to increased estimation variance for the non-LOS sites.

### 6.8 Conclusions

This chapter outlined the results of a field demonstration of the UAX receiver system described in Chapter 5 using the principles developed in Chapters 2–4. The main purpose was to verify the proper operation of the system, and to identify its shortcomings in practical operation. As a side result, this study provides preliminary findings on the feasibility of transmitter localization by AOA estimation, and also on the nature of the angle spectrum in the test area.

Concerning system performance, it was found that the receivers performed as expected. However, it was noted that the element-to-element noise floor variation induced by the minor differences between the receivers turns out to limit the ability of the system to resolve high-rank effects, such as angle spread. This issue was overlooked in the original specifications, but must be considered in future work. For this reason, the field effort focussed on the performance of the low-order discrete AOA estimators ML₁ and ML₂, MUSIC, and Capon.

The field measurements included a variety of situations. For LOS sites, the measurements yielded single AOA estimates with low bias and RMSE, and seemed to be consistent with the presence of a strong LOS component. In the case of building shadowing, it was found that AOA estimation statistics were difficult to interpret,
in one case (Site 12) being simply biased; whereas in another case (Site 25), it was unrelated to the true bearing. However, if one considers the likely scattering mechanisms, these results can be explained. For example, one should expect signals from Site 12 to be scattered by buildings to the south, resulting in the observed bias in the AOA statistics. Furthermore, one should expect that the occlusion of the LOS path for Site 25 would result in high-order, wide-angle multipath becoming the primary means for signals to travel from the mobile to the array. Similarly, the behaviors observed during testing from the Suburban and Complex Scattering sites, while puzzling from a statistical viewpoint, seem to be consistent with the scattering mechanisms predicted by electromagnetic theory.
CHAPTER 7

CONCLUSIONS

This dissertation has described the design of a system for angle spectrum estimation at a UHF mobile radio base station. The system was built and evaluated in field conditions at 460 MHz. The system uses a 7-element "Y"-shaped array of monopoles above a ground plane, shown in Figures 3.9 and 3.10, and a novel array receiver, shown in Figures 5.7 and 5.12, to sample the signals incident on the array. This data is then analyzed using conventional AOA estimators – ML, MUSIC, and the Capon Estimator – assuming a single discrete AOA. The quality of the estimate is assessed using a simple "residual power" metric which provides an indication of how well the discrete AOA estimate describes the data obtained from the array. If warranted, a higher-order ML estimate is then used to identify significant multipaths or cochannel interference.

In Chapters 3–5, various aspects of the system design were considered and optimized for best performance commensurate with cost constraints. For example, the performance of the array in the elevation plane was considered. To mitigate errors due to local scattering, the array was implemented using vertically-polarized monopoles over a ground plane. An electromagnetic analysis of this array showed that the resulting design provides the desired low-elevation response, but also degrades the
gain along the horizon. An improved design using convex rolled edge terminations was analyzed and shown to be significantly better, although not implemented in this study.

Calibration was identified as a major system design issue. Calibration requirements were derived from a simulation study using the array designed in Chapter 3, taking into account mutual coupling computed using the Moment Method. A new approach to calibration was proposed in which the array response (neglecting mutual coupling), the mutual coupling, and the receiver complex gains are computed separately. Because the array designed in Chapter 3 is rigid, its response can be computed in advance. The receiver complex gains are computed "on-the-fly"; i.e., simultaneously with the data collection, using CW signals injected at the antenna terminals. Finally, the mutual coupling is estimated using a novel procedure in which a model of the open-circuit impedance matrix, computed using the Moment Method, is varied using a single parameter to optimize the bias for a small number of AOA estimates corresponding to known transmitter locations. An estimate of orientation error (bias due to an unknown rotation of the array with respect to the experimental coordinate system) is generated as side information. In Chapter 6, it was shown that this technique reduces the total range of biases observed from over 13° for the uncalibrated array to 3° for AOAIs in the calibrated field-of-view, and was able to identify a 3° error in the orientation of the array.

A novel wideband digital array receiver was designed and evaluated for this study, as described in Chapter 5. The receiver specifications are reported in Figure 5.1. The central feature of the array receiver is a subsystem, shown in Figures 5.7 and 5.12, which accepts inputs from two elements of the array, and outputs a digital
representation of the signals received by those two elements in a form which makes them easy to manipulate in post-processing software. The receiver uses a single analog frequency conversion to a 25 MHz or 35 MHz intermediate frequency (IF), which is digitized and further manipulated using a series of multiply-free "\( F_s/4 \)" downconversions and downsampling. Individual narrowband signals are extracted from the 2.5 MHz passband using an iterative estimation-subtraction technique to remove all undesired signals, followed by a spectral shift to zero-IF.

In the field evaluation described in Chapter 6, both LOS and non-LOS cases were considered. The results were summarized for various scenarios in Figures 6.1, 6.2, 6.3, and 6.4. It was shown that transmitter localization to within a few degrees is possible assuming one incident plane wave \((M = 1)\) in most cases. Increasing model order \((M > 1)\) appeared to improve the bias only in one case. A few examples of complex terrain scattering were observed, including evidence of multipath reflections from buildings. The results also suggest that angle spread in this experiment was less than 5° in LOS cases, and less than 20° in non-LOS cases. This is consistent with the findings of other field studies. However, it was also observed that the array receiver generates a non-white spatial noise covariance which is easily mistaken for angle spread; thus these values should be considered to be upper bounds.

As a result of this work, many possibilities have been identified which deserve additional attention. A few are summarized below.

- It was found that range was limited to under 2 km due to limited transmit power (0.5 W), high environmental noise figure (12 dB), and 3 dB loss due to the untreated ground plane edges. The ability to detect and resolve weak multipaths was also limited by this condition. Increasing transmit power to
just 5 W would yield an immediate 10 dB improvement in SNR and, assuming path loss proportional to $R^{-3.8}$, would approximately double the range. This transmit power is possible in the 460 MHz band, but only with a special license. Also, SNR can be improved by an additional 5 dB by terminating the ground plane using the elliptical rolled edge termination proposed in Chapter 3.

- Another useful modification to the experiment would be to transmit on multiple adjacent channels simultaneously. (This appears to be allowed under FCC rules.) For example, if the mobile were to transmit on 5 of the 7 adjacent frequencies simultaneously, this would increase the SNR by a factor of 5 (recall that the total noise bandwidth for this system is 2.5 MHz regardless of the bandwidth of interest). Alternatively, this could yield 5 independent observations at the original SNR. Since all 5 frequencies would fall within 150 kHz, they all would experience approximately the same propagation channel.

- The data sets collected in this study can be analyzed using techniques which are robust to the non-white spatial noise covariance. As mentioned previously, a candidate ML-based technique is given in [27]. Although hardware modifications might be considered to improve the noise imbalance, this is probably not an economical or practical solution.

- Should additional funding for hardware development become available, provision for longer observation time and faster transfer to PC storage should be considered. Recall that the channel coherence time was estimated to be at least 20 ms; therefore the observation time can be increased by a factor of about 20 (from 802 µs) with confidence that the channel remains invariant over the
observation time. The required memory buffers would be 320K samples deep and probably could not be implemented economically using FIFOs. A suitable custom wideband memory buffer based on DSP-controlled static RAM is described in [75], and would support near-real-time transfer to PC RAM as well as a certain amount of on-the-fly signal processing. This would further allow many more trails per unit time, which in turn would simplify the logistics of the field experiment and allow better statistical characterization.

- It would be useful to obtain an independent set of measurements to characterize the scattering mechanisms identified in this study with increased resolution. For example, for each transmit site evaluated using the compact array system, one could simultaneously measure the angle spectrum using an aperture probing technique. At the existing ESL site, approximately 40 wavelengths is available in the East–West direction and about 10 wavelengths is available in the North–South direction for aperture probing. Each such measurement would take far longer than the channel coherence time and would be vulnerable to various other errors. On the other hand, these measurements would be simple to calibrate and would allow the use of the classical Fourier method, which is model-independent. Also, recall that UAX provides an additional receiver that might be used for this purpose. Joint measurements in this fashion would provide a useful independent check of the compact array results.
BIBLIOGRAPHY


[59] W.D. Burnside. EE712 Microwave Optics (Course Notes). The Ohio State University.


